**Capstone Project Ideas**

I select 3 datasets from Kaggle competition. Although I prefer 2nd and 3rd datasets, I am not sure whether I can handle the two projects. Really need your feedback on choosing dataset.

1. Bike Sharing Demand ( <https://www.kaggle.com/c/bike-sharing-demand/data> )

Bike sharing systems are a means of renting bicycles where the process of obtaining membership, rental, and bike return is automated via a network of kiosk locations throughout a city. Using these systems, people are able rent a bike from a one location and return it to a different place on an as-needed basis. Currently, there are over 500 bike-sharing programs around the world.

The data generated by these systems makes them attractive for researchers because the duration of travel, departure location, arrival location, and time elapsed is explicitly recorded. Bike sharing systems therefore function as a sensor network, which can be used for studying mobility in a city. In this competition, participants are asked to combine historical usage patterns with weather data in order to forecast bike rental demand in the Capital Bikeshare program in Washington, D.C.

1. Walmart Store Sales Forecasting

( <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting/data> )

One challenge of modeling retail data is the need to make decisions based on limited history. If Christmas comes but once a year, so does the chance to see how strategic decisions impacted the bottom line.

In this dataset, historical sales data for 45 Walmart stores located in different regions are provided. Each store contains many departments, and participants must project the sales for each department in each store. To add to the challenge, selected holiday markdown events are included in the dataset. These markdowns are known to affect sales, but it is challenging to predict which departments are affected and the extent of the impact.

Want to work in a great environment with some of the world's largest data sets? This is a chance to display your modeling mettle to the Walmart hiring teams.

1. Facebook: Human or Robot?

( <https://www.kaggle.com/c/facebook-recruiting-iv-human-or-bot/data> )

In this dataset, you'll be chasing down robots for an online auction site. Human bidders on the site are becoming increasingly frustrated with their inability to win auctions vs. their software-controlled counterparts. As a result, usage from the site's core customer base is plummeting.

In order to rebuild customer happiness, the site owners need to eliminate computer generated bidding from their auctions. Their attempt at building a model to identify these bids using behavioral data, including bid frequency over short periods of time, has proven insufficient.

The goal of this project is to identify online auction bids that are placed by "robots", helping the site owners easily flag these users for removal from their site to prevent unfair auction activity.