Homework 4 - Data is in google drive folder called HW 4

[https://drive.google.com/drive/folders/1j5Si6FoD91JLOV7h-7kXLg8hLYSHAgV0?usp=sharing](https://drive.google.com/drive/folders/1j5Si6FoD91JLOV7h-7kXLg8hLYSHAgV0?usp=sharing)

[Links to an external site.](https://drive.google.com/drive/folders/1j5Si6FoD91JLOV7h-7kXLg8hLYSHAgV0?usp=sharing)

Analysis and Launch plan for Retail Startup

Goals for homework;

1. Design and communicate an analytical plan for launch of targeted retail locations – determine the best 10 markets to compete effectively against an established competitor.

The retail concept is that of a Target ‘light’ a multi-purpose home and grocery store that will compete with Wal-Mart & Target for middle income to high income households.

It is understood that competitive sales data represents either Wal-Mart or Target stores.

The goal is to identify the top 10 markets in an analytical and data driven way. Presenting the data that drove the recommendation.

1. Conduct analysis on provided datasets, seek out and analyze additional relevant datasets that can help determine launch strategy

1. Write up your analysis with a clear recommendation on which markets to choose for launch of new store.

Submission is a written document (can be word and/or powerpoint), include clear description of what you did, visualizations as needed and a data driven recommendation of what markets to choose.

Reach out with questions anytime.

DMA Data set (PI-18803 DMA.xlsx);

You are provided with a demographic dataset at the DMA (Designated Market Area aka major city or metropolitan area). You can use this data to analysis the population, household, income and ethnicity of the various DMA’s as potential locations.

MSA Excel files;

MSA (Metropolitan Statistical Areas) are similar in nature to DMA’s (note; they are not 1 to 1 likeness, but the purposes of this analysis they should be assumed to be the same). These files provide input into consumer expenditure in major categories. You may make a qualitative judgement on which categories (and consequently) which spend levels you believe to be the most attractive. Explain all assumptions. It is possible that not all markets represented in the store data will be present in the MSA files. This should not automatically disqualify a market, but analysis should take place to determine if other data points could be sourced to support your decision.

Retail Data information;

You are provided with historical sales data for 45 competitive stores located in different major cites, some major markets have more than one store. Each store contains a number of departments, but the goal is to compare overall store performance and determine a ranking of markets most desirable to launch in.

In addition, the competitor runs several promotional markdown events throughout the year. These markdowns precede prominent holidays, the four largest of which are the Super Bowl, Labor Day, Thanksgiving, and Christmas. The weeks including these holidays are weighted five times higher in the evaluation than non-holiday weeks. An optional part of the challenge presented by this analysis is modeling the effects of markdowns on these holiday weeks in the absence of complete/ideal historical data to ensure that the impact of these markdowns is accounted for in the recommendation.

Stores data-set with DMA.xls

This file contains information about the 45 stores, indicating the type and size of store and location.

Sales data-set.csv

This is the historical training data, which covers to 2010-02-05 to 2012-11-01. Within this file you will find the following fields:

Store - the store number

Dept - the department number

Date - the week

Weekly\_Sales - sales for the given department in the given store

IsHoliday - whether the week is a special holiday week

features.csv

This file contains additional data related to the store, department, and regional activity for the given dates. It contains the following fields:

Store - the store number

Date - the week

Temperature - average temperature in the region

Fuel\_Price - cost of fuel in the region

MarkDown1-5 - anonymized data related to promotional markdowns that the competitor is running. MarkDown data is only available after Nov 2011, and is not available for all stores all the time. Any missing value is marked with an NA.

CPI - the consumer price index

Unemployment - the unemployment rate

IsHoliday - whether the week is a special holiday week

For convenience, the four holidays fall within the following weeks in the dataset (not all holidays are in the data):

Super Bowl: 12-Feb-10, 11-Feb-11, 10-Feb-12, 8-Feb-13

Labor Day: 10-Sep-10, 9-Sep-11, 7-Sep-12, 6-Sep-13

Thanksgiving: 26-Nov-10, 25-Nov-11, 23-Nov-12, 29-Nov-13

Christmas: 31-Dec-10, 30-Dec-11, 28-Dec-12, 27-Dec-13

Note: the retail dataset has been taken from a Kaggle competition run by Wal-Mart – that can be found here. [https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting/data](https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting/data)

[Links to an external site.](https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting/data)

Please note for the purposes of this assessment, the goal of the analysis has changed from the Wal-Mart competition and as such you should assume this sales data can represent either Wal-Mart or Target.

Business Questions:

What are the best locations based on sales in each region?

What are the best sales weeks based on the department? Are they a holiday week?

Has the holiday weekend affected the CPI?

Has the unemployment rate affected the CPI?

Has thetemperature affected the CPI?

Has the holiday weekend affected the markdowns?

How has the variable “type” affect the dataset?