Using Operations Research & Analytics to Increase the Effectiveness of Service Allocation to Families with Infants in Out of Home Care Due to Substance Abuse in the Texas Child Welfare System

A Major Qualifying Project Proposal submitted to the faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of a Bachelors in Science

Authored by:

Claire Barrameda

Mark Calnan

Jillian Clemente

Jackson Conroy

Advised by:

Professor Andrew C. Trapp, Ph.D.

Professor Emily M. Douglas, Ph.D.

Abstract

This project is an extension of a 2018-2019 academic year MQP that conducted an initial investigation on how to improve service allocation in the foster care system. The 2018-2019 MQP team focused on improving service allocation to children and families across the United States who have entered the foster care system due to substance abuse. Our team narrowed the scope of the 2018-2019 project to better understand the impacts of the factors and services had on children and families. We narrowed the scope of the project by improving service allocation to infants from urban areas of Texas who were placed into foster care as a result of parental substance abuse. Through the use of predictive analytics, we determined the impact that services and other factors had on a child’s length of stay in foster care. Based on those findings, we used prescriptive analytics to develop a mechanism that reallocates services with the goal of minimizing the total amount of time children spend in care. Our results demonstrate that there is an opportunity to improve service allocation by examining not only the details within a child's case, but also the environmental factors surrounding the child’s case.

Acknowledgements

We would like to thank everyone who contributed their time and efforts to help make this project successful. First off, we would like to thank Michael Dineen from the Cornell University College of Human Ecology for providing us with the necessary data to complete this project and for the time he took to help us understand the data. Additionally, we would like to thank Dr. Melinda Gushwa from Simmons University School of Social Work for providing her expertise with interpreting our numerical results and providing insights. Furthermore, we would like to thank Siamak Najafi, James Kingsley, and Ermal Toto from Worcester Polytechnic Institute Information Technology Center for their help setting up the server for us to complete our work on.

Finally, we would like to thank our advisors Professor Trapp and Professor Douglas for their guidance, support, and encouragement throughout the year. Our project would not have been a success without their help.

Authorship

Chapter 1: Introduction	
Section 1.1	Jackson Conroy
Section 1.2	Mark Calnan, Jillian Clemente
Section 1.3	Claire Barrameda, Jackson Conroy
Chapter 2: Background	
Section 2.1	Jillian Clemente
Section 2.2	Mark Calnan
Section 2.3	Jackson Conroy
Section 2.4	Claire Barrameda
Chapter 3: Methodology	
Sections 3.1 - 3.2	Claire Barrameda
Section 3.3	Jackson Conroy, Claire Barrameda
Chapter 4: Results	
Section 4.0	Jackson Conroy
Chapter 5: Alternative Objective Functions	
Sections 5.1-5.2	Claire Barrameda
Chapter 6: Total Unimodularity	
Sections 6.1-6.2	Claire Barrameda
Chapter 7: Conclusion	
Section 7.1	Jackson Conroy
Section 7.2	Mark Calnan
Section 7.3	Jillian Clemente
Chapter 8: IE Reflection	
Section 8.1 - 8.2	Mark Calnan
Section 8.3 - 8.4	Jillian Clemente

Table of Contents

Table of Figures

Table of Tables

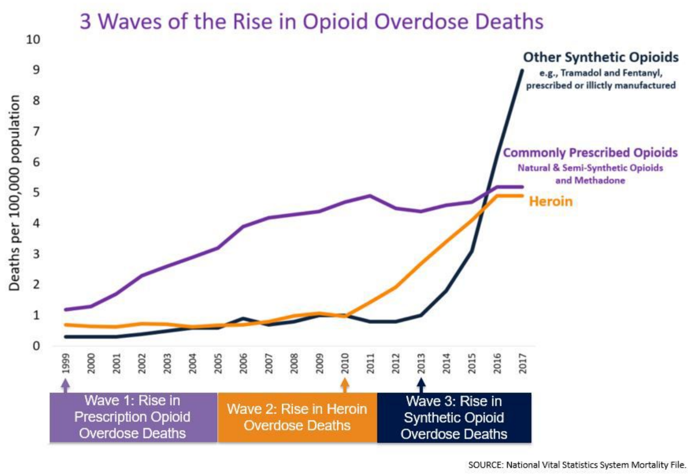
Chapter 1: Introduction

This chapter introduces the scope of our project. We will describe the current state of the opioid epidemic and how it affects the United States foster care system, followed by a brief summary of the past research on this topic and what we plan to do to further the investigation. Key terms related to the foster care system, the opioid crisis, and the mathematics utilized in the study can be found in the glossary in Appendix A.

1.1 National Opioid Epidemic

From 1999 to 2017, over 400,000 people have died from opioid drug overdoses in the United States (Scholl, 2018). This trend has been classified as the national opioid epidemic and its impact has become so severe that it is now considered a public health crisis. The outbreak of the epidemic can be broken into three major waves. The first wave occurred in the late 1990s when pharmaceutical companies began prescribing opioid pain relievers (Kolodny, 2015). The companies assured the public that these pain relievers were not addictive, and this led to widespread use. By the time the public became aware of the addictive nature of the drugs, opioid use, both legal and illegal, was rampant in communities throughout the United States. The second wave began in 2010 with an increase in heroin use, resulting in a higher rate of heroin overdoses. As legislation against prescription opioids began to set in, prescription opioids became harder to obtain, and heroin became a cheaper alternative. Approximately 80% of heroin users misused prescription opioids before turning to heroin (Jones, 2013). The third wave began around 2013 with a significant rise in the use and overdose of synthetic opioids. The most notable synthetic opioid is fentanyl which is illegally manufactured as an alternative to the other forms of opioids. Figure 1 below illustrates the rise of each wave of the opioid epidemic.

Figure 1: Three Waves in the Rise in Opioid Overdose Deaths



The opioid epidemic in the United States has had severe consequences on communities and the nation alike. Opioid related deaths reached an all time high in the United States in 2017 with over 47,000 reported cases of overdoses (Scholl, 2018). As a result of these deaths and widespread addiction, the epidemic has had a major economic burden on the United States. By analyzing a variety of factors such as healthcare, criminal justice, and substance abuse treatment, research found that the annual economic impact of the opioid epidemic in the United States was $78.5 billion (Florence, 2016). Nearly 25% of these economic costs are financed by public sources. As a result, the opioid epidemic has put significant strain on hospitals, governmental programs and services, as well as families across America.

1.2 Effects of Opioid Epidemic on the Foster Care System

The opioid epidemic has profoundly impacted the United States foster care system. Foster care is the temporary, out of home placement of children who are unable to remain in their homes due to conditions that threaten their well-being, such as in cases of child abuse or neglect (Foster Care, n.d.). In 2016, approximately 92,000 children were removed from their homes because at least one parent had a substance abuse issue (Administration for Children and Families, 2018). Historically, periods of drug epidemics have caused spikes in foster care populations. Notable increases have occurred during the crack cocaine epidemic in the 1980s, the methamphetamine crisis in the 1990s, and most recently, the opioid crisis (Radel, 2018).

The United States foster care system faced about 674,000 cases of child maltreatment in 2017 (Administration for Children and Families, 2018). While this number equates to roughly 1% of all the children living in the US, only 20% of these cases resulted in a child entering the foster care system. The remaining cases received support services through their state and county agencies. Children can be removed from their homes and placed into foster care for reasons such as physical abuse from a caregiver, drug and alcohol abuse, behavioral issues caregivers cannot manage, disabilities, death of a parent with no new caregiver in place, or neglect (AACAP, 2018). Neglect is defined as a caretaker's inability to take care of a child's physical, emotional, educational, or medical needs, and is the most common reason children enter the system (Bradley, 2017). Children removed can be placed in one of three options: kinship, with a sponsored family, or in specialized care (AACAP, 2018). Each type fulfils the specific needs the child has when being removed in order to provide the child with the safest temporary solution. When a child first enters the system of foster care, the child welfare system determines a desired case goal. The case goal is what the child, family, and caseworkers strive to achieve in order to establish permanency for the child (Annie E. Casey Foundation, 2019). Federal policy as well as the professional literature on child welfare designate reunification as the preferred permanent solution for children, whenever possible. However, there are other case goal options such as permanent placement with a relative, foster care family, or emancipation. These different ways to exit the system are unique to each case, but final placement is determined to be the safest and least restrictive placement for the child.

While the mortality and overdose rates among opioid abusers is frequently highlighted, one of the most profound impacts of the opioid epidemic is on children of opioid abusers. One study found that the number of children entering the foster care system due to drug abuse has doubled since 2000 (Meinhofer, 2019). The researchers at Weill Cornell Medicine used the Adoption and Foster Care Analysis and Reporting System (AFCARS) to analyze five million instances of children entering the foster care system. In 2017, they found that 1.2 million children entered the system due to parental drug abuse as the primary reason, a 21% increase from 2000 (Neilson, 2019).

The opioid epidemic affects the foster care system in three major ways. First, it leads to an increase in the number of children entering the foster care system (Collier, 2018). The opioid epidemic brings kids into the system in two ways: the death of a parent or the resulting neglect, abuse, or other maltreatment resulting from drug abuse. In the case of death, reunification is not possible, and the state has to explore other avenues of care such as kinship, temporary foster care placement, group home placement, or adoption. In the case of unfitness to care for a child, Child Protective Services (CPS) will remove the child from the home and place him into foster care. The state's foster care system will work with the child and family to develop a plan for reunification with the family or other permanency options. Regardless of how a child enters the foster care system, it is a traumatic experience that has long lasting physical, emotional, and psychological effects on a child (Collier, 2018).

In addition to increases in child placements, the opioid epidemic has profoundly strained the resources and services within each state's foster care system. The United States Department of Health and Human Services found that the current opioid epidemic has strained state's foster care resources more than previous epidemics such as the methamphetamine crisis in the 1990s or the crack cocaine epidemic in the 1980s because of multi-generational drug abuse (Radel, 2018). During the crack cocaine and methamphetamine epidemics, family members and community institutions shielded children from the consequences of parental drug abuse. With the opioid crisis, caseworkers, judges, and other influential foster care personnel have found that drug abuse among an abuser's extended family has become more common. As a result, children are forced into foster care or group home placement, which strains the resources of the state's foster care system. With limited resources within each state, children and families are subjected to poorly funded programs, long waiting lists to receive services, or in some cases, unable to receive the services they need (Radel, 2018).

The opioid epidemic not only affects the foster care system as a whole, but also has a profound impact on the children themselves. Possibly the most vulnerable segment are newborn babies with mothers suffering from opioid addiction. “From 1992 to 2012, the overall proportion of pregnant admissions remained at 4%; however, admissions of pregnant women reporting prescription opioid abuse increased substantially from 2% to 28%” (Martin, 2014). A child who is conceived from a mother who is addicted to opioids often suffers from neonatal opioid withdrawal syndrome (NOWS). These infants are hospitalized for an average of 16 days after birth, compared to the 2.1 days that healthy newborn babies spend. Since 2000, US hospitals have reported a 20% increase in NOWS in live births. The long-term effects of NOWS include developmental issues, academic and behavioral challenges, and increased risk of future addiction (Martin, 2014). Children born with NOWS not only face severe developmental issues, but also they are often born to a mother who struggles with opioid addiction. These mothers are more likely to have poor nutrition, decreased access to the health care system, and increased exposure to violence. As a result, these children are often neglected and born into an unsafe home environment (Conradt, 2018). The opioid crisis has drastically impacted children of all ages (Martin, 2014). In 2014, the National Survey on Drug Use and Health found that of all the adults suffering from opioid addiction, 820,000 of them have at least one child from the ages 0-18 (Feder, 2018). These children are more likely to have a lower socioeconomic status and increased difficulties in academics, social settings, and family functions than kids who live with parents who do not suffer from opioid abuse. Additionally, they are subjected to higher risk of parental abuse or neglect (Martin, 2014). The opioid epidemic has increased child placement in foster care, which has strained state’s foster care resources, leading to continual stress, trauma, neglect, and abuse of children within the system (Collier, 2018).

1.3 Project Overview

The overburdening of the foster care system is not only a major social issue, but an economic, operations, and business issue as well. In 2018, WPI students Diefendorf, Doherty, Tropeano, and Yagoobi applied an analytical approach to resource distribution within the foster care system. Using predictive modelling and optimization, they aimed to determine what foster care services had the most positive impacts for children and how those services should be allocated to minimize the overall length of stay in the system.

1.3.1 Project Methodology and Results from 2018

Diefendorf, Doherty, Tropeano, and Yagoobi analyzed data from the National Child Abuse and Neglect Data System (NCANDS) and the Adoption and Foster Care Analysis and Reporting System (AFCARS) to assess the effectiveness of the foster care services with regards to reducing the amount of time a child spends in the system (Diefendorf et al., 2019). They combined the two data sets, linking records pertaining to the same child, indicated by a Child ID consistent across both data sets. After sorting out incomplete data and inconsistencies, there were roughly 147,000 children and 60 data points for each case for the years 2010-2015 (Diefendorf et al., 2019).

The group constructed linear regression models to determine the impacts of each service offered. After generating all second degree interaction terms, their data set included over 1,700 features. They began by applying LASSO (Least Absolute Shrinkage and Selection Operator), elastic net, and ridge regression to shrink the number of features and using cross validation to determine which shrinkage method was the best. The models were trained on data from 2010 to 2014 and were cross validated using data from 2015. They ultimately found 53 statistically significant factors which included a mix of primary and interaction factors.

Next, they implemented an integer nonlinear optimization model with the objective of minimizing the total number of days spent by children in the system. The decision variables were whether or not a child/family will receive a given service. Services had expected impacts: adding or subtracting a number of days from the child's total time in the system. The scale of the impact is determined by the coefficients from the linear regression model.

The results of the optimization model are displayed in Table 1. The results contain a conservative estimate and most likely estimate. The conservative estimate was developed by using the worst-case scenario for each individual regression coefficient within the optimization model. The average savings per child were calculated by using a value of $70 to estimate the average cost per child per day in foster care (National Council for Adoption, 2011).

Table 1: Results from 2018-2019 Predictive Model

	Conservative Estimate	Most Likely Case
Total Days in Care	6,327,684	5,913,547
Actual Days in Care	6,411,901	6,411,901
Total Days in Care Saved	84,217	498,347
Average Days in Care Saved	5.2	31
Average Savings per Child ($)	$361	$2,138
Total Annual Savings (Substance Abuse Cases)	$36,589,536	$216,518,233
Total Annual Savings (Entire Foster Care)	$98,890,367	$585,185,233

Some shortcomings of this research include the factors used to predict total days in care. While the model had a relatively high RSQ value of 0.707, it was using predictors such as number of removals and length of most recent stay. These factors should not be treated as predictors because they are outcomes rather than predictors or a child's experience in foster care. This analysis also lacked some potentially important explanatory variables, such as family income, urbanicity of the town, and other factors known to influence the outcomes of foster care cases. A major limitation of the regression is the censoring of data: the NCANDS and AFCARS data does not indicate the amount or frequency of a service children receive, but rather is a binary indicator of whether or not the child has ever received that service. Predicting outcomes is inherently more difficult because the data often underestimates the amount of services received.

1.3.2 Overview of Project Extension

This project sought to expand upon the previous work by improving the accuracy of their predictive modeling and implementing alternative integer linear programming problems. We enhanced the predictive analysis by incorporating environmental data and by narrowing the scope of our research to a specific set of children. We chose to narrow the scope of our project to children under the age of two because they represent the most vulnerable subset of the foster care population. These children are under constant supervision of a caretaker and have limited ability to ask for help or to physically protect themselves. Additionally, we chose to focus on children within urban cities in Texas to limit the discrepancies in how data is reported from state to state in NCANDS and AFCARS. Finally, we chose Texas because it was the state with the most data for children under the age of two. Ultimately, the goal of this project is to determine the optimal allocation of services to provide the greatest benefits to children and families in the foster care system. The previous research group aimed to minimize the total number of days a child spends in the system. However, the negative impacts of removal cannot be captured in a single metric such as days spent in care. Research suggests that other factors play a larger role in the child's well-being. To investigate this, we experimented with an alternative goal of minimizing the longest time any one child spends in care and balancing this new goal with the original goal.

Chapter 2: Background

This chapter outlines the background research conducted for our project. We will describe an overview of the foster care system and some of the challenges associated with it, the legal processes from removing a child through the final permanency hearing, the environmental factors correlated to children entering the system, and finally, a literature review of similar studies conducted.

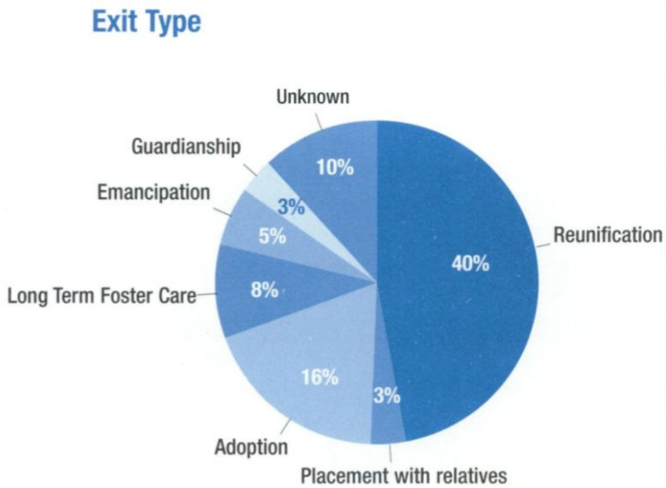
2.1 Overview of the Foster Care System

The foster care and child welfare systems promote the well-being of a child by ensuring his safety, helping establish permanency, and strengthening families through programs and services (How the Child Welfare System Works, 2013). The federal government sets guidelines that all states must follow regardless of their individual regulations and operations of the child welfare system. The foster care system in each state differs in physical structure, but the majority of states follow the same general model. Within each state's Department of Health and Human Services, there is a division of social services that oversees the child welfare system. A child and family's involvement with the foster care system starts with a report to the state's Department of Child Protective Services (CPS). Next, the child will be removed and placed either with a relative in kinship, with a sponsored foster family, or in specialized care. Kinship is often the preferred placement since there is already an established relationship between the child and caregiver, but it is not always a viable option. After kinship care is placement with foster families, which despite the added foreignness can still provide a sense of normalcy for the child compared to other options (Annie E. Casey Foundation, 2019). The final placement type is specialized care such as group homes. This setting is primarily for extreme cases where a child has special needs such as behavioral issues, teenage pregnancy, or disabilities that make it difficult to find suitable kinship or foster caregiver. Group homes are managed by trained staff and are often dedicated to children sharing the same special needs (Group and Residential Care, 2018).

2.1.1 Current State of the Foster Care System

The number of children in the foster care system has steadily increased from 2013 to 2017. The number of children in foster care has increased by about 3% a year, leading to nearly 443,000 children in foster care by 2017 (Children's Bureau, 2018). When a child is placed into the system, a caseworker is assigned to represent the child (Administration for Children and Families, 2018). The caseworker is required to go to court to remove the child and transfer custody to the state. The desired end goal of the foster care system is to provide the child with a safe and permanent solution, preferably through reunification. Children can spend a range of time in the system. Around 50% of the cases last less than a year, 40% for one to three years, and the remaining 10% last longer than three years in care. The longer the child is in the system, the more time they will likely spend moving from home to home, resulting in a multitude of negative experiences. They will be less likely to have a consistent education, form strong relationships, or experience growing up in a healthy family setting. Children moved far from home also face increased difficulty visiting their parents; this will amplify the difficulties of being in foster care. In the end, about half of the children who enter the system are reunified with their original caregivers. Others will be adopted, but unfortunately some children exit the system only when they turn 18 through what is known as emancipation. Children who leave the system through aging out tend to face significantly more financial struggles such as homelessness and joblessness when they are thrown into the adult world from the foster care system (Annie E. Casey Foundation, 2019). Figure 2 shows the percentages of the different exit types for children.

Figure 2: Exit Types for Children Leaving the Foster Care System



Adapted from Meeting the Challenges of Contemporary Foster Care (2001)

2.1.2 Services Provided by the Foster Care System

Children and adults within the foster care system are offered a variety of services to aid in reaching their case goal (Child Welfare Information Gateway, 2013). For cases deemed no risk to moderate risk, families are offered some services within the community and some from the foster care system. Services range from counseling, therapy, and support groups for the lower risk cases to voluntary in-home services of the same nature. Any case classified as a high risk case will result in the child being removed so that both he and the parents can receive the necessary services they need. These services could include child care, parent education, and higher intensity counseling for the most severe cases. For the children who age out of the system, the state responsible for their case will provide services to help with the transition from foster child to independent adult. This may include financial services such as Medicaid coverage, housing, and tuition to a state institution or training and other support groups for how to be independent (FindLaw's team, 2019). Table 2 lists some of the potential services and a description of what they are.

Table 2: Available Foster Care Services and Description

Service	Recipient	Description
Adoption Services	Child	Services provided to assist in the adoption of the child.
Case Management Services	Child	Services aimed to deter, reduce, or eliminate dependence on substances.
Counseling Services	Child	Provides the family additional guidance during tough times.
Daycare Services	Child	Services aimed to deter, reduce, or eliminate dependence on substances.
Education Services	Family	Services aimed to deter, reduce, or eliminate dependence on substances.
Employment Services	Family	Services aimed to deter, reduce, or eliminate dependence on substances.
Family Planning Services	Family	Services aimed to deter, reduce, or eliminate dependence on substances.
Family Preservation Services	Family	Provides the family services to prevent an out-of-home placement.
Family Support Services	Family	Community based groups designed to alleviate stress and promote parental competencies and behaviors that will increase the ability of families to successfully nurture their children.
Health and Home Health Services	Family	Community based groups where parents can collaborate with others experiencing the same events.
Home Based Services	Family	Community based groups where parents can collaborate with others experiencing the same events.
Housing Services	Family	Services or activities designed to assist individuals or families in locating, obtaining or retaining suitable housing.
Information and Referral Services	Family	Community based groups where parents can collaborate with others experiencing the same

		events.
Legal Services	Family	Community based groups where parents can collaborate with others experiencing the same events.
Mental Health Services	Child	Services to help overcome issues involving emotional disturbances or maladaptive behaviors.
Other Services	Child	Community based groups where parents can collaborate with others experiencing the same events.
Pregnancy and Parenting Services	Family	Community based groups where parents can collaborate with others experiencing the same events.
Respite Services	Family	Provides temporary care of the child to alleviate stress on the caretaker.
Special Services Disabled	Child	Community based groups where parents can collaborate with others experiencing the same events.
Special Services Juvenile Delinquent	Child	Community based groups where parents can collaborate with others experiencing the same events.
Substance Abuse Services	Family	Services aimed to deter, reduce, or eliminate dependence on substances.
Transitional Living Services	Child	Community based groups where parents can collaborate with others experiencing the same events.
Transportation Services	Family	Community based groups where parents can collaborate with others experiencing the same events.

2.1.3 Challenges of the Foster Care System

The foster care system faces many challenges in tending to the children and families it serves. One of these challenges is that the cases are beginning to have more complex needs

(Chipungu & Bent-Goodley, 2004). Children coming into the system after being neglected require different services than those coming in who were abused. Neglect is a general, catch-all category, as neglect can be a symptom of other types of maltreatments that are harder to identify. Since neglect cases cover a range of circumstances, cases of neglect are difficult to assist in and match to useful services. There is often a mismatch of the services offered to families and what they actually need, so cases are not resolved as efficiently as possible. Another challenge is finding and retaining foster parents who are willing to host children. Foster parents are tasked with being physically and emotionally available for a child they house, something that is increasingly more difficult than any full-time job. Parenting is an around-the-clock, seven days a week, year round job. Fostering a child may also require arranging visits with the child's parents or caseworker, as well as other appointments such as court hearings. Foster parents are not the only members of the workforce who are difficult to retain. About 90% of child welfare agencies reported struggling with insufficient staffing for the volume of casework they handle, and only one third of those staff members are trained social workers. It is estimated that the turnover rate of caseworkers is between 30 and 40% nationwide, and the average time someone spends as a caseworker is just two years (Employment and the Child, 2010). The desire to reach a final solution faster is increasingly more apparent within the families, but the resources available to do so are just not enough. As a result, the system as a whole struggles to satisfy the increasing complexity of the cases without having foster parents and caseworkers available to help them (Chipungu & Bent-Goodley, 2004).

2.1.4 Vulnerability in the Foster Care System

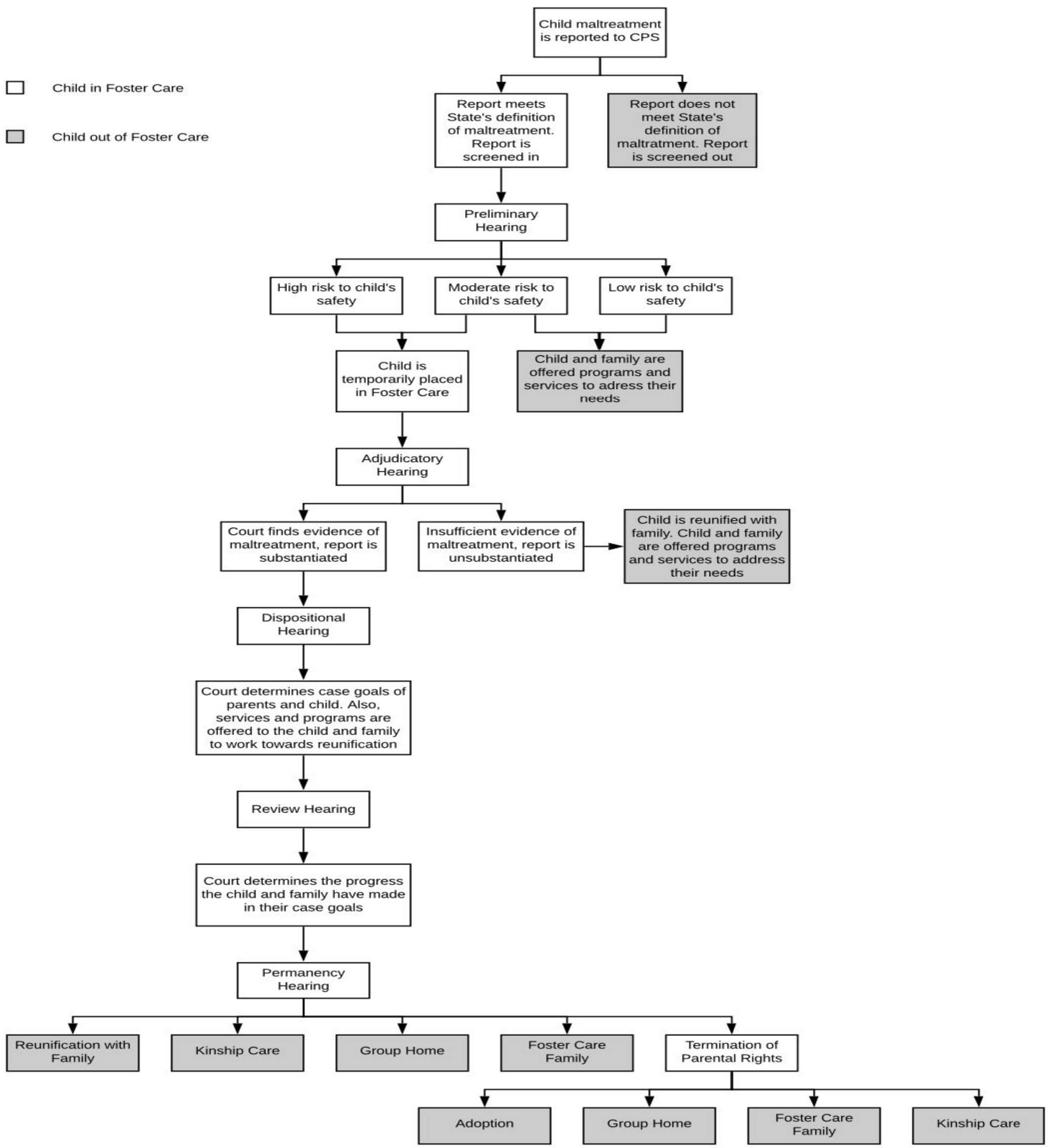
Of the 265,000 children and youth who entered the foster care system in FY 2017, 19% of them were infants under the age of one (Williams & Sepulveda, 2019). Abuse and neglect is more likely to induce long-term conditions in infants than in older children, since early life experiences are crucial in the development of a child. Maltreatment at this stage often leads to altered brain functioning, and its effects can be carried into adulthood. Infants and toddlers are twice as likely to enter the foster care system than older children in general, and two of the most common reasons for this is due to neglect or parental drug abuse. It is estimated that 68% of infants and toddlers enter the system due to neglect, compared to 59% for children over the age of 1. Similarly, parental drug abuse is a risk factor in 46% of cases for children under the age of 1, compared to 30% for children over the age of 1. Infants are much more vulnerable to neglect than older children as they are totally dependent on their caregivers (Child protection intervention, n.d.). They need constant attention and care, but when a parent is struggling with substance abuse, the ability to care for the child greatly decreases, thus preventing the infant from receiving the care it needs (Williams & Sepulveda, 2019). Furthermore, 4% of infants who enter the foster care system have been exposed to drugs themselves, typically through prenatal exposure from a mother's use of drugs. Prenatal exposure can lead to withdrawal symptoms after birth, requiring intensive neonatal care (Dysart, 2018). Overall, infants are an especially vulnerable group of children and face unique challenges not experienced in other age groups.

2.2 The Legal Processes of the Foster Care System

All child welfare systems share the common goal of promoting the safety of children by establishing permanency, and strengthening families through programs and services (How the Child Welfare System Works, 2013). CPS and social workers from the child welfare system conduct an initial investigation to assess the risk to the child's safety. If they determine that the risk to the child's safety is low, CPS will allow the child to remain in the current living situation. The child and family will receive programs and services to address the issues. If CPS and child welfare workers determine that there is a high risk to the child's safety, they will remove the child and place them into temporary foster care. The child and parents will then attend an adjudicatory hearing in which the judge determines whether or not the allegations against the parents are substantiated. If the judge does not find enough evidence, the report is unsubstantiated, the child is returned to his/her parents, and the family is offered services to address its needs. If the judge finds sufficient evidence, the report is substantiated and further court proceedings are required to determine the permanency of the child.

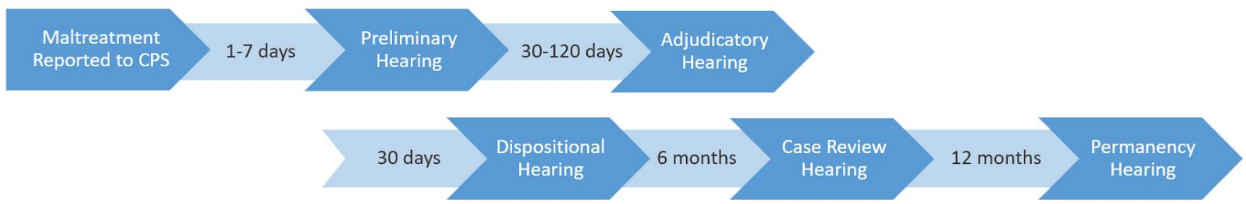
Following the adjudicatory hearing is the dispositional hearing. During this hearing, the court determines the case goals for the family. Additionally, social workers decide the programs and services the child and parents need to assist them in achieving their goals. The child and parents are not summoned to court again until the review hearing. During the review hearing, the judge and social worker assess the progress the child and parents have made towards their case goals. The court may require that the child or parents receive more services or attend more programs. Additionally, if the judge determines that the parents are digressing from their case goals, they may begin the process of the termination of parental rights. The final hearing the child and family attend is the permanency hearing. During this hearing, the court decides the living placement that will best promote the safety and well being of the child until they age out of the system (How the Child Welfare, 2013). Figure 3 provides an overview of the legal process of the foster care system.

Figure 3: Legal Process of the Foster Care System



The legal process of the foster care system contributes to the amount of time a child will stay in the foster care system. In an effort to reduce a child's length of stay, the Children's Bureau, the federal governing body of the child welfare system, has implemented federal regulations that require all states to hold a permanency hearing within 12 months of the removal of a child. The maximum time between the hearings prior to the permanency hearing varies greatly from state to state. Figure 4 exhibits the time it takes for a case to reach a permanency hearing.

Figure 4: Duration of Court Proceedings



Because states follow different standards for the time between court hearings, and because children may be in foster care for the duration of the legal process, we can expect there may be differences between states in the number of days a child spends in care. Furthermore, caseworkers are required to complete the process, as they execute investigations, determine case goals, and help match families to services. The process could be slowed down due to a heavy caseload that strains both the CPS workforce and the legal system.

Summary statistics for states can be found in the annual Child Maltreatment report published by the U.S. Department of Human Health and Services. They include the average response time, CPS workforce size, and CPS caseload which are potential indicators of the state's processing capacity and speed. Table 3 provides descriptions of these metrics.

Table 3: State Summary Metrics

Metric	Description
Average Response Time	The time from the CPS's receipt of a referral to the initial face-to-face contact with the alleged victim
CPS Workforce Size	The number of full-time equivalent workers operating in a state
CPS Caseload	The national number of reports per worker: the total of completed reports for the reporting states, divided by the total number of investigation and alternative response workers

2.3 Environmental Factors Influencing Foster Care Trends

Extensive research has shown that community, county, and state level indicators of risk factors are often associated with child maltreatment rates (McGuinness, 2007). These factors are largely outside the control of the children and their families, but still play an important role in the likelihood of children entering the foster care system. This section will explore the effect of environmental factors such as urbanicity, poverty rate, collectivist versus individualist cultures, and racial-ethnic diversity in the United States foster care system.

2.3.1 Urbanicity

Research has suggested that urbanicity influences the risk of children entering or successfully exiting the foster care system. In a study conducted by the Chapin Hall Center for Children, urbanicity was a major indicator in how adolescents left the foster care system. The study examined exit outcomes for 1.3 million children in 12 different states from urban communities (Hislop, 2002). They found that children from urban communities had higher adoption rates, higher runaway rates, and lower permanency rates, as compared to children from rural communities. This research was backed by a related study that examined nearly three million children that entered the system in 21 different states from 2009-2015 (Wulcyzn, 2017). They found that children in large urban counties achieved permanency at lower rates, reached the age of majority in the system more frequently, and had higher runaway rates, as compared to children from suburban or rural communities. Explanations for why these differences occur between different levels of urbanicity vary. One theory is that higher placement rates in urban areas occur because delinquency culture is more accepted and social disorganization is more common (Hislop, 2002). Additionally, different levels of urbanicity vary in terms of foster care service accessibility and availability, social connections, and poverty rates (Carlson, 2009). Together, these factors create different risk levels and outcomes for children from different urban settings.

It is important to note that previous research has typically classified counties into three urban levels: primary-urban, secondary-urban, and non-urban/rural. There may be opportunities to expand this classification using the National Center for Health Statistics (NCHS) Urban Rural Classification Scheme for counties, which has classified each county in the United States (Ingram, 2014). The classification breaks counties first into either metropolitan or non-metropolitan. Metropolitan counties are then broken into Large Central Metros, Large Fringe Metros, Medium Metros, and Small Metros. Non-metropolitan counties are broken down into either Micropolitan or non-core metros.

2.3.2 Poverty Rate

One of the most widely researched environmental factors is the effect of poverty and economic disadvantage on the foster care system. Unsurprisingly, child maltreatment and foster care entry rates are significantly higher among children living in poverty (McGuinness, 2007). It is estimated that half of the children in the foster care system live in households whose income falls below 50% of the poverty level (Barth, 2006). While individual economic status plays a role, it is also important to look at how poverty rates within communities affect children's likelihood of entering the foster care system. Another study researched the interactive effects of individual and community poverty on child maltreatment rates from 946 families in Ohio. They found that children who lived in higher poverty neighborhoods had higher rates of maltreatment with respect to their individual economic status. Additionally, they found that there does not appear to be a compounding effect of being poor and living in a poor neighborhood, and that living in a non-poor neighborhood does not protect children from maltreatment (Maguire-Jack & Font, 2017). While conducted on a relatively small sample, this research suggests that the impact of living in a community with high poverty rates, not just individual economic status, plays a role in the likelihood of children entering the foster care system.

2.3.3 Collectivism and Individualism

Collectivist and individualist cultures may impact a child's risk of entering the foster care system. Collectivism is a measure of how close-knit a community is. Members of a community with a high collectivist culture would prioritize the well-being of the entire community and look out for one another. Conversely, individualism is a measure of how independent and self-reliant members of a community are. Members of a community with a high individualistic culture would often keep to themselves and prioritize their own needs over others. To classify these ideologies in the United States, Vandello and Cohen (1999) developed a ranking system that scored each state from 1-100 with a high score indicating more collectivist states and a low score indicating more individualistic states. Using this ranking system, researchers were able to use predictive analytics to explain what factors impact foster care entry rates (Macgill, 2015). Using data from ten different sources on a national level, the study found that collectivism was the single most efficient predictor of foster care entry rates among states. States with low levels of collectivism and high levels of individualism had higher average foster care entry rates than states with high levels of collectivism . These findings are explained because collectivist communities put an emphasis on social support and interdependence, and would be more likely to help prevent children in their communities from entering the foster care system (Macgill, 2015).

2.3.4 Racial-Ethnic Diversity

Another environmental factor that can impact a child's risk for entering the child welfare system is the racial-ethnic diversity of their community. Racial-ethnic diversity is a measure of the homogeneity of a community. Research was conducted to measure how racial-ethnic diversity impacts the likelihood of child welfare involvement (Klein, 2014). The study used data from Los Angeles County, California, and examined how neighborhood diversity affects maltreatment referral rates for black, hispanic, and white children. Even after accounting for different control factors, they found that black, hispanic, and white children in communities with high racial-ethnic diversity were significantly more likely to be reported to child protective services and enter the foster care system (Klien, 2014). Klien and Merrit theorize that these results occurred because with greater racial-ethnic diversity, cultural differences and racism are more prevalent. These differences decrease social cohesion among neighbors and, in turn, decrease the community's capacity to enforce norms regarding acceptable parenting (Klien, 2014). These findings are similar to those of Russell and Macgill (2015) regarding collectivism and individualism. Higher racial-ethnic diversity tends to result in more individualistic communities which leads to higher likelihood of children entering the foster care system.

2.4 Literature Review of Operations Research in Foster Care

There is a multitude of data driven analytics of the foster care system, however the majority is fairly basic statistical analysis designed to measure the importance of a single factor. Most commonly, research in foster care aims to determine the impact of a single factor on a particular outcome measurement, such as how a specific service affects reunification rates. Furthermore, to limit the variance due to environmental factors, research typically focuses on a narrow demographic, such as a small subset of families in the same state, with the same case goals, or receiving the same service. Many researchers aim to predict the outcome of a child, whether it be the risk of entering the system, the chances of reunification, or the time expected for reunification. Our research intends to estimate the effects of a wide range of environmental factors and Foster Care services on all cases related to substance abuse. Despite the difference in scale, many techniques are transferable. This section will outline past research and relevant methodologies used to measure impacts on reunification and removal rates, specifically research related to substance abuse cases.

2.4.1 Use of Data Analytics to Predict Foster Care Outcomes

Oftentimes, researchers want to predict the significance of a single factor, such as drug use in a child's home county or whether or not a family received a service. For example, in 2019, Quast, Bright, and Delcher examined the correlation between levels of opioid prescriptions in a region and the rate of child removals per 1000 children. They pulled in data from California's Prescription Drug Monitoring Program, which allowed them to calculate the percentage of people with unreasonably high opioid prescriptions of over 90 morphine milligram equivalent units (MME). They combined this data with the NCANDS and AFCARS data on child removals and limited their sample set to children in California who were removed due to either neglect or parental drug use. All data was aggregated to county level (Quast et al, 2019). While the study did not find significant correlations between opioid availability and child removal rates, their findings were consistent with other research which found positive relations between other measures of opioid use.

Another study from 2012 by Brook, McDonald, and Yan examined the impacts of the Strengthening Families Program (SFP) offered to families with parents who struggled with drug or alcohol abuse. SFP was launched in Kansas and is currently offered statewide. It is a 14 week program that focuses on skills training, child development, and family training, but only one week of the program focuses directly on drug abuse. The program is aimed to aid in cases in which the family goal is reunification and the child had been placed in out-of-home care (Brook et al., 2012). Brook, McDonald, and Yan compared a group of 214 SFP participants to 423 non-participants whose cases also included substance abuse by parents and were seeking reunification, but were not referred to SFP. They found that at the start of the program, the reunification rates between the two groups did not differ. However, after a year of participation in the program, the rates began to diverge and participants reached reunification faster than non-participants. These metrics were determined through survival analysis of the two groups, in other words, calculating the percent of families in each group that reached reunification after 90, 180, 360, 720, and 1080 days after removal.

A similar study aimed to determine the time a child spends in out-of-home care before leaving the foster care system permanently (Shaw, 2010). This study conducted a survival analysis utilizing 15 years of NCANDS and AFCARS data. Shaw implemented a two-phase model to eliminate the impacts of censoring in the data. He determined that standard survival models over estimated the amount of time required for a child to reach permanency. His findings were in keeping with previous work, and most estimates were found to be statistically significant.

2.4.2 Data Used in Foster Care Analysis

The most commonly used data sets are National Child Abuse and Neglect Data System (NCANDS) and the Adoption and Foster Care Analysis and Reporting System (AFCARS). NCANDS and AFCARS are voluntary data collection systems gathered from all 50 states to examine trends in child abuse and negelct across the country. The data sets are managed by Cornell University. NCANDS was established in 1988 and collects data voluntarily state by state. This data includes information from all records of reported neglect, abuse, or maltreatment. It contains demographic information on the child and parents, types of maltreatment reported, child risk factors such as disabilities or behavioral issues, parental risk factors such as domestic violence and substance abuse, and services provided such as pregnancy and parenting services, legal services, or housing services. AFCARS contains case-level data, some of which is repeated from the NCANDS data set. It includes more generalized information on a child's case such as reason for removal, number of removals, days spent in care, current case goal, and how the child was discharged from foster care if applicable.

Some limitations of the joint data set include the aggregation scheme used in record keeping. Services in NCANDS are simply marked '1' if the child/family has received the service and '0' if not. There is no further indicator of the frequency, quantity, or timing of the service. This form of data censoring is a hindrance to more accurate modeling, but is not biased since it applies to all cases without exception. Another limitation is the masking of values to protect the identities of children. For each data set, certain values such as the county identification code are masked for cases in counties with fewer than 1,000 cases. This type of censoring is informed, since it is dependent on the volume of cases from a county, which could be used as a predictive factor. One approach to handling this data is combining all masked cases into a single pseudo county, as done by Quast, Bright, and Delcher (2019).

2.4.3 Accounting for Environmental Factors

When assessing the effects of a single factor, it is necessary to still include other environmental factors that may contribute to the outcome. For example, when determining the effects of opioid prescriptions on the rate of child removals in a county, Quast, Bright, and Delcher introduced other factors such as average weekly pay of residents in the county, level of urbanicity of the county, and the number of jail bookings per 100 residents. These are factors previously found to impact the rate of child removals and will account for some of the variation from county to county. They then used linear regression to predict the rate of removals using the opioid prescription metrics. To further account for urbanicity as an environmental factor, they fit three separate linear models for the three levels of urbanicity (urban, suburban, and rural). This is similar to introducing interaction factors with urbanicity, since it allows for different coefficients of other factors depending on the urbanicity. Ultimately the study found the rate of overprescription had a positive, but statistically insignificant impact on the rate of child removals.

In their assessment of the Strengthening Families Program, Brook, McDonald, and Yan (2012) took a different approach to accounting for environmental factors. First, they narrowed their study to only include children who were discharged from the system due to emancipation or children who were reunified before SFP was initiated. Next, they matched the cases of participants to their non-participant counterparts with the same time in foster care, child gender, child age, and ethnicity. They determined there were no significant differences in the characteristics of the two groups, meaning any significant differences between the reunification rates of the two groups should be attributed to participation in SFP.

Researchers commonly use survival rates when attempting to analyze exit rates from the foster care system, as was done by Brook, McDonald, and Yan (2012). However, survival analysis assumes there is no informed censoring. Censoring occurs when a value is partially known, such as "age is over 65" or "subject made no more than $15 per hour". Informed censoring occurs when data is censored for reasons which could relate to the outcome being analyzed (Shaw, 2010).

A standard survival analysis would not take into account that the goals of some cases were not reunification, but rather emancipation, adoption, or other placements. Instead, a basic survival model assumes that all case goals are reunification. Moreover, different case goals are the result of different circumstances, meaning that there may be significant characteristic differences between the populations of cases aiming for reunification and those not. To account for this, Shaw created a two phase model in which the impact of informed censoring is removed using weights on outcomes before survival analysis is conducted. This is a time intensive way of further including environmental factors as predictors (Shaw, 2010).

2.4.4 Alternative Metrics of Factors

Another interesting technique is the use of alternative metrics for a given factor. The Quast, Bright, and Delcher study implemented several different measures of opioid prescription rates in order to determine which, if any, were predictors of removal rates. In one variety of the model, they shifted the overprescription rates back a quarter, to assess if overprescription rates had a lagging effect on child removal rates. In other words, they predicted if the over prescription rates from one quarter are correlated with child removal rates of the following quarter. In other varieties, they used the overall prescription rates which were the percentage of people who were prescribed opioids, prescription rates among females only, rate of multiple prescriptions, and rate of overlapping opioid and benzodiazepine prescriptions. All were found to be statistically insignificant. This is confirmation that overprescription rate is most likely not a factor in child removal rates for a county, since none of the measures were statistically significant.

This is a technique which could be implemented in our analysis, for both the predictive factors and dependent variables. For environmental factors, we will be including measures of urbanicity, economic state, and prevalence of collectivism. There are many ways of measuring these factors, such as different indices for urbanicity (Ingram, 2014). For measuring the economic state of a county, we can use the average salary, unemployment rates, or even the minimum wage of a county. Using shrinkage methods, we can determine which of these measures is the best predictor. Moreover, we can utilize different dependent variables. In place of predicting the total number of days spent in foster care, we can attempt to predict the number of removals a child will undergo, or the number of living placements they will go through before reaching permanency.

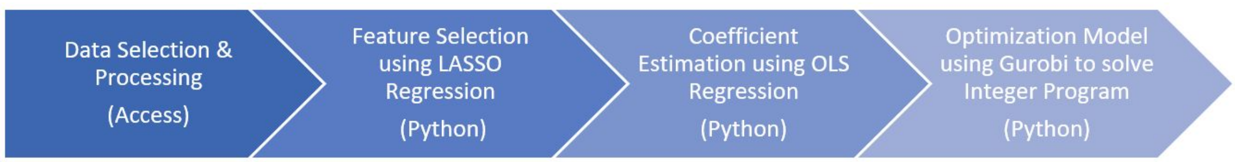
2.4.5 Shortcomings of Literature Review

Literature has shown that statistical analysis can be used to predict the importance of environmental factors in determining a child’s experience with the foster care system. Some research has been done on the effects of services, such as the Strengthening Families Program in Kansas. However, these studies are limited in that they tend to investigate a small set of factors or a single factor at a time. Furthermore, these studies do not consider the foster care system as a whole, which leads to limitations in the recommendations yielded from their work. While past research has examined the impacts of services provided by the foster care system, there are no recommendations on how to distribute these services. Given the increasing number of children in the foster care system and growing needs for services, it is important to distribute services as efficiently as possible.

Chapter 3: Methodology

This project consisted of two phases: running linear regression to estimate the impacts of services and running an integer linear program to optimize the allocation of the services. The first phase included all data compilation, processing, and cleaning, followed by feature selection and linear regression. The second phase included the development of the algebraic model and the implementation in Gurobi.

Diagram of Methodology Breakdown:



3.1 Data Selection

The core data sets used were from the National Child Abuse and Neglect Data System (NCANDS) and the Adoption and Foster Care Analysis and Reporting System (AFCARS). NCANDS captures all children who appeared in a child abuse or neglect report and contains data on the risk factors pertaining to each child or parent, as well as any services which were given to that family. AFCARS only contains children who entered into the foster care system and it holds information such as a child's length of stay, permanency goal, and removal reason. A full list of variables contained in each data set can be found in the codebooks cited in the references.

Our primary goal is to estimate the impacts of services on a child's length of stay to see how many days they add or subtract. Naturally, whether or not a child received these services were included as factors in our predictive model. It is important to note that NCANDS data only reflects whether or not a child received a service, not the quantity or frequency with which they received it. Therefore, service columns have binary entries, '1' if the child received the service, '0' if not. Any blanks were filled in as 0, assuming that a child had not received a service. In last year's analysis, the services offered were broken into two categories: child level services and family level services. Child level services are those directed at aiding one specific child in a family. For example, Daycare Services or Mental Health Services which are tailored to the child receiving it. Family services are directed towards the caregiver or the family as a whole, and thus benefits the entire family. For example, Employment Services is considered a family level service since it is aimed at assisting caregivers in finding work, but ultimately can benefit the family as a whole by improving financial stability. Table 4 below indicates which services are categorized as child and family level.

Table 4: Categorization of Services

Category	Services
Child Level	Adoption, Case Management, Counseling, Daycare, Mental Health, Other Services, Special Services Disabled, and Transitional Living Services
Family Level	Education, Employment, Family Planning, Family Preservation, Family Support, Health and Home Health, Home Based, Housing, Informational and Referral, Legal, Pregnancy and Parenting, Respite, Substance Abuse, and Transportation Services

It's important to note that, in reality, the distinction between child level and family level services is not black and white. Some services can be designed for either individual children or the family as a whole, such as Educational Services which is defined as "services provided to the victim and or family to improve knowledge or daily living skills and to enhance cultural opportunities". In the case of our subset, Educational Services will primarily be directed toward the caregivers, but for other population subsets, this service could vary from child to family level on a case to case basis. Moreover, some services that are directed toward individual children can result in benefits for the caretaker and family at large. For example, if a child with behavioral issues was provided with daycare services, the parent would be receiving temporary relief, but so would any siblings. This relief for the parent could, for example, allow them more time to work, improving the financial stability of the family and mental wellness of the parents, which again benefits the entire family. Due to the customization of services and the interconnectedness of families, it is less than ideal to classify services as strictly child or family level. However, for simplicity of our model and analysis, we only considered these two classifications.

As research has shown, many environmental factors are found to be predictors of a child's ability to reach reunification, risk of entering the system, or other aspects of their experience in the foster care system. To account for the impacts of environmental factors, we incorporated risk factors from both NCANDS and AFCARS such as caretaker conditions, reason for removal, child conditions, and permanency discharge type. Moreover, we joined in additional environmental factors pertaining to the county where a child's case is handled. We joined in county level data on crime rates, unemployment, poverty, median household income, and racial ethnic diversity. This data was found through the U.S. Census Bureau and the Bureau of Labor Statistics and joined into the core data set using the Federal Information Processing Standard (FIPS) codes associated with each county. The combined data sets contained 137 unique columns. The full list of factors and their sources can be found in Appendix B. The data from NCANDS and AFCARS was selected and joined in MS Access using the SQL script in Appendix C.

Last year's analysis examined any children who entered the system between 2010 and 2016 and where substance abuse was a present risk factor. This data set included 147,000 children. For our analysis, we only used a subset of this data: cases from urban counties in Texas where the child was under 2 years of age and has reached permanent discharge from foster care. These filters were chosen with the goal of increasing consistency within the data set. Ultimately, our data set contained 3,173 cases.

Our first restriction on the data set was that all children must have been discharged into a permanent placement. This eliminates cases where children are still in the foster care system and are still receiving services and spending more time in care. The goal in this is to eliminate cases where the length of stay does not accurately reflect the total amount of time a child will spend in care before leaving the system.

Next, we narrowed our research to cases in Texas. The goal in doing this was to eliminate unwanted variance seen in data sets that span multiple states. Different states may follow different guidelines in allocating services, have different budgets, availability, and restrictions, and different interpretations of the service. Furthermore, states report data voluntarily and can choose to only report subsets of their data. These variations between states make predicting off a broad data set more difficult, so we narrowed our analysis to one state. Texas was chosen because it had the most cases.

Finally, we restricted our data set to children from urban counties who were under 2 years old. Urbanicity and age are two major factors that play into the time a child is expected to spend in care. For example, families in urban areas may experience more difficulty receiving services since the population is more dense and there is more competition for resources. This can lead to an increased time in care for families that have to wait on services. Moreover, urbanicity and age can influence which services might be appropriate for a case. For example, families with infants are much more likely to benefit from Daycare Services than from Independent and Transitional Living Services, which helps foster children transition to living on their own. Looking at infants in particular narrows down some of the risk factors that may be seen in a child, such as behavioral issues or juvenile court records.

3.2 Feature Selection and Regression

Before reducing the size of our data set, we first generated interaction terms to account for the compounded effects of combining a particular set of risk factors or services together. Our interaction terms included the interaction of environmental factors with other environmental factors, environmental factors with services, and services with other services. This resulted in a data set with 6,427 features. The complete data set was then split into a training and testing data set using random sampling, with 80% of the data in the training set. All feature selection and models were generated using the training set exclusively. The Python scripts for the following LASSO and OLS regression process can be found in Appendix C.

The first measure taken to reduce the width of the data set was to eliminate columns that lacked variance, and therefore would not be practical to compute regressions on. We removed any columns where the variance was under 0.09. When applying the variance reduction strategy to the binary service columns, this means columns cannot contain more than 90% of the same number. Any columns with more than 90% of entries as ‘1’ or 90% of entries as ‘0’ would be eliminated. This brought the data set to 816 features.

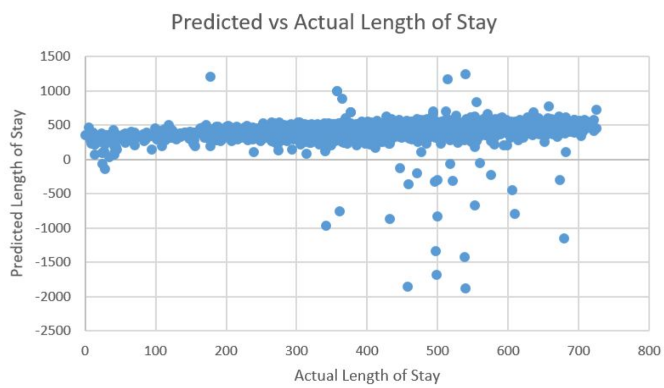
To further reduce the data set, we used least absolute shrinkage and selection operator (LASSO) regression. The LASSO model eliminated more than half the features, reducing the data set to 157 features. These 157 features were then input into an ordinary least squares (OLS) linear regression, with no penalty for additional features. The corresponding training and testing scores of the two models can be found in Table 5 below.

Table 5: Results of the Training and Testing Scores

Model	Features Input	Training RSQ	Testing RSQ
LASSO	816	0.587	0.603
OLS	318	0.522	0.492

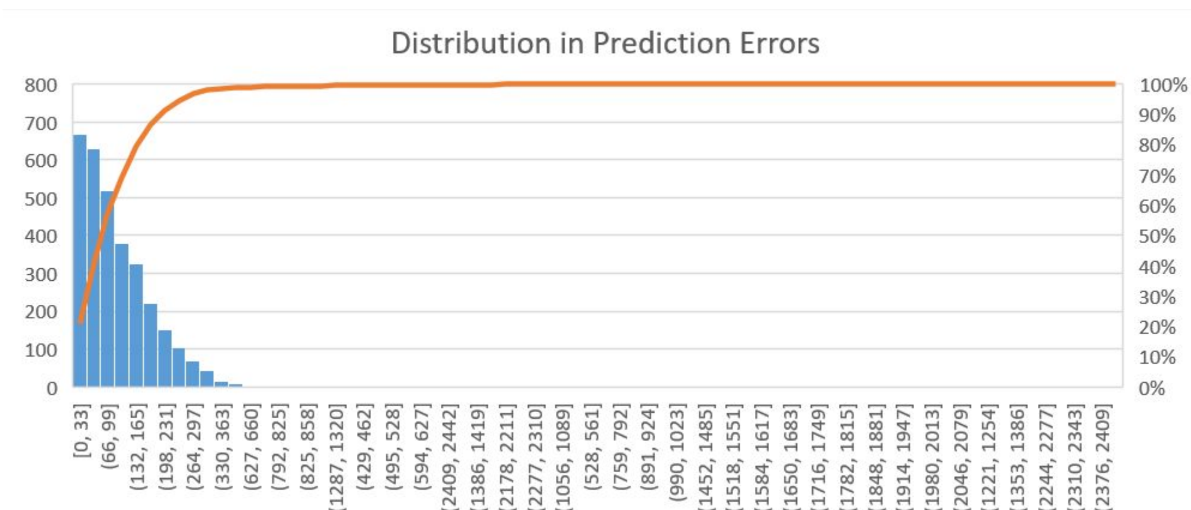
Of the 157 factors input into the OLS regression, 38 were found to have statistically significant coefficients, using the p-value ≤ 0.05 threshold. Originally, we only used these significant 38 factors in the optimization model, but the results were severe underestimates of the actual length of stays. Instead, we decided to use the complete set of 157 features from the LASSO regression in our optimization mode. A closer look into the predicted length of stay for each child versus the actual length of stay for each child reveals the accuracy of the predictions based on this full 157 feature set. It is evident that prediction accuracy decreases as the length of stay increases, with a tendency to underestimate length of stay. The model predicts negative length of stays for 25 cases, most of which have actual lengths of stay over 400 days. Figure 5 highlights the discrepancy between the predicted and actual length of stay for each child.

Figure 5: Predicted versus Actual Length of Stay



Overall the errors on estimations generally follow a folded normal distribution, with over 60% of predictions falling within 99 days of the true value. Figure 6 below shows the distribution of errors in predictions. While some errors are relatively large, these make up only a small portion of the predictions.

Figure 6: Difference between Predicted and Actual Length of Stay



The full results from the LASSO regression with their corresponding coefficients and p-values are presented in Appendix D. Of the 157 factors from LASSO, there were 38 factors which included services in interaction terms and no services appearing by themselves. There were 22 unique services, only seven of which were found to be statistically significant. Table 6 below contains the names and coefficients of the significant service features.

Table 6: Significant Service Factors and Coefficients from OLS Regression

Service Factor	Coefficient	P-Val
Family Planning Services * Poverty Rate 2017	7.019	0.001
Pregnancy and Parenting Services * Unemployment Rate 2017	-11.397	0.003
Housing Services * Monthly FC Payment	-9.045	0.005
Case Management Services * Monthly FC Payment	-9.541	0.007
Home Based Services * Monthly FC Payment	6.472	0.018
Respite Services * Monthly FC Payment	6.670	0.024
Daycare Services * Current Placement Setting	-1.069	0.035

** Note that features that are not originally binary (Monthly FC Payment and Poverty Rate) are normalized before conducting analysis. These columns have an average of 0 and variance of 1.*

It is important to note that not all services are associated with a decrease in days. Any services with positive coefficients indicate that children who receive those services are expected to spend more time in care. One possible explanation for this is that there is a wait time associated with receiving a service. It could be that a family needs Home Based Services, but there is a wait list or other time delays associated with receiving the service. In this case, the coefficient reflects less on the effectiveness of the service and more on the availability or timeliness with which it is delivered. Another possibility is that services with positive coefficients are serving as proxies to reveal the severity of a case. For example, Transportation Services are provided for families that require assistance in reaching other services, medical care, or employment. As a result Transportation Services might be capturing the effects of other hardships the family faces such as financial, medical, or employment hardships. From this holistic perspective, Transportation Services itself is not causing an increased length of stay for children, but rather serving as a proxy to environmental factors which increase length of stay.

Using the regression coefficients from the predictive model, we calculated how many days we would expect each child to spend in care. The predicted total days in care over all children was 1,274,406 days. This figure is about 0.3% lower than the actual total 1,278,050 days children spent in care. While the predicted days in care is slightly lower than the actual days in care, these discrepancies were expected due to the predictive model's RSQ of 0.492. The relatively low RSQ is likely a result of the coarseness of the dataset and the lack of other important features in the data set.

3.3 Optimization Model

The goal of this study is to construct an optimization model that would minimize the total number of days children spend in foster care by determining an optimal reallocation of services related to foster care. The number of days a child spends in care is estimated using the coefficients determined by the final OLS regression model. The decision variables are whether or not a child receives a service. The model sums over all children the estimated days spent in care based on the services provided and the environmental factors involved in each particular case. The objective function is split into two parts: (1) child level services and (2) family level services. This is because child and family level services have different effects on the objective value and are subject to different constraints. In order for a child to reap the benefits of a child level service, the child must expend one unit of that service since these services are tailored to each child. However, when a household receives a family service, every child in it benefits without consuming more units of the service. For example, suppose Employment Services was associated with a 10 day reduction in length of stay. If a single parent with three children receives Employment Services, we would expect each of her children to spend 10 fewer days in care. However, it is not necessary to expend three units of Employment Services since there is only one parent. In other words, family services benefit each child while only consuming at most unit per household.

To construct the model, we first define the sets of factors and cases. Factors are divided into environmental factors and service factors, which is further divided into child level and family level service factors. Other sets include the set of all children and households in the data set, which are represented by Child ID and Report ID numbers respectively.

Table of Variable Sets:

Set	Definition
D	The set of all environmental factors. Indexed by d.
T	The set of all child level services to be allocated. Indexed by t.
F	The set of all family level services to be allocated. Indexed by f.
S	The set of all services. $S = T \cup F$. Indexed by s.
H	The set of all households (i.e. families). Indexed by h.
C	The set of all children. Indexed by c.
C_h	The set of all children in household h.

Any coefficients indicating the impact of a factor or interaction is a fixed number, determined through predictive modeling. The value of an environmental factor for a given child, such as whether or not a child's goal is adoption, is also fixed. These values are retrieved from the NCANDS and AFCARS data. The coefficients from regression and the environmental factor values are the only constants which appear in the objective function. The limit constants l_s appear in the right hand side vector of the integer linear program.

Table of Constants:

Constants	Definition
r_m	The coefficient associated with factor m. $m \epsilon \{S \cup D\}$
r_{mn}	The coefficient associated with interaction of factors m and n. $m, n \epsilon \{S \cup D\}$
b_{cd}	The value of an environmental variable d for child c.
l_s	The number of units of service s to allocate

The model will determine which children receive which services. Since we are creating an integer linear program, we cannot directly include the interaction of two services in the objective function, since the term would be nonlinear. Instead, for the interaction of two services, we introduce an auxiliary variable which indicates whether or not a child received both services. Through linearization constraints, described below, these auxiliary variables are defined to be equal to the product of its two service factors. Resources are distributed on the child level or family level, depending on the service. For any family level services, an additional auxiliary variable is introduced to indicate whether or not the entire family receives a service. If a family is given a service, each child in the family benefits from the service without consuming more of the resource.The Table of Decision Variables summarizes the decision variables in this integer linear program, followed by the objective function and constraints.

Table of Decision Variables:

Decision Variable	Definition
x_{ct}	Indicates whether or not child c receives child level service $t \epsilon T$
x_{hf}	Indicates whether or not the children in household h receive family level service $f \epsilon F$
$x_{cs\bar{s}}$	Indicates whether or not child c receives both service s and $\bar{s} \epsilon S$
x_{csd}	The product of an environmental factor times x_{cs}, i.e., 0 when the child does not receive service s and b_{cd} when the child does receive service s.

Objective Function:

$$minimize\left\{\sum_{c\epsilon C}\left\{\begin{array}{c}\sum_{d\epsilon D} r_d b_{cd} + \\ \sum_{t\epsilon T} r_t x_{ct} \\ \sum_{t\epsilon T}\sum_{\bar{t}\epsilon T} r_{t\bar{t}} x_{ct\bar{t}} + \\ \sum_{d\epsilon D}\sum_{\bar{d}\epsilon D} r_{d\bar{d}} b_{cd} b_{c\bar{d}} + \\ \sum_{t\epsilon T}\sum_{d\epsilon D} r_{td} b_{cd} x_{ctd}\end{array}\right\} + \sum_{h\epsilon H}\left\{\begin{array}{c}\sum_{f\epsilon F}\sum_{c\epsilon C_h} r_f x_{hf} + \\ \sum_{f\epsilon F}\sum_{d\epsilon D}\sum_{c\epsilon C_h} r_{fd} b_{cd} x_{hfd} + \\ \sum_{f\epsilon F}\sum_{t\epsilon T}\sum_{c\epsilon C_h} r_{fd} x_{hft} + \\ \sum_{f\epsilon F}\sum_{d\epsilon D}\sum_{c\epsilon C_h} r_{fd} b_{cd} x_{hfd}\end{array}\right\}\right\}$$

Subject to the following constraints:

$\sum_{c\epsilon C} x_{cs} = l_s$	$\forall\ s\ \epsilon\ S$	(1)
$\sum_{c\epsilon C_h} x_{cf} \leq 1$	$\forall\ f\ \epsilon\ F$	(2)
$x_{cs\bar{s}}\ \leq\ x_{cs}$	$\forall\ s, \bar{s}\ \epsilon\ S$	(3)
$x_{cs\bar{s}}\ \leq x_{c\bar{s}}$	$\forall\ s, \bar{s}\ \epsilon\ S$	(4)
$x_{cs\bar{s}}\ \geq x_{c\bar{s}} + x_{cs} - 1$	$\forall\ s, \bar{s}\ \epsilon\ S$	(5)
$x_{cs}\ \epsilon\ \{0,1\}$	$\forall\ s\ \epsilon\ S$	(6)
$x_{cs\bar{s}}\ \epsilon\ \{0,1\}$	$\forall\ s, \bar{s}\ \epsilon\ S$	(7)

Constraint (1) forces the model to allocate exactly the same number of units for each service type. Constraint (2) limits each family to only receiving one unit of any family service. Constraints (3) through (5) are used to force an interaction term to be the product of its components, i.e. to linearize a nonlinear term. Constraints (6) and (7) force all decision variables to be binary, notation used in the original data set. Appendix E shows how many units of each type of service were available for reallocation in the optimization model. The script for the whole optimization model can be found in Appendix F.

Chapter 4: Results

This chapter covers the results of our study. We will discuss the results of the optimization model and a comparison of the actual days a child spends in care versus the optimized days in care. We also discuss how different constraints applied would lead to a different outcome for the model.

4.1 Optimization Results

Using the results of the predictive model, we ran an optimization model to reallocate the services offered to each child with the goal of minimizing total days spent in care. While no service was found to be a significant predictor on its own, 21 different services appeared in interaction terms with environmental factors. The exact same number of each service allocated within the entire dataset was required to be reallocated in the optimization model. In total, the model's run time was just over an hour long. Of this run time, the vast majority was spent initiating the decision variables, adding constraints, and building the objective function. The time taken to actually solve the integer program was under one second. This indicates that there is ample room for improvement in reducing the run time of our model. A comparison of the actual, predicted, and optimized days in care can be found in Table 7.

Table 7: Results from Predictive and Optimization Model

Measure	Total	Average per Child
Actual Days in Care	1,278,050	403.3
Predicted Days in Care	1,274,406	402.8
Optimized Days in Care	1,229,162	387.5
Days in Care Saved	44,794	15.3

Optimized reallocation of the services was found to reduce the total time in foster care by 44,794 days. For the 3,173 children in the dataset, this figure saves an average of 15.3 days in care per child when compared to the predicted days in care. It is important to note that the total number of each service reallocated in our model is the exact same as the total number of services used by the 3173 children in our dataset. One of the greatest impacts of this constraint is that the model is forced to allocate all services even if they appear in positive regression coefficients. A positive regression coefficient would translate to a positive number of days added in care. As previously discussed in Section 3.5, this may not be an indication that the service is not effective, but rather that the children receiving this service tend to face other challenges which add to time in care. If we remove this constraint and allow the optimization model to allocate less of each service, the new optimized total length in care is 1,204,251 days or an average length of stay of 379.5 days in care. Compared to the predicted days in care, this figure saves an average of 23.2 days in care per child. However, if we choose not to allocate all available resources, we might be decreasing the accuracy of our predictions as well as withholding services that, despite increasing time in care, are necessary for the wellbeing of a family.

Another constraint we explored in the optimization model is requiring the minimum length of stay for each child to be zero days in care. It is possible that our model is reallocating services in a way that would leave some children with a negative length of stay in care. In practice, this outcome would not be realistic or make sense. When we added a constraint that the minimum days in care for each child must be greater than zero, the model solution was now infeasible. This occurs as a result of the inaccuracy of our regression coefficients from the predictive model. While the predictive model was very close to the actual value for the overall days in care, it was less accurate on the individual child level. In total, the regression model predicted that 25 cases would have a negative length of stay. With the constraint of a minimum time in care of zero days, the optimization model was reallocating services to these children to add days in care and make their total non-negative. The result was that the model was unable to allocate enough services to these 25 cases that resulted in positive days in care for each case. As a result, we did not include this constraint in our final optimization model. After optimizing allocation, it is expected that 28 children will have negative length of stay, which is only a handful of thousands of cases.

Chapter 5: Alternative Objectives

The current objective function aims to minimize the total days in care (Goal 1). This objective focuses on creating the greatest reduction in length of stay without regard for which children are benefitting, or suffering, from the reallocation. An alternative objective is to minimize the maximum days any single child spends in care (Goal 2). This means prioritizing the most extreme cases, even at the expense of increasing the length of stay for other children. This chapter will explore this alternative objective function as well as a multi-criteria integer linear program which balances the two objectives.

5.1 Background on Balancing Models

Utilitarianism is the theory that the morally right decision is the one with the highest expected social utility (Harsanyi, 1985). In the context of our project, Goal 1 embodies utilitarianism by aiming to maximize the benefits reaped from a finite set of sources. However, in a social context such as this, Hooker and Williams argue that equity should also be considered when utility leads to large disparities in resource distribution (2012). Equity, or egalitarianism, seeks to reach equality for all people, which can be interpreted in two ways in the context of allocating foster care resources. The first is that children should not be treated differently regardless of their circumstances. For example, preference should not be given to a child with medical conditions even if his or her need for a particular service may be greater. This interpretation of equity is seen in the objective function of Goal 1, which only differentiates by the impacts of a service and does not differentiate by needs. The other interpretation of equity is essentially the opposite: resources should first be given to those most in need, with the goal of making their conditions more equal to those of better off children.

5.2 Minimizing Maximum Days in Care

The goal of this integer linear program is to minimize the maximum days in care, or the minimax. To construct a minimax problem, we can introduce a decision variable z which represents the maximum length of stay out of all the children in the data set. We add the constraint that the length of stay for any child is no more than z and set the objective function to minimize z (Williams, 1999). The algebraic model is formulated as follows, using the same notational conventions as described in Chapter 3.

$$minimize\ (z)$$

Subject to:

$$\left(\sum_{d \epsilon D} r_d b_{cd} + \sum_{s \epsilon S} r_t x_{cs} + \sum_{d \epsilon D} \sum_{\bar{d} \epsilon D} r_{d\bar{d}} b_{cd} x_{cd\bar{d}} + \sum_{d \epsilon D} \sum_{s \epsilon S} r_{ds} x_{cds} + \sum_{s \epsilon S} \sum_{\bar{s} \epsilon S} r_{s\bar{s}} b_{s\bar{s}} x_{cs\bar{s}}\right) \leq z\ \forall\ c\ \epsilon\ C \tag{1}$$

$$\textstyle\sum_{c \epsilon C} x_{cs} = l_s \qquad \forall\ s\ \epsilon\ S \tag{2}$$

$$\textstyle\sum_{c \epsilon C_h} x_{cf} \leq 1 \qquad \forall\ f\ \epsilon\ F \tag{3}$$

$$x_{cs\bar{s}} \leq x_{cs} \qquad \forall\ s, \bar{s}\ \epsilon\ S \tag{4}$$

$$x_{cs\bar{s}} \leq x_{c\bar{s}} \qquad \forall\ s, \bar{s}\ \epsilon\ S \tag{5}$$

$$x_{cs\bar{s}} \geq x_{c\bar{s}} + x_{cs} - 1 \qquad \forall\ s, \bar{s}\ \epsilon\ S \tag{6}$$

$$x_{cs}\ \epsilon\ \{0,1\} \qquad \forall\ s\ \epsilon\ S \tag{7}$$

$$x_{cs\bar{s}}\ \epsilon\ \{0,1\} \qquad \forall\ s, \bar{s}\ \epsilon\ S \tag{8}$$

Constraint (1) defines z as the maximum of all of the length of stays. Constraint (2) and (5) through (9) are identical to those in the algebraic model for Goal 1. Note that a new set C_h'' was introduced in Constraint (3). This set contains exactly one child from each family which acts as representative for the whole family, replacing the need for auxiliary variables. All children in a family are required to receive the same family level services as the representative child c'. Since we are allocating services to either every or no child in a family, this maintains the desired behavior in the objective function that all children in the same family receive the benefits or drawbacks of a family level service. However, additional children in the same family should not consume more units of a family service. To account for this, the number of units given out is only determined by counting how many children from set C_h' received a service. In other words, a family with multiple children can have each child receive a family level service, but only the first child will consume a service and count against the total units available. Overall, the constraints enforce the same behavior as in the previous model from Chapter 3, with the only change being the addition of decision variable z. This is the structure that was followed in the optimization script, which can be found in Appendix G.

Table 8 below compares the minimum total length of stay and the minimum maximum length of stay from before optimization and after optimization with different objective functions.

Table 8: Comparison of Metrics Before and After Optimization

Metric	Average Length of Stay	Maximum Length of Stay
True Value from Data Set	403.3	725
Predicted Value from OLS Model	402.8	1251.3
Value after Minimizing Total Length of Stay	387.5	994.3
Value after Minimizing Maximum Days in Care	402.2	960.6

The true maximum days in care from our data set is actually 725. Our predictive model, before any optimization, estimates that the maximum length of stay is 1,251.3 days and the average is 402.8 days. In reality, this maximum of 1,251 is impossible since our subset only looks at children who reached permanency under the age of two years old, or 730 days. However, due to the inaccuracy of the regression coefficients, some predicted lengths of stay exceed 730 days. After optimization to minimize the maximum days of care, the maximum days in care is 960.6 and the average length of stay is 402.2 days. This is a 290.7 day reduction in the maximum and 0.6 day reduction in the average. These results suggest that resources can be successfully reallocated to more severe cases without hurting the overall average. This concept will be further explored in the next section.

5.3 Balancing Objective Functions

A third integer linear program was designed to consider both objectives at the same time. The goal is to reduce both the average days in care and the maximum days in care simultaneously (Goal 3). In this objective function, we've replaced the total days in care with average days in care, which is the total days in care divided by the number of children. This was done in order to compare lengths of stay on a child level: the length of stay for an average child, versus the length of stay for the most extreme case. Reducing the maximum and the average are potentially conflicting objectives because Goal 1 prioritizes the general whole while Goal 2 prioritizes the extreme cases. Placing more emphasis on Goal 1, with only some consideration for Goal 2 may yield different results than if both goals are considered equally important. To investigate the relationship between these goals, each goal is weighted with constants β and $1-\beta$ respectively, where $\beta \in [0,1]$. Larger alphas place a greater importance on Goal 2, minimizing the maximum days in care. The algebraic model is formulated as follows, using the same variable naming conventions as described in Chapter 3. For any value of $\beta \in [0,1]$,

$$minimize \left(\beta z + (1-\beta)\frac{L}{|C|}\right)$$

$$L \ = total\ days\ in\ care =$$

$$\sum_{c\epsilon C}\left\{\sum_{d\epsilon D} r_d b_{cd} + \sum_{s\epsilon S} r_t x_{cs} + \sum_{d\epsilon D}\sum_{\bar{d}\epsilon D} r_{d\bar{d}} b_{cd} x_{cd\bar{d}} + \sum_{d\epsilon D}\sum_{s\epsilon S} r_{ds} x_{cds} + \sum_{s\epsilon S}\sum_{\bar{s}\epsilon S} r_{s\bar{s}} b_{s\bar{s}} x_{cs\bar{s}}\right\}$$

$$|C| = number\ of\ children$$

Subject to:

$$\left\{\begin{matrix} \sum_{d\epsilon D} r_d b_{cd} + \\ \sum_{t\epsilon T} r_t x_{ct} \\ \sum_{t\epsilon T}\sum_{\bar{t}\epsilon T} r_{t\bar{t}} x_{ct\bar{t}} + \\ \sum_{d\epsilon D}\sum_{\bar{d}\epsilon D} r_{d\bar{d}} b_{cd} x_{cd\bar{d}} + \\ \sum_{t\epsilon T}\sum_{d\epsilon D} r_{td} b_{cd} x_{ctd} \end{matrix}\right\} + \left\{\begin{matrix} \sum_{f\epsilon F}\sum_{c\epsilon C_h} r_f x_{hf} + \\ \sum_{f\epsilon F}\sum_{d\epsilon D}\sum_{c\epsilon C_h} r_{fd} b_{cd} x_{hfd} + \\ \sum_{f\epsilon F}\sum_{t\epsilon T}\sum_{c\epsilon C_h} r_{fd} x_{hft} + \\ \sum_{f\epsilon F}\sum_{d\epsilon D}\sum_{c\epsilon C_h} r_{fd} b_{cd} x_{hfd} \end{matrix}\right\} \leq z \quad \forall\, c\,\epsilon\, C \tag{1}$$

$\sum_{c\epsilon C} x_{cs} = l_s \qquad \forall\, s\,\epsilon\, S$ (2)

$\sum_{c\epsilon C_h{}'} x_{cs} = l_s \qquad \forall\, s\,\epsilon\, F$ (3)

$x_{cf} = x_{c\prime f} \qquad c, c'\,\epsilon\, h, \forall\, f\,\epsilon\, F, \forall\, h\,\epsilon\, H$ (4)

$x_{cs\bar{s}} \leq x_{cs} \qquad \forall\, s, \bar{s}\,\epsilon\, S$ (5)

$x_{cs\bar{s}} \leq x_{c\bar{s}} \qquad \forall\, s, \bar{s}\,\epsilon\, S$ (6)

$x_{cs\bar{s}} \geq x_{c\bar{s}} + x_{cs} - 1 \qquad \forall\, s, \bar{s}\,\epsilon\, S$ (7)

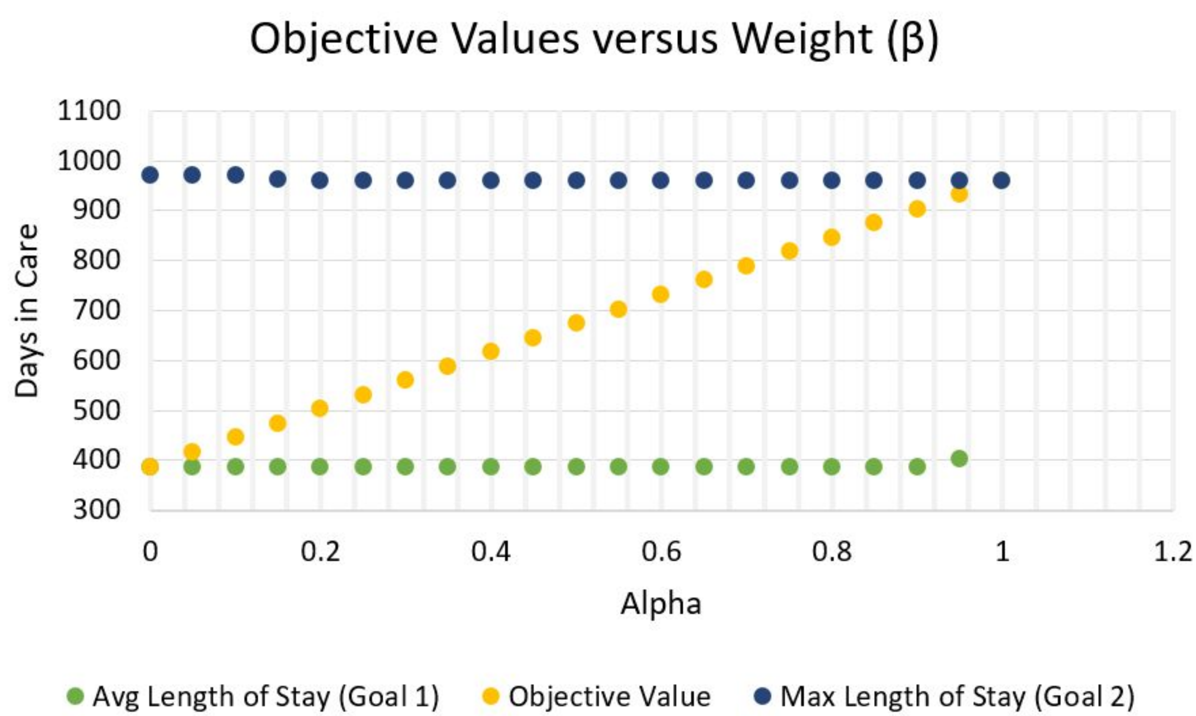
$x_{cs}\,\epsilon\, \{0,1\} \qquad \forall\, s\,\epsilon\, S$ (8)

$x_{cs\bar{s}}\,\epsilon\, \{0,1\} \qquad \forall\, s, \bar{s}\,\epsilon\, S$ (9)

Constraint (1) defines z as the maximum of all the length of stays. Constraint (2) through (9) are identical to those in the algebraic model for Goal 1. The script for this model can be found in Appendix G.

This integer linear program was run with beta values from 0 to 1 in increments of 0.05. Figure 8 below shows the resulting objective function values as well as the values for Goal 1 and Goal 2 separately. It was found that when $\beta \leq 0.10$, placing more weight on average days in care, that average days in care dominated the whole objective function. Note that on Figure 8, the flat line for a Max Length of Stay of 970.4 days spans from $\beta \in [0, 0.10]$. It isn't until more weight ($\beta \geq 0.2$) is placed onto minimizing the Max Length of Stay that we see the model begin to allocate resources towards reducing this metric. For $\beta \geq 0.2$, the model will minimize Max Length of Stay to 960.6, which was found to be the optimal value for this metric. Similarly, the average days in care is minimized to its optimal value of 387.699 when beta is at most 0.10. When $\beta \geq 0.2$ and the model begins considering Max Length of Stay more, the average is increased to 387.705. This slight increase in the average translates to a mere 14 additional days added to the expected total across all children.

Figure 8: Objective Values versus Weight (β)



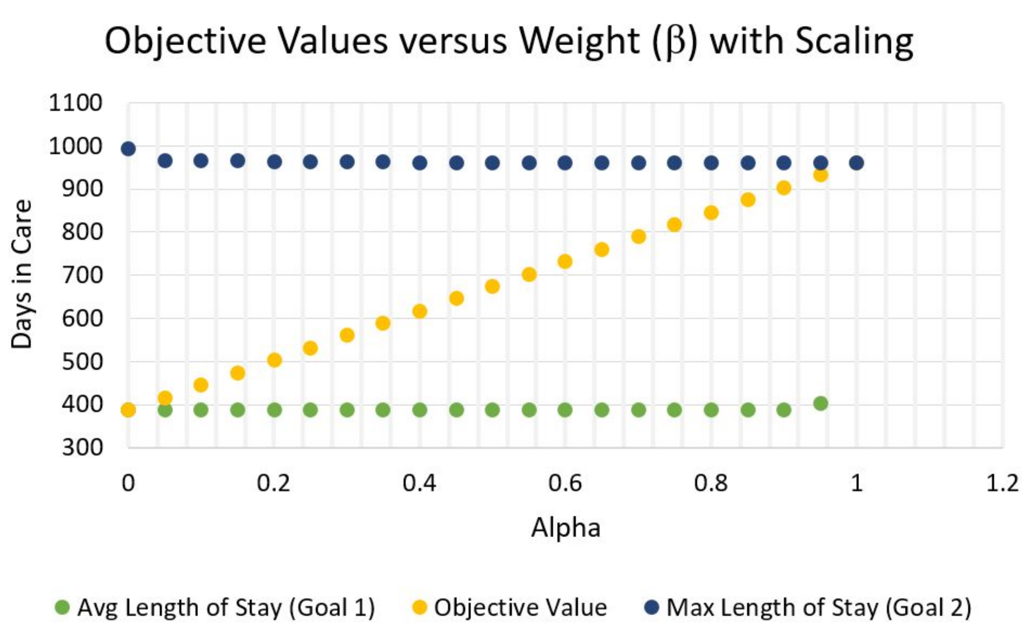
Between β = 0.1 and β = 0.2 are the transitional points where neither Goal 1 nor Goal 2 are optimized, but rather a compromise is reached. When β = 0.15, the max length of stay was reduced to 963.3 while the average length of stay was set to 387.701.

The same analysis was conducted with an additional coefficient to account for the difference in scale between the max days in care and the average. The modified objective function was:

$$minimize\ \left(\beta z * \frac{387.7}{960.6} + (1-\beta)\frac{L}{|C|}\right)$$

The additional constant is the ratio between the optimal values for average days in care and maximum days in care. It reduces the weight of the maximum on the overall objective value, but ultimately yields similar results to the model without the constant. The results can be seen below in Figure 9 which shows objective values by beta. The data in Figure 9 is in Appendix H.

Figure 9: Objective Values versus Weight (β) with Scaling



For most values of beta, the objective values are unchanged. For any β ≥ 0.4, maximum days in care is prioritized and minimized to 960.6. For β < 0.4, the values decrease from 994.3 to 963.3

as beta increases. More values of beta result in compromises between the two objectives, where neither reach their optimal values but are still significantly reduced from their original values.

One way to interpret these results is that these two objectives do not compete with each other significantly, but rather these objectives work in tandem. It is possible to optimize one metric without making major sacrifices on the other. This means that adding in Goal 2 to create a multi-criteria model is acting as a second set of guidelines for distributing services. Using only Goal 1 allows for flexibility on which children receive which services, since the model does not consider which specific children it is reducing the length of stay for. Reducing five days from Child 1 is the same as reducing five days from Child 2. Just as reducing ten days from Child 1 is the same as reducing five days from Child 2 and five days from Child 3. In Goal 1, only the total matters. Adding in Goal 2 simply provides criteria for differentiating between children when deciding which to give services to.

Another explanation for the lack of variance in objective values is that Goal 2 has a fixed objective value. That is, the maximum days in care cannot be reduced below the value 960.6. This could be a result of cases where the environmental factors are associated with large increases to length of stay. It could be that regardless of services, an expected length of stay is abnormally high and the model is incapable of improving upon that number. To correct for this would require improved coefficients from improved analysis or data.

Chapter 6: Total Unimodularity and Integrality

The integer linear program discussed in Chapter 3 is constrained to have all decision variables take on binary values 0 or 1. Another way of defining this constraint is that decision variables take on integer values ranging from 0 to 1. This forces the model to determine the optimal objective value with an integral solution.

In particular cases, integrality does not need to be an added constraint, but rather comes naturally through existing constraints. A linear program is considered integral if, when any optimum exists, at least one has a solution where all decision variables are integers. This occurs when the constraint matrix A is totally unimodular (TU) and the right hand side vector, b, is also integral. One important consequence of integrality is the computational difficulty of solving. Integer programs are generally NP hard, which means they are in a class of problems which are computationally challenging to solve. However, if one can show that the constraint matrix is totally unimodular and the right hand side vector is integral, then it is possible to relax the integer variables to be continuous, and solve a linear program. Linear programs are a simpler class of problems which are significantly easier to solve computationally.

In this chapter, we'll analyze the case in which no interaction terms between two services are included. Since our predictive analysis found this to be the case, this condition reflects the integer linear programs used in this project, as described in Chapter 3 and Chapter 5. We will illustrate that when there are no interactions between two services, the resulting integer linear program can be relaxed to be a linear program.

6.1 Structure of Constraint Matrix

A standard linear program is formulated with a constraint matrix A, a column vector b containing decision variables, and a right hand side vector with constants. The relationship follows the form:

$$Ax \leq b$$

For this analysis, we'll be considering the constraint matrix created following the algebraic model described in Chapter 3. To describe the constraint matrix, we'll first consider a small subset of children and factors, then expand the form to describe a general set of *n* children and *m* factors. Suppose you have a data set with three children c_1, c_2, and c_3 where c_1 and c_2 come from the same family. Suppose we are reallocating two types of child level services, t_1 and t_2, and one family service f_1. Suppose there are l_{t1}, l_{t2}, and l_f units of each service available for allocation, respectively. The right hand side vector would be $[2, 1, 1]^T$ and the vector of decision variables would be:

$$x = \left[x_{c_1t_1}, x_{c_1t_2}, x_{c_1f}, x_{c_2t_1}, x_{c_2t_2}, x_{c_2f}, x_{c_3t_1}, x_{c_3t_2}, x_{c_3f}\right]^T$$

For each decision variable, there will be a corresponding column in the constraint matrix. Each row of the constraint matrix represents a unique constraint, either setting a limit on the number of units of a service to allocate or defining the relationship between two decision variables. Let α represent a row from the constraint matrix with entries indexed by the same subscripts as the decision variables:

$$\alpha = \left[\alpha_{c_1t_1}, \alpha_{c_1t_2}, \alpha_{c_1f}, \alpha_{c_2t_1}, \alpha_{c_2t_2}, \alpha_{c_2f}, \alpha_{c_3t_1}, \alpha_{c_3t_2}, \alpha_{c_3f}\right]$$

For this particular set of three children and three services, there will be four constraints. Three constraints will set the limits on how many units of t_1 and t_2 , and f_1 can be distributed, enforcing Constraint (1) from the algebraic model in section 3.3. A fourth constraint will limit the family with two children to receive at most one unit of family service f, enforcing Constraint (2). This constraint is not necessary for families with only one child, such as the family containing c_3 since that child is already limited to receiving at most one unit of any service. The constraints will be:

$$Ax \leq b$$

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{c_1t_1} \\ x_{c_1t_2} \\ \dots \\ x_{c_3f} \end{bmatrix} \leq \begin{bmatrix} l_{t1} \\ l_{t2} \\ l_f \\ 1 \end{bmatrix} \quad \begin{matrix} Constraint\ (1) \\ Constraint\ (1) \\ Constraint\ (1) \\ Constraint\ (2) \end{matrix}$$

The general form of the constraint matrix is as follows. Suppose there are n children and m services (both child and family). Let m_t be the number of child services and m_f be the number of family services. The decision variables will be:

$$x = \left[\, x_{c_1s_1}, x_{c_1s_2}, \dots, x_{c_1s_m}, x_{c_2s_1}, x_{c_2s_2}, \dots, x_{c_2s_m}, \quad \dots \quad , x_{c_ns_1}, x_{c_ns_2}, \dots, x_{c_ns_m} \right]^T$$

Each row of the constraint matrix A will take the form:

$$\alpha = \left[\, \alpha_{c_1s_1}, \alpha_{c_1s_2}, \dots, \alpha_{c_1s_m}, \alpha_{c_2s_1}, \alpha_{c_2s_2}, \dots, \alpha_{c_2s_m}, \quad \dots \quad , \alpha_{c_ns_1}, \alpha_{c_ns_2}, \dots, \alpha_{c_ns_m} \right]$$

1. For each child service t there exists a row α with entries defined by:

$$\alpha_{c_is} = 1 \quad if\ s = t \quad \forall\, i \in [1, 2, \dots, n]$$

$$\alpha_{c_is} = 0 \quad if\ s \neq t \quad \forall\, i \in [1, 2, \dots, n]$$

2. For each family service f there exists a row α with entries defined by:

$$\alpha_{c_is} = 1 \quad if\ s = f\ and\ c_i \in C_h' \quad \forall\, i \in [1, 2, \dots, n]$$

$$\alpha_{c_is} = 0 \quad else \quad \forall\, i \in [1, 2, \dots, n]$$

3. For each child c and for each family service f there exists a row α with entries:

$$\alpha_{c_i s} = 1 \quad if\ s = f\ and\ c_i = c \qquad \forall\, i \in [1, 2, \ldots, n]$$
$$\alpha_{c_i s} = 0 \quad else \quad \forall\, i \in [1, 2, \ldots, n]$$

6.2 Proof of Total Unimodularity

In a paper from *Linear Inequalities and Related Systems*, Seymour describes special cases of matrices which are totally unimodular. Our constraint matrix has the special property of consecutive-ones, which is a sufficient, but not necessary, condition for a matrix to be totally unimodular. Any matrix A' is said to have the consecutive-ones property if A is a 0-1 matrix in which for every row, the ones appear consecutively (Seymour, 1980). It has also been proven that if a matrix A' is TU, then any matrix resulting from the permutation of its columns is also TU. These two facts can be applied to demonstrate that our constraint matrix A is TU.

First, we begin by showing that our matrix only has binary entries. This is trivial and comes directly from the definitions of our rows which sets each entry to be either 0 or 1. Next, we reorder the columns such that all ones appear consecutively in a column. To do this, we group the columns which correspond with the same service together. Using the example from Section 6.1, the resulting reordering would look like:

$$\alpha = \alpha_{c_1 t_1}, \alpha_{c_2 t_1}, \alpha_{c_3 t_1}, \alpha_{c_1 t_2}, \alpha_{c_2 t_2}, \alpha_{c_3 t_2}, \alpha_{c_1 f}, \alpha_{c_2 f}, \alpha_{c_3 f}$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{c_1 t_1} \\ x_{c_1 t_2} \\ \cdots \\ x_{c_3 f} \end{bmatrix} \leq \begin{bmatrix} l_{t1} \\ l_{t2} \\ l_f \\ 1 \end{bmatrix} \quad \begin{matrix} Constraint\ (1) \\ Constraint\ (1) \\ Constraint\ (1) \\ Constraint\ (2) \end{matrix}$$

Notice how all columns pertaining to service t_1 are together, followed by t_2 and f. In this example, the matrix is already ordered to have consecutive-ones. However, in more complicated sets where children are less ordered, it might be necessary to take one more step. Without disrupting the sorting for services, group decision variables pertaining to children in the same family together. For example, if c_1 and c_3 were siblings, we would order the columns to have rows with order:

$$\alpha = \alpha_{c_1t_1}, \alpha_{c_3t_1}, \alpha_{c_2t_1}, \alpha_{c_1t_2}, \alpha_{c_3t_2}, \alpha_{c_2t_2}, \alpha_{c_1f}, \alpha_{c_3f}, \alpha_{c_2f}$$

Ordering the columns by these rules will result in consecutive-ones across rows regardless of the number of children, families, and services included in the integer program. Since altering the order of columns does not change whether a matrix is TU or not, we've shown that the constraint matrix is TU regardless of the order in which columns appear. Combining this with the fact that all entries in the right hand side vector *b* are integers, we have sufficient conditions for the integer program to be integral. This means the integer program can be relaxed to a linear program with continuous variables and is now simpler to solve.

This proof shows that in the case where there are no interaction terms between two services, the integer program can be relaxed to a linear program since no linearization constraints are required for the nonlinear terms. In our predictive analysis, we found that there were no significant interactions of services, thus no linearization constraints were required. It is possible that the constraint matrix for a model with nonlinear interaction terms is TU. However, the constraint matrix would be more complicated, because introducing linearization constraints results in entries of -1, 0, and 1, as well as rows with more than two non-zero entries. This means the consecutive-ones property would not apply and other theorems for proving TU are not readily applicable.

Chapter 7: Conclusion

The chapter of our report covers the conclusions we have drawn from our study. We discuss our results and what they mean in a broader context beyond data analytics. We also highlight the limitations of our study and recommend some future ideas for studies that can further our research on the foster care system.

7.1 Interpretation of Results

The results of our regression analysis revealed several interesting takeaways. The first major takeaway is that non-service factors play an important role in predicting how long a child spends in care. Of the 157 coefficients from our regression analysis, every coefficient includes a non-service factor in some form (either by themselves or within an interaction term). The five most common non-service factors and the number of coefficients they are represented in can be seen in Table 9.

Table 9: Non-service factors in regression coefficients

Non-service Factor	Number of Instances
Monthly FC Payments	38
Median Household Income	21
Poverty Rate	19
Crime Rate	18
Unemployment Rate	18

The most common non-service factor, Monthly FC Payments, appeared in 24% of all factors. Monthly FC Payments is the amount of money given to a family each month for foster care maintenance. It captures the total amount received from Federal, State, county, municipality, tribal, or private aid organizations (AFCARS, 2019). Children that receive higher Monthly FC

Payments generally need more financial support and are more likely to be impoverished. For this reason, the factor Monthly FC Payments may be a proxy for the poverty level of a child's family. While poverty rates based on the county average is a factor already included in our dataset, Monthly FC Payments may be giving insight into the individual child's family's financial status. For this reason, it makes sense that Monthly FC Payments is such an impactful factor in predicting the length of stay for children. After Monthly FC Payments, the next four most common factors depicted in Table 9 are all environmental factors. In total, 114/157 coefficients or 73% of coefficients from the regression analysis include an environmental factor in some form. This emphasizes the influence of poverty and geography on a child's length of stay in the foster care system.

A second finding from the regression analysis is that no services were found to be significant on their own. Instead, services were only found to be significant in interaction terms with environmental factors. This is important because it suggests that services may not have a uniform benefit to all children. Different children with different risk levels or in different communities may benefit from receiving a service more than other children. For instance, the regression coefficient for the interaction between 'Pregnancy and Parenting Services' and 'Unemployment Rate' is negative. When the optimization model reallocates this service, the greatest benefit will occur when it is allocated to children in communities with higher unemployment rates. This can be interpreted that Pregnancy and Parenting Services are more effective in communities with higher unemployment rates. While all children may benefit from receiving this service, some children are expected to benefit more than others.

7.2 Broader Implications

The days in care saved from reallocating foster care services can have a significant impact on the children, their families, and the foster care system as a whole. On average, our model reduced the average time a child spends in care with respect to the predicted days in care by 15.3 days. This means a period of over two weeks less that each child has to spend in the foster care system, away from his or primary caregiver. Particularly for infants, two weeks less in care can be extremely beneficial. It gives the child the opportunity to be reunited with their caregiver and more time to develop a parent-child relationship that can be vital to a child's development.

In addition to the benefits for families, this reduction in days in care translates to a huge financial savings for the welfare system. It's estimated that in the United States, it costs $70 per day to keep a child in the foster care system (National Council for Adoption, 2011). If this figure is applied to our findings, the average savings per child is $1,071 for just 15.3 days. Applied to the 3,173 children in our data set, the total savings would be approximately $3,400,000 (see Table 10).

Table 10: Total Savings by Days in Care Reduction

	Total	**Average per Child**
Cost of Predicted Days in Care	$89,208,420	$28,196
Cost of Optimized Days in Care	$86,041,340	$27,125
Savings	$3,167,080	$1,071

While the state of Texas does not have exact data available on the average cost per child per day, the figure can still be estimated. In 2016, Texas spent $1,558,371,303 on child welfare services (Child Welfare League of America, 2019). Of that total, $730,329,297 was from federal funds and $828,042,006 was from state and local funds. In 2016, Texas had 30,577 children in foster care (Fostering Success Foundation, 2019). If the total cost of child welfare is divided by the number of children in care, Texas on average spent $50,965 per child in 2016. From the reallocation of services in our optimization model, we were able to reduce the total days in care for infants in the dataset by 3.8% (402.8 days to 387.5 days). If we attempt to apply that value to the entire foster care system in Texas and assume a 3.8% reduction in the number of days in care translates to a 1% reduction in costs, the average savings would be \$509.65 per child. If that savings is applied to 30,577 children in the Texas foster care system, the total savings in the state of Texas would be \$15,583,568.

7.3 Limitations

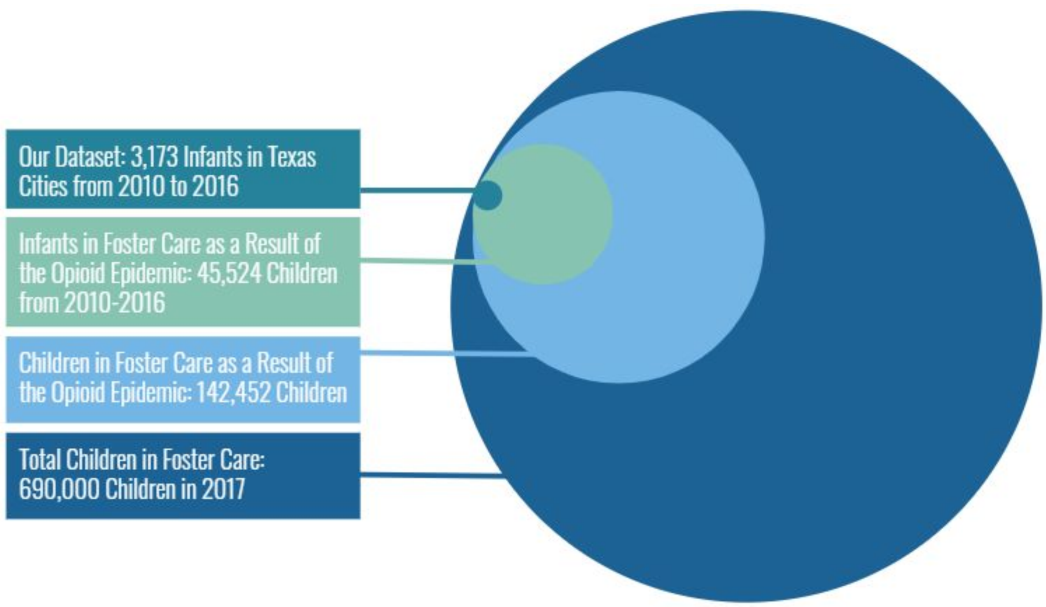
There are several limitations of our project that are important to acknowledge. The first limitation deals with the dataset that we worked with. Our dataset uses cases found in both NCANDS and AFCARS. Although these are national data collection tools, states are not required to report to them. As a result, our data does not capture all of the foster care cases, but just those that have been reported to NCANDS and AFCARS. Additionally, our dataset is filtered to only capture the youngest children within the Texas foster care system as a result of parental/caregiver substance abuse. Our model reduces the length of stay for infants in urban areas of the Texas child welfare system that have been placed in out-of-home care due to substance abuse issues.

Furthermore, the NCANDS and AFCARS data we received was not granular data. Many of the fields within either the NCANDS or AFCARS datasets were aggregated from collections of previous child reports. Within NCANDS, there was no indication of the number of times a child or family received a service. The NCANDS data only indicates whether or not a child or family had ever received a service while engaged with the child welfare system. As a result, the regression models did not know the frequency or magnitude with which a family received a service and only predicted the impacts of receiving versus not receiving. This means the predictive models had less precise information to work with and the coefficients they generated do not take frequency or magnitude into consideration. Ultimately, our coefficients are an estimate of the impacts from receiving a service and children, who in reality receive different degrees of a services, are expected to experience the same exact benefits or drawbacks

7.4 Future Work

The design of this study and its results can be used for future studies of the United States foster care system. By selecting just infants from large central-metro counties in Texas, we limited the dataset dramatically to a subset of cases with less variance in environmental conditions. However, this model can be applied to any subset of children in the foster care system across the United States. The model can explore subsets of children in different age groups, regions, risk factors, etc. Figure 10 demonstrates the scope of our project in comparison to all children in foster care as a result of the opioid epidemic and to the foster care system as a whole.

Figure 10: Project Scope in Greater Context



An example of an interesting comparison to our scope could be done within children of different age groups but also within urban counties of Texas Urbanicity. This could aid with determining how services are allocated to everyone from those areas and how your age can influence which services are most effective. Overall, it could better help the child welfare workers know what each child needs most to benefit them, and what combination of services provide the best results in decreasing the length of stay a child has.

Another comparison could be done on infants from Texas across all urbanicities. This would be beneficial because some services are more relevant for more urban areas than rural areas. Transportation Services, for example, would likely be better suited for children in rural areas where services are not typically within walking distances, as opposed to urban areas where public transportation is more common. By recognizing any common services across all six urbanicities and comparing their usefulness to each other, researchers could determine which will have a more positive impact on certain areas and should be more heavily allocated to those urbanicities than ones where children will not benefit as much.

Rather than solely focusing on age or Texas, future studies could be extended to different infants in Urbanicity 1 locations across the nation. These studies could be beneficial because the child welfare system is different from state to state, and what one state excels at could be another state's weakness. By comparing a small subset of cases from Urbanicity 1 locations from California, Texas, and New York, for example, the CWS could easily determine how resources are allocated differently and provide constructive feedback on what each state could do better to help children be more quickly returned to normalcy. Something to take into consideration when comparing states to each other is the fact that each has its own jurisdiction within the CWS. There is the possibility that each has a system so unique from the others that they cannot be compared, rendering a state-by-state comparison interesting but unable to draw conclusions from.

Overall, the data collected by NCANDS and AFCARS as well as the environmental factors incorporated can all be used to slice the data into comparable subsections. Child age, urbanicity, and state are clearly determining factors of the types of services a child receives in care, but the environmental factors are contributors as well. Environmental slices that can be considered are whether the child entered the system due to abuse versus neglect, whether they reside in a high or low poverty area, or if there are a lot of crimes versus no crime in their neighborhood. These comparisons will provide valuable insights for the child welfare system to use when matching children and services.

Instead of inputting the results from the regression analysis into an optimization model, one could use the results to determine the services that have the greatest impact or those that are in high demand. After talking with Adam Schaffer from the Harvard Kennedy School and Dr. Melinda Gushwa from Simmons University School of Social Work about child welfare service allocation and budgeting we found that a major issue within any state's child welfare system is inability to properly forecast the demand for a given service each year. We noticed this trend within our when analyzing the coefficients that were derived from our regression analysis. Our model interprets all negative coefficients as a benefit to reducing a child's length of stay. While all positive coefficients add additional days to a child's length of stay. This would suggest a state should invest more in the services with negative coefficients. However, from our discussion with Dr. Gushwa, we determined that just because a coefficient is positive, it does not necessarily mean it does not provide a benefit to the child or family. A coefficient could be positive because it is in high demand and the state did not budget enough resources towards it. Thus, the child or family had to wait for the service which increased their length of stay within the system. An extension of forecasting the demand of services to properly budget for them would require more granular data than the dataset we worked with. One would need to know how much of a service was budgeted for in a given year and the number of times that service was distributed in a given year.

Chapter 8: Industrial Engineering Reflection

In the end, this project allowed us to apply what we learned in our classes to a real-world problem presented to us. Throughout the year, we were able to consider many different factors in the context of the project and limit the variables to better suit our research question. We were able to learn more than just the concepts we were taught in our courses, and grew as a team and as individuals. This chapter covers our reflection of the project.

8.1 Design of Project Scope

We decided to scope our project as an extension of last year's project. From the beginning, we planned to examine the connection between substance abuse and the United States child welfare system. By utilizing different mathematical techniques than last year's project, we hoped to improve upon their findings. As we analyzed our dataset, we struggled to gain meaningful insight into the impact different social services had on a child within the system. After deliberating as a group and with our advisors, we were able to narrow our project scope to improving the allocation of services to infants within urban areas of the Texas child welfare system. We chose to scope our project around infants because they are an extremely vulnerable population since they need constant supervision from their parents or guardians. Additionally, by focusing on infants, we were able to gain a better sense of how long they had been in the child welfare system. Also, we chose to narrow the scope of our project to Texas because it had the most infants cases within our dataset. The final way in which we narrowed our dataset was by cities within Texas with an urbanicity score one because we wanted to limit the discrepancies between how different counties report data within our dataset.

Once we narrowed the scope, the goal of our project became clearer. We were going to develop an optimization model to improve the allocation of social services to families in urban areas within Texas struggling with substance abuse and who have infants in out-of-home care. To develop an effective optimization model, we had to filter our dataset to match the scope of our project. In addition to filtering the data by state, urbanicity, and child age, we had to remove services that were not pertinent for infants. Fortunately, during the initial analysis of the data, we had written scripts in python to split the data by these factors. Although our initial analysis led us to narrow the scope of our project, a majority of the scripts we developed were able to be reused or altered for the new dataset. Our thorough data exploration allowed us to quickly adapt to a new project scope and meet the project deadline.

8.2 Constraints in the Design

One of the major constraints in the design of our project was not having an official sponsor. Without a sponsor, we did not have a clear scope for our project. For the first five months of our project, we believed that the goal of our project was to improve upon the previous year's findings. To expand upon last year's project, we incorporated environmental factors such as poverty rate, urbanicity, and crime rate to the existing dataset. Additionally, we split the data into different subsets and analyzed the impact the services had within those subsets. After five months of exploring and analyzing the data, we could not derive meaningful results from the dataset. There was not enough continuity within the dataset to determine the impact different services had on a child within the child welfare system. As a result, we decided to narrow the scope of our project approximately one month before the deadline. Even though this put us under a significant time constraint, we believe that we were able to develop an impactful project.

Another challenge in the design of the project was the data that we were given. To ensure that we were making progress throughout the duration of the project, we worked with the data set from last year's project. This proved to be challenging because a lot of the fields within the dataset were aggregated in a way that was unclear to us. As a result, we requested an updated dataset with granular data within each field. The updated dataset would allow us to not only generate new data fields for each child, but also gain a better understanding of the complexity of each child's case. Unfortunately, by the time we received the updated dataset and the script that was used to aggregate the previous year's dataset, we were occupied with adapting the existing dataset from last year to develop a regression model to fit our narrowed project scope.

8.3 Experience Acquiring and Applying New Knowledge

Our project introduced us to many new data analysis techniques that we did not originally have a strong understanding of. Half of the group had a basic understanding of coding with Python, but most of the coding aspect came from online research. Online resources like Stack Overflow were extremely helpful for specific coding needs and when errors occurred that we were unable to solve ourselves. In addition to Python, we were introduced to Gurobi and SQL. SQL was used to join the data sets together and Gurobi handled the optimization. None of the courses we took were focused on Gurobi and SQL, so reaching out to Professor Trapp was important for answering our questions and aiding us when we had trouble continuing with the optimization model.

Another skill we learned was how to effectively conduct research. Our group did not have an extensive background of the Child Welfare System, so we had to learn as much as we could ourselves. The internet was a helpful resource because it contained answers to many of the questions we had, but we found the CWS to be an incredibly complicated and messy system as a whole. Some of the most valued points of contact during our confusion were the people we reached out to. Professor Douglas is an expert in the field, and the people she knew from her work such as Dr. Melinda Gushwa were some of the most helpful assets to this project. Another professional was Michael Dineen, a statistician and data archivist, from Cornell. He is the one who works directly with the data from the CWS, and was able to clarify whatever data-related questions we posed. Finally, the group who conducted this study last year was also available for questions, and they provided us with valuable feedback on where they struggled and what to improve upon for our project.

A skill we can apply to our careers after graduation is how to effectively communicate with another person. During our research, we had to reach out to many different people, most of whom we had no previous relations with. We learned the importance of having set agendas and questions before contacting someone else, as it makes it easier to communicate our ideas if the other person had a chance to prepare for what we wanted to discuss. We received clearer responses when we were best able to articulate our needs through both written and spoken words. Overall, we feel as though we became better at defining exactly what we were looking for when we reached out for help and believe we can carry these skills forward in the future.

8.4 Teamwork

Our team took a few weeks to develop a routine that was effective for our success. At the beginning of each term, we constructed a Gantt Chart with our team meetings, weekly advisor meetings, and what deliverables we want to accomplish over the course of those seven weeks. This helped us set a timeline for each week and stay on track to achieve all our goals. Our first term as a team was spent mostly determining each person's strengths so that we could focus on what makes us a stronger team rather than teaching people things they did not have any background in. Meeting three times a week for those seven weeks led us to fall naturally into our roles on the team without establishing someone's set position.

After getting to know each other for a term, we found it beneficial for our team to foster an environment open to inclusion and feedback within the group. When someone had an idea, everyone listened to what was being said and it was either incorporated or built upon by other members to better fit our project. When someone was confused about something, members would take the time to explain it using words or whiteboard drawings so that everyone was on the same page. This was helpful for the mathematical models because only one group member was also majoring in mathematics, and some of the terminology was unfamiliar to the other members.

Our advisors were also critical to our success as a team. The weekly meetings enabled us to ask questions about anything that troubled us that week and develop further research questions for the coming weeks. They were always open for communication and were quick to provide us with their knowledge or a related contact to help us when we needed it. In the end, we determined that our goal for this project was to improve upon the research from the previous group and discover something new that could be applied for future studies on the United States foster care system.

Bibliography

AACAP. (2018, October). Foster Care. Retrieved from https://www.aacap.org/AACAP/Families_and_Youth/Facts_for_Families/FFF-Guide/Foster-Care-064.aspx

AFCARS Foster Care Annual File Code Book (2019). Retrieved from https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/afcars-foster-care-file-codebook.pdf

A Guide to Dependency for Arizona Youth in Foster Care, A Guide to Dependency For Arizona Youth in Foster Care (2007). Retrieved from http://www.azchildren.org/MyFiles/PDF/GuidetoDependency.pdf

A Handbook for Youth in Foster Care, A Handbook for Youth in Foster Care (0AD). Retrieved from https://www.michigan.gov/documents/dhs/DHS-Pub-0089_353116_7.pdf

Administration for Children and Families. (2018, January 28). Child Abuse and Neglect Data Release. Retrieved October 10, 2019, from https://www.acf.hhs.gov/media/press/2019/child-abuse-neglect-data-released.

Barth, R. P., Wildfire, J. P., & Green, R. P. (2006). *American Journal of Orthopsychiatry*, *Vol 76(3)*, 358–368. Retrieved from https://psycnet.apa.org/doiLanding?

Bradley, J. (2018, November 8). Understanding the Different Forms of Child Neglect. Retrieved from https://www.highspeedtraining.co.uk/hub/forms-of-child-neglect/

Brook, J., McDonald, T., & Yan, Y. (2012, April). Analysis on the impact of Strengthening Families Program on family reunification in child welfare. *Children and Youth Services Review 34*, pp. 691-695.

Carlson, K. T., & Slovak, K. L. (2008). Gender and Violence Exposure in a Study of Rural Youth. *Journal of Public Child Welfare*, *1*(4), 67–89. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/15548730802118298?journalCode=wpcw20

Children's Bureau. (2018). The Afcars Report . The AFCARS Report . Retrieved from https://www.acf.hhs.gov/sites/default/files/cb/afcarsreport25.pdf

Child protection intervention with high-risk infants. (n.d.). Retrieved from https://www.communities.qld.gov.au/resources/childsafety/practice-manual/intervention-with-high-risk-infants.pdf

Child Welfare Information Gateway. (2013). How the child welfare system works. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau. Retrieved from https://www.childwelfare.gov/pubPDFs/cpswork.pdf

Child Welfare League of America. (2019, April). Texas' Children at a Glance. Retrieved February 20, 2020, from https://www.cwla.org/wp-content/uploads/2019/04/Texas-2019.pdf

Chipungu, S. S., & Bent-Goodley, T. B. (2004). Meeting the Challenges of Contemporary Foster Care. *The Future of Children*, *14*(1), 74–93. Retrieved from https://www.jstor.org/stable/pdf/1602755.pdf?refreqid=excelsior:59db0152940fb6091b494ed583d581ea

Collier, L. (2018). Young Victims of the Opioid Crisis. Retrieved from https://www.apa.org/monitor/2018/01/opioid-crisis

Conradt, E., Crowell, S., & Lester, B. (2018, August 15). Early life stress and environmental influences on the neurodevelopment of children with prenatal opioid exposure. Retrieved from https://www.sciencedirect.com/science/article/pii/S2352289518300183.

Diefendorf, B., Doherty, R., Tropeano, R., Yagoobi, K. (2018). Improving U.S. Foster Care Service Allocation. Major Qualifying Project. Worcester Polytechnic Institute.

Dysart, Kevin C., et al. "Prenatal Drug Exposure - Pediatrics." Merck Manuals Professional Edition, Merck Manuals, www.merckmanuals.com/professional/pediatrics/metabolic,-electrolyte,-and-toxic-disorders-in-neonates/prenatal-drug-exposure.

Employment and the Child Welfare System. (2010). Retrieved from https://cdn.ymaws.com/www.naswtx.org/resource/resmgr/imported/Child_Welfare_and_Social_Workers_White_Paper.pdf

Feder, K. A., Mojtabai, R., Musci, R. J., & Letourneau, E. J. (2018, July). U.S. adults with opioid use disorder living with children: Treatment use and barriers to care. Retrieved from https://www.sciencedirect.com/science/article/pii/S0740547218301892

FindLaw's Team. (n.d.). Aging Out of Foster Care. Retrieved from https://family.findlaw.com/foster-care/aging-out-of-foster-care.html

Florence, C. S., Zhou, C., Luo, F., & Xu, L. (2016). The Economic Burden of Prescription Opioid Overdose, Abuse, and Dependence in the United States, 2013. *Medical Care*, *54*(10), 901–906. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5975355/

Foster Care. (n.d.). Retrieved from https://www.childwelfare.gov/topics/outofhome/foster-care/

Foster Care Numbers Up for Fifth Straight Year. (2019, January 18). Retrieved from https://www.nacac.org/2019/01/18/foster-care-numbers-up-for-fifth-straight-year/

Fostering Success Foundation. (2019, March 31). National and Texas Foster Care Statistics. Retrieved February 20, 2020, from https://fosteringsuccessfoundation.org/resources/foster-care-statistics/?gclid=EAIaIQobChMIssihz7PH5wIVJIVaBR14egT7EAAYASAAEgJ9mPD_BwE

Group and Residential Care. (2018). Retrieved from https://www.childwelfare.gov/topics/outofhome/group-residential-care/

Harsanyi, John C. "Rule Utilitarianism, Equality, and Justice." Social Philosophy and Policy 2.2 (1985): 115-27. Retrieved from https://www.cambridge.org/core/journals/social-philosophy-and-policy/article/rule-utilitarianism-equality-and-justice/4238DA68B43694794D1E22A35BEECE94.

Heller, I.; Tompkins, C.B.Gh (1956), "An Extension of a Theorem of Dantzig's", in Kuhn, H.W.; Tucker, A.W. (eds.), Linear Inequalities and Related Systems, Annals of Mathematics Studies, 38, Princeton (NJ): Princeton University Press, pp. 247–254

Herrenkohl, E. C., Herrenkohl, R. C., & Egolf, B. P. (2003). The Psychosocial Consequences of Living Environment Instability on Maltreated Children. *American Journal of Orthopsychiatry, 73(4)*, 367-380. http://dx.doi.org/10.1037/0002-9432.73.4.367

Hislop, K. B., & Wulczyn, F. (2002). Foster care dynamics in urban and non-urban counties. Retrieved from https://aspe.hhs.gov/basic-report/foster-care-dynamics-urban-and-non-urban-counties#Exit2

How the Child Welfare System Works (2013). Retrieved from https://www.childwelfare.gov/pubPDFs/cpswork.pdf#page=1&view=Introduction

Ingram, D. D., & Franco, S. J. (2014). 2013 NCHS urban-rural classification scheme for counties. *Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics*.

Jones, C. M. (2013). Heroin use and heroin use risk behaviors among nonmedical users of prescription opioid pain relievers – United States, 2002–2004 and 2008–2010. *Drug and Alcohol Dependence*, *132*(1-2), 95–100. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23410617

Klein, S., & Merritt, D. H. (2014). Neighborhood racial & ethnic diversity as a predictor of child welfare system involvement. *Children and Youth Services Review*, *41*, 95–105. Retrieved from https://library.childwelfare.gov/cwig/ws/library/docs/gateway/Record?rpp=10&upp=0&m=1&r=1&w= NATIVE('recno=93898')

Kolodny, A., Courtwright, D. T., Hwang, C. S., Kreiner, P., Eadie, J. L., Clark, T. W., & Alexander, G. C. (2015). The Prescription Opioid and Heroin Crisis: A Public Health Approach to an Epidemic of Addiction. *Annual Review of Public Health*, *36*(1), 559–574. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25581144

Life in Foster Care, A Guide for Idaho Youth, Life in Foster Care, A Guide for Idaho Youth (0AD). Retrieved from http://healthandwelfare.idaho.gov/Portals/0/Children/AdoptionFoster/FosterGuideFINAL_plain.pdf

Maguire-Jack, K., & Font, S. A. (2017). Intersections of individual and neighborhood disadvantage: Implications for child maltreatment. *Children and Youth Services Review*, *72*, 44–51. Retrieved from https://library.childwelfare.gov/cwig/ws/library/docs/gateway/Record?rpp=10&upp=0&m=1&r=1&w= NATIVE('recno=110315')

Martin, C. E., Longinaker, N., & Terplan, M. (2014, July). Recent trends in treatment admissions for prescription opioid abuse during pregnancy. Retrieved from https://www.sciencedirect.com/science/article/pii/S074054721400144

Mcguinness, T. M., & Schneider, K. (2007). Poverty, Child Maltreatment, and Foster Care. *Journal of the American Psychiatric Nurses Association*, *13*(5), 296–303. Retrieved from https://journals.sagepub.com/doi/pdf/10.1177/1078390307308421

National Council for Adoption. (2011). Adoption Advocate. *A Publication of the National Council for Adoption,35*. Retrieved from https://www.adoptioncouncil.org/images/stories/NCFA_ADOPTION_ADVOCATE_NO35.pdf.

NCANDS Child File Code Book (2020). Retrieved from https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/ncands-child-file-codebook.pdf

Newton, R. R., Litrownik, A. J., & Landsverk, J. A. (2000). Children and youth in foster care: disentangling the relationship between problem behaviors and number of placements. *Child Abuse & Neglect*, *24*(10), 1363–1374. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0145213400001897

Neilson, S. (2019, January). More kids are getting placed in foster care because of parents' drug use. Retrieved from https://www.npr.org/sections/health-shots/2019/07/15/741790195/more-kids-are-getting-placed-in-foster-care-because-of-parents-drug-use

Putting it All Together, A Handbook for Youth in Foster Care, Putting it All Together, A Handbook for Youth in Foster Care (2009). Retrieved from http://www.fixcas.com/scholar/Illinoisfoster.pdf

Quast, T., Bright, M. A., & Delcher, C. (2019). The relationship between foster care entries and high-dose opioid prescribing in California. *Addictive Behaviors*, *93*, 52–58. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S030646031830916X

Radel, L., Baldwin, M., Crouse, G., Ghertner, R., & Waters (2018). Substance use, the Opioid Epidemic, and the Child Welfare system: Key findings from a mixed method study. Retrieved from https://bettercarenetwork.org/sites/default/files/SubstanceUseChildWelfareOverview.pdf

Russell, J., & Macgill, S. (2015). Demographics, policy, and foster care rates; A Predictive Analytics Approach. *Children and Youth Services Review*, *58*, 118–126. Retrieved from https://www.sciencedirect.com/science/article/pii/S0190740915300554

Scholl, L., Seth, P., Kariisa, M., Wilson, N., & Baldwin, G. (2018). Drug and Opioid-Involved Overdose Deaths — United States, 2013–2017. *MMWR. Morbidity and Mortality Weekly*

Report, *67*(5152). Retrieved from https://www.cdc.gov/mmwr/volumes/67/wr/mm675152e1.htm

Schrijver, Alexander. Theory of Linear and Integer Programming. Wiley, 1989.

Seymour, P. D. (1980), "Decomposition of regular matroids", Linear Inequalities and Related Systems, Journal of Combinatorial Theory (B), 28, Elsevier, pp. 305–359

Shaw, T. V. (2010). Reunification from Foster Care: Informing Measures Over Time. *Children and Youth Services Review* 32, pp. 475-481.

State Statutes of Abuse and Neglect, State Statutes of Abuse and Neglect (2019). Retrieved from https://www.childwelfare.gov/pubPDFs/define.pdf

U.S. Department of Human Health and Services. (2018). *Child Maltreatment 2017. Child Maltreatment 2017*. Children's Bureau. Retrieved from https://www.acf.hhs.gov/sites/default/files/cb/cm2017.pdf

Vandello, J. A., & Cohen, D. (1999). Patterns of individualism and collectivism across the United States. *Journal of Personality and Social Psychology*, *77*(2), 279–292. Retrieved from https://psycnet.apa.org/record/1999-03699-004

Williams, Paul H. Model Building in Mathematical Programming. 4th ed., John & Wiley, 1999.

Williams, S. C., & Sepulveda, K. (2019, April 3). Infants and toddlers are more likely than older children to enter foster care because of neglect and parental drug abuse. Retrieved from https://www.childtrends.org/infants-and-toddlers-are-more-likely-than-older-children-to-enter-foster-care-because-of-neglect-and-parental-drug-abuse

Wulczyn, F., Hurr, S., & Schmits, F. (2017). Understanding the Differences in How Adolescents Leave Foster Care. *Chapin Hall at the University of Chicago*. Retrieved from https://fcda.chapinhall.org/wp-content/uploads/2017/12/Chapin-Hall_Understanding-the-Differences-in-How-Adolescents-Leave-Foster-Care_12.26_DEF.pdf

Appendices

Appendix A: Glossary Terms

Adoption and Safe Families Act (ASFA): Legislation requires timely permanency planning for children and emphasizes that the child's safety is the paramount concern.

Adoption and Foster Care Analysis and Reporting System (AFCARS): Collects case level information from state and tribal title IV-E agencies on all children in foster care and those who have been adopted with title IV-E agency involvement.

Case Goal: The desired end result of a case.

Caseworker: A type of social worker who is employed by a government agency, non-profit organization, or other group to take on the cases of individuals and provide them with advocacy, information or other services.

Child Abuse Prevention and Treatment Act (CAPTA): Provides Federal funding and guidance to States in support of prevention, assessment, investigation, prosecution, and treatment activities and also provides grants to public agencies and nonprofit organizations, including Indian Tribes and Tribal organizations, for demonstration programs and projects.

Child Maltreatment: Any act, intentional or not, that results in harm, potential for harm, or threat of harm to a child.

Child Protective Services (CPS): A branch of your state's social services department that is responsible for the assessment, investigation and intervention regarding cases of child abuse and neglect.

Collectivism: A social culture where individuals prioritize the group's interests over themselves

Department of Health and Human Services (DHHS): A cabinet level department of the U.S. federal government with the goal of protecting the health of all Americans and providing essential human services.

Emancipation: When a minor legally gains independence from his/her guardians.

Federal Information Processing Standard (FIPS): Unique codes to identify the geographic location for children in the database

Fentanyl: A synthetic opioid which is illegally manufactured as an alternative to the other forms of opioids

Foster Care: A temporary service provided by States for children who cannot live with their families. Children in foster care may live with relatives or with unrelated foster parents. Foster care can also refer to placement settings such as group homes, residential care facilities, emergency shelters, and supervised independent living.

Foster Care Services: Services provided to children and their families in foster care to help children reach their case goal. Services may include counseling, therapy, support groups, child care, parent education, etc.

Guardian Ad Litem: An objective, impartial person whom the court appoints to act as a representative for the minor children in contested custody proceedings.

Individualism: A social culture where individuals prioritize own interests over the rest of the group/community

Kinship Care: Care of children by relatives or close family friends. Kinship care is often preferred to alternative care options because of the child's close relations to the caretaker

Linear Regression: A linear approach and the most common form of predictive modeling

National Opioid Epidemic: Public health crisis in the United States that began in the 1990's. Over 400,000 people have died over the last 20 years as a result of overdosing from opioids

NCANDS: A voluntary data collection system that gathers information from all 50 states about reports of child abuse and neglect. NCANDS was established in response to the Child Abuse Prevention and Treatment Act of 1988. The data are used to examine trends in child abuse and neglect across the country

Neglect: A form of child abuse and is a deficit in meeting a child's basic needs, including the failure to provide adequate health care, supervision, clothing, nutrition, housing as well as their physical, emotional, social, educational and safety needs.

Neonatal Opioid Withdrawal Syndrome (NOWS): Medical issues that occur in a newborn who was exposed to addictive opiate drugs while in the mother's womb.

Optimization: The selection of the best element/outcome using a set of criteria and constraints

Permanency: A long term placement for a child such as the original caregivers, adoptive parents, long term foster care, or a group home.

Predictive Modeling: A process that involves using data and statistics to forecast outcomes

Prescription Opioids: Legal form of opioids that are used for pain relief. Pharmaceutical companies pushed these opioids in the 1990's which set off the national opioid epidemic

Random Forest: A form of regression modeling that consists of a large number of individual decision trees that operate as an ensemble. Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes the model's prediction.

Reunification: The most common case goal for foster care children that involves reuniting a child with its original family.

United States Children's Bureau: The federal body that governs the foster care system on a national level. The main role of the Children's Bureau is to pass legislation and support programs and services at the state level with federal funding

Urbanicity: The degree to which a geographical area is urban.

Appendix B: Predictive Factors and Sources

Factor Name	**Source**
Child Race Native American	AFCARS
Child Race Asian	AFCARS
Child Race Black	AFCARS
Child Race Pacific Islander or Hawaiian	AFCARS
Child Race White	AFCARS
Child Race Unable to Determine	AFCARS
Child Hispanic Origin	AFCARS
Ever Adopted	AFCARS
Voluntary Removal	AFCARS
Court Ordered Removal	AFCARS
Removal Physical Abuse	AFCARS
Removal Sexual Abuse	AFCARS
Removal Neglect	AFCARS
Removal Parent Alcohol Abuse	AFCARS
Removal Parent Drug Abuse	AFCARS
Removal Child Alcohol Abuse	AFCARS
Removal Child Drug Abuse	AFCARS
Removal Child Disability	AFCARS
Removal Child Behavior	AFCARS
Removal Parents Died	AFCARS
Removal Parents Jail	AFCARS
Removal Parent Coping	AFCARS
Removal Abandonment	AFCARS
Removal Relinquishment	AFCARS
Removal Inadequate Housing	AFCARS
Current Placement Setting	AFCARS
Placed Out of State	AFCARS

Goal Reunification	AFCARS
Goal Kinship	AFCARS
Goal Adoption	AFCARS
Goal Long Term Foster Care	AFCARS
Goal Guardianship	AFCARS
Goal Not Yet Established	AFCARS
Goal Missing	AFCARS
Caretaker Married Couple	AFCARS
Caretaker Unmarried Couple	AFCARS
Caretaker Single Female	AFCARS
Caretaker Single Male	AFCARS
Foster Caretaker NA	AFCARS
Foster Caretaker Married Couple	AFCARS
Foster Caretaker Unmarried Couple	AFCARS
Foster Caretaker Single Female	AFCARS
Foster Caretaker Single Male	AFCARS
Discharge Reunification	AFCARS
Discharge Kinship	AFCARS
Discharge Adoption	AFCARS
Discharge Emancipation	AFCARS
Discharge Guardianship	AFCARS
Discharge Transfer	AFCARS
Discharge Runaway	AFCARS
Discharge Missing	AFCARS
Title IV Foster Care Payments	AFCARS
Title IV Adoption Assistance	AFCARS
Title IV AFDC	AFCARS
Title IV Child Support	AFCARS
Title IV Medicaid	AFCARS
Title XIX	AFCARS
State Support Only	AFCARS

Monthly FC Payment	AFCARS
Length of Stay Total	AFCARS
Child Age	AFCARS
Rural Urban Continuum Code	AFCARS
In FC During FY	AFCARS
Child Awaiting Adoption	AFCARS
Parental Rights Terminated	AFCARS
Child Over Age	AFCARS
FIPS_SCC	NCANDS
Child Male	NCANDS
Caretaker Married	NCANDS
Caretaker Married Parent and Step Parent	NCANDS
Caretaker Cohabitating Couple	NCANDS
Caretaker Unknown Couple	NCANDS
Caretaker Single Mother and Adult	NCANDS
Caretaker Single Father and Adult	NCANDS
Caretaker Non-parent Relative	NCANDS
Caretaker Non-relative	NCANDS
Caretaker Group Home	NCANDS
Caretaker Other	NCANDS
Caretaker Unknown	NCANDS
Military	NCANDS
Child Alcohol	NCANDS
Child Drug	NCANDS
Child Mental Retardation	NCANDS
Child Emotional Problem	NCANDS
Child Visual or Hearing Problems	NCANDS
Child Learning Disability	NCANDS
Child Physical Disability	NCANDS
Child Behavioral Problem	NCANDS
Child Medical Problem	NCANDS

Caretaker Alcohol Abuse	NCANDS
Caretaker Drug Abuse	NCANDS
Caretaker Mental Retardation	NCANDS
Caretaker Emotional Problems	NCANDS
Caretaker Visual or Hearing Problems	NCANDS
Caretaker Learning Disability	NCANDS
Caretaker Physically Disabled	NCANDS
Caretaker Medical Problems	NCANDS
Domestic Violence	NCANDS
Inadequate Housing	NCANDS
Financial Problem	NCANDS
Public Assistance	NCANDS
Post Investigation Services	NCANDS
Court Appointed Representative	NCANDS
Family Support Services	NCANDS
Family Preservation Services	NCANDS
Foster Care Services	NCANDS
Adoption Services	NCANDS
Case Management Services	NCANDS
Counseling Services	NCANDS
Daycare Services	NCANDS
Education Services	NCANDS
Employment Services	NCANDS
Family Planning Services	NCANDS
Health and Home Health Services	NCANDS
Home Based Services	NCANDS
Housing Services	NCANDS
Transitional Living Services	NCANDS
Informational and Referral Services	NCANDS
Legal Services	NCANDS
Mental Health Services	NCANDS

Pregnancy and Parenting Services	NCANDS
Respite Services	NCANDS
Special Services Disabled	NCANDS
Special Services Juvenile Delinquent	NCANDS
Substance Abuse Services	NCANDS
Transportation Services	NCANDS
Other Services	NCANDS
Unemployment Rate	US Census Bureau
Poverty Rate	US Census Bureau
Average Response	NCANDS
County	NCANDS
2013 NHCS Urbanicity Score (1-6)	US Census Bureau
Racial Ethnic Diversity Score	US Census Bureau
Unemployment Rate 2017	Bureau of Labor Statistics
Poverty Rate 2017	Bureau of Labor Statistics
Median Household Income 2017 ($)	Bureau of Labor Statistics
Crime Rate	US Census Bureau

Appendix C: SQL Script

```
SELECT
AFCARS.ID AS [Child ID AFCARS],
AFCARS.FY AS FY,
AFCARS.St AS STATE_ABRV,
AFCARS.STATE AS STATE_SC,
 AFCARS.AMIN AS [Child Race Native American],
AFCARS.ASIAN AS [Child Race Asian],
AFCARS.BLKAFRAM AS [Child Race Black],
AFCARS.HAWAIIPI AS [Child Race Pacific Islander or Hawaiian],
AFCARS.WHITE AS [Child Race White],
AFCARS.UNTODETM AS [Child Race Unable to Determine],
IIF(AFCARS.HISORGIN=1,1,0) AS [Child Hispanic Origin],
AFCARS.EVERADPT AS [Ever Adopted],
IIF(AFCARS.MANREM=1,1,0) AS [Voluntary Removal],
IIF(AFCARS.MANREM=2,1,0) AS [Court Ordered Removal],
AFCARS.PHYABUSE AS [Removal Physical Abuse],
AFCARS.SEXABUSE AS [Removal Sexual Abuse ],
AFCARS.NEGLECT AS [Removal Neglect],
AFCARS.AAPARENT AS [Removal Parent Alcohol Abuse],
AFCARS.DAPARENT AS [Removal Parent Drug Abuse],
AFCARS.AACHILD AS [Removal Child Alcohol Abuse],
AFCARS.DACHILD AS [Removal Child Drug Abuse],
AFCARS.CHILDIS AS [Removal Child Disability],
AFCARS.CHBEHPRB AS [Removal Child Behavior],
AFCARS.PRTSDIED AS [Removal Parents Died],
AFCARS.PRTSJAIL AS [Removal Parents Jail],
AFCARS.NOCOPE AS [Removal Parent Coping],
AFCARS.ABANDMNT AS [Removal Abandonment],
AFCARS.RELINQSH AS [Removal Relinquishment],
AFCARS.HOUSING AS [Removal Inadequate Housing],
AFCARS.CURPLSET AS [Current Placement Setting],
IIF(AFCARS.PLACEOUT=1,1,0) AS [Placed Out of State],
IIF(AFCARS.CASEGOAL=1,1,0) AS [Goal Reunification],
IIF(AFCARS.CASEGOAL=2,1,0) AS [Goal Kinship],
IIF(AFCARS.CASEGOAL=3,1,0) AS [Goal Adoption],
```

```
IIF(AFCARS.CASEGOAL=4,1,0) AS [Goal Long Term Foster Care],
IIF(AFCARS.CASEGOAL=6,1,0) AS [Goal Guardianship],
IIF(AFCARS.CASEGOAL=7,1,0) AS [Goal Not Yet Established],
IIF(AFCARS.CASEGOAL=8,1,0) AS [Goal Missing],
IIF(AFCARS.CTKFAMST=1,1,0) AS [Caretaker Married Couple],
IIF(AFCARS.CTKFAMST=2,1,0) AS [Caretaker Unmarried Couple],
IIF(AFCARS.CTKFAMST=3,1,0) AS [Caretaker Single Female],
IIF(AFCARS.FOSFAMST=0,1,0) AS [Foster Caretaker NA],
IIF(AFCARS.FOSFAMST=1,1,0) AS [Foster Caretaker Married Couple],
IIF(AFCARS.FOSFAMST=2,1,0) AS [Foster Caretaker Unmarried Couple],
IIF(AFCARS.FOSFAMST=3,1,0) AS [Foster Caretaker Single Female],
IIF(AFCARS.FOSFAMST=4,1,0) AS [Foster Caretaker Single Male],
IIF(AFCARS.DISREASN=1,1,0) AS [Discharge Reunification],
IIF(AFCARS.DISREASN=2,1,0) AS [Discharge Kinship],
IIF(AFCARS.DISREASN=3,1,0) AS [Discharge Adoption],
IIF(AFCARS.DISREASN=4,1,0) AS [Discharge Emancipation],
IIF(AFCARS.DISREASN=5,1,0) AS [Discharge Guardianship],
IIF(AFCARS.DISREASN=6,1,0) AS [Discharge Transfer],
IIF(AFCARS.DISREASN=7,1,0) AS [Discharge Runaway],
IIF(AFCARS.DISREASN=99,1,0) AS [Discharge Missing],
AFCARS.IVEFC AS [Title IV Foster Care Payments],
AFCARS.IVEAA AS [Title IV Adoption Assistance],
AFCARS.IVAAFDC AS [Title IV AFDC],
AFCARS.IVDCHSUP AS [Title IV Child Support],
AFCARS.XIXMEDCD AS [Title IV Medicaid],
AFCARS.SSIOTHER AS [Title XIX ],
AFCARS.NOA AS [State Support Only],
AFCARS.FCMntPay AS [Monthly FC Payment],
AFCARS.LifeLOS AS [Length of Stay Total],
AFCARS.AgeAtEnd AS [Child Age],
AFCARS.RU13 AS [Rural Urban Continuum Code],
AFCARS.Served AS [In FC During FY],
AFCARS.IsWaiting AS [Child Awaiting Adoption],
AFCARS.IsTPR AS [Parental Rights Terminated],
AFCARS.AgedOut AS [Child Over Age],

NCANDS.rptfips AS FIPS_SCC,
IIF(NCANDS.chsex=1,1,0) AS [Child Male],
```

```
IIF(NCANDS.chlvng=0,1,0) AS [Caretaker Married],
IIF(NCANDS.chlvng=1,1,0) AS [Caretaker Married Parent and Step Parent],
IIF(NCANDS.chlvng=4,1,0) AS [Caretaker Cohabitating Couple],
IIF(NCANDS.chlvng=5,1,0) AS [Caretaker Unknown Couple],
IIF(NCANDS.chlvng=8,1,0) AS [Caretaker Single Mother and Adult],
IIF(NCANDS.chlvng=9,1,0) AS [Caretaker Single Father and Adult],
IIF(NCANDS.chlvng=10,1,0) AS [Caretaker Non-parent Relative],
IIF(NCANDS.chlvng=11,1,0) AS [Caretaker Non-relative],
IIF(NCANDS.chlvng=12,1,0) AS [Caretaker Group Home],
IIF(NCANDS.chlvng=88,1,0) AS [Caretaker Other],
IIF(NCANDS.chlvng=99,1,0) AS [Caretaker Unknown],
NCANDS.chmil AS Military,
NCANDS.cdalc AS [Child Alcohol],
NCANDS.cddrug AS [Child Drug],
NCANDS.cdrtrd AS [Child Mental Retardation],
NCANDS.cdemotnl AS [Child Emotional Problem],
NCANDS.cdvisual AS [Child Visual or Hearing Problems],
NCANDS.cdlearn AS [Child Learning Disability],
NCANDS.cdphys AS [Child Physical Disability],
NCANDS.cdbehav AS [Child Behavioral Problem],
NCANDS.cdmedicl AS [Child Medical Problem],
NCANDS.fcalc AS [Caretaker Alcohol Abuse],
NCANDS.fcdrug AS [Caretaker Drug Abuse],
NCANDS.fcrtrd AS [Caretaker Mental Retardation],
NCANDS.fcemotnl AS [Caretaker Emotional Problems],
NCANDS.fcvisual AS [Caretaker Visual or Hearing Problems],
NCANDS.fclearn AS [Caretaker Learning Disability],
NCANDS.fcphys AS [Caretaker Physically Disabled],
NCANDS.fcmedicl AS [Caretaker Medical Problems],
NCANDS.fcviol AS [Domestic Violence],
NCANDS.fchouse AS [Inadequate Housing],
NCANDS.fcmoney AS [Financial Problem],
NCANDS.fcpublic AS [Public Assistance],
NCANDS.postserv AS [Post Investigation Services],
NCANDS.cochrep AS [Court Appointed Representative],
NCANDS.famsup AS [Family Support Services],
NCANDS.fampres AS [Family Preservation Services],
NCANDS.fostercr AS [Foster Care Services],
```

```
NCANDS.adopt AS [Adoption Services],
NCANDS.casemang AS [Case Management Services],
NCANDS.counsel AS [Counseling Services],
NCANDS.daycare AS [Daycare Services],
NCANDS.educatn AS [Education Services],
NCANDS.employ AS [Employment Services],
NCANDS.famplan AS [Family Planning Services],
NCANDS.health AS [Health and Home Health Services],
NCANDS.homebase AS [Home Based Services],
NCANDS.housing AS [Housing Services],
NCANDS.transliv AS [Transitional Living Services],
NCANDS.inforef AS [Informational and Referral Services],
NCANDS.legal AS [Legal Services],
NCANDS.menthlth AS [Mental Health Services],
NCANDS.pregpar AS [Pregnancy and Parenting Services],
NCANDS.respite AS [Respite Services],
NCANDS.ssdisabl AS [Special Services Disabled],
NCANDS.ssdelinq AS [Special Services Juvenile Delinquent],
NCANDS.subabuse AS [Substance Abuse Services],
NCANDS.transprt AS [Transportation Services],
NCANDS.othersv AS [Other Services]

FROM AFCARS,
NCANDS
WHERE AFCARS.StFCID=NCANDS.StFCID
AND AFCARS.AgeAtEnd <= 1
AND AFCARS.RU13 <=1
AND AFCARS.DISREASN <> 0
AND (AFCARS.St = 'TX')
AND AFCARS.LifeLos is not null;
```

Appendix D: Regression Script

```
import pandas as pd
import numpy as np
import math

from sklearn.model_selection import train_test_split
from sklearn.linear_model import Lasso
from sklearn.metrics import r2_score
import statsmodels.api as sm

from collections import OrderedDict
import os

# define the three inputs:
labelName = 'Length of Stay Total' # the name of the feature you're trying to predict
fileName = 'DataSet_AUS_InteractionTerms.csv' # the entire data set to train and test with
os.chdir(r'\\research.wpi.edu\BUS\atrapplab\MQP_Foster_Care_Research_2019\MQP_2020\Stat
e_AUS') # the location to find and save files

def manualSetUp(data):
    print("shape before filtering: ", data.shape)
    # drop columns with irrelevant or non-numerical data,
    # including all terms involving Post Investigation Services and Foster Care Services
      dropList = ['Child ID AFCARS', 'Child ID_ST', 'County', 'STATE_ABRV', 'FIPS_SCC',
'STATE_SC', 'FY',
              'Foster Care Services', 'Post Investigation Services']
    dropList = ['County', 'STATE_ABRV', 'FIPS_SCC', 'STATE_SC', 'FY',
              'Foster Care Services', 'Post Investigation Services']
    cols = list(data.columns)
    postInv = [k for k in cols if ('Post Investigation Services' in k)]
    fosterCare = [k for k in cols if ('Foster Care Services' in k)]
    data = data.drop(dropList + postInv + fosterCare, axis=1)

    # drops columns where length of stay is blank
    data = data.dropna(subset=[labelName])
```

```
    # drop columns lacking with variance under p(1-p)
    data = removeNonVarCols(data=data, p=0.10)
    data.fillna(value=0, inplace=True)
    print("shape after filtering: ", data.shape)
    return data

def removeNonVarCols(data, p=0.10, verbose = False):
    # drop columns lacking with variance under p(1-p)
    # for binary columns, this means if under 10% of entries are 0, the column will be dropped
    cols = list(data.columns)
    varDict = {}
    for str in cols:
        varDict[str] = data[str].var()
    cutoff = p - (p * p)
    keepList = [key for key in varDict if varDict[key] > cutoff]

    if verbose:
        noVarList = [key for key in varDict if varDict[key] <= cutoff]
            print(len(noVarList), ' columns were dropped. Their variances were under the ', cutoff, '
cutoff.')
        print(noVarList)

    data = data[keepList]
    return data

def linearRegressionRegularized(data, verbose=True):
    # split into training and testing
    train_data = data.sample(frac=0.8, random_state=0)
    test_data = data.drop(train_data.index)

    X_train = train_data.drop([labelName], axis=1)
    y_train = train_data[labelName]
    X_test = test_data.drop([labelName], axis=1)
    y_test = test_data[labelName]

    # run lasso regression to get list of features to keep
    lasso = Lasso()
    lasso.fit(X_train, y_train)
```

```
if verbose:
    print("-------------------\nLASSO REGRESSION:")
    print("training score:", lasso.score(X_train, y_train))
    print("test score: ", lasso.score(X_test, y_test))
    print("number of features used: ", np.sum(lasso.coef_ != 0))
cdf = pd.DataFrame(lasso.coef_, X_train.columns, columns=['Coefficients'])
cdf.to_csv('lasso_coefficients.csv')
cdf = cdf[cdf['Coefficients'] != 0]
keepCols = cdf.index.values

# drop features before running OLS
X_train = X_train[keepCols]
X_test = X_test[keepCols]

# run ordinary least squares linear regression
X_train = sm.add_constant(X_train)
mod = sm.OLS(y_train, X_train)
fitMod = mod.fit()
if verbose:
    print("-------------------\nOLS REGRESSION:")
    train_predictions = fitMod.predict(X_train)
    train_rsq = r2_score(y_train, train_predictions)
    print("training score:", train_rsq)
    # cross validate with testing set
    X_test = sm.add_constant(X_test, has_constant='add')
    predictions = fitMod.predict(X_test)
    test_rsq = r2_score(y_test, predictions)
    print('testing score: ', test_rsq)

# get coefficients in table
sumryTbl = fitMod.summary2().tables[1]
coefDF = pd.DataFrame(data=sumryTbl)
coefDF.drop(['Std.Err.', 't', '[0.025', '0.975]'], axis=1, inplace=True)
coefDF.columns = ['coef_', 'pval_']
coefDF.to_csv(('coeff_table.csv'), index=True)

# compile summary of the final OS model
```

```
        cvSum = {'Predicting': labelName, 'OLS Test RSQ': [test_rsq], 'Num Training Rows':
[len(y_train)],
                 'Average LOS Training': train_data[labelName].mean(), 'Average LOS Testing':
test_data[labelName].mean()}
    summaryDF = pd.DataFrame(data=cvSum)
    summaryDF.to_csv(('lr_model_summaries.csv'), index=False)

def getOptInputs(origDF, reducedDF):
        # prep the inputs for the optimization model: CompleteDataSet.csv, ServiceLimits.csv,
EnvFactors.csv
    # join in the record ID table, which indicates the family a child comes from
    # alternatively, include this column when pulling data from Access

    df = origDF
                                                                  ridTable              =
pd.read_csv(r'\\research.wpi.edu\BUS\atrapplab\MQP_Foster_Care_Research_2019\MQP_2020\
RID_Table.csv', encoding='ISO 8859-1')
    df = pd.merge(df, ridTable, how='left', on=['Child ID AFCARS', 'STATE_ABRV'])

df.to_csv(r'\\research.wpi.edu\BUS\atrapplab\MQP_Foster_Care_Research_2019\MQP_2020\O
ptimization CSVs\CompleteDataSet.csv', index=False)
    getServiceLimits(df)
    getListEnvFactors(reducedDF)

def getServiceLimits(data):
    # outputs a .csv with a table of service names and number of units allocated
    # for child services, this is done by simply summing on each service column
    # for family services, all of which are on the list 'fam' below, the limit is the
    # count of how many families received the srvice.

    serviceCols = [k for k in list(data.columns) if ('ervice' in k) and not ('*' in k)]
    fam = ['Family Support Services', 'Family Preservation Services', 'Education Services',
            'Employment Services', 'Family Planning Services', 'Health and Home Health Services',
'Home Based Services',
            'Housing Services', 'Informational and Referral Services', 'Legal Services', 'Pregnancy and
Parenting Services',
          'Respite Services', 'Substance Abuse Services', 'Transportation Services']
```

```
    limitDict = {}
    servData = data[serviceCols+['Report ID']].groupby(by='Report ID').max()
    for service in serviceCols:
        if service in fam:
            lim = servData[service].sum()
            limitDict[service] = lim
            print('FAMILY SERVICE: ', service, ' has limit ', lim)
        else:
            limitDict[service] = len(data.loc[data[service] != 0])
    limitDF = pd.DataFrame(data=limitDict, index=[0])

limitDF.to_csv(r'\\research.wpi.edu\BUS\atrapplab\MQP_Foster_Care_Research_2019\MQP_20
20\Optimization CSVs\ServiceLimits.csv', index=False)

def getListEnvFactors(data):
    # outputs a .csv with a list of all the non-service factors

    caseCols = [k for k in list(data.columns) if not ('ervice' in k) and not ('*' in k)]
    caseDF = pd.DataFrame(data=[], columns=caseCols)

caseDF.to_csv(r'\\research.wpi.edu\BUS\atrapplab\MQP_Foster_Care_Research_2019\MQP_20
20\Optimization CSVs\EnvFactors.csv', index=False)

def main():
    orig_data = pd.read_csv(fileName)
    cleaned_data = manualSetUp(orig_data)
    getOptInputs(orig_data, cleaned_data)
    linearRegressionRegularized(cleaned_data)

main()
```

Appendix E: Full List of Factors & Coefficients from OS Regression

Service Factor	**Coefficient**	**P-Val**
Constant	305.3269734	5.94E-47
Monthly FC Payment * Goal Adoption	43.10322153	3.63E-09
Monthly FC Payment * Goal Kinship	20.29216206	5.46E-08
Discharge Reunification * Goal Reunification	64.11813934	1.49E-07
Monthly FC Payment * Goal Reunification	15.80643826	4.47E-05
Placed Out of State * Unemployment Rate 2017	8.209463367	0.000213199
Poverty Rate	0.040582622	0.000387382
Monthly FC Payment * Goal Guardianship	9.150339904	0.000456226
Removal Sexual Abuse * Median Household Income 2017 ($)	-12.47072317	0.000869722
Family Planning Services * Poverty Rate 2017	7.018585463	0.001295024
Monthly FC Payment * Goal Not Yet Established	-5.846805646	0.001451393
Removal Sexual Abuse * Crime Rate	12.28282886	0.001935501
Child Age * Median Household Income 2017 ($)	22.70907589	0.002436458

Pregnancy and Parenting Services * Unemployment Rate 2017	-11.3972003	0.003361497
Housing Services * Monthly FC Payment	-9.044622259	0.005380222
Child Age * Unemployment Rate 2017	23.13306302	0.005439073
Caretaker Married Couple * Crime Rate	-6.10035463	0.006488205
Title IV Foster Care Payments * Monthly FC Payment	-11.70469553	0.006945064
Case Management Services * Monthly FC Payment	-9.540618383	0.006959019
Child Age * Current Placement Setting	2.406413127	0.00727772
Poverty Rate 2017	-5.609566804	0.008249308
Removal Parent Drug Abuse * Poverty Rate 2017	-5.609566804	0.008249308
Court Ordered Removal * Poverty Rate 2017	-5.609566804	0.008249308
Monthly FC Payment * Removal Physical Abuse	-10.57284141	0.009130504
Child Male * 2013 NHCS Urbanicity Score (1-6)	-11.15724818	0.010937976
Removal Child Disability * 2013 NHCS Urbanicity Score (1-6)	15.39654814	0.013624808
Caretaker Non-parent Relative * Median Household Income 2017 ($)	5.62510777	0.016109926
Home Based Services * Monthly FC Payment	6.472398392	0.018198268

Respite Services * Monthly FC Payment	6.670272252	0.023695171
Monthly FC Payment * Crime Rate	-13.69682848	0.026518729
Child Age * Goal Adoption	35.79802191	0.030738963
Daycare Services * Current Placement Setting	-1.068632013	0.035301057
Discharge Adoption * Unemployment Rate 2017	24.99411437	0.040076019
Monthly FC Payment * Placed Out of State	-4.83334029	0.042514645
Goal Adoption	16.84170488	0.045605366
Title IV Medicaid * Goal Adoption	16.84170488	0.045605366
Monthly FC Payment * Caretaker Visual or Hearing Problems	3.493109447	0.046599516
Monthly FC Payment * Current Placement Setting	-11.18791621	0.04930949
Housing Services * Median Household Income 2017 ($)	4.722972007	0.053106797
Unemployment Rate	-3.375832308	0.056775008
Parental Rights Terminated	27.28926738	0.057536598
Inadequate Housing * Median Household Income 2017 ($)	-4.796026433	0.058711683
Title XIX * Crime Rate	3.353150478	0.063000727
Discharge Transfer * Poverty Rate 2017	-4.332873399	0.063466006

Child Age * Poverty Rate 2017	15.17574248	0.06418091
Child Male * Goal Reunification	14.87665109	0.066089474
Removal Relinquishment * Crime Rate	3.923420895	0.071283451
Monthly FC Payment * Caretaker Mental Retardation	-4.570163292	0.074176437
Child Age * Discharge Adoption	29.29413558	0.084665637
Child Visual or Hearing Problems * Median Household Income 2017 ($)	-3.748744693	0.096389353
Child Hispanic Origin * Goal Reunification	11.80799883	0.097250471
Substance Abuse Services * Monthly FC Payment	-6.648251524	0.101468954
Monthly FC Payment * Removal Child Disability	-5.655718449	0.107996664
Discharge Reunification * 2013 NHCS Urbanicity Score (1-6)	-11.49868514	0.114979695
Current Placement Setting * 2013 NHCS Urbanicity Score (1-6)	-0.624550827	0.132766118
Public Assistance * Goal Reunification	11.38780369	0.141719307
Case Management Services * Child Male	9.685124766	0.15198521
Caretaker Alcohol Abuse * 2013 NHCS Urbanicity Score (1-6)	5.668476241	0.163850866
Respite Services * Unemployment Rate 2017	-3.947277664	0.164313684

Title IV Foster Care Payments * Goal Reunification	12.81947845	0.173597631
Child Visual or Hearing Problems * Current Placement Setting	2.341793881	0.173678311
Child Male * Removal Physical Abuse	9.909909441	0.175884258
Pregnancy and Parenting Services * Current Placement Setting	1.45644246	0.180668671
Removal Abandonment * Crime Rate	-3.784197218	0.191411085
Removal Child Disability * Poverty Rate 2017	10.54491907	0.202120418
Discharge Reunification * Median Household Income 2017 ($)	-8.08109406	0.205847456
Goal Reunification	6.520824206	0.217690668
Title IV Medicaid * Goal Reunification	6.520824206	0.217690668
Family Support Services * Median Household Income 2017 ($)	3.00747472	0.220443549
Monthly FC Payment * Child Race Unable to Determine	2.767730933	0.222938216
Child Race White * Median Household Income 2017 ($)	-3.190214076	0.225192923
Special Services Disabled * Unemployment Rate 2017	-3.14649956	0.227960244
Monthly FC Payment * 2013 NHCS Urbanicity Score (1-6)	-4.130117814	0.250814592

Caretaker Physically Disabled * Median Household Income 2017 ($)	-2.252510119	0.250817298
Title IV Child Support * Current Placement Setting	1.00517282	0.264915326
Removal Child Behavior * Crime Rate	-1.920157238	0.271456231
Title IV Foster Care Payments * Current Placement Setting	0.69217814	0.274372227
Monthly FC Payment * Child Male	-3.331191936	0.275523941
Informational and Referral Services * Unemployment Rate 2017	-3.106031236	0.277837682
Removal Parent Alcohol Abuse * Poverty Rate 2017	-2.60869287	0.285644251
Caretaker Unknown * Median Household Income 2017 ($)	-5.666871757	0.292536128
Case Management Services * State Support Only	6.858803082	0.297444973
Monthly FC Payment * Caretaker Learning Disability	-2.318255249	0.308838907
Caretaker Group Home * Median Household Income 2017 ($)	-2.559809751	0.313774585
Caretaker Medical Problems * Unemployment Rate 2017	2.40365244	0.317961884
Monthly FC Payment * Discharge Transfer	-2.553757281	0.319453005
Removal Neglect * Median Household Income 2017 ($)	-2.249858792	0.331581807
Title IV AFDC * Current Placement Setting	-0.775210896	0.333497636

Title IV Foster Care Payments * Public Assistance	7.146993475	0.333940675
Child Race White * Current Placement Setting	-0.496635791	0.353727153
Monthly FC Payment * Discharge Guardianship	2.619230645	0.35657107
Title IV AFDC * Monthly FC Payment	-1.885762433	0.36919529
Title IV Child Support * Median Household Income 2017 ($)	3.643630637	0.394928473
Monthly FC Payment * Child Race Black	-2.135304878	0.409333696
Legal Services * Crime Rate	1.717461226	0.414313161
Monthly FC Payment * Removal Abandonment	-2.251946903	0.418470329
Other Services * Crime Rate	1.484059446	0.423733869
Monthly FC Payment * Caretaker Medical Problems	-2.044056104	0.425388606
Case Management Services * 2013 NHCS Urbanicity Score (1-6)	3.372720751	0.429406287
Removal Child Disability * Unemployment Rate 2017	-7.446086471	0.438587131
Monthly FC Payment * Caretaker Other	-1.473707669	0.441340271
Discharge Adoption * Median Household Income 2017 ($)	-7.222461062	0.447168359
Mental Health Services * Current Placement Setting	-0.619128457	0.448160035
Monthly FC Payment * Inadequate Housing	-1.912745927	0.459662393

Average Response	-0.107979104	0.462670426
Removal Physical Abuse * Median Household Income 2017 ($)	3.440588291	0.464335384
Child Race Pacific Islander or Hawaiian * Median Household Income 2017 ($)	-1.395905354	0.470209903
Daycare Services * Poverty Rate 2017	1.631245053	0.476628022
Monthly FC Payment * Removal Parent Coping	-1.766045997	0.479166899
Inadequate Housing * Current Placement Setting	0.384103293	0.482791103
Monthly FC Payment * Removal Child Alcohol Abuse	-1.592585962	0.485783005
Removal Physical Abuse * Crime Rate	3.249232388	0.487319986
Removal Child Drug Abuse * Crime Rate	2.624975509	0.489243737
Military * Crime Rate	1.311264899	0.505885358
Transportation Services * Current Placement Setting	-0.921047849	0.532006732
Domestic Violence * Crime Rate	-1.257604956	0.533167413
Adoption Services * Crime Rate	-2.514701491	0.538187214
Discharge Adoption * Poverty Rate 2017	-5.597363215	0.579680336
Child Medical Problem * Poverty Rate 2017	-2.742771185	0.628261541
Removal Parents Jail * Unemployment Rate 2017	3.118758774	0.648486389

Child Race Unable to Determine * Current Placement Setting	1.093217539	0.658176686
Health and Home Health Services * Unemployment Rate 2017	-0.986579201	0.661766869
Home Based Services * Unemployment Rate 2017	-1.312112391	0.665812193
Family Preservation Services * Crime Rate	-4.152778545	0.6694089
Child Male * Current Placement Setting	0.172561928	0.671919264
Monthly FC Payment * Caretaker Non-parent Relative	0.843208238	0.676486252
Monthly FC Payment * Caretaker Married Couple	1.005105783	0.681503583
Counseling Services * Poverty Rate 2017	-4.407270308	0.683540447
Title IV Foster Care Payments * 2013 NHCS Urbanicity Score (1-6)	2.511736584	0.688312403
Counseling Services * Unemployment Rate 2017	4.350431918	0.701318875
Title IV Child Support * Poverty Rate 2017	-1.478857576	0.721182724
Monthly FC Payment * Removal Parent Alcohol Abuse	-0.857552974	0.724537875
Title IV Foster Care Payments * Poverty Rate 2017	2.960327548	0.734229133
Removal Inadequate Housing * Median Household Income 2017 ($)	-1.485124854	0.748364664
Title IV Adoption Assistance * Unemployment Rate 2017	0.581457771	0.752183372

State Support Only * Goal Adoption	2.339093375	0.75374306
Family Preservation Services * Unemployment Rate 2017	-2.74156284	0.775518587
Monthly FC Payment * Discharge Reunification	-0.99376332	0.776920475
Transportation Services * Monthly FC Payment	-0.717518364	0.778309098
Removal Parent Coping * Median Household Income 2017 ($)	-1.31655275	0.802785748
Case Management Services * Current Placement Setting	0.159286846	0.809586486
Caretaker Unknown * Poverty Rate 2017	1.321091697	0.822020352
Child Medical Problem * Median Household Income 2017 ($)	-0.704939976	0.852820205
Education Services * Current Placement Setting	-0.241866306	0.853438272
Employment Services * Crime Rate	1.17916693	0.876559048
Child Medical Problem * Crime Rate	-0.784294286	0.888857061
Title IV Foster Care Payments * Unemployment Rate 2017	1.321204437	0.894582937
Removal Parent Coping * Poverty Rate 2017	-0.4940218	0.92421858
Substance Abuse Services * Unemployment Rate 2017	-0.882361233	0.931913963
Substance Abuse Services * Crime Rate	0.815079176	0.93359569
Home Based Services * Current Placement Setting	0.103263132	0.938292291

Public Assistance * Child Age	-0.455389216	0.94215632
Removal Inadequate Housing * Crime Rate	-0.268758374	0.953101152
Employment Services * Poverty Rate 2017	0.321760174	0.966880434
Adoption Services * Median Household Income 2017 ($)	0.144041163	0.968219981
Removal Parents Jail * Poverty Rate 2017	0.157209718	0.981474007
Monthly FC Payment * Caretaker Emotional Problems	-0.000292947	0.999888925

Appendix F: Number of Services Available

Services	Number Available
Case Management Services	2209
Daycare Services	745
Family Preservation Services	316
Counseling Services	302
Substance Abuse Services	280
Pregnancy and Parenting Services	212
Mental Health Services	167
Informational and Referral Services	115
Special Services Disabled	110
Home Based Services	92
Legal Services	83
Respite Services	60
Transportation Services	46
Health and Home Health Services	35
Family Planning Services	17
Employment Services	13
Housing Services	13
Family Support Services	12
Education Services	12
Other Services	9
Adoption Services	7

Appendix G: Optimization Python Script

```
from gurobipy import *
import pandas as pd
# Model is divided into five main components: (1) Inputting files, (2) building lists for variables with regression coefficients, (3) building optimization model, (4) creating decision variables, and (5) adding constraints

# -----------
# Input Files
# -----------

# Read entire dataset
complete_data_set = pd.read_csv('Optimization CSVs\CompleteDataSet.csv', low_memory=False)

# Read file with 'Child Services' and create list
child_services_df = pd.read_csv('Optimization CSVs\ChildServices.csv')
child_services = child_services_df.columns.values.tolist()

# Read file with 'Family Services' and create list
family_services_df = pd.read_csv('Optimization CSVs\FamilyServices.csv')
family_services = family_services_df.columns.values.tolist()

# Read file with 'Non-service variables' (environmental factors or variables that are not services) and create list
Non_Service_Variables_df = pd.read_csv('Optimization CSVs\EnvFactors.csv')
Non_Service_Variables = Non_Service_Variables_df.columns.values.tolist()

# Read file that defines the amount of each service to allocate
service_limits_df = pd.read_csv('Optimization CSVs\ServiceLimits.csv')

# Read file with regression coefficients and create list
regression_coefficients_df = pd.read_csv('Optimization CSVs\RegressionCoefficients.csv')
included_variables = regression_coefficients_df.columns.values.tolist()

# Create list of all different families
```

```
familyList = []
for x in complete_data_set['Report ID']:
    if x not in familyList:
        familyList.append(x)

# Create dictionary where key is family ID and value is number of children in the family
familyLength = {}
for x in familyList:
    sum = 0
    for y in complete_data_set['Report ID']:
        if x == y:
            sum = sum + 1
            familyLength.update({y : sum})

# -----------
# BUILD LIST FOR EACH PERMUTATION
# -----------

# Create list of 'Child Services' that appear in any regression coefficients (List CS)
CS = []
for variable in included_variables:
    for service in child_services:
        if service in variable:
            if service not in CS:
                CS.append(service)
print("CS List: " + str(CS))

# Create list of 'Family Services' that appear in any regression coefficients (List FS)
FS = []
for variable in included_variables:
    for service in family_services:
        if service in variable:
            if service not in FS:
                FS.append(service)
print("FS List: " + str(FS))

# Create list of 'Child Services x Child Services' with regression coefficients (List CSCS)
CSCS = []
```

```
for variable in included_variables:
    for service1 in child_services:
        for service2 in child_services:
            if service1 != service2:
                if service1 in variable and service2 in variable and ' * ' in variable:
                    CCInteraction = [service1, service2]
                    CCInteractionReverse = [service2, service1]
                     if CCInteraction not in CSCS and CCInteractionReverse not in CSCS:  # Check
reverse isn't included
                        CSCS.append(variable.split(' * '))
print("CSCS List: " + str(CSCS))

# Create list of 'Child Services x Family Services' with regression coefficients (List CSFS)
CSFS = []
for variable in included_variables:
    for service1 in child_services:
        for service2 in family_services:
            if service1 in variable and service2 in variable and ' * ' in variable:
                CFInteraction = variable.split(' * ')
                CSFS.append(CFInteraction)
print("CSFS List: " + str(CSFS))

# Create list of 'Family Services x Family Services' with regression coefficients (List FSFS)
FSFS = []
for variable in included_variables:
    for service1 in family_services:
        for service2 in family_services:
            if service1 != service2:
                if service1 in variable and service2 in variable and ' * ' in variable:
                    FFInteraction = [service1, service2]
                    FFInteractionReverse = [service2, service1]
                         if FFInteraction not in FSFS and FFInteractionReverse not in FSFS:  # Check
reverse isn't included
                        FSFS.append(variable.split(' * '))
print("FSFS List: " + str(FSFS))

# Create list of 'Non-service Variables x Child Services' with coefficients (List NSCS)
NSCS = []
```

```
service_interactions = CSCS + CSFS + FSFS
for variable in included_variables:
    for service in child_services:
        if service in variable and ' * ' in variable and variable not in service_interactions:
            individual_terms = variable.split(' * ')
            individual_terms.insert(0, service)
            individual_terms = list(dict.fromkeys(individual_terms))  # Deletes duplicates
            NSCS.append(individual_terms)
print("NSCS List: " + str(NSCS))

# Create list of 'Non-service Variables x Family Services' with coefficients (List NSFS)
NSFS = []
for variable in included_variables:
    for service in FS:
        if service in variable and ' * ' in variable and variable not in service_interactions:
            individual_terms = variable.split(' * ')
            individual_terms.insert(0, service)
            individual_terms = list(dict.fromkeys(individual_terms))  # Deletes duplicates
            NSFS.append(individual_terms)
print("NSFS List: " + str(NSFS))

# Create list of 'Non-service Variables' with regression coefficients. . . .  . . . . Constant value
NS = []
for variable in included_variables:
    for nsVariable in Non_Service_Variables:
        if nsVariable == variable:
            NS.append(variable)
print("NS List: " + str(NS))

# Create list of 'Non-service Variables x Non-service Variables' with coefficients. . .Constant
Value
NSNS = []
for nsVariable1 in Non_Service_Variables:
    for nsVariable2 in Non_Service_Variables:
        if nsVariable1 + " * " + nsVariable2 in included_variables:
            NSInteraction = [nsVariable1, nsVariable2]
            NSNS.append(NSInteraction)
print("NSNS List: " + str(NSNS))
```

```
# --------------
# BUILD MODEL
# --------------

# Define the Model: Create Gurobi optimization model named FosterCare
mod = Model('foster')

# Create a list of all child IDs
children = complete_data_set['Child ID AFCARS'].tolist()
count_row = complete_data_set.shape[0]

# Set Index: Change the index of the optimization dataframe to be Child ID
complete_data_set.set_index('Child ID AFCARS', inplace=True)

# Method to remove spaces in a string. This is used to create variable names below without spaces

def remove(string):
    string = str(string)
    return string.replace(" ", "")

# ----------
# DECISION VARIABLES
# ----------

# Create decision variables (s) to determine if a child receives a child level service
s = {}
for child in children:
    for service in CS:
        s[child, service] = mod.addVar(vtype=GRB.BINARY, name='s(' + str(child) + ')(' + remove(service) + ')')

# Create decision variables (f) to determine if a family receives a family level service
f = {}
for family in familyList:
    for service in FS:
```

```
            f[family, service] = mod.addVar(vtype=GRB.BINARY, name='f(' + str(child) + ')(' +
remove(service) + ')')

# Create decision variables (j) to assign a value to each 'Child Service' regression coefficient
j = {}
for child in children:
    for variable in included_variables:
        if variable in CS:
                    j[child, variable] = mod.addVar(obj=regression_coefficients_df[variable][0],
vtype=GRB.BINARY,
                                name='j(' + str(child) + ')(' + remove(variable) + ')')

# Create decision variables (k) to assign a value to each 'Family Service' regression coefficient

k = {}
for family in familyList:
    for variable in included_variables:
        if variable in FS:
            objective = regression_coefficients_df[variable][0] * familyLength[family]
            k[family, variable] = mod.addVar(obj=objective, vtype=GRB.BINARY,
                                name='k(' + str(child) + ')(' + remove(variable) + ')')

# Create decision variables (m) to assign a value to each 'Child Service x Child Service'
regression coefficient\
m = {}
for child in children:
    for interaction in CSCS:
        for service in CS:
            if service == interaction[0]:
                m[child, interaction[0], interaction[1]] = mod.addVar(
                        obj=regression_coefficients_df[interaction[0] + ' * ' + interaction[1]][0],
vtype=GRB.BINARY,
                name='m(' + str(child) + ')(' + str(interaction[0]) + ')(' + str(interaction[1]) + ')')

# Create decision variables (n) to assign a value to each 'Child Service x Family Service'
regression coefficient
n = {}
for child in children:
```

```
    for interaction in CSFS:
        for service in CS:
            if service in interaction[0]:
                n[child, interaction[0], interaction[1]] = mod.addVar(
                                obj=regression_coefficients_df[interaction[0] + ' * ' + interaction[1]][0],
vtype=GRB.BINARY,
                    name='n(' + str(child) + ')(' + str(interaction[0]) + ')(' + str(interaction[1]) + ')')
        for service in FS:
            if service == interaction[0]:
                n[child, interaction[0], interaction[1]] = mod.addVar(
                                obj=regression_coefficients_df[interaction[0] + ' * ' + interaction[1]][0],
vtype=GRB.BINARY,
                    name='n(' + str(child) + ')(' + str(interaction[0]) + ')(' + str(interaction[1]) + ')')

# Create decision variables (p) to assign a value to each 'Family Service x Family Service'
regression coefficient
p = {}
for family in familyList:
    for interaction in FSFS:
        for service in FS:
            if service == interaction[0]:
                    objective = (regression_coefficients_df[interaction[0] + ' * ' + interaction[1]][0]) *
familyLength[family]
                p[family, interaction[0], interaction[1]] = mod.addVar(
                    obj= objective, vtype=GRB.BINARY,
                    name='p(' + str(child) + ')(' + str(interaction[0]) + ')(' + str(interaction[1]) + ')')

# Create decision variables (q) to assign a value to each 'Child Service x Non-service Variable'
regression coefficient
obj_NSCS_value = 0
q = {}
for child in children:
    for service in CS:
        for interaction in NSCS:
            if service == interaction[0]:
                if interaction[0] + ' * ' + interaction[1] in included_variables:
                                obj_NSCS_value = regression_coefficients_df[interaction[0] + ' * ' +
interaction[1]][0] * \
```

```
                    complete_data_set[interaction[1]][child]
                q[child, interaction[0], interaction[1]] = mod.addVar(obj=obj_NSCS_value,
vtype=GRB.BINARY,
                                              name='q(' + str(child) + ')(' + str(
                                                interaction[0]) + ')(' + str(
                                                interaction[1]) + ')')

        else:
                    obj_NSCS_value = regression_coefficients_df[interaction[1] + ' * ' +
interaction[0]][0] * \
                    complete_data_set[interaction[1]][child]
                q[child, interaction[0], interaction[1]] = mod.addVar(obj=obj_NSCS_value,
vtype=GRB.BINARY,
                                              name='q(' + str(child) + ')(' + str(
                                                interaction[0]) + ')(' + str(
                                                interaction[1]) + ')')

# Create decision variables (r) to assign a value to each 'Family Service x Non-service Variable'
regression coefficient
obj_NSFS_value = 0
r = {}
for family in familyList:
  for service in FS:
    for interaction in NSFS:
      if service == interaction[0]:
        if interaction[0] + ' * ' + interaction[1] in included_variables: #If service comes first
          obj_NSFS_value = 0
          for child in children:
            if complete_data_set["Report ID"][child] == family:
                                                    obj_NSFS_value = obj_NSFS_value +
regression_coefficients_df[interaction[0] + ' * ' + interaction[1]][0] * \
                    complete_data_set[interaction[1]][child]
                r[family, interaction[0], interaction[1]] = mod.addVar(obj=obj_NSFS_value,
vtype=GRB.BINARY,
                                              name='r(' + str(family) + ')(' + str(
                                                interaction[0]) + ')(' + str(
                                                interaction[1]) + ')')
        else:
```

```
                obj_NSFS_value = 0
                for child in children:
                    if complete_data_set["Report ID"][child] == family:
                        obj_NSFS_value = obj_NSFS_value + \
                                regression_coefficients_df[interaction[0] + ' * ' + interaction[1]][0] *
\
                                complete_data_set[interaction[1]][child]
                r[family, interaction[0], interaction[1]] = mod.addVar(obj=obj_NSFS_value,
vtype=GRB.BINARY,
                                                    name='r(' + str(family) + ')(' + str(
                                                        interaction[0]) + ')(' + str(
                                                        interaction[1]) + ')')

# Assign Value to constant term: 'Non-Service Variable' (t)
constant_value = 0
for child in children:
        constant_value = constant_value + regression_coefficients_df['const'][0]    # Assigns
y-intercept value to each child
    for variable in NS:
        constant_value = constant_value + (
            regression_coefficients_df[variable][0] * complete_data_set[variable][child])

# Assign Value to constant term: 'Non-Service Variable x Non-Service Variable' (t)

for child in children:
    for variable in Non_Service_Variables:
        for interaction in NSNS:
            if variable == interaction[0]:
                constant_value = constant_value + (
                    regression_coefficients_df[interaction[0] + ' * ' + interaction[1]][0] *
                                                    complete_data_set[interaction[0]][child]    *
complete_data_set[interaction[1]][child])

constant_list = [1]
t = {}
for items in constant_list:
    t[items] = mod.addVar(obj=constant_value, vtype=GRB.BINARY, name='ConstantTerm')
```

```
mod.update()

# ----------
# CONSTRAINTS
# ----------

# 1. Constant Term stays constant for items in constant_list:
    mod.addConstr(t[items] == 1)
# 2. For each service, the number of Child Services allocated is EQUAL to the original number of Child Services

for service in CS:
    mod.addConstr(quicksum([s[child, service] for child in children]) == service_limits_df[service][0])

# 3. For each service, the number of Family Services allocated is EQUAL to the original number of Family Services

for service in FS:
    mod.addConstr(quicksum([f[family, service] for family in familyList]) == service_limits_df[service][0])

# 4. For each child, the number of Child Services (s) allocated is equal to the amount of auxiliary variable j

for child in children:
    for variable in included_variables:
        if variable in CS:
            mod.addConstr(j[child, variable] == s[child, variable])

# 5. For each child, the number of Family Services (f) allocated is equal to the amount of auxiliary variable k

for family in familyList:
    for variable in included_variables:
        if variable in FS:
            mod.addConstr(k[family, variable] == f[family, variable])
```

```
# 6. For each child, the number of Child Services (m) allocated is equal to the amount of auxiliary variable m

for child in children:
    for interaction in CSCS:
        for service in CS:
            if service == interaction[0]:
                mod.addConstr(m[child, interaction[0], interaction[1]] <= s[child, interaction[0]])
                mod.addConstr(m[child, interaction[0], interaction[1]] <= s[child, interaction[1]])
                mod.addConstr(
                    m[child, interaction[0], interaction[1]] >= s[child, interaction[0]] + s[child, interaction[1]] - 1)

# 7. For each child, the number of Child Services (m) and Family Services (f) allocated is equal to the amount of auxiliary variable n

for child in children:
    for interaction in CSFS:
        for service in CS:
            if service == interaction[0]:
                mod.addConstr(n[child, interaction[0], interaction[1]] <= s[child, interaction[0]])
                mod.addConstr(n[child, interaction[0], interaction[1]] <= f[child, interaction[1]])
                mod.addConstr(
                    n[child, interaction[0], interaction[1]] >= s[child, interaction[0]] + f[child, interaction[1]] - 1)
        for service in FS:
            if service == interaction[0]:
                mod.addConstr(n[child, interaction[0], interaction[1]] <= f[child, interaction[0]])
                mod.addConstr(n[child, interaction[0], interaction[1]] <= s[child, interaction[1]])
                mod.addConstr(
                    n[child, interaction[0], interaction[1]] >= f[child, interaction[0]] + s[child, interaction[1]] - 1)

# 8. For each child, the number of Family Services (f) allocated is equal to the amount of auxiliary variable p

for family in familyList:
    for interaction in FSFS:
```

```
        for service in FS:
            if service == interaction[0]:
                mod.addConstr(p[family, interaction[0], interaction[1]] <= f[family, interaction[0]])
                mod.addConstr(p[family, interaction[0], interaction[1]] <= f[family, interaction[1]])
                mod.addConstr(
                        p[family, interaction[0], interaction[1]] >= f[family, interaction[0]] + f[family,
interaction[1]] - 1)

# 9. For each child, the number of Child Services (s) allocated is equal to the amount of auxiliary
variable q

for child in children:
    for interaction in NSCS:
        for service in CS:
            if service == interaction[0]:
                mod.addConstr(q[child, interaction[0], interaction[1]] == s[child, interaction[0]])

# 10. For each child, the number of Family Services (f) allocated is equal to the amount of
auxiliary variable r
for family in familyList:
    for interaction in NSFS:
        for service in FS:
            if service == interaction[0]:
                mod.addConstr(r[family, interaction[0], interaction[1]] == f[family, interaction[0]])

# 11. Each child must spend a minimum of zero days in care. (Non-negative)
# Calculate the total days in care for each child with the service decision variables
# (Only calculates NS, NSNS, NSCS and NSFS because they are the only permutations with
values

eachChildsStay = {}
for child in children:
    los = 0
    los = regression_coefficients_df['const'][0]
    for x in NS:
        los = los + complete_data_set[x][child] * regression_coefficients_df[x][0]
    for y in NSNS:
        los = los + (complete_data_set[y[0]][child] * complete_data_set[y[1]][child] *
```

```
                regression_coefficients_df[y[0] + ' * ' + y[1]][0])
    for z in NSCS:
        los = los + (s[child, z[0]] * complete_data_set[z[1]][child] *
                regression_coefficients_df[z[0] + ' * ' + z[1]][0])
    for u in NSFS:
        for fam in familyList:
            if complete_data_set['Report ID'][child] == fam:
                los = los + (f[fam, u[0]] * complete_data_set[u[1]][child] *
                        regression_coefficients_df[u[0] + ' * ' + u[1]][0])
    mod.addConstr(los >= 0)

mod.update()

# ---------
# WRITE AND SOLVE MODEL
# ---------

mod.write('foster.lp')
mod.optimize()
mod.write('foster.sol')

print("The total number of days in care is " + str(mod.objVal))
print("The number of children in this set: " + str(count_row))
print("The average length of stay in care is " + str(mod.objVal / count_row))
```

Appendix H: Balancing Objectives Python Script

```
from gurobipy import *
import pandas as pd
import os
import re
import datetime
import timeit
import numpy as np

def splitTermsToFactors(terms_raw):
    terms = [k.replace('" * "', '*') for k in terms_raw]
    terms = [re.split(r'"\*"', k) for k in terms]
    return terms

def getFactorVal(child, factor):
# this function returns the value of a factor depending on whether it's
# a fixed constant from the environmental data
# or a decision variable, meaning a service
    if factor in EnvList:
        val = caseData[factor][child]
    elif factor in ServList:
        val = x[child, factor]
    else:
        print('FAILURE to categorize factor:', factor)
    return val

def getFactorVal_Opt(child, factor):
# this function returns the value of a factor depending on whether it's
# a fixed constant from the environmental data
# or a decision variable, meaning a service AFTER optimization assignment
    if factor in EnvList:
        val = caseData[factor][child]
    elif factor in ServList:
        val = x[child, factor].X
    else:
        print('FAILURE to categorize factor:', factor)
```

```
    return val

def printStats():
    maxLos = 0
    minLOS = 3000
    avgLOS = sum(LOS_List) / len(LOS_List)
    sumLOS_pos = 0
    count_pos = 0
    for los in LOS_List:
        if los > maxLos:
            maxLos = los
        if los < minLOS:
            minLOS = los
        if los > 0:
            sumLOS_pos += los
            count_pos += 1
    print('ALPHA:', alpha)
    print('max days in care: ', maxLos)
    print('min days in care: ', minLOS)
    print('avg days in care:', avgLOS)
    print(LOS_List)
    return minLOS, maxLos, avgLOS, sumLOS_pos/count_pos

# start timer
start = timeit.default_timer()
print(datetime.datetime.now())

# Read in data, create lists of children, families, and first children
if True:

os.chdir(r'\\research.wpi.edu\BUS\atrapplab\MQP_Foster_Care_Research_2019\MQP_2020\Opt
imization CSVs')

    # Read entire data set, create column for constant
    caseData = pd.read_csv('CompleteDataSet.csv', low_memory=False)
    caseData['const'] = 1
    caseData.fillna(value=0, inplace=True)
```

```
    # create list of all children IDs
    ChildList = caseData['Child ID AFCARS'].tolist()

    # Create list of all families (report IDs) and remove dupliactes
    FamList = caseData['Report ID'].tolist()
    FamList = list(set(FamList))

    famIDs = caseData['Report ID'].tolist()
    childIDs = caseData['Child ID AFCARS'].tolist()

    # make a dictionary of representative children {fam : first child from a family}
    # convert the values to a list to get a list of first children
    FirstChildDict = {}
    index = -1
    for fam in famIDs:
        index += 1
        FirstChildDict[famIDs[index]] = childIDs[index]
    FirstChildList = list(FirstChildDict.values())

    # create a dictionnary of all children and families {child id: family id}
    ChildFamDict = {k: v for k, v in zip(childIDs, famIDs)}

    # create a dictionary of families and all children in family {fam: [list of children]}
    FamChildDict = {}
    for fam in famIDs:
        list = []
        for child in childIDs:
            if ChildFamDict[child] == fam:
                list.append(child)
        FamChildDict[fam] = list

    # set index of data set to be Child ID for quick searching later
    caseData.set_index('Child ID AFCARS', inplace=True)

# Read in files, create lists of raw terms (not interactions) and limits
if True:
    # Read file with 'Child Services' and create list
    childServDF = pd.read_csv('ChildServices.csv')
```

```
    ChildServList = childServDF.columns.values.tolist()

    # Read file with 'Family Services' and create list
    famServDF = pd.read_csv('FamilyServices.csv')
    FamServList = famServDF.columns.values.tolist()
    FamServList.remove('Post Investigation Services')

    # Read file with 'Non-service variables' (environmental factors or variables that are not
services) and create list
    envDF = pd.read_csv('EnvFactors.csv')
    EnvList = envDF.columns.values.tolist()
    EnvList.append('const')

    # create list of all services & a list of all terms (as in raw factors, no interaction terms)
    ServList = ChildServList + FamServList
    TermList = ServList + EnvList

    # Read file with regression coefficients and create list
    r = pd.read_csv('RegressionCoefficients_New.csv')
    SignifTermList = r.columns.values.tolist()
    SignifTermList = splitTermsToFactors(SignifTermList)

    # Read file that contains the service limits (number of units to allocate)
    serviceLims = pd.read_csv('ServiceLimits.csv')
    # serviceLims = pd.read_csv('ServiceLimits_SimpleCount.csv')

resultsList = []  # initialize list to store the results of each run
for alpha in np.linspace(0, 1, 11):  # use 11 for 1/10 step size: (0, 0.1, 0.2, ... 1)
    print('RUNNING OPTIMIZATION FOR ALPHA =', alpha)

    # initialize model and decision variables, set objective
    if True:
        # Initialize Gurobi optimization model named FosterCare
        mod = Model('FosterCare')

        # Initialize decision variables (s) to determine if a child receives a service (child or family
level)
        x = {}
```

```
        count = 0
        for child in ChildList:
            for serv in ServList:
                count += 1
                    x[child, serv] = mod.addVar(vtype=GRB.BINARY, name=str(child) + '_' +
serv.replace(' ', ''))
        print('successfully initiated', count, 'decision variables')

        max = mod.addVar(vtype=GRB.INTEGER, name="max")
        net = mod.addVar(vtype=GRB.INTEGER, name="net")
        mod.update()

        # mod.setObjective(max, GRB.MINIMIZE)       # goal is to minimize the maximum days in
care
        # mod.setObjective(net, GRB.MINIMIZE)       # goal is to minimize net days in care
         mod.setObjective((alpha*max) + ((1-alpha) * net/len(ChildList)))  # goal is minimize both
Max LOS and Avg LOS


    # calculate the LOS of a child and add the constraint that is must be less than maxLOS
    if True:
        LOS_List = []
        for child in ChildList:
            LOSc = 0
            for term in SignifTermList:
                if len(term) == 1:
                    # if terms is not an interaction
                    LOSc += r[term[0]][0] * getFactorVal(child, term[0])
                elif len(term) == 2:
                    # if term is an interaction
                    if (term[0] in ServList) and (term[1] in ServList):
                        # if the term is an interaction of two services
                        # additional constraints are needed to linearize
                        print('Warning: Introducing interaction term!')
                        LOSc += r[term[0] + term[1]][0] * x[child, term]
                        mod.addConstr(x[child, term] <= x[child, [term[0]]])
                        mod.addConstr(x[child, term] <= x[child, [term[1]]])
                        mod.addConstr(x[child, term] >= x[child, [term[0]]] + x[child, [term[0]]] - 1)
```

```
                if (term[0] in EnvList) or (term[1] in EnvList):
                    # one or both factors in interaction term are non-service
                    # no linearization required, at least one factor is a constant
                        LOSc += r[term[0] + ' * ' + term[1]][0] * getFactorVal(child, term[0]) *
getFactorVal(child, term[1])
            else:
                print('FAILURE to categorize by length: \n', term)
            LOS_List.append(LOSc)
        mod.addConstr(LOSc - max <= 0)
        mod.addConstr(LOSc >= -1876.535383)
    print('\ndefined LOS for each child')

    # calculate net LOS and add variable to cap it
    mod.addConstr(quicksum(los for los in LOS_List) <= net)

    for serv in ServList:
        if (serv in ChildServList):
            # for any child service the total number of units given out must EQUAL the limit
                mod.addConstr(quicksum([x[child, serv] for child in ChildList]) ==
serviceLims[serv][0])
            # print('constrained child service', serv, '=', serviceLims[serv][0])

        if serv in FamServList:
        # children in same family receive same amount of service
            for fam in FamList:
                for child in FamChildDict[fam]:
                    mod.addConstr(x[child, serv] - x[FirstChildDict[fam], serv] == 0)
                        mod.addConstr(quicksum([x[child, serv] for child in FirstChildList]) ==
serviceLims[serv][0])
            # print('constrained family service ', serv, ' == ', serviceLims[serv][0])
    print('made family level limit constraints\n')

    # run optimization and print results
    if True:
        mod.update()
        # mod.write('foster.lp')
        mod.optimize()
```

```
        # mod.write('foster.sol')
        print('PRINTING RESULTS FOR ALPHA =', alpha)
        print('----------------------\nSTATS FROM OPT MODEL')
        print("The obj val is: ", mod.objVal)
        print('z is: ',max.X)
        print('net is: ', net.X)
        print("The number of children in this set: ", len(ChildList))
        print("The obj val / # children is: ", mod.objVal/len(ChildList))

    # calculate the LOS of a child post optimization, save to list
    if True:
        LOS_List = []
        for child in ChildList:
            LOSc = 0
            for term in SignifTermList:
                if len(term)==1:
                # if terms is not an interaction
                    LOSc += r[term[0]][0]*getFactorVal_Opt(child, term[0])
                elif len(term)==2:
                # if term is an interaction
                        LOSc += r[term[0]+' * '+term[1]][0] * getFactorVal_Opt(child, term[0]) *
getFactorVal_Opt(child, term[1])
                else:
                    print('FAILURE to calculate LOS after assignment: \n', term)
            LOS_List.append(LOSc)
        print('----------------------\nSTATS FROM POST OPT (MANUAL)')
        minLOS, maxLos, avgLOS, avgLOS_noNegatives = printStats()
        resultsAlpha = {'ALPHA':[alpha], 'OBJ VAL': [mod.objVal], 'POST MIN LOS': [minLOS],
'POST MAX LOS': [maxLos],
                    'POST AVG LOS': [avgLOS], 'POST AVG - AS 0': [avgLOS_noNegatives],
'Max':max.X, 'net': net.X}
        resultsDF_alpha = pd.DataFrame(data=resultsAlpha)
        resultsList.append(resultsDF_alpha)

    stop = timeit.default_timer()
    print('Time: ', stop - start)
    resultsDF = pd.concat(resultsList)
    resultsDF.to_csv('optimizationResultsByAlpha_'+str(alpha)+'.csv')
```

```
    print(resultsList)

# calculate the LOS of a child before optimization, save to list
if False:
    LOS_List = []
    for child in ChildList:
        LOSc = 0
        for term in SignifTermList:
            if len(term) == 1:
                # if terms is not an interaction
                LOSc += r[term[0]][0] * caseData[term[0]][child]
            elif len(term) == 2:
                # if term is an interaction
                LOSc += r[term[0] + ' * ' + term[1]][0] * caseData[term[0]][child] * caseData[term[1]][child]
            else:
                print('FAILURE to calculate LOS after assignment: \n', term)
        LOS_List.append(LOSc)

    print('-----------------------\nSTATS FROM PRE-OPT')
    minLOS, maxLos, avgLOS, avgLOS_noNegatives = printStats()
     # resultsAlpha.update({'PRE MIN LOS': minLOS, 'PRE MAX LOS': maxLos, 'PRE AVG LOS': avgLOS, 'PRE AVG - AS 0': avgLOS_noNegatives})
```

Appendix I: Balancing Objectives Data

Beta	Avg Length of Stay (Goal 1)	Max Length of Stay (Goal 2)	Objective Value
0.000	387.691	994.303	387.691
0.050	387.691	965.557	416.585
0.100	387.691	965.557	445.478
0.150	387.691	965.557	474.371
0.200	387.692	963.255	502.805
0.250	387.692	963.255	531.583
0.300	387.692	963.255	560.361
0.350	387.692	963.255	589.139
0.400	387.697	960.647	616.877
0.450	387.697	960.647	645.524
0.500	387.697	960.647	674.172
0.550	387.697	960.647	702.819
0.600	387.697	960.647	731.467
0.650	387.697	960.647	760.114
0.700	387.697	960.647	788.762
0.750	387.697	960.647	817.409
0.800	387.697	960.647	846.057
0.850	387.697	960.647	874.704
0.900	387.697	960.647	903.352
0.950	387.697	960.647	931.999
1.000	402.179	960.647	960.647