Capstone 2 Project Proposal

Davy Gia Nguyen

My main goal for this capstone project is to apply and enhance my data science skills from the Springboard program by predicting total sales for every product and store by analyzing several time-series datasets consisting of daily sales data from 1C Company, one of the largest Russian software firms.

I hope to gain the business insights and data intelligence that a project like this would require so I can assist future clients with similar goals. Since sales prediction and forecasting are paramount to almost all kinds of businesses, I hope to bring my analysis and insights to a level that would help a client have full confidence in making these crucial executive decisions regarding future sales predictions.

The time-series data from 1C Company is composed of 5 datasets with all the items data, item categories data, sales training data, test data, and shops data.

My approach to reaching this goal of predicting total sales for every product and store in the next month is to perform in-depth exploratory data analysis, apply several models such as linear regression, random forest, and extreme boosting gradient regression and compute the root mean squared error to evaluate the performance of each to decide the best method for accurate prediction.

My deliverables include all the datasets and the necessary code, analysis, and insights provided in my github.