Project Deliverable 0

Our project is mostly involved with analyzing caseworker’s note and tries to find a relation between the backgrounds of the youths and the re-enrollment in the program.

We also will do a sentimental analysis on caseworkers’ notes, trying to find how caseworkers’ use of languages affect this re-enrollment. And current difficulties are: as caseworkers’ notes are sometimes colloquial and words are not complete, language usage by different races are different, so having a NLP analysis on notes are somewhat difficult.

As this project starts in summer, there are some preprocessing works done by the previous team: data cleaning, visualization, and topic modeling, and they have done some graphical analysis on the data.

Our goal is to continue doing NLP studies on the caseworkers’ notes and trying to find specific standards of whether the caseworker is successful or not by evaluating his/her notes.