

Ideas For Capstone Project 1

Please provide feedback on the following:

- Is this project useful to demonstrate real-life data science skills?
- Would companies care about this project?
- Is this project within the scope of Capstone 1 or are there significant obstacles ahead that I'm not aware of?

Project 1:

- **Link:** <https://www.kaggle.com/c/restaurant-revenue-prediction>
- **Data Science Problem:** Linear Regression
- **Goal:** To create a model that increases the effectiveness of investments in new restaurant sites
- **Clean Data:** No, missing P1 – P37 values and City names need cleaning
- **Data Columns:** Open Data, City, City Group, Type (of Restaurant), P1 – P37 (Representing different demographics), Revenue

Project 2:

- **Link:** <https://www.kaggle.com/c/web-traffic-time-series-forecasting>
- **Data Science Problem:** Time Series / Regression
- **Goal:** Create a model to predict the future views of a Wikipedia webpage
- **Clean Data:** Missing traffic values and will require data wrangling to pull together tables and columns
- **Data Columns:** Date and Traffic Value

Project 3:

- **Link:** <https://www.kaggle.com/c/talkingdata-mobile-user-demographics>
- **Data Science Problem:** Supervised Classification (and / or Linear Regression)
- **Goal:** Build a model that predicts a mobile users demographics based off of app usage, geolocation, and mobile device properties
- **Clean Data:** No, there are missing category values, unclean phone brand / device models, and will require data wrangling
- **Data Columns:** Age, Gender, Age Group, Category (of App), Device Model, Phone Brand, Time and Location of Event