**Project Proposal**

**Executive Summary:**

**Business Objectives:**

This project will answer two questions:

* How is juvenile incarceration influenced by the Foster Care System?
* Does the age group of children exiting Foster Care influence juvenile incarceration.

**Background:**

The Foster Care System is a necessary government program meant to help children who are orphaned or removed from parental care; however, there are unfortunate outcomes in the system. There is a Foster Care to Prison pipeline. There are studies showing 80% of inmates in prison have spent time in foster care. We are checking if there is also a link between the Foster Care System and juvenile incarceration. We are also determining if the age group of children exiting the Foster Care has any effect on the juvenile incarceration numbers.

**Scope:**

* The programs we are willing to use are:
  + Python
  + R Studio
  + GitHub/GitKraken
  + Trello
  + Zoom
  + Slack
  + Tableau
  + Microsoft Office
  + Jupyter Notebook
  + GPower
* Any programs not listed will not be used
* All skills needed for this project have been acquired. There will be no need for any additional skills.
* Willing to work at least 20 hours a week to finish the project. If needed, willing to work more than 20 hours to finish.
* If needed, changing data or adding necessary data is encouraged and permitted
* We are looking to use Independent Chi-square and analyses.
* Machine Learning is not one of the analyses that will be completed.
* Any statistical analyses that are needed will be completed or applied.
* Big data is not required to complete this project.

**Functional Requirements:**

* Python/ R Studio
* Statistical data on children in the Foster Care System by age from 2001-2019. All raw data is from:

**Personnel Requirements:**

Douglas Bell – 20 hours

Rushelle Phillips – 20 hours

Rachel Korman – 20 hours

Joseph Raetano – 20 hours

Group meetings will be at least 1 hour a week.

Meetings with Joseph(instructor) 30 minutes a week.

Following team meetings will be planned at the end of the current sprints planned meeting. Instructor meetings will be played by ear.

A Saturday through Friday week will be adopted for this project. Saturdays will be the start of a sprint with Fridays being the end of the sprint. To best adapted to the websites time line.

**Delivery Schedule:**

* Week 1: Project Planning
  + Form a group
  + Agree on the topic for the project
  + Turn in both a primary dataset and secondary dataset
  + Formulate at least two evaluation questions for our dataset.
  + Meet on a zoom call
  + Collaborate on the Project Proposal
  + December 17th is the cut off for this Sprint
* Week 2: Data Wrangling - Douglas (s.m.)
  + From December 18th through December 31st no work is required to be done.  If available or willing, during these two weeks progression on wrangling the data is encouraged. Doing so will put the project ahead of schedule
  + January 1st officially begins this Sprint
  + Starting January 1st with the data wrangled is preferred. At minimum this data begins the wrangling process.
  + January 7th is the cut off for this Sprint.  All wrangling must be done or progess is behind.
  + Import datasets into R
  + Combine the two datasets into one dataframe
  + Drop unnecessary columns
  + Reformat any data structure (.int, .str, etc)
  + Upload codes to GitHub
* Week 3: Data Exploration - Rushelle (s.m.)
  + January 8th begins this sprint
  + This sprint ends January 14th
  + Check normality of each dataset
* Week 4: Data Analysis - Rachel (s.m)
  + January 15th begins this sprint
  + January 21st ends this sprint
  + Start with eval question on. Begin the predetermined analysis, Independent Chi-square.
  + For eval question two perform the predetermined analysis, Independent Chi-square.
* Week 5: Data Visualization - Douglas (s.m)
  + Gather all data and results
  + Work in Tableau to get graphs and charts for the PowerPoint from the results from the Data Analysis week.
  + Each member will
* Week 6: Data Reporting and Presentation - Rushelle (s.m)

**Other Requirements:**

**Assumptions:**

* Python/R Studio
* GitHub/GitKraken
* Zoom
* Slack communications notifying us on new messages

**Limitations:**

* Scheduling both a group meeting and instructor meeting with members in different time zones. Being flexible and understanding with others and their personal time and limitations.
* Little knowledge (improper training) on GitHub/GitKraken and Trello. Leading to small mistakes taking up too much time to resolve. Willingness to look up web tutorials or videos with clear instructions. Asking instructor or mentors for clarification on those websites/programs.
* Lack of commitment. As a group we need to encourage each team member.

**Risks:**

* Work schedule interfering with delivery schedule. Ask work supervisor to work with the groups schedule
* Poor Internet service/Internet service failing due to weather. Find a friend or family to their internet access. Or a library’s internet access. Mobile access is also a work around.
* Family matters (ex. death of family member). Be understanding and willing to take on team member’s tasks, if needed.