**Entrepreneur Research**

**Introduction/Background:**

This project is a continuation of Project I with two different team members. Our main objective is still the same ***“to assist entrepreneurs to identify potential locations for potential businesses”***. However, we would like to dynamically analyze and predict the potential locations for potential businesses based on client input on the web form.

This is a high-level analysis, national and state level, which may lead them to conduct more granular level research, county and city level, before coming to a decision???

**Scope:**

The following questions will be answered based on Census datasets within the past 5-10 years:

* What is the current U.S. population on national and state level?
* What is the demographic for all the states?
  + - Gender
    - Race
    - Household Size & Median Income
    - Education Level
* What is the crime rate for all the states?
* What type of businesses are in each state?
* What is the tax rate for each state?
* What is the consumer spending on retail category in each state?

**Team Members: 😊**

**entr\_researchers = (‘jane\_wallace’, ‘jose\_’, ‘alicia\_’, ‘charleen\_carr’,)**

**Team Members Availability:**

**Resources:**

* Potential Dataset:
  + - <https://www.yelp.com/dataset/download>
    - <https://www.census.gov/data/developers/data-sets.html>
    - <https://data.world/rickyhennessy/startup-names-and-descriptions>
    - <https://www.data.gov/developers/>
    - <https://developers.google.com/>
* Actual Dataset:
  + - [https://catalog.data.gov](https://catalog.data.gov/)
    - https://www.census.gov
    - [https://www2.census.gov](https://www2.census.gov/programs-surveys/susb/tables/2014/)
    - https://factfinder.census.gov

**Period of Performance:**

* **Start Date:**Saturday, April 13, 2019
* **Due Date:**Thursday, May 9, 2019

**Place of Performance:**

* ½ of work -> classroom on T/Th/S
* ¼ of work -> a group meeting on zoom
* ¼ of work -> individual assigned tasks

**Work Requirements:**

* Create a Jupyter Notebook describing the **\*\*data exploration and cleanup\*\*** process
* Create a Jupyter Notebook illustrating the **\*\*final data analysis\*\***
* Use Matplotlib to create a total of 6-8 visualizations of your data (ideally, at least 2  
  per "question" you ask of your data)
* Save PNG images of your visualizations to distribute to the class and instructional  
  team, and for inclusion in your presentation
* Optionally, use at least one API, if you can find an API with data pertinent to your  
  primary research questions
* Create a write-up summarizing your major findings. This should include a heading  
  for each "question" you asked of your data, and under each heading, a short  
  description of what you found and any relevant plots.

**Schedule/Milestones:**

|  |  |  |
| --- | --- | --- |
| **Date** | **Activities** | **Milestone** |
| **Start-up Week** (April, 6-12) | Team Formation & Project Planning | Project Plan & Flow Chart |
| **1st Week** (April, 13-19) | Research, Data Cleaning/Calculation, & Load to Database | Data Cleaning Files & Database |
| **2nd Week** (April, 20-26) | Data Analysis & Web Design | Website & Dynamic Charts & Map Files |
| **3nd Week** (April 27 - May 3) | Data Analysis & Prediction with Machine Learning | Web Files with Demo |
| **4nd Week** (May, 4-9) | Presentation & Practice | Project Plan Result & PP Slide |

**Acceptance Criteria:**

* A 10-minute, formal presentation
* Detail Explanation:
  + - The questions you and your group found interesting, and what motivated you to answer them
    - Where and how you found the data that you used to answer these questions
    - The data exploration and cleanup process (accompanied by your Jupyter Notebook)
    - The analysis process (accompanied by your Jupyter Notebook)
    - Your conclusions. This should include a numerical summary as well as  
      visualizations of that summary
    - Discuss the implications of your findings. This is where you get to have an  
      open-ended discussion about what your findings "mean".

**Other Requirements:**

* To be annouced

**Project Summary:**

* Actual Project Breakdown
  + - 2 Weeks -> Research & Dataset Selection
    - 1 Week -> Data Cleaning
    - 1 Week -> Data Analysis & Presentation
* Over-estimated Scope -> Re-Scope
  + - Research Finding
    - Huge Learning Curve
    - Team Dynamic
* Project Artifacts
  + - Presentation: Entrepreneurs Research.pptx
    - Data Cleaning Code: Group\_Data\_Clean.ipynb
    - Data Analysis Code: Group\_Data\_Analysis.ipynb
* Lesson Learned
  + - Flexible Team Player

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Acceptance:**

Approved by:

Date:

<Approvers Name> <Approvers Title>