**PROJECT 3 – GROUP 1**

*Collaborators:*

Andrew Belfiglio, Katy Phillips, Shikha Patel, Jacob Field,

Travis Guadamuz Ruth, Peter Matsui, Michelle Wiley

*Project Guidelines:*

For Project 3, you will work with your group to tell a story using data visualizations. Here are the specific requirements:

1. Your visualization must include a
   * Python Flask-powered API,
   * HTML/CSS,
   * JavaScript, and
   * at least one database (SQL, MongoDB, SQLite, etc.).
2. Your project should fall into one of the following three tracks:
   * A combination of web scraping and Leaflet or Plotly
   * A dashboard page with multiple charts that update from the same data
   * A server that performs multiple manipulations on data in a database prior to visualization (**must be approved**)
3. Your project should include at least one JS library that we did not cover.
4. Your project must be powered by a dataset with at least 100 records.
5. Your project must include some level of user-driven interaction (e.g., menus, dropdowns, textboxes).
6. Your final visualization should ideally include at least three views.

*Topic:*

Analyze department of education data for the state of PA to determine if the “achievement rate” is comparable to the expenditure per pupil. Drill down into the districts of interest and look at additional mitigating factors that could also affect success rate and spend (ex. tax/income base, median household income, geographic considerations, student population)

*Possible Questions to Consider:*

* How can achievement be quantified? Statewide testing data - High School reading and math? Standardized test scores? Graduation rates?
* Are there areas where expenditure may be more but achievement rate lower?
* Do IEP and ESL achievement rates follow the same path as standard classes in schools/districts?
* Does the number of students or number of students per teacher play a role in achievement status or expenditure per pupil?
* Assumption: Spend is the same across all grades per pupil if budget is not broken down by school level.
* Why should everyone in the state care about educational achievements and expenditure?

*Project Outline:*

1. Choose Topic
2. Find data
3. Clean data
4. Flask, HTML, JavaScript, DataBase requirements
5. Visualizations
6. Dashboard
7. Analyze
8. Presentation

*Action Items:*

|  |  |  |  |
| --- | --- | --- | --- |
| Task | Responsible | Status | Due |
| Review test score and budget data. Small write up explaining the different columns in the spreadsheet | Katy | In Progress | R – 10/12 |
| County taxes and median income data in PA | Shikha  Michelle | In Progress | R – 10/12 |
| Geographic data for PA | Travis | In Progress | R - 10/12 |

*Possible Data Resources:*

* Department of Education - Data and Reporting
  + <https://www.education.pa.gov/DataAndReporting/Pages/default.aspx>
    - Enrollment Reports (2004-2023)
    - SAT Scores (2001-2019) and ACT Scores (2011-2019)
    - School Locale – “Urban/Rural” Classification of Schools and LEAs
* National Center for Education Statistics
  + <https://nces.ed.gov/>
    - Students per school district (2021-2023)
* School district boundaries:
  + <https://data.pa.gov/Geospatial-Data/Pennsylvania-School-Districts-Boundaries/s629-r52w>
* AFR Data: Summary Level Expenditure, revenue and tax information (2011-2022)
  + <https://www.education.pa.gov/Teachers%20-%20Administrators/School%20Finances/Finances/AFR%20Data%20Summary/Pages/AFR-Data-Summary-Level.aspx>
* Nation’s Report Card API
  + <https://educationdata.urban.org/documentation/>
* School District of Philadelphia
  + <https://www.philasd.org/performance/programsservices/open-data/>
* PA Dept of Community and Economic Development
  + https://munstats.pa.gov/Reports/ReportInformation2.aspx?report=CountyTaxSummary\_Dyn\_Excel

*Project Resources:*

* Medium - Flask Article
  + <https://towardsdatascience.com/talking-to-python-from-javascript-flask-and-the-fetch-api-e0ef3573c451>
* JavaScript Libraries
  + <https://underscorejs.org/> - easier to manipulate data
  + <https://lodash.com/> - data utility library
  + <https://www.algolia.com/> - autofill addresses\* (use for schools?)
  + <https://www.chartjs.org/> - chart library
* Map subplots in Plotly
  + <https://plotly.com/python/map-subplots-and-small-multiples/>
* Pictorial charts
  + <https://www.highcharts.com/docs/chart-and-series-types/pictorial>

*Random Idea Tracker for Inspiration:*

* Map based visualizations - State to District to Individual schools?
* Creates story – same type of graphs showing different views
* Single source data
* Dashboard by year – dropdown menu
* Historical Graph Over Time
* Hover data of either spend or achievement over map of school district/county
* Could this information help influence policy makers?
* Cluster analysis wealth bands?
* Choropleth map by county or school district view
* Markers in the county for the school district or summary statistics card if cannot segment by school district