ORIE4741 Project Proposal

Sales In Stormy Weather

Team members: Faisal Alkaabneh (fma34) and Lawrence Hu (ljh238).

Overview

Walmart operates 11,450 stores in 27 countries, managing inventory across varying climates and cultures. Extreme weather events, like hurricanes, blizzards, and floods, can have a huge impact on sales at the store and product level.

In this project we take an attempt to solve the challenge of accurately predicting the sales of 111 potentially weather-sensitive products (like umbrellas, bread, and milk) around the time of major weather events at 45 of Walmart's retail locations.

For instances, we may expect an uptick in the sales of umbrellas before a big thunderstorm, but it's difficult for replenishment managers to correctly predict the level of inventory needed to avoid being out-of-stock or overstock during and after that storm. Walmart relies on a variety of vendor tools to predict sales around extreme weather events, but it's an ad-hoc and time-consuming process that lacks a systematic measure of effectiveness.

The ultimate goal of our project hence is to help Walmart better predict sales of weather-sensitive products to keep their valued customers out of the rain. The dataset is available through: https://www.kaggle.com/c/walmart-recruiting-sales-in-stormy-weather/data

As far as the collected data is concerned, we will be using the following data to achieve our goal of getting a high accurate prediction model:

- Weather data following the National Oceanic and Atmospheric Administration (NOAA) standards: a guide to understand the data related to the weather. NOAA weather information is provided for each station and day.

- Sales data for all stores & dates in the training set, the set also include information about each item.

Objective

Our task is to predict the amount of each product sold around the time of major weather events. For the purposes of this project, we want predict the units sold for a window of ± 3 days surrounding each storm. We are expecting heavy use of modeling techniques and regressions to get a very accurate model as well as different testing strategies to confirm the validity of different models before deciding the best model to use.