# STORY TELLING -2

## MAVEN TOYS

GAYATRI KRISHNA

21BDA16

# INTRODUCTION


**Software Used:** Jupyter Notebook

- The dataset is about Maven Toy Store that contains information on order details of customers for orders of a toy store in Mexico. It includes information about products, stores, daily sales transactions.
- It consists of 829262 observations and 14 columns. There are 50 stores in 29 different cities of Mexico.
- There are stores which are just 5 years old whereas some are 29 years old but maximum are 7 years old.
- From the dataset, we can say that first store for the Maven Toys was opened in the year 1992 and the latest store for the Maven Toys was opened in the year 2016.
- There 5 broad categories of products in which Toys has the maximum count.
- There are 4 Locations in these cities where the stores are located: Residential, Airport, Commercial and Downtown

# OBJECTIVE OF THE STUDY 

The goal of this analysis is find out how their stores are performing in Mexico and also to generate some insight on Sales.

- To perform basic analysis with the given dataset.
  a) Number of stores by location.
  b) The various product categories.
  c) Location wise unit sold.
  d) Number of stores over the years

- To find store location and product category.
- To find the Top 10 Products Sold.
- To find the top 10 revenue generating Product.
- To find which store locations is generating the most of the revenue and profit for the store.
- To find the ratio of profit to sales for different store locations.
- Profit over the years.

# DATA CLEANING

- The Product_Cost and Product_Price columns have '$' symbol, making it a string data type. As a result we can't do any statistical analysis for these two columns. To remove the dollar sign we can use the list comprehension with the .strip method.

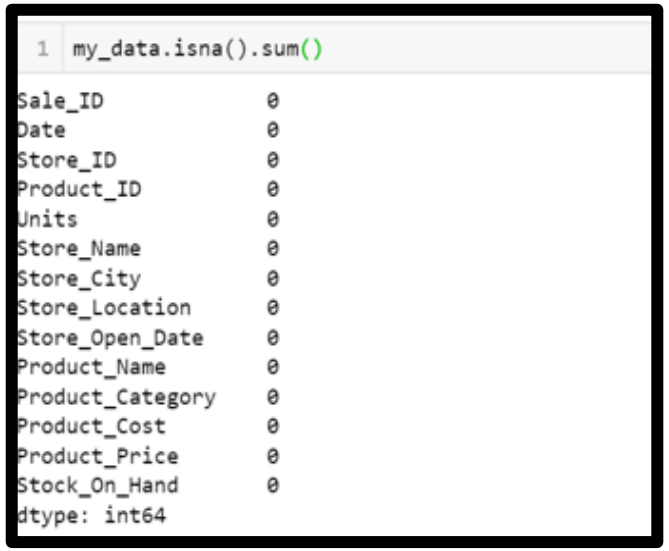
| Product_Cost | Product_Price |
|---|---|
| $9.99 | $12.99 |
| $9.99 | $15.99 |
| $3.99 | $6.99 |
| $11.99 | $15.99 |
| $34.99 | $39.99 |

| Product_Cost | Product_Price |
|---|---|
| 9.99 | 12.99 |
| 9.99 | 15.99 |
| 3.99 | 6.99 |
| 11.99 | 15.99 |
| 34.99 | 39.99 |

- The dataset consists of Dates (ie) Date of the sale transaction and Date when the store was opened. We see that the dtype of the these dates are 'object'. So we use the datetime package to convert these dates from object to date.

# CONTINUATION…

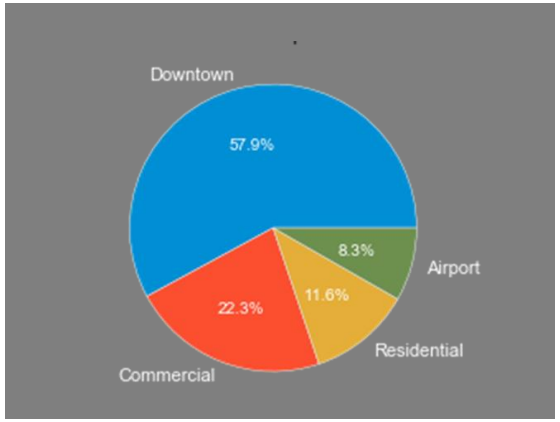
- We see that missing values are present only in the Stock_On_Hand column. About 6232 observations are missing in this column out of 829262 observations.
- The missing values present in the column ‘Stock on Hand’ is replaced with zeros.
- After replacing the missing values , we cross check if anymore missing values are present in the dataset.
- Python has a function ‘is.na()’ which is  used to detect missing values. We see that there is no missing values present in the toys dataset. The sum is zero for all the columns.



# ANALYSIS and VISUALIZATION

## Number of Stores by Location

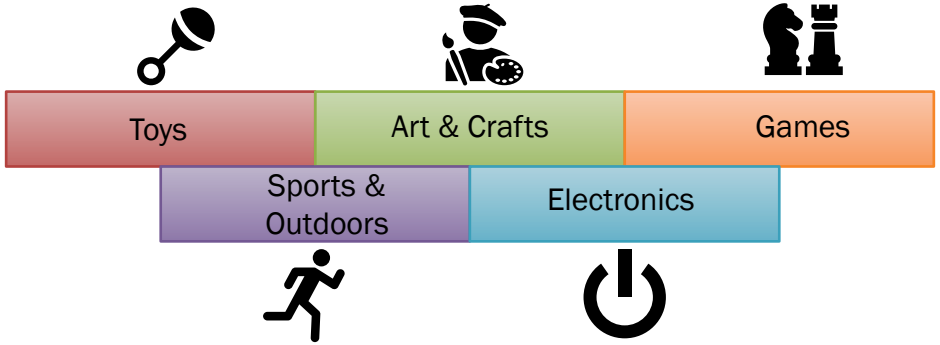
- From the adjacent pie chart we see that among the 4 locations, Downtown has the number of stores of Maven Toys.
- The least number of stores for Maven Toys is in airport.




## The Product Category

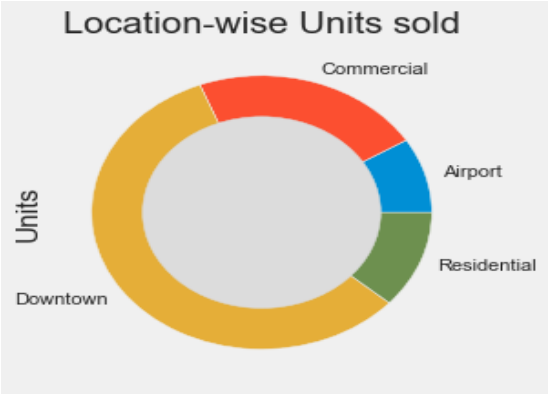
- We see that there are 5 categories of products: Toys, Art and Crafts, Games, Sports and outdoors, Electronics
- Among the 5 categories, toy have the highest frequency followed by Art and Crafts.

| | index | Frequency |
|---|---|---|
| 0 | Toys | 221227 |
| 1 | Art & Crafts | 220673 |
| 2 | Games | 157006 |
| 3 | Sports & Outdoors | 131331 |
| 4 | Electronics | 99025 |



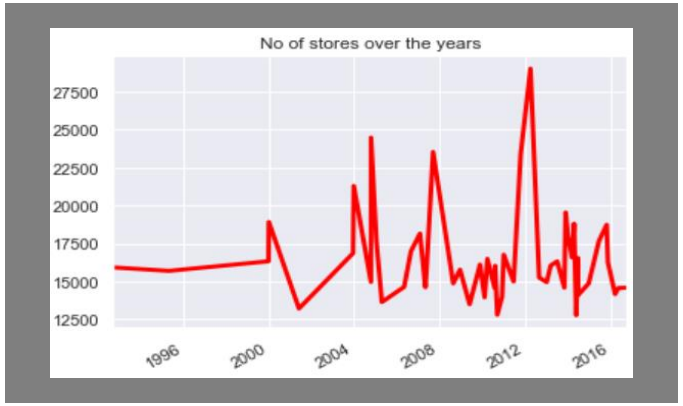
## Location wise units sold

- We know that there are 4 locations in this dataset. The units sold each location is shown in the adjacent donut chart.
- We see that location Downtown has the highest number of units sold, followed by Commercial.



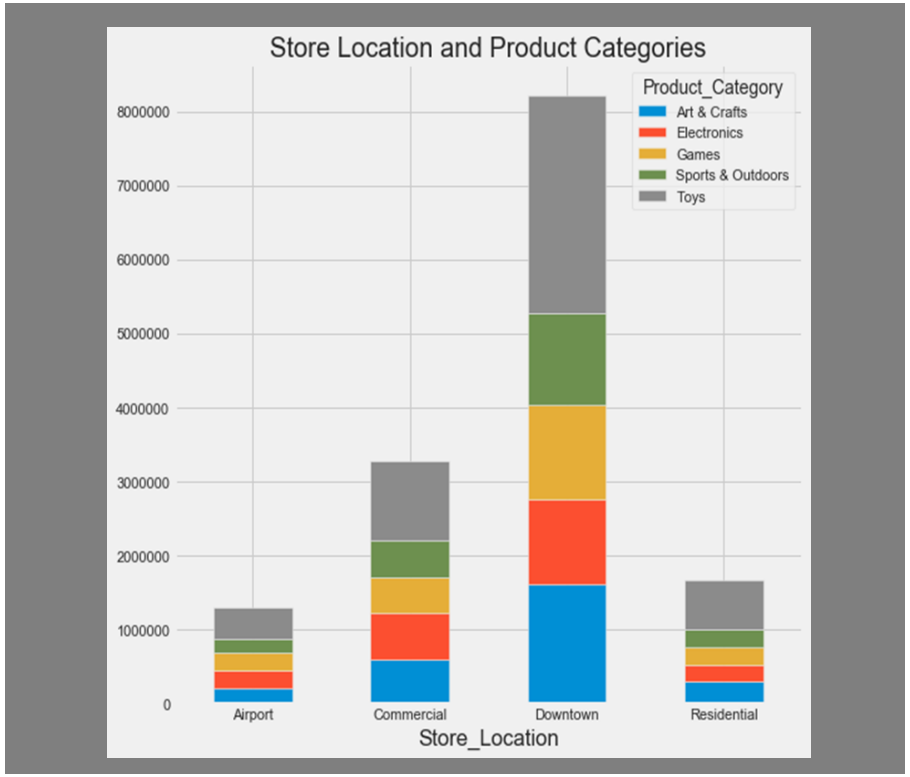

## Number of Stores over the years

- The above line graph tells us the number of stores that were opened in different years. The first store was opened in the year 1992 and the latest store was opened in 2016.
- From the line graph we see that 2012 had the highest number of shops opened (ie) 29024 stores.



## Store Location and Product Category

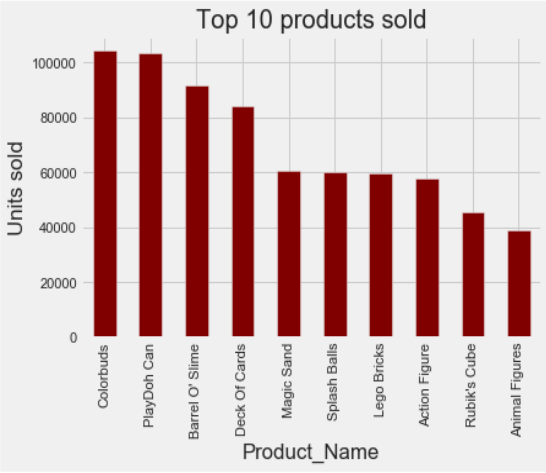
- The below graph is a stack plot that shows the total number of product_categories :- (Art&Crafts, Electronics, Games, Sports and outdoors and toys) sold in various Store_locations. The legends are indicated at the right top corner.
- We know that Downtown has the highest number of units sold. Among the 5 categories, Downtown sells more toys.



## Top 10 Products Sold

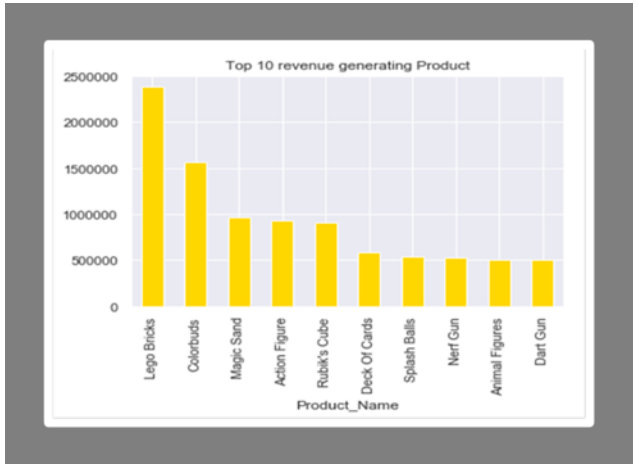
- The dataset has a column name Product_Name (ie) it tells us the name of the Products. There are 35 different products in this dataset.
- We observe that Colorbuds occupy the top position.

104368 units are sold.

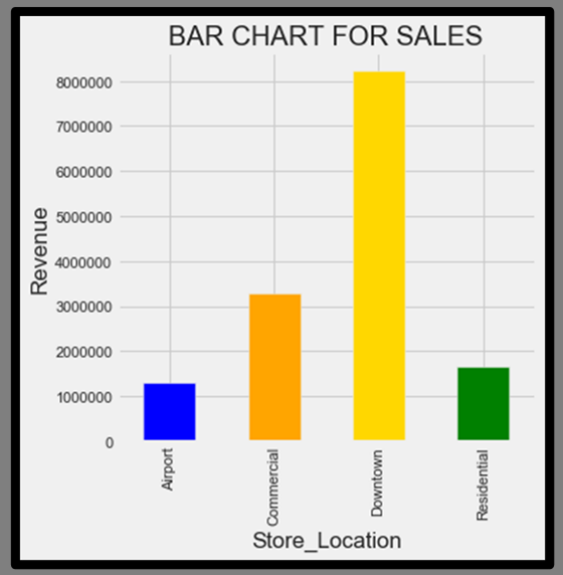


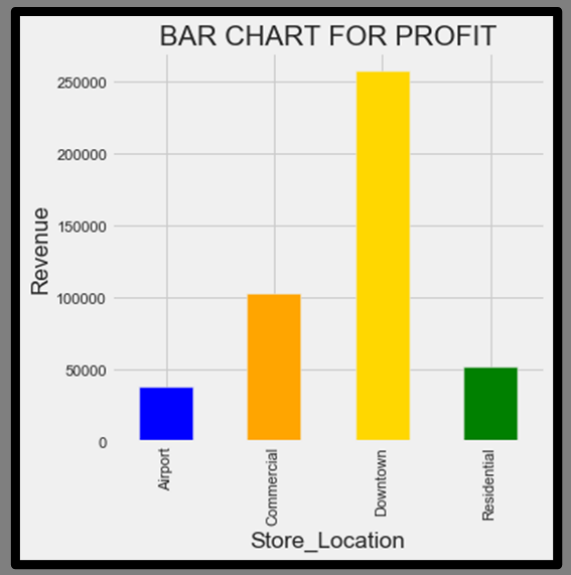

## Top 10 revenue generating Product

- To find the top 10 revenue generating products, we need two columns from the data namely *sales and* Product_Name. For plotting, we need a set of values from the data to be arranged in a particular manner.
- We can see that the Lego Bricks, Colorbuds, Magic Sand are the top 3 product in revenue worth $2.39M, $1.56M, and $9.68M. From this it can be concluded that these product are in demand and we should have surplus quantity in stock.



## Location wise Revenue and Profit




| | Store_Location | sales | Profit |
|---|---|---|---|
| 0 | Downtown | 8.219596e+06 | 257136.0 |
| 1 | Commercial | 3.279140e+06 | 102646.0 |
| 2 | Residential | 1.656114e+06 | 52041.0 |
| 3 | Airport | 1.289723e+06 | 38113.0 |

- With the help of the profit and sales columns we can now find the location wise revenue and profit.
- We see Store located in Downtown area has generated the most of profit as well as revenue whereas store located in Airport area has generated least sales as well as revenue as compared to other locations.

## Profit to Sales Ration Analysis for Store Location

- When we calculated profit to sales ratio and we observe that the store located in Residential area are having highest ratio value as 0.03142.
- This means that store located in Residential area is selling the high profit margin product. If we increase our number of stores in Residential area then the profit may increase even more.

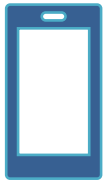
```
#Profit to Sales Ration Analysis for Store Location
top_location1['Profit_to_Sales'] = top_location1['Profit']/top_location1['sales']
top_location1['Profit_to_Sales_in%']= top_location1['Profit']*100/top_location1['sales']
top_location1
```

| | Store_Location | sales | Profit | Profit_to_Sales | Profit_to_Sales_in% |
|---|---|---|---|---|---|
| 0 | Downtown | 8.219596e+06 | 257136.0 | 0.031283 | 3.128329 |
| 1 | Commercial | 3.279140e+06 | 102646.0 | 0.031303 | 3.130272 |
| 2 | Residential | 1.656114e+06 | 52041.0 | 0.031424 | 3.142356 |
| 3 | Airport | 1.289723e+06 | 38113.0 | 0.029551 | 2.955132 |

## Profit over the years

From the below line graph we see that there is decline in the profit from the 2014. The profit is decreased in 2016 from 2015, so company should work on the some lose areas.

# SUGGESTIONS



Increase Sale: Products from electronic category gives a great profit despite the less sale. So management should come up with marketing strategies that will increase the demand for this category.

Target : Company should increase the number of stores in airport area which may results into increase in profit also.

Stock : Colorbuds, Lego Bricks and Magic Stand are the three products whose demand is high and we should have surplus quantity in stock.

# THANK YOU