ABC Pharmacy  
Business capability analysis and data model

12/15/2019

# Executive summary

## Overview

This document provides results on the data analysis performed on the historical data for ABC pharmacy, describes the data model associated, findings and business opportunities. It provides product performance across spectrum of product portfolio as well store performance across various states.

## Business Goals

* Identify investment opportunities on the products that are performing, so as to increase the quality of those products. Maintain appropriate inventory and adjust the supply and demand of the products.
* Sales revenue prediction and set appropriate sales targets.
* High and low performing stores
* Identify the seasonality associated with any products so that appropriate stocks can be maintained.
* Better business decisions based on historical data.

## Data Description

Data provides insights in the following for ABC Pharmacy

* Major product categories, product categories, products sub categories, product segments and finally product.
* Pharmacies across various states
* Sales transactions for various products and across pharmacies.

## Methodological summary

Overall approach has been to combine the data from various tables and group them across various parameters like state, product, pharmacy, major product category etc. and then calculate the total quantity or the sales revenue which is quantity multiplied by the sale amount for that product. In some case data has been exported to CSV and then the table is used to display data in this document. Also for better data visualization graphs for excel are build using the CSV file data exported from the R queries.

## Results

* Data Span across 6 months from Jan 2016 to June 2016
* Sales is restricted to only 6 states NJ,PA,NY,CT,MA,ME,DE
  + NJ State has highest sales revenue of $6,413,468.42
  + DE has lowest sales revenue of $7,050.33
  + Other states where there are stores do not generate any revenue, these states include PR, RI, NH, VI, VT
  + There are 991 stores across all states that do not generate any sales revenue, so are probably closed or not operational.
* Sales revenue maximum in March 2016 of $3,707,182
* 23% of the total products are generating any sales revenue.
* Top 3 Products by Quantity

|  |  |
| --- | --- |
| Product | Quantity |
| DME SALES | 32583 |
| GENERICQS1ITEM | 21352 |
| VITAMINS/SUPPLEMENTS | 19201 |

* Top 3 Products By Sales Amount

|  |  |
| --- | --- |
| Products | Sales Value $ |
| MONEY ORDER | $ 1,662,010.26 |
| VITAMINS/SUPPLEMENTS | $ 756,299.96 |
| STRUTZ PRO | $ 745,804.90 |

* Top 3 Pharmacy stores by sales revenue

|  |  |
| --- | --- |
| Pharmacy Store | Pharmacy Sales Revenue $ |
| GNP PHARMACY #453 | $ 2,116,067.26 |
| GNP PHARMACY #200 | $ 1,743,900.83 |
| GNP PHARMACY #680 | $ 1,701,263.19 |

## Recommendations

* Product portfolio needs to be looked and appropriate products need to be decommissioned.
* Stores without any revenue generation should be closed, there are closed to 991 such stores.
* Products which are related to Money like money order are high revenue generating so more investment can be done on those kinds of products to increase the sales revenue further.

# Core report

## Overview

This report provides the insights on the data analysis performed on the data of ABC pharmacy which includes their sales transactions from Jan to June 2016. It articulates business goals and questions associated with those goals as well as how data can help in achieving the business goals. Based on data analysis it provides recommendations for additional business opportunities and how the losses can be minimized. Based on the historical trends, patterns ABC pharmacy would be able to make appropriate predictions for future. It provides product and store performance which could critical for success of the business in future.

## Business goals

* Investments on type of products
  + Identify investment opportunities on the products that are performing exceptionally well, so as to increase the quality of those products.
* Adjust demand of supply of high volume products
  + Maintain appropriate inventory and adjust the supply and demand of the products.
  + Identify the seasonality associated with any products so that appropriate stocks can be maintained.
* Sales Revenue Targets
  + Sales revenue prediction and set appropriate sales targets for stores
* Store Performance
  + High and low performing stores
  + Identify geographically which states have higher sales revenue and which have lower.
* Total sales revenue for 6 months of transactions.

## Business questions

1. What’s total sales revenue for ABC Pharmacy for Jan 2016 to June 2016?
2. What’s percentage gain/loss in sales revenue every month for Jan 2016 to June 2016?
3. Which are the best-selling product?
4. Which major product categories perform better?
5. Are some products categories getting sold in greater quantities based on seasonality?
6. What are the average number of products sold per order, per month?
7. What’s state wise sales revenue?
8. Which store has maximum performance based on sales revenue?
9. What’s average sale per store?
10. What are number of stores state wise?
11. Non Performing stores - If there is need to minimize the losses, then which stores ABC pharmacy should be closing? (Note: Stores with least sales revenue can be closed)
12. For any particular store which product generates maximum revenue?
13. For any particular store which product does not generate any revenue or least revenue?
14. For any particular store which month has maximum sales?

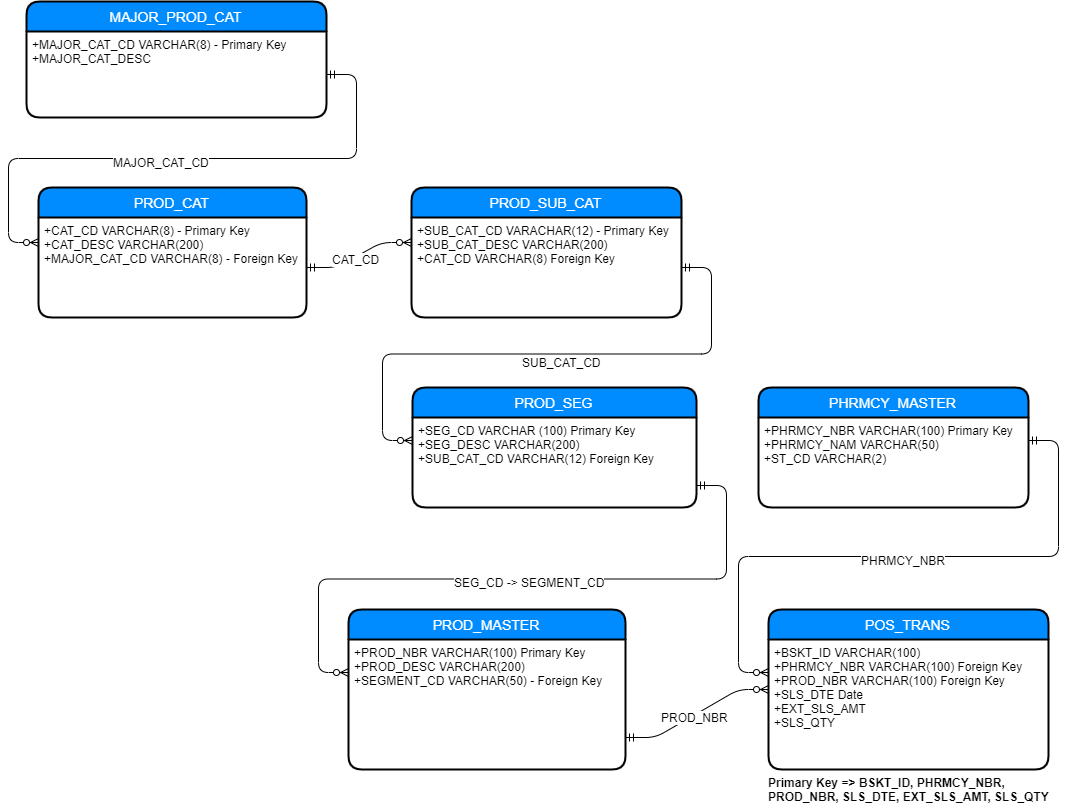
## Data description

## Data quality

* Overall data seems to be appropriate with appropriate and consistent data types.
* Duplicate check
  + Performed duplicate check for following and did not find any duplicates
    - MAJOR\_PROD\_CAT
    - PROD\_CAT
    - PROD\_SUB\_CAT
    - PROD\_SEG
    - PROD\_MASTER
    - PHRMCY\_MASTER
  + In POS\_TRANS checked duplicates based on BSKT\_ID, PROD\_NBR, PHRMCY\_NBR, there are duplicates based on that so added SLS\_DTE, EXT\_SLS\_AMT and SLS\_QTY as part of unique identifier.
* Sales quantity is POS\_TRANS is negative which seems to be legit as they could be returns in the store.
* CSV file related to MAJOR\_PROD\_CAT had an empty column so had to clean it.
* From PROD\_MASTER there is no need to have the below columns as they can be derived based on SEG\_CD.
  + MAJOR\_CAT\_CD
  + CAT\_CD
  + SUB\_CAT\_CD
* In Pharmacy master there is pharmacy which is assigned to state with State code = ‘Z’ which is invalid.

## Data model

Below is the data model for ABC pharmacy data



## Methodological summary

R has been used along with tidyverse package to answer majority of the business questions. For visualization have created charts using excel charts, alternatively can be created using R plotting libraries as well.

* For calculating total sales revenue for ABC pharmacy for Jan 2016 to June 2016 had to just sum the line item total for each order
* For percentage gain and loss month on month, first calculate the sales revenue for each month hand from Jan 2016 to June 2016 and then export data to excel and then use percentage calculation by subtracting current month minus previous month salary and diving it by previous month sales.
* For identifying top selling products by quantity had to group by product number and description and then do the sum of the sales quantity
* For identifying top selling products by $ value had to group by product number and description and then do the sum of the sales quantity multiplied by price each.
* For Product Categories performance joined transaction, product, product segment, product sub category, product category and product major category, and then grouped on Major product category and summed line item total which is sales quantity into sales amount.
* For identifying seasonality associated with major Product Categories joined transaction, product, product segment, product sub category, product category and product major category, and then grouped on Major product category, month and year of transaction date and summed line item total which is sales quantity into sales amount. Once the data is available export to csv and used excel to generate graphs and analysis.
* Calculating averages was pretty straight forward for per order had to group by basket id and per month had to group on the month of the transaction date.
* State wise sales revenue again was grouping by the state code in the pharmacy table.
* For Identifying top performing stores had to calculate the sales revenue generated by each store and then sort by sales revenue in descending order, then choose the first 10 which would be top performing stores.
* For Average sale per store, calculated the sales revenue across all stores then find the average across all stores.
* Number of stores across state is pretty straightforward, Use the Pharmacy master table and group by state code, this will give count of pharmacies across various states.
* Non Performing stores
  + Stores with no sales revenue – Join pharmacy master and pos transaction table, to calculate store revenue and then apply filter where there is no sales revenue for store.
  + Stores with < 10K sales revenue – Same as above but instead of no sales revenue condition use the sales revenue < 10K condition.
* Data Analytics for store.
  + GNP PHARMACY #453 – This store is chosen for analytics as it has maximum sales revenue due the time frame from Jan 2016 to June 2016
  + For selected pharmacy identify top performing products – For this join transaction and product master with selected store filter and then group by product and calculate sales revenue for each product.
  + Worst Performing products for selected store – Same as above with filter condition of sales revenue by product less than 10$.
  + Month on month sales revenue for store – From transaction table filter selected store and group by the month and year of transaction date.

## Results

### Total Sales revenue

* Total sales revenue for ABC Pharmacy from Jan 2016 to June 2016 is $17,283,853.00

### Percentage Gain/Loss month on month

Below table show the month on month sales

|  |  |  |
| --- | --- | --- |
| Month And Year | Monthly Sales $ | Month on Month Sales Revenue  (Increase/Decrease %) |
| Jan-16 | $ 3,026,923.89 |  |
| Feb-16 | $ 2,900,311.95 | -4% |
| Mar-16 | $ 3,707,182.09 | 28% |
| Apr-16 | $ 2,304,948.19 | -38% |
| May-16 | $ 2,523,208.54 | 9% |
| Jun-16 | $ 2,821,277.92 | 12% |

### Best Selling products

#### By Quantity

Below are 10 best-selling products quantity wise

|  |  |  |
| --- | --- | --- |
| PROD\_NBR | PROD\_DESC | PRODUCT\_QTY\_ORDER |
| 90400000000 | DME SALES | 32583 |
| 92000000000 | GENERICQS1ITEM | 21352 |
| 90800000000 | VITAMINS/SUPPLEMENTS | 19201 |
| 90300000000 | DME | 14580 |
| 99100000066590000000006 | CANDY OPEN DEPARTMENT | 14146 |
| 90600000000 | CARDS | 13323 |
| 1820025008 | 01820025008 | 12057 |
| 40003000753 | MIDWEST FASTENER | 11098 |
| 90900000000 | STAMPS | 10250 |
| 99100000120141000000010 | CANDY & BEVERAGE | 8708 |

#### By Sales Revenue

Below are 10 best-selling products by sales revenue

|  |  |  |
| --- | --- | --- |
| PROD\_NBR | PROD\_DESC | PRODUCT\_VALUE\_ORDER |
| 99100000770330000000003 | MONEY ORDER | $ 1,662,010.26 |
| 90800000000 | VITAMINS/SUPPLEMENTS | $ 756,299.96 |
| 98650000000000185441572 | STRUTZ PRO | $ 745,804.90 |
| 91000000000 | LOTTO | $ 480,597.98 |
| 98650000000000944904150 | WHEELED WALKER BRA/SEAT | $ 474,030.50 |
| 90400000000 | DME SALES | $ 472,319.23 |
| 79936677762 | GIFT CARD VISA $20-$500 | $ 426,615.30 |
| 92000000000 | GENERICQS1ITEM | $ 378,584.54 |
| 90600000000 | CARDS | $ 357,964.00 |
| 90300000000 | DME | $ 288,135.41 |

### Non Performing products

There are close to 145,761 unique products which have not been sold in six months’ time frame of Jan 2016 to June 2016

### Major Product Categories Performance

Below are the sales revenue by product categories

|  |  |
| --- | --- |
| **Product Categories** | **Sales Revenue $** |
| HEALTH CARE | $ 6,311,436.82 |
| GENERAL MERCHANDISE | $ 5,380,539.27 |
| HOME HEALTH CARE | $ 2,071,468.19 |
| GREETING CARDS | $ 1,115,704.89 |
| EDIBLES | $ 973,702.30 |
| PERSONAL CARE | $ 641,856.92 |
| BEAUTY | $ 378,114.97 |
| PHOTO | $ 211,492.51 |
| DIABETES | $ 172,144.30 |
| MISC | $ 27,392.41 |
| **Total** | **$ 17,283,852.58** |

### Seasonality associated with any major product category sales

Health care product category seems to be consistently higher is sales revenue, there is unusual spike in month of March 2016 where the sales revenue has jumped dramatically by $704,502.72

March month seems to highest sales revenue month with health care as major product category having maximum sales revenue, let’s see the products in that major category that performed better.

|  |  |
| --- | --- |
| PROD\_DESC | March Sales Revenue |
| STRUTZ PRO | $ 486,103.66 |
| VITAMINS/SUPPLEMENTS | $ 241,266.68 |
| TEGADERM 4X4 3/4 1626 | $ 137,635.20 |
| ACIDOPHILUS SUPER PROBIO 5BILLN CAP 60CT | $ 54,624.93 |
| GNP U/PAD SPR 30X36 PAD 4X10 | $ 22,785.42 |
| GLOVE VINYL P/F L/F MD GLV 100 | $ 22,019.54 |
| SALONPAS PATCH HOT CAPSICUM UNS 50X1CT | $ 7,035.41 |
| FLORASTOR 250MG CAPSULE 50CT | $ 5,979.68 |
| GLOVE VINYL P/F L/F LG GLV 100 | $ 5,931.51 |

Now if we compare the above list with top 10 selling products by sales revenue, following are the products which are there in above but not in top 10 products list

* TEGADERM 4X4 3/4 1626
* ACIDOPHILUS SUPER PROBIO 5BILLN CAP 60CT
* GNP U/PAD SPR 30X36 PAD 4X10
* GLOVE VINYL P/F L/F MD GLV 100
* SALONPAS PATCH HOT CAPSICUM UNS 50X1CT
* FLORASTOR 250MG CAPSULE 50CT
* GLOVE VINYL P/F L/F LG GLV 100

So this does in indicate that is some seasonality pattern associated with some products.

### Averages

* Per Order average number of products sold = 2 products (1.82)
* Per Order average sale value = $27.8
* Per Month average number of products sold = 188661
* Per month average sales revenue = $2,880,642

### State wise sales revenue

|  |  |
| --- | --- |
| ST\_CD | STATE\_SALES\_REVENUE |
| NJ | $ 6,413,468.42 |
| PA | $ 4,277,291.77 |
| NY | $ 2,733,151.59 |
| CT | $ 2,665,387.18 |
| MA | $ 798,851.27 |
| ME | $ 388,652.02 |
| DE | $ 7,050.33 |
| Total | $ 17,283,852.58 |

### Best Performing stores

Below are 10 best performing stores based on sales revenue

|  |  |  |
| --- | --- | --- |
| PHRMCY\_NBR | PHRMCY\_NAM | PHRMCY\_NBR\_SALES\_REVENUE |
| 4416100399456673861 | GNP PHARMACY #453 | $ 2,116,067.26 |
| 2487426938853675539 | GNP PHARMACY #200 | $ 1,743,900.83 |
| 61520549788616420 | GNP PHARMACY #680 | $ 1,701,263.19 |
| 3009693108150153253 | GNP PHARMACY #269 | $ 1,437,316.97 |
| 2759406693434064370 | GNP PHARMACY #232 | $ 574,577.54 |
| 1174450154022548624 | GNP PHARMACY #18 | $ 522,080.09 |
| 6991356705459241502 | GNP PHARMACY #795 | $ 519,509.13 |
| 657046048504326744 | GNP PHARMACY #739 | $ 471,032.51 |
| 7003025686214903268 | GNP PHARMACY #798 | $ 412,827.33 |
| 8506230257184703229 | GNP PHARMACY #990 | $ 397,611.05 |

### Average Sale per store

Average Sales revenue per store is $163,055

### Number of Stores per State

|  |  |
| --- | --- |
| ST\_CD | NUMBER\_OF\_STORES |
| NY | 390 |
| NJ | 288 |
| PA | 196 |
| PR | 97 |
| MA | 51 |
| CT | 42 |
| ME | 21 |
| DE | 4 |
| RI | 3 |
| NH | 2 |
| VI | 1 |
| VT | 1 |
| Z | 1 |

Even though following states have the stores, but there is no revenue generated from those stores which indicates that stores are probably closed down.

* PR
* RI
* NH
* VI
* VT

### Non Performing stores

1. Stores with no sales revenue
   1. There are 991 stores for which no revenue has been generated in span for Jan 2016 to June 2016
2. State wide stores with no revenue, which indicates which are not operational or closed

|  |  |
| --- | --- |
| ST\_CD | NON\_PERF\_PHARMACY\_COUNT |
| CT | 34 |
| DE | 3 |
| MA | 47 |
| ME | 10 |
| NH | 2 |
| NJ | 256 |
| NY | 374 |
| PA | 162 |
| PR | 97 |
| RI | 3 |
| VI | 1 |
| VT | 1 |
| Z | 1 |
| Total | 991 |

1. Below are 10 stores which have revenue of less than 10K, so in case if there is a need to minimize losses and closed certain stores then these stores could be the candidates

|  |  |  |
| --- | --- | --- |
| PHRMCY\_NBR | PHRMCY\_NAM | PHRMCY\_NBR\_SALES\_REVENUE |
| 2109913237959930897 | GNP PHARMACY #150 | 100 |
| 1864727887909581358 | GNP PHARMACY #116 | 151.77 |
| 2557779967183367219 | GNP PHARMACY #211 | 260.98 |
| 7352666603740569574 | GNP PHARMACY #835 | 610.69 |
| 767790889581973829 | GNP PHARMACY #881 | 1075.47 |
| 4474219946559825502 | GNP PHARMACY #458 | 2100.6 |
| 7020056645499029586 | GNP PHARMACY #800 | 2420.52 |
| 8932873020424827010 | GNP PHARMACY #1053 | 3325.08 |
| 6456621789121903095 | GNP PHARMACY #725 | 4920.59 |
| 8146743720163240233 | GNP PHARMACY #942 | 6315.58 |
| 2460130771898242917 | GNP PHARMACY #196 | 6381.91 |
| 3204676104489970153 | GNP PHARMACY #297 | 7050.33 |
| 4754955394504903019 | GNP PHARMACY #500 | 9574.51 |
| 7872808829300808721 | GNP PHARMACY #902 | 9713.96 |

### Data analytics for a selected store

In this case we are going to consider the store with maximum sales revenue which is

|  |  |  |
| --- | --- | --- |
| PHRMCY\_NBR | PHRMCY\_NAM | PHRMCY\_NBR\_SALES\_REVENUE |
| 4416100399456673861 | GNP PHARMACY #453 | $ 2,116,067.26 |

* For GNP PHARMACY #453 following are the top 10 performing products.

|  |  |  |
| --- | --- | --- |
| PROD\_NBR | PROD\_DESC | Sales Revenue |
| 99100000770330000000003 | MONEY ORDER | $ 1,662,010.26 |
| 99100000234116000000009 | MONEYGRAM | $ 263,872.18 |
| 1820025008 | 01820025008 | $ 34,315.38 |
| 99100000081989000000040 | STAMPS | $ 18,392.50 |
| 99100000179742000000050 | SYRINGE & NEEDLE | $ 13,510.48 |
| 30087510191 | TYROS 1 INFANT FORMULA | $ 11,050.00 |
| 99100000120141000000010 | CANDY & BEVERAGE | $ 10,921.33 |
| 2610000573 | NEWPORT 100'S | $ 10,623.60 |
| 1820025002 | NAT LT 15PK 25OZ CAN | $ 7,479.65 |
| 99100000643311000000027 | CELL CARDS | $ 6,295.42 |

* For GNP PHARMACY #453 there are 822 products which have sales revenue less than 822.
* Month on Month sales revenue for the store.

|  |  |  |
| --- | --- | --- |
| Year | Month | Sales Revenue $ |
| 2016 | Jan | $ 351,803.07 |
| 2016 | Feb | $ 370,287.13 |
| 2016 | Mar | $ 413,442.72 |
| 2016 | Apr | $ 361,776.36 |
| 2016 | May | $ 306,808.08 |
| 2016 | Jun | $ 311,949.90 |

March 2016 had maximum sales revenue for the store GNP PHARMACY #453

## Conclusions and recommendations

### Conclusion

* Sales revenue month on month does not seem to be consistent, it is positive and negative at time, for 6 months data there is an effective 1% of sales revenue growth.
* Health care as product category yields maximum sales revenue.
* Money related to products like money order, lotto (lottery) gift cards visa 250 - 500$ have higher sales revenue.
* 991 stores not generating any sales revenue
* There are certain states where stores are not generating no sales revenue.
* There is no seasonality as such observed in the dataset which indicates the sudden spike in demand for particular set of products.
* Out of 189052 products 145761 do not have any sales revenue in the 6 months’ time frame, so only effectively only 23% of products are revenue generating.

### Recommendations

* Money order as product is generating maximum sales revenue, so more investments can made on that product, apart from those there 9 other product listed in document above.
* All non performing stores at least with no sales revenue can be closed.
* More portfolio of products can be introduced in Health care category as it is high sales revenue yielding category.
* Average sales revenue for store which is generating any revenue is $163,055 for 6 months, so every newly open store should have a target around the same with +-20% margin.
* Only 23% of products are revenue generating, so remaining product portfolio must be analyzed and appropriate products need to be decommissioned.