**Project Title**

An analysis of US education and spending and student success in K-12

**Team Members**

1. Ana Saavedra
2. Analiza Asuncion
3. Ben Pollock
4. Eric Wetzel
5. Sahar Alaei

**Description**

The goal of this analysis is to determine whether K-12 school funding has an impact on student outcomes, specifically as they relate to standardized test scores and student graduation rates.

**Research Questions to Answer**

* Do districts with greater expenditures per student produce students who perform better on standardized tests (math and reading)? - Sahar
* Do districts with greater expenditures per student have higher graduation rates? - Analiza
* Is there a spending “sweet spot” – a point beyond which investing additional resources in students does not meaningfully affect scores and graduation rates? - Ben
* Does district size affect total spend or average spend per student? -Ana

**Datasets to be Used**

* districts.csv (the finances of elementary and high school data organized based on school districts)
* states.csv (the finances of elementary and high school data organized based on school states)
* naep.csv (a summary of data from the NAEP (National Assessment of Educational Progress))
* Public\_Schools.csv (all Public elementary and secondary education facilities in the USA)

**Rough Breakdown of Tasks**

* Exporting and importing the required datasets about US education system
* Cleaning and formatting datasets using Pandas
* Extracting and Merging needed data from collected datasets based on the needs and information decided to be visualized
* Analysis of data to answer the questions listed in the section “Research Questions to Answer” using VBA, Python, or/and Pandas.
* Creating a Jupyter Notebook describing the data exploration and cleanup process.
* Visualizing data and extracted information in different graphs such as pie, bar, scatter, line, and predicting some info for the future based on the available data using Python Matplotlib.
* Saving the PNG images of the visualizations.
* Developing the heatmap of school spends on gmaps using Python API
* Creating a Jupyter Notebook to illustrate the final data analysis.
* Writing the summary
* Creating the presentation slides (around 10 slides)

Visuals

* Heatmap of school spend on gmap
* Regression analysis on average expenditure per student and test scores
* Regression analysis on average expenditure per student and graduation rates
* Line graph of diminishing marginal utility (revenue/spend vs. outcome metric)

Bonus

Drill down into poorly funded schools that are high-performing and vice-versa

District spend over time