Bachelor Of Technology

Computer Science and Engineering (Data Science)

311 Program in collaboration with Virginia Tech. USA

Mini-Project Presentation

SVKM's NMIMS

Mukesh Patel School of Technology Management and Engineering, JVPD, Vile Parle, West

Mumbai- 400056

A PRESENTATION ON

"SALES FORECASTING FOR WOWMART"

By

NIHARIKA AMRITKAR

SHREYAS DESAI

KAUNCHI JAIN

Faculty Mentors:

Prof. Ami Munshi, Prof. Mahesh Mali and Prof. Khinal Parmar

Contents

SVKM'S
NMIMS
Deemed-to-be UNIVERSITY

Introduction

For our project, we are taking help of WoW-Mart, a fictional company and using it's sales related dataset to carry out the operations. Our project aims at predicting sales for the store's outlets in various locations and regions and helping the management staff to take decisions like whether a particular region, or outlet in general is proving profitable for them or not and whether they should continue operating there.

Our Experiment will also help WoW-Mart to help them in planning their supplies, incase a specific region or store is expecting heavy demand, then more supplies can be directed towards it, and if some store is expecting fall in demand, then it can either be shut or less resources can be directed towards it.

Literature Survey

We surveyed 3 research papers on Moving Average:

1) Ms. Savita Satav, Dr. Netra Apte: “The Moving Average Crossover Strategy: A Study”, JMME, Volume 10, No. 03, July 2020, pp- 141- 146
2) Puchong Praekhaow: “Determination of Trading Points using Moving Average Methods”, June 2010.
3) Seng Hansun: “A Novel Research of New Moving Average Method in Time Series Anaysis”, August 2014

SVKM'S NMIMS Deemed-to-be UNIVERSITY

Problem Statement

To forecast sales with the help of Moving Averages on the basis of Stores, Locations and Regions for comparative analysis.

To identify crossovers and MAD for prediction of trend and volatility in sales for the next quarter.

SVKM'S
NMIMS
Deemed-to-be UNIVERSITY

System Description

Hardware requirements:

- Functional Computer with sufficient RAM

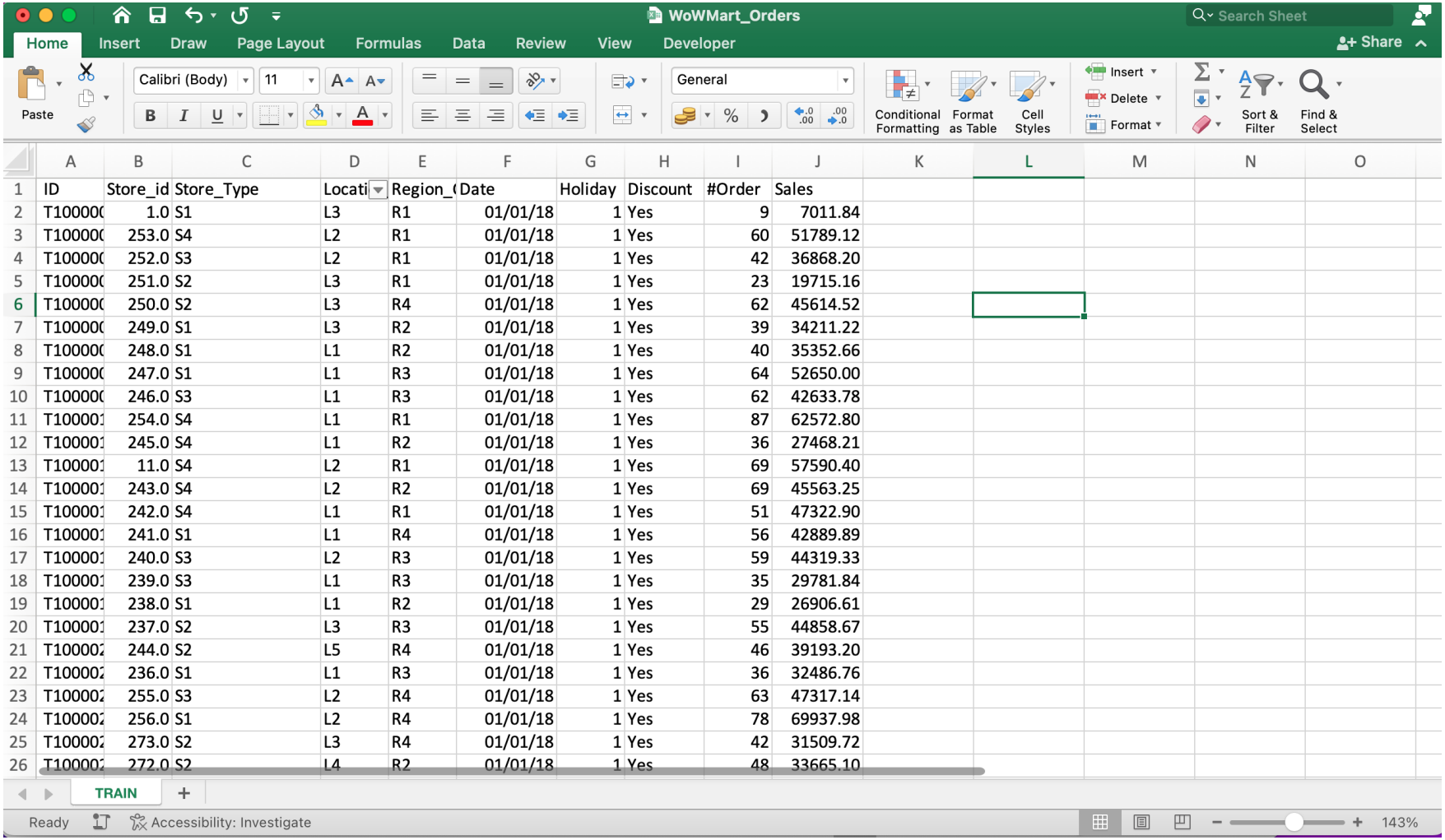
Software requirements:

- Jupyter Notebook
- SAS
- Excel
- Microsoft Presentation, Microsoft Word

Algorithm & Flowchart

- Obtaining dataset through statistic means and organizing it
- Cleaning irrelevant data and managing null values
- Specifying the purpose of the analysis and opting for suitable method to analyze the dataset
- Working through the programming algorithm to develop the fundamentals
- Implementing the analysis technique using python
- Importing dataset in the python notebook and working with the fundamentals developed
- Making program flexible to different types of statistical data
- Running the code and checking if the code works for all types of Inputs
- Error management
- Concluding the finding of the program
- Analyzing the findings of the program
- Commenting on the analysis with regard to sales and predicting the trend with the help of MAD, Crossovers, Variance and Standard Deviation

Our Dataset



WoWMart_Orders

	A	B	C	D	E	F	G	H	I	J
1	ID	Store_id	Store_Type	Locati	Region_(	Date	Holiday	Discount	#Order	Sales
2	T100000(	1.0	S1	L3	R1	01/01/18	1	Yes	9	7011.84
3	T100000(	253.0	S4	L2	R1	01/01/18	1	Yes	60	51789.12
4	T100000(	252.0	S3	L2	R1	01/01/18	1	Yes	42	36868.20
5	T100000(	251.0	S2	L3	R1	01/01/18	1	Yes	23	19715.16
6	T100000(	250.0	S2	L3	R4	01/01/18	1	Yes	62	45614.52
7	T100000(	249.0	S1	L3	R2	01/01/18	1	Yes	39	34211.22
8	T100000(	248.0	S1	L1	R2	01/01/18	1	Yes	40	35352.66
9	T100000(	247.0	S1	L1	R3	01/01/18	1	Yes	64	52650.00
10	T100000(	246.0	S3	L1	R3	01/01/18	1	Yes	62	42633.78
11	T100001	254.0	S4	L1	R1	01/01/18	1	Yes	87	62572.80
12	T100001	245.0	S4	L1	R2	01/01/18	1	Yes	36	27468.21
13	T100001	11.0	S4	L2	R1	01/01/18	1	Yes	69	57590.40
14	T100001	243.0	S4	L2	R2	01/01/18	1	Yes	69	45563.25
15	T100001	242.0	S4	L1	R1	01/01/18	1	Yes	51	47322.90
16	T100001	241.0	S1	L1	R4	01/01/18	1	Yes	56	42889.89
17	T100001	240.0	S3	L2	R3	01/01/18	1	Yes	59	44319.33
18	T100001	239.0	S3	L1	R3	01/01/18	1	Yes	35	29781.84
19	T100001	238.0	S1	L1	R2	01/01/18	1	Yes	29	26906.61
20	T100001	237.0	S2	L3	R3	01/01/18	1	Yes	55	44858.67
21	T100002	244.0	S2	L5	R4	01/01/18	1	Yes	46	39193.20
22	T100002	236.0	S1	L1	R3	01/01/18	1	Yes	36	32486.76
23	T100002	255.0	S3	L2	R4	01/01/18	1	Yes	63	47317.14
24	T100002	256.0	S1	L2	R4	01/01/18	1	Yes	78	69937.98
25	T100002	273.0	S2	L3	R4	01/01/18	1	Yes	42	31509.72
26	T100002	272.0	S2	L4	R2	01/01/18	1	Yes	48	33665.10

TRAIN

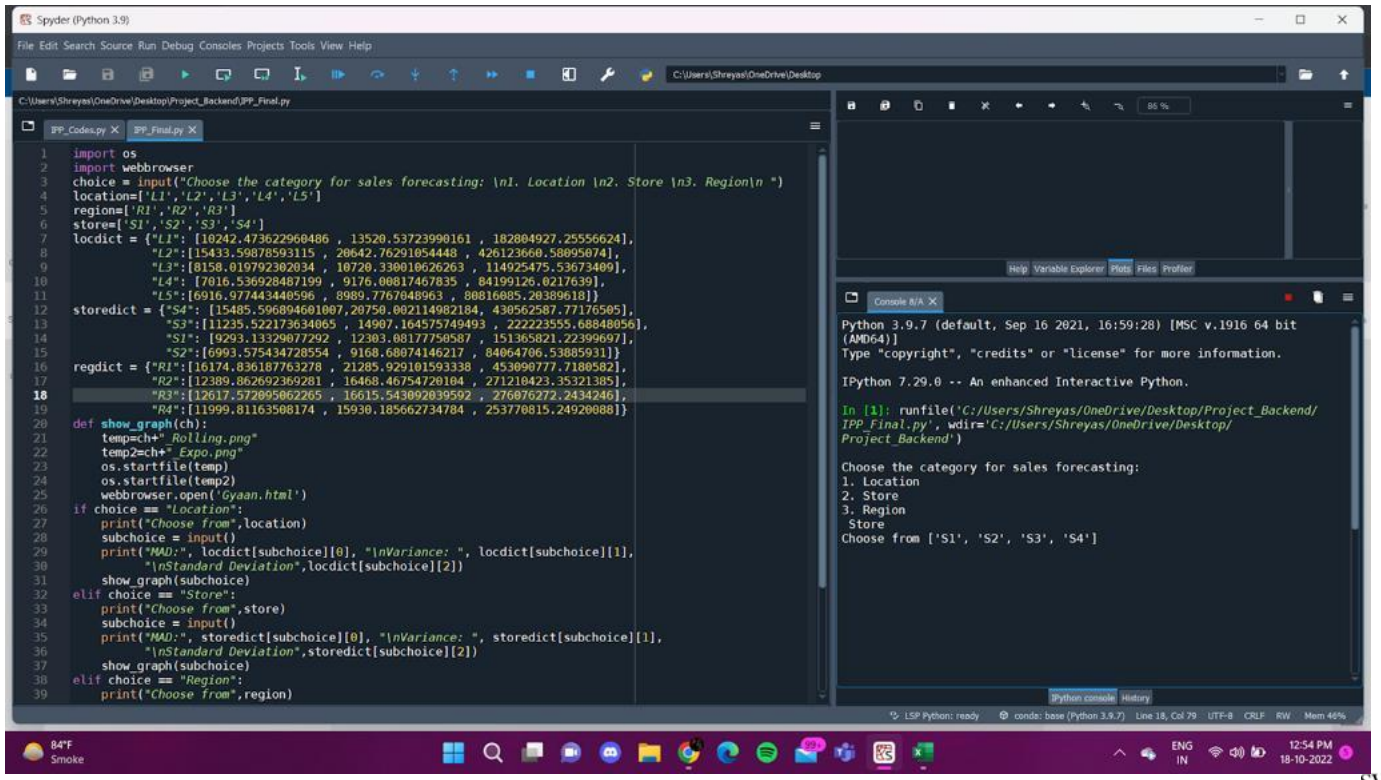
Ready Accessibility: Investigate 143%

Data Cleaning

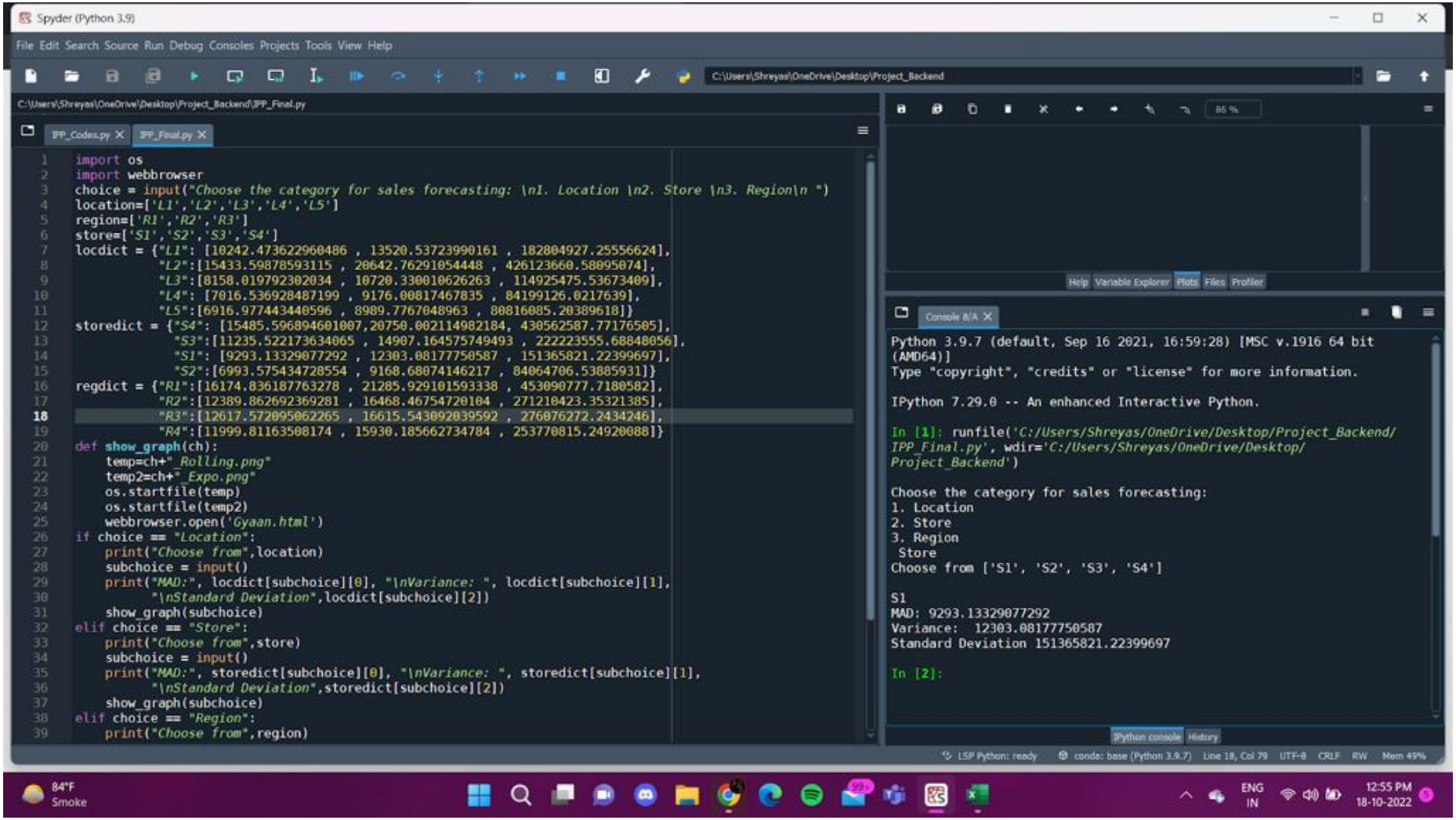
- Discounts, Order Id, Holidays and Order number was removed using python.
- A new variable was created and the useful information was stored in it.
- The variable was later converted to .csv file, which is the data we worked on.

Project Outlook

- Let’s look at the sales of store type S1 and predict how will it perform in the next quarter.



SVKM'S NMIMS Deemed-to-be UNIVERSITY



We chose Store from menu and then chose S1 for the results

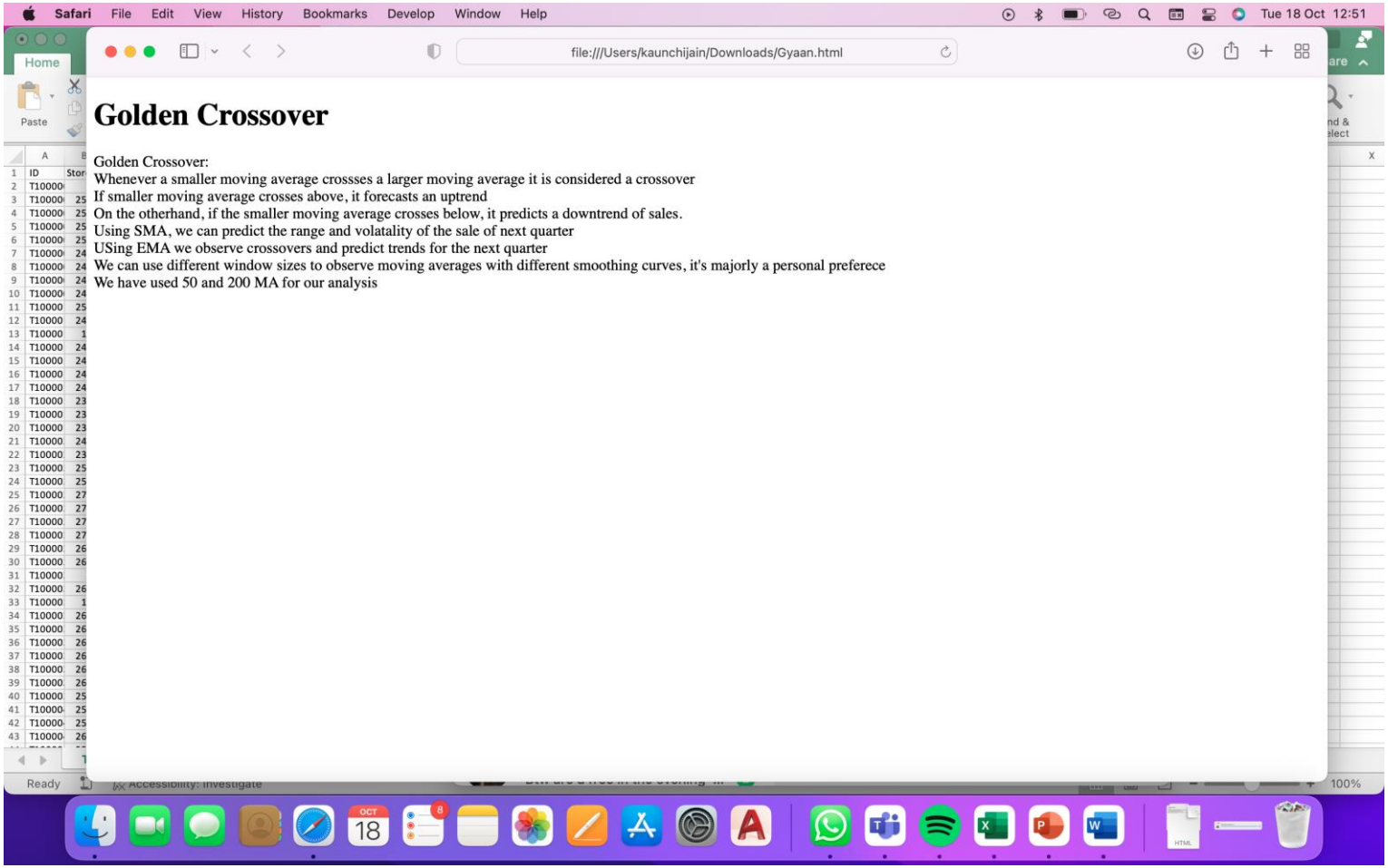
Results:

- MAD is derived
- Standard Deviation is calculated
- Variance is calculated
- 2 graphs are generated:

1. Simple Moving Average with 1 window size
2. Exponential Moving Average with 2 window sizes

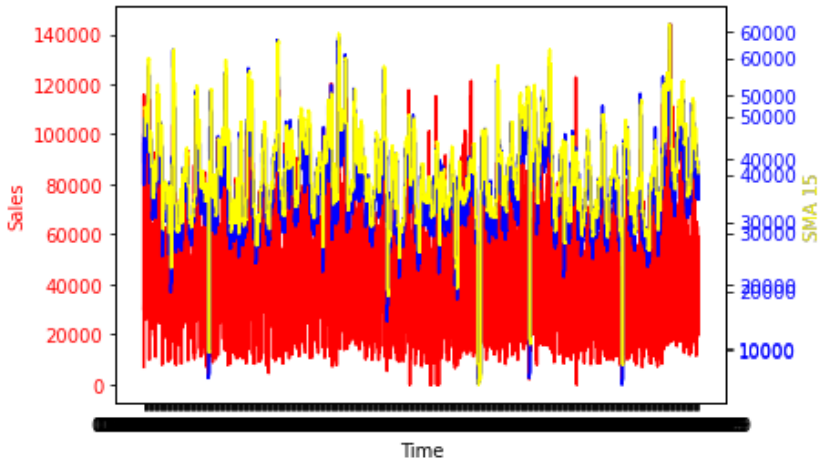
- Manual to analyze the results is popped up

Manual using HTML

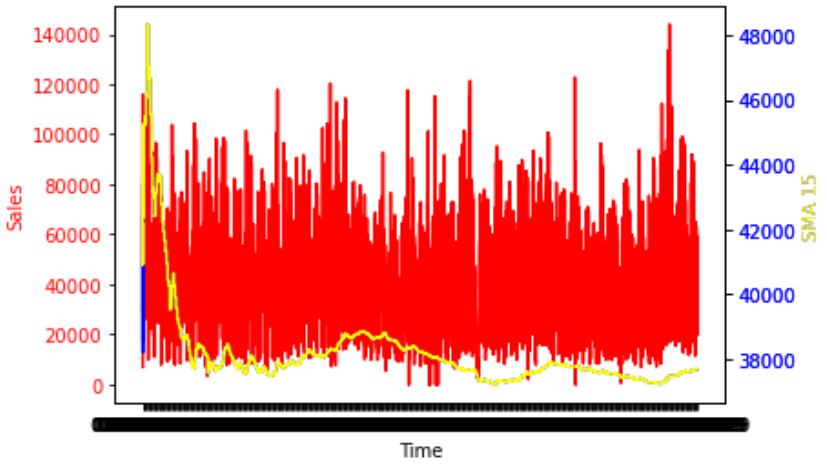


Graphs

- SMA



- EMA



Concepts Used: IBSAM Statistics

- Moving Average

1. Simple Moving Average
2. Exponential Moving Average

- Standard Deviation
- Variance
- Mean Absolute Deviation
- Crossovers in Moving Averages of 2 window sizes
- Forecasting: Uptrend & Downtrends

Implementation: Driver code

```
choicedf = region_list.rolling(50)
choicedf1 = region_list.rolling(200)


moving_average=choicedf.mean()
moving_average_list = moving_average.tolist()
moving_average_list = np.array(moving_average_list)


moving_average1=choicedf1.mean()
moving_average_list1 = moving_average1.tolist()
moving_average_list1 = np.array(moving_average_list1)
```

Calculating values for SMA

```
fig, ax1 = plt.subplots()

ax1.set_xlabel('Time')
ax1.set_ylabel('Sales', color = 'red')
ax1.plot(specific_region['Date'], specific_region['Sales'], color = 'red')
ax1.tick_params(axis ='y', labelcolor = 'red')

ax2 = ax1.twinx()

ax2.set_ylabel('SMA', color = 'blue')
ax2.plot(specific_region['Date'],moving_average_list , color = 'blue')
ax2.tick_params(axis ='y', labelcolor = 'blue')

ax3 = ax1.twinx()

ax3.set_ylabel('SMA', color = 'yellow')
ax3.plot(specific_region['Date'], moving_average_list1, color = 'yellow')
ax3.tick_params(axis ='y', labelcolor = 'blue')

plt.show()
```

Plotting 2 values on y-axis and one value on x axis i.e. dates

Main code

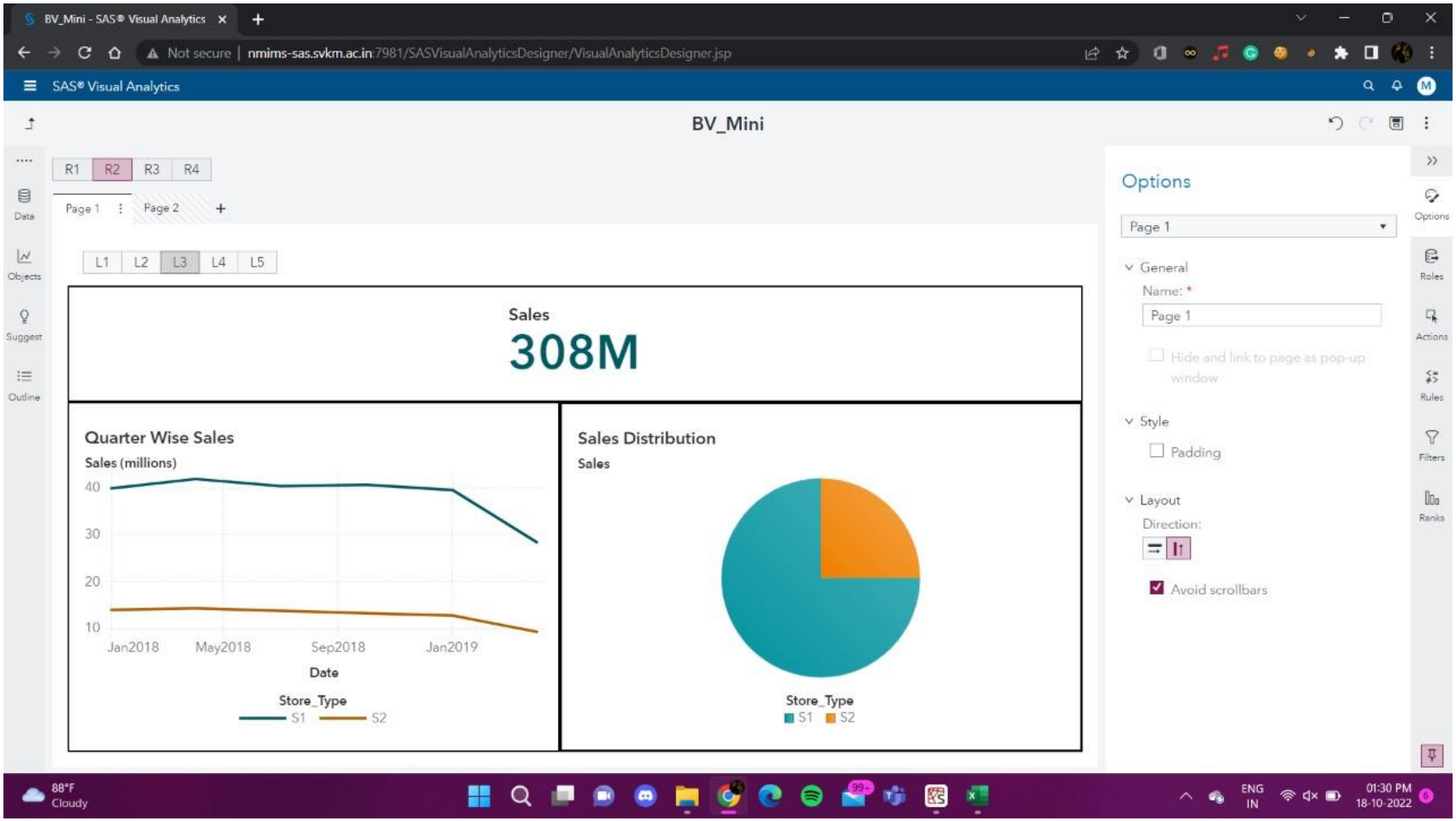
```
def show_graph(ch):
    temp=ch+"_Rolling.png"
    temp2=ch+"_Expo.png"
    os.startfile(temp)
    os.startfile(temp2)
    webbrowser.open('Gyaan.html')
if choice == "Location":
    print("Choose from",location)
    subchoice = input()
    print("MAD:", locdict[subchoice][0], "\nVariance: ", locdict[subchoice][1],
          "\nStandard Deviation",locdict[subchoice][2])
    show_graph(subchoice)
elif choice == "Store":
    print("Choose from",store)
    subchoice = input()
    print("MAD:", storedict[subchoice][0], "\nVariance: ", storedict[subchoice][1],
          "\nStandard Deviation",storedict[subchoice][2])
    show_graph(subchoice)
elif choice == "Region":
    print("Choose from",region)
    subchoice = input()
    print("MAD:", regdict[subchoice][0], "\nVariance: ", regdict[subchoice][1],
          "\nStandard Deviation",regdict[subchoice][2])
    show_graph(subchoice)
else:
    print("Invalid Choice")
```

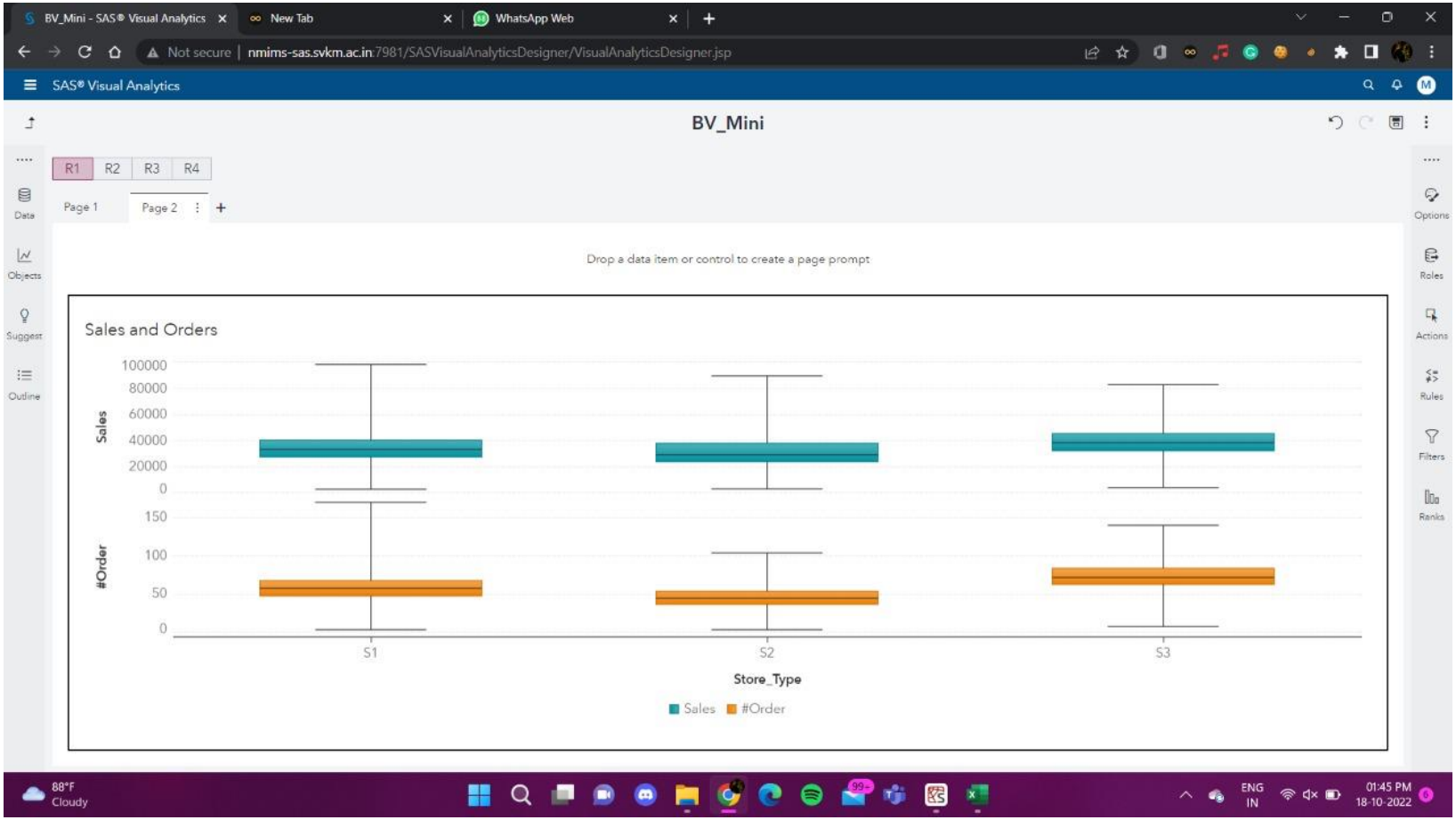
Taking input for required analysis and using code flexibility to predict sales for the same

Concepts used: Python

- Numpy
- Functions
- Pandas
- Matplotlib
- Lambda functions
- Data Structures

Visualization of Data





Directed Page Link

Concepts Used: Business Visualization

- Action Prompts
- Pie Chart
- Line Chart
- Box Plot
- Information Graphics
- Hidden Pages and page links

Result and Analysis

- Whenever a smaller moving average crosses a larger moving average, it is considered a crossover.
- If smaller moving average crosses above, it forecasts an uptrend.
- On the other hand, if the smaller moving average crosses below, it predicts a downtrend of sales.
- Using SMA, we can predict the range and volatility of the sale of next quarter.
- Using EMA we observe crossovers and predict trends for the next quarter.
  We can use different window sizes to observe moving averages with different smoothing curves, it's majorly a personal preference.
- We have used 50 and 200 MA for our analysis.
- If MAD is larger, the sales are more volatile.
- We can use variances and standard deviation to predict sales for the next term.

References

file:///Users/kaunchijain/Downloads/Gyaan.html

https://www.geeksforgeeks.org/python-pandas-dataframe-rolling/

https://www.geeksforgeeks.org/append-extend-python/

https://www.researchgate.net/publication/233988919_Determination_of_Trading_Points_using_the_Moving_Average_Methods

https://www.researchgate.net/publication/301547550_A_Novel_Research_of_New_Moving_Average_Method_in_Time_Series_Analysis

https://core.ac.uk/download/pdf/153776849.pdf

SVKM'S NMIMS Deemed-to-be UNIVERSITY

THANK YOU

SVKM'S
NMIMS
Deemed-to-be UNIVERSITY