WALMART SALES FORECASTING

DBDA.X402. (2) Business Intelligence Solutions

NITHYAA K
NOVEMBER 2, 2018

1. Introduction

The project has been done for the fulfilment of the course DBDA.X402. (2) Business Intelligence Solutions. The data which formed part of the core analysis was the Walmart Sales Forecasting dataset from Kaggle. The data contained weekly sales of various departments within different stores over different period.

2. Data

The dataset covers the weekly sales of 45 Walmart stores which are categorized as Type A, B and C depending on the store size with as many as 81 departments from 2010-02-05 to 2012-11-01. The data set has three files: train.csv, features.csv and store.csv

3. Business Problem

For this project, we are analyzing the sales revenue over the recorded time frame in all the departments of the three store types. We are going to find out the following:

1. The departments that are generating maximum revenue in all the store types.
2. The departments that shows poor performance
3. The sales revenue during the holiday season. 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

4. Dashboard Overview

Total Sales by Type

C

B

A

6.74bn
Total Sales

45
Total Number of Stores

Type	Number of Departments
A	81
B	80
C	66
Total	**81**

Top 10 performing departments

Dept ●2 ●8 ●13 ●38 ●40 ●72 ●90 ●91 ●92 ●95

0.4bn

0.35bn
0.31bn

0.3bn

0.20bn
0.2bn   0.16bn 0.17bn
         0.12bn
                        0.14bn
0.1bn   0.07bn 0.10bn 0.10bn
                               0.08bn
                        0.03bn   0.05bn 0.05bn
                               0.01bn 0.02bn
0.0bn
          A   B   C

Yearly Total Sales Comparison

● Total Sales(2010) ● Total Sales(2011) ● Total Sales(2012)

$0.3bn

$0.2bn

January February March April May June July August September October November December

Under-performing departments in Type C

Dept ●19 ●24 ●29 ●33 ●34 ●35 ●44 ●55 ●71 ●99

1,000
          832
                715
     516
                   62
          11   175
0   2   7
-500
                -518
          C

Under-performing departments in Type B

Dept ●39 ●43 ●45 ●47 ●51 ●54 ●60 ●77 ●78 ●99

0.4M   0.49M

0.22M
0.2M

0.00M   0.02M
0.0M
          -0.01M
          B

Under-performing departments in Type A

Dept ●39 ●43 ●45 ●47 ●51 ●54 ●60 ●77 ●78 ●99

0.4M   0.41M
        0.33M
0.2M 0.20M   0.26M   0.35M
                  0.06M
0.00M 0.00M   0.02M   0.00M 0.03M   0.09M
0.0M   0.00M 0.00M
       2010   2011   2012

4.1 Exploratory Analysis

1. Top 10 departments

The above chart shows the top performing departments (limited to 10) across the three store types A, B and C. The department 92 has done well in the group A. Even in the group C, where the total sales is the least, the department 92 pulls in more revenue compared to the other departments.

Top 10 performing departments

Dept ●2 ●8 ●13 ●38 ●40 ●72 ●90 ●91 ●92 ●95

0.4bn

0.3bn

0.2bn

0.1bn

0.0bn
          A   B   C
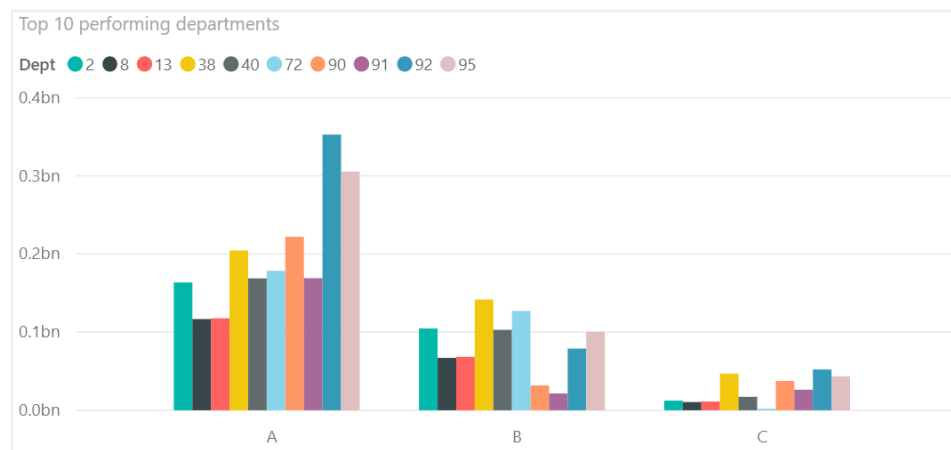
The departments 38, 92 and 95 has consistently shown increased sales over the years. Please note that the dataset does not contain data for the months November and December which is when the sales peak.

Top 10 performing departments

Dept ●2 ●8 ●13 ●38 ●40 ●72 ●90 ●91 ●92 ●95
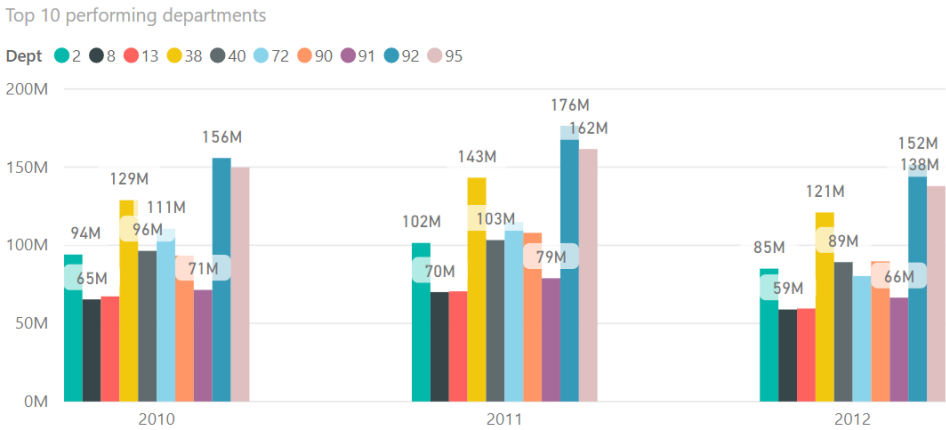
2. Top 10 Stores

In order to classify the stores based on the store size, we added a column named Size_Class and categorized the stores depending on the size.
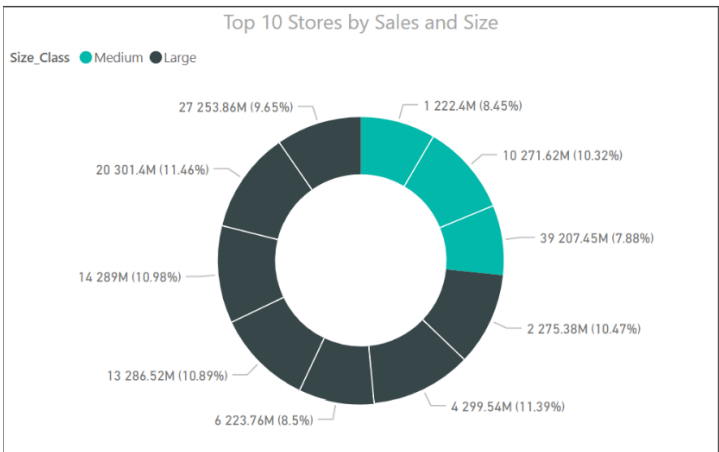
The above chart shows that Store 20 and Store 4 generated 301 million USD and 299 million USD respectively on sales revenue. The top 10 stores with maximum sales revenue is a mixture of large and medium sized stores. Store 10 with 271.62 million USD is the largest revenue generating medium sized store. Together, these 10 stores contributed 39% to the total sales.

3. Yearly Sales Comparison

In order to compare yearly sales, we created new measures to calculate the total sales for the years 2010, 2011 and 2012.

Yearly Total Sales Comparison

● Total Sales(2010) ● Total Sales(2011) ● Total Sales(2012)
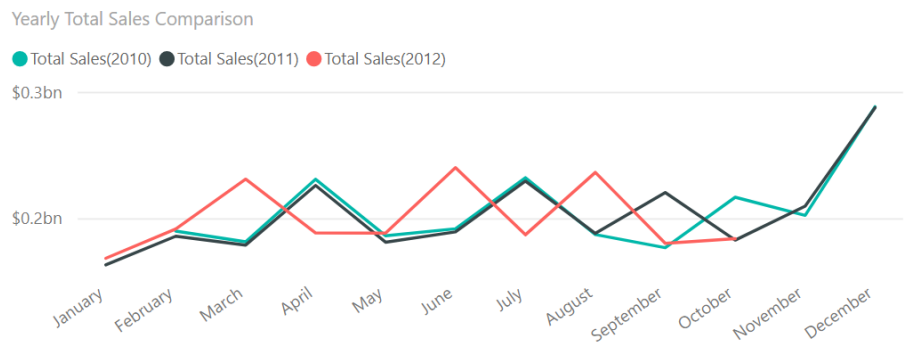
The chart below shows that the total sales revenue of the year 2012 is on the rise with 2 billion USD even without the data for the peak holiday season when the stores show an increased sale.

Total Sales by Type

2.00bn

Total Sales

45

Total Number of Stores

Year
☐ 2010
☐ 2011
■ 2012

Yearly Total Sales Comparison

● Total Sales(2010) ● Total Sales(2011) ● Total Sales(2012)

4. Departments with low sales

The chart shows that departments 39, 43, 47, 51 and 78 are the consistently least revenue generating departments in Type A stores over the recorded time frame.

Under-performing departments in Type A
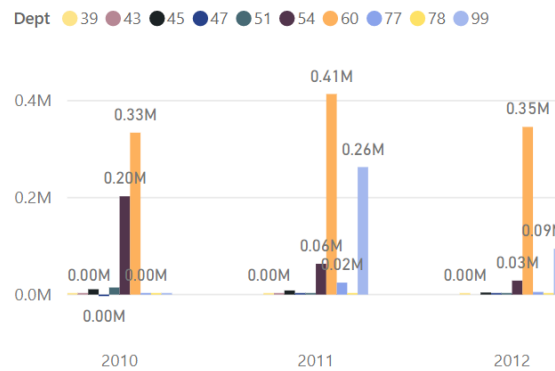
The departments 39, 47, 77, 78 and 79 are the least revenue generating departments in Type B stores. The department 47 generated negative revenue - - 5K USD and -2K USD during the years 2011 and 2012, respectively.


Under-performing departments in Type B
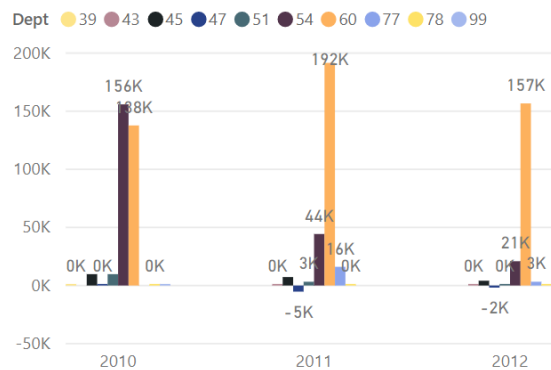
The department 71 from the Type C stores was consistently showing negative revenues during 2011 and 2012. The departments 35 is another under-performing one in Type C.

Under-performing departments in Type C
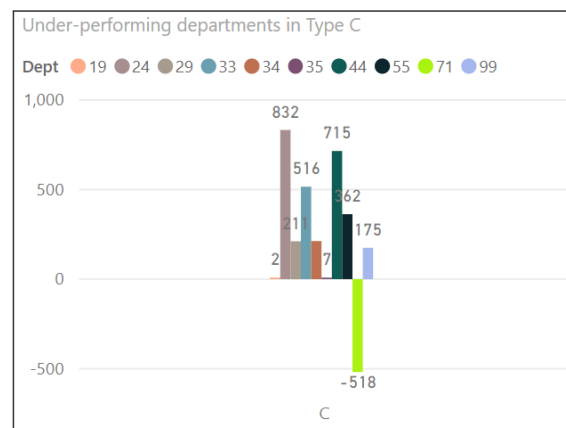
5. Stores with negative revenue

When drilled down the weekly sales for negative revenue over all the stores, it is shown that department 47 is unpopular among 37 of the 45 stores. The department from the store 35 has least sales revenue with -3.57K USD, closely followed by the store 1 with -2.56K USD and store 10 with -2.33K USD.


Departments generating negative revenue

Stores with negative revenue from dept 47
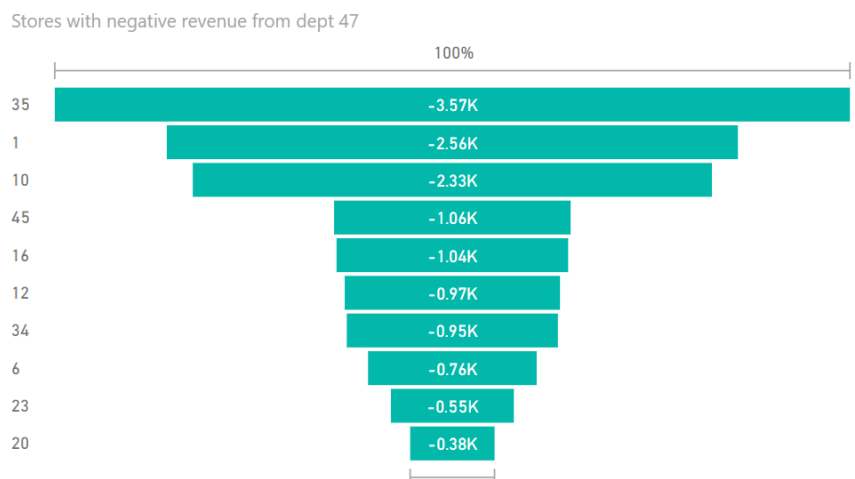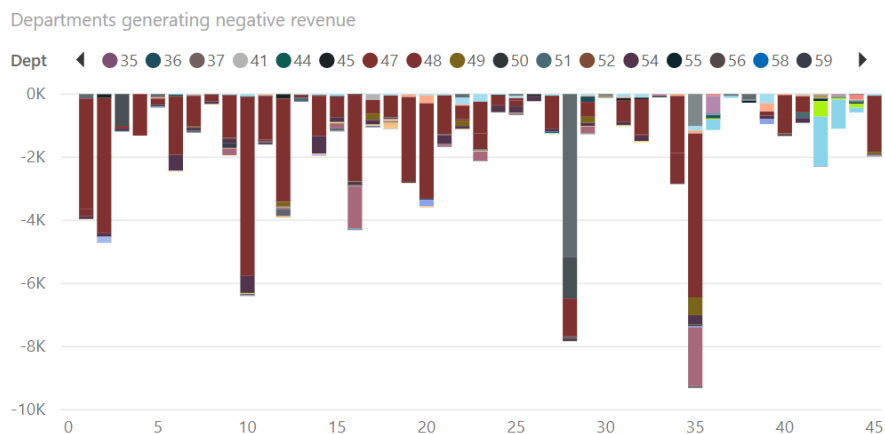
6. Departments with maximum revenue

The departments 92 generates the highest revenue of all the departments over the recorded time frame, closely followed by department 95 and 38. Together the three departments contributed 20% to the total sales revenue.
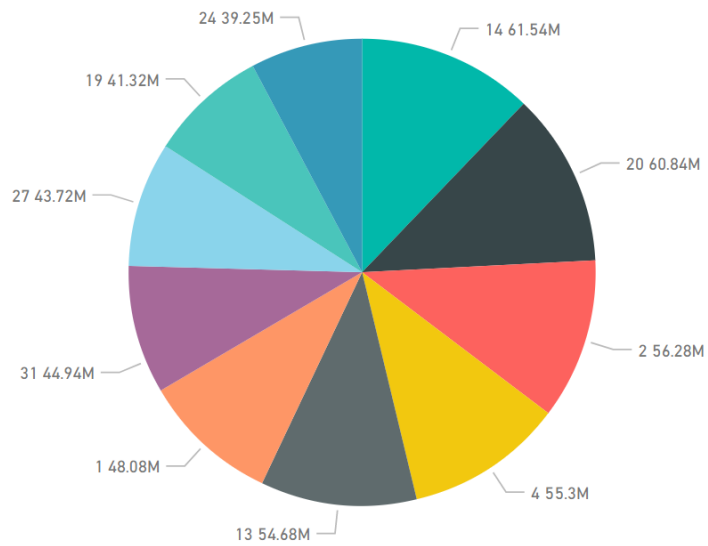
Departments with maximum revenue

Dept ●2 ●8 ●13 ●38 ●40 ●72 ●90 ●91 ●92 ●95
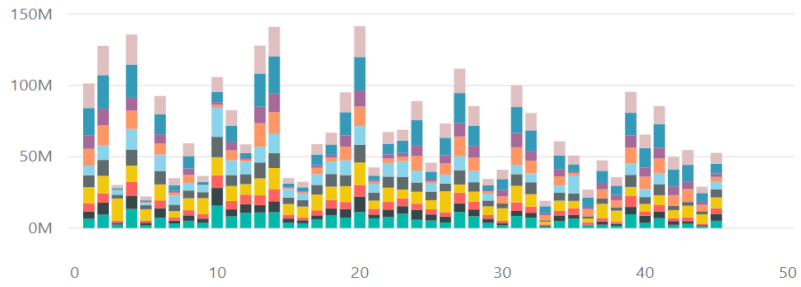
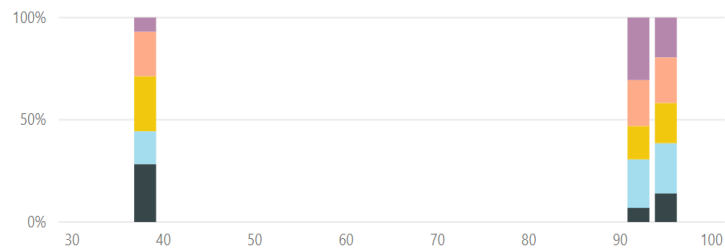Stores with max revenue from dept 38, 92 and 95

Store ●14 ●20 ●2 ●4 ●13 ●1 ●31 ●27 ●19 ●24

The chart below shows the store 33 has low revenue in department 38. The department 92 and 95 has their lowest revenue in the store 5.

Stores where department 38, 92 and 95 can be better

Store ●5 ●7 ●15 ●29 ●33

7. Pre-Holiday Sales

For simplifying the calculation of sales during Holiday season, we used the weekly sales of the two weeks prior to every holiday. We have created new measures for calculating the total sales two week prior to every holiday.

The dataset mentions that Walmart runs promotions before every holiday. The dataset has feature called Markdown (1-5) which is available after Nov 2011. The data is not complete; however, the sales has increased considerably compared to the previous year.
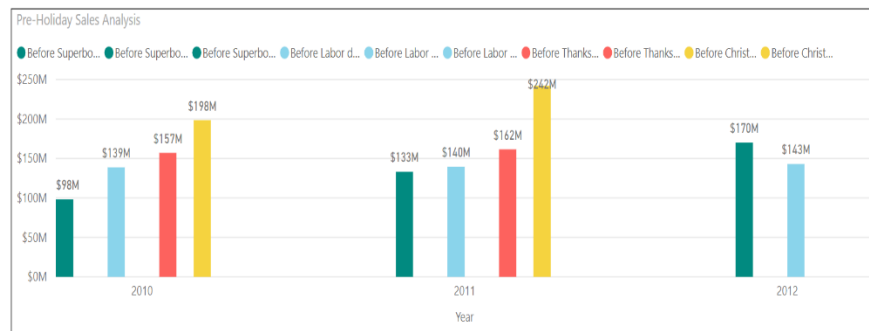
The month of December shows the maximum total sales because of the Pre-Christmas sales which is followed by November.

4.2   Key Performance Indicators

The key performance indicators we have chosen for our dashboard are listed below:

1. Weekly Sales by department
2. Weekly Sales by year
3. Weekly Sales by Stores and Size

4. Store type by department
5. Total yearly Sales by month

4.3 Findings

Based on the analysis, listed below are my findings:

1. The department 47 has been consistently unpopular and generating negative revenue. The products in that department should be discontinued and replaced with competitive products.
2. The department 6 in store 28 is generating negative revenue and since it is not generating such low revenue in other stores, we need to consider the running promotions or better placement in the store.
3. The products in department 38 of the store 33 should be reconsidered as the department is one of the top 3 departments in all the stores.
4. Likewise, the products in department 92 and 95 of the store 5 should be replaced with high performing products that belong in that department.
5. The store should also run promotions before the summer since the sales are high in the month of May and June. The promotions can boost up the sales revenue.
6. The department 71 in the Type C stores can be discontinued as it continues to generate negative revenue.