

## CAPSTONE PROJECT 1: IDEAS

### Project 1: Consumer & Visitor Insights for neighborhoods

#### Insights:

- What are the most popular brands in a neighborhood? Are there regional preferences for some brands over others?
- How do people travel between neighborhoods? How does distance traveled compare for suburban and urban communities?
- What times do people visit certain census block groups (ex. Manhattan during the day vs. night)?
- Which neighborhoods are the most mobile? Which neighborhoods receive the most outside visitors?

Dataset: <https://www.kaggle.com/safegraph/census-block-group-american-community-survey-data/home>

This dataset will be combined with census data to get demographic insights alongside these visit patterns

Dataset: <https://www.kaggle.com/muonneutrino/us-census-demographic-data/home>

### Project 2: H1b Visa Petitions:

#### Insights:

- Is the number of petitions with Data Science job title increasing over time?
- Which industry has the most number of Data Scientist positions?
- Which employers file the most petitions each year?

Dataset: <https://www.kaggle.com/nsharan/h-1b-visa/home> (dataset from 2011 to 2016)

<https://www.kaggle.com/jonamjar/h1b-data-set-2017> (dataset from 2016 to 2017)

### Project 3: TMDB Box office Prediction

Dataset: <https://www.kaggle.com/c/tmdb-box-office-prediction/data>

### Project 4: Home credit default risk

Can you predict how capable each applicant is of repaying a loan?

#### Insights:

- Top industry of the seller
- Reason for previous application rejection
- Types of organization who applied for loan

Dataset: <https://www.kaggle.com/c/home-credit-default-risk/data>