Proposal

I will cover sales from 2010-02-05 to 2012-11-01, in the file WalmartStoresales.

Design:

1)What store has the most sales?

2)What holidays have the most sales? what are the least?

3)What seasons have the most sales? what are the least?

Data:

The data contains 6435 observations with 8 attributes.

Names of columns: Store - the store number

Date - the week of sales

Weekly_Sales - sales for the given store

 Holiday_Flag - whether the week is a special holiday week 1 – Holiday week 0 – Non-holiday week
Temperature - Temperature on the day of sale

Fuel_Price - Cost of fuel in the region

CPI – Prevailing consumer price index Unemployment -

 Prevailing unemployment rate

Tools:

NumPy

Pandas

Datetime

Seaborn

Matplotlib

Scikit-learn