Springboard Capstone Project Ideas

**Lending Club Loan Data Analysis.**

This could be a classic data science project to begin with. Lending Club is the world’s largest online marketplace connecting borrowers and investors. Though this project we could conduct a set of exploratory analysis followed by some machine learning implementation to identify some of the loans and borrowers characteristic features that could serve as good predictors of loan defaults.

LC website hosts complete loan data for all loans issued through the 2007-2015, including the current loan status (Current, Late, Fully Paid, etc.) and latest payment information. The file is a matrix of about 890 thousand observations and 75 variables.

<https://www.lendingclub.com/info/download-data.action>

**Customer Churn Analysis**

“Customer churn refers to when a customer ceases relationship with a company. Customer attrition is a key metric of business health. In most industries, the cost of acquiring new customers exceeds the cost of marketing to existing ones. So monitoring attrition on a regular basis can help the companies’ project future growth or declines.” Finding publicly available dataset is difficult. The closest dataset I could get on Kaggle is this which is very huge 22GB. This could be a good project to work with provided we could find closely aligned dataset publicly available.

<https://www.kaggle.com/c/acquire-valued-shoppers-challenge/data>

**Housing Price Prediction**

Predicting house prices would be an interesting project to work with as well. We could get a sample dataset from Zillow/Redfin. We can make use of US crime rate dataset and other demographic dataset to base our analysis and predictions on.

<https://www.zillow.com/research/data/>

<https://www.redfin.com/blog/data-center>