Final Project Proposal

Narrative Track

Team members:
Xinwen Liu (xinwenl), Xinyu Lin (xinyulin), Yuxi Luo (yuxiluo), Shaobo Guan (shaobog)
Project link:
https://github.com/CMU-IDS-2020/fp-vectorization
Dataset link:
https://www.kaggle.com/donorschoose/io

Proposal:

Our team chooses the narrative track for this final project. We will analyze the donation statistics from *DonorsChoose.org* using the database provided by them on Kaggle. Our essential question is: what is the current state of funding for public schools. The initiative comes from news like "Why K-12 Education Needs More Federal Stimulus Funding", in which we learn that during the pandemic, public schools need more funding to provide quality education that will play an important role in students' learning outcomes. We would like to explore the question through two perspectives: the potential donors, and the platform.

From the potential donors' perspective, we would like to help them understand the challenges that public schools face and what their needs are. We also want to explore the dataset and find the most efficient way for them to find the projects they are interested in making donations. With the detailed data provided by the platform, we plan to display interactive visualizations to reveal useful information for potential donors. First of all, we plan to explore a geographical distribution of previous donation records, which could help to understand the donating demand across different regions. Next, it is also vital to see how the patterns of donations changed over the past years, for a good understanding about the historical trend. Furthermore, we plan to demonstrate detailed statistics (e.g., average / median / variance of individual donations) and attributes (e.g., types of physical resource donation and their proportions) inferred by the dataset. Through the visualization of donation statistics, the platform should be able to attract more potential donors, support more classrooms, and make wiser decisions in future information displays.

Further, we want to help the platform and the teachers in the platform to strategize their projects, so that they could get fundings in a timely manner. We plan to perform statistical analysis on the average duration of teacher's projects to get fully funded, and then explore the factors that can improve the speed for the projects to get funded faster. In this way, we can make suggestions to teachers so that their projects can be effective quicklier. For example, we plan to build a statistical model to analyze the project descriptions, and explore the relationship between the quality of a project description and the speed of that project to get fully funded. We will define the ways to illustrate the quality of the project descriptions (e.g., the length of the project description, the keywords that occur often in the description, etc.), and then explore the possible correlations.