**Final Project Description**

* Due Apr 21 by 9pm
* Points 0
* Submitting a website url

**Final project overview**

**The best neighborhood in Pittsburgh**

For the final project you will work as part of a team to answer the question: ***What is the best neighborhood in Pittsburgh?***Using data from the [WPRDCLinks to an external site.](https://data.wprdc.org/dataset/" \t "_blank)

To answer this question you need to do the following:

1. Come up with a team name!

* Team Awesome, Team Bogus, The Conquistadors…

1. As a group, come up with some ways of defining and measuring "bestness". This doesn't have to be a serious metric, it could be whimsical or secretly "worstness.”

* Most bars per square foot
* Most restaurants per square foot
* Most televisions per square foot
* Most asbestos permits
* Highest concentrations of parking meters
* Highest concentration of smart trash containers
* Landslide prone areas
* Ballfields
* Number of farmer’s markets
* Dog licenses
* Lowest number of covid related deaths
* Illegal dump sites
* Number of public swimming pools
* Number of bridges
* …

1. Each group member will be responsible for one submetric each.

* Matthew - crosswalks
* Chen – bus stops
* John – parking meters

1. Use at least 3 datasets in your argument. The easiest way to meet this requirement is for each member to choose their own unique dataset to analyze their submetric.
   1. Note that if you are working in a group of 2, you still need at least 3 datasets. One of you will have to analyze two datasets, or you will have to work together on analyzing a third dataset.

* Matthew -
* Chen -
* John -

1. Combine your sub-metrics into a single metric as you see fit.
2. Create a git repository to store your data, notebooks and code.

* https://github.com/fluxcapacitor88mph/CMPINF0010\_FinalProjectGroup32

**Notebook structure**

It is best to think about your final project as a data-driven report. You will need to put everything into a Jupyter notebook with the following structure:

* **Introduction:** Introduce the project, and your approach, talk about the process of how you came up with the metric and some alternatives you may have explored.
* **The Metric:** Describe your metric, and what features are you measuring. What datasets are you using?
* **The Best Neighborhood:** Apply the metric from the previous section to determine the best neighborhood in Pittsburgh. Beyond just executing code, provide a narrative about why you think this is the best neighborhood. Incorporate a data visualization, perhaps to rank all of the neighborhoods or show a neighborhood’s best-ness over time. The key is to make a data driven argument.
* **Conclusion:** Reflect on how the data-driven determination of “best neighborhood” is the same or different from your personal favorite neighborhood. Each member of the group should write their own response to this.

When your group presents your findings to your lab section (see below), the presentation will

**The Github repo**

There are some key things that need to be included in your repository. Think of this project as a combination of all of the things we learned in skills labs this semester, with a focus on the last few weeks. So, there are some minimum requirements of files and information you will want to put in your group repo.

* The final report Jupyter notebook
* Each group member's personal notebook(s) (Show your work)
* A Readme that contains:
  + Team name
  + Team members and their email addresses
  + Description and links to the datasets used
  + A one sentence overview of your repository, meant for someone who is not familiar with the project

**Final project timeline**

**Submitting your assignment**

**Please make sure your repository is public before you submit.**

Each group’s Github repository should contain:

* A Combined notebook / Presentation Notebook with a conclusion paragraph from each group member.
* A Jupyter Notebook from each group member with their dataset, analysis, and answer. This will allow us to verify that each team member pulled his/her weight.
* A Readme with the following information:
  + Team name
  + Team member's names and email addresses
    - Additionally, your README should tell us which individual analysis notebooks belong to which team members
  + Canvas group number
  + Description and Links to the datasets used
  + A one-sentence overview of your repository, meant for someone who is not familiar with the project