1. **Project Title :Consumer & Visitor Insights For Neighborhoods**

**Consumer Insights & Visitation Data At The Census Block Group Level**

A Census Block Group (CBG) is the most granular level the US Census Bureau reports data on, and covers ~1500 households.

SafeGraph derived the popularity of a CBG or distances traveled to a CBG by analyzing a large panel of GPS movement data.

We also combined this GPS data with our dataset of 5 million building footprints for Points-of-Interest in the U.S. ([SafeGraph Places](https://www.safegraph.com/)) to determine visits to places like stores or restaurants. We used these visit counts to derive consumer insights such as top brands (ex. McDonald's) visited in a CBG.

https://www.kaggle.com/safegraph/visit-patterns-by-census-block-group

1. **Project Title : Predict Future Sales**

You are provided with daily historical sales data. The task is to forecast the total amount of products sold in every shop for the test set. Note that the list of shops and products slightly changes every month. Creating a robust model that can handle such situations is part of the challenge.

## File descriptions

* **sales\_train.csv** - the training set. Daily historical data from January 2013 to October 2015.
* **test.csv** - the test set. You need to forecast the sales for these shops and products for November 2015.
* **sample\_submission.csv** - a sample submission file in the correct format.
* **items.csv** - supplemental information about the items/products.
* **item\_categories.csv**  - supplemental information about the items categories.
* **shops.csv**- supplemental information about the shops.

## Data fields

* **ID** - an Id that represents a (Shop, Item) tuple within the test set
* **shop\_id** - unique identifier of a shop
* **item\_id** - unique identifier of a product
* **item\_category\_id** - unique identifier of item category
* **item\_cnt\_day** - number of products sold. You are predicting a monthly amount of this measure
* **item\_price** - current price of an item
* **date** - date in format dd/mm/yyyy
* **date\_block\_num** - a consecutive month number, used for convenience. January 2013 is 0, February 2013 is 1,..., October 2015 is 33
* **item\_name** - name of item
* **shop\_name** - name of shop
* **item\_category\_name** - name of item category

Data URL: https://www.kaggle.com/c/competitive-data-science-predict-future-sales/data/

1. **Project Title : Human Protein Atlas Image Classification**

Proteins are “the doers” in the human cell, executing many functions that together enable life. Historically, classification of proteins has been limited to single patterns in one or a few cell types, but in order to fully understand the complexity of the human cell, models must classify mixed patterns across a range of different human cells.

Images visualizing proteins in cells are commonly used for biomedical research, and these cells could hold the key for the next breakthrough in medicine. However, thanks to advances in high-throughput microscopy, these images are generated at a far greater pace than what can be manually evaluated. Therefore, the need is greater than ever for automating biomedical image analysis to accelerate the understanding of human cells and disease.

https://www.kaggle.com/c/human-protein-atlas-image-classification