DATA SCIENCE WITH AIML

Submitted by :	Submitted to:
1.Shivam purohit	->Rishu Trivedi
2.Vaibhav	LNB mentor
3.Divyansh Singh	

Project: Demand Forecasting

Industry: E Commerce, Retail

1. Project   Overview

This project aims to develop an advance demand Forecasting model to maximum facilitates the retailers. For this purpose it is necessary to optimize retail inventory management and supply chain operations. For this purpose, we have to analyse sales history, promotional events, seasonal trends and economic indicators. The projects objective is to accurately predict product demand and minimise issues such as excess inventory and stockouts.

   To develop this model, we consider the following aspects:

a. Data Collection – Collect various time period data to validate maximum authenticity.

b. Data processing – To find an algorithm between them.

c. Model Development – Considering various aspects we develop a model to detect the maximum possibility areas of demand creation.

Contents

1.Introduction to Pandas

Pandas is an open-source python package built on top of Numpy developed by Wes McKinney**.** It is used as one of the most important data cleaning and analysis tool. It provides fast, flexible, and expressive data structures.

*Pandas is derived from the term "**Pan**el-**da**ta-**s**" an econometrics term for data sets include observations over multiple time periods for the same individuals. -source Wikipedia*

3. How to import Pandas?

```
import pandas as pd
```

By using the above command you can easily import pandas library.

4.Pandas Data Structures

Pandas deals with three types of data structures :

- Series

- DataFrame

- Panel

a)Series is a one-dimensional array-like structure with homogeneous data. The size of the series is immutable(cannot be changed) but its values are mutable.

What is NumPy?

NumPy is a Python library used for working with arrays.

It also has functions for working in domain of linear algebra, fourier transform, and matrices.

NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely.

NumPy stands for Numerical Python.

Why Use NumPy?

In Python we have lists that serve the purpose of arrays, but they are slow to process.

NumPy aims to provide an array object that is up to 50x faster than traditional Python lists.

The array object in NumPy is called `ndarray`, it provides a lot of supporting functions that make working with `ndarray` very easy.

Arrays are very frequently used in data science, where speed and resources are very important.

Which Language is NumPy written in?

NumPy is a Python library and is written partially in Python, but most of the parts that require fast computation are written in C or C++.

Import NumPy

Once NumPy is installed, import it in your applications by adding the `import` keyword:

```
import numpy
```

Now NumPy is imported and ready to use.

Example Get your own Python Server

```
import numpy

arr = numpy.array([1, 2, 3, 4, 5])

print(arr)
```

Try it Yourself »

NumPy as np

NumPy is usually imported under the `np` alias.

alias: In Python alias are an alternate name for referring to the same thing.

Create an alias with the `as` keyword while importing:

```python
import numpy as np
```

Now the NumPy package can be referred to as np instead of numpy.

Example

```python
import numpy as np

arr = np.array([1, 2, 3, 4, 5])

print(arr)
```

Checking NumPy Version

The version string is stored under __version__ attribute.

Example

```python
import numpy as np

print(np.__version__)
```

Matplotlib Pyplot

❮ PreviousNext ❯

Pyplot

Most of the Matplotlib utilities lies under the pyplot submodule, and are usually imported under the plt alias:

```python
import matplotlib.pyplot as plt
```

Now the Pyplot package can be referred to as plt.

Example Get your own Python Server
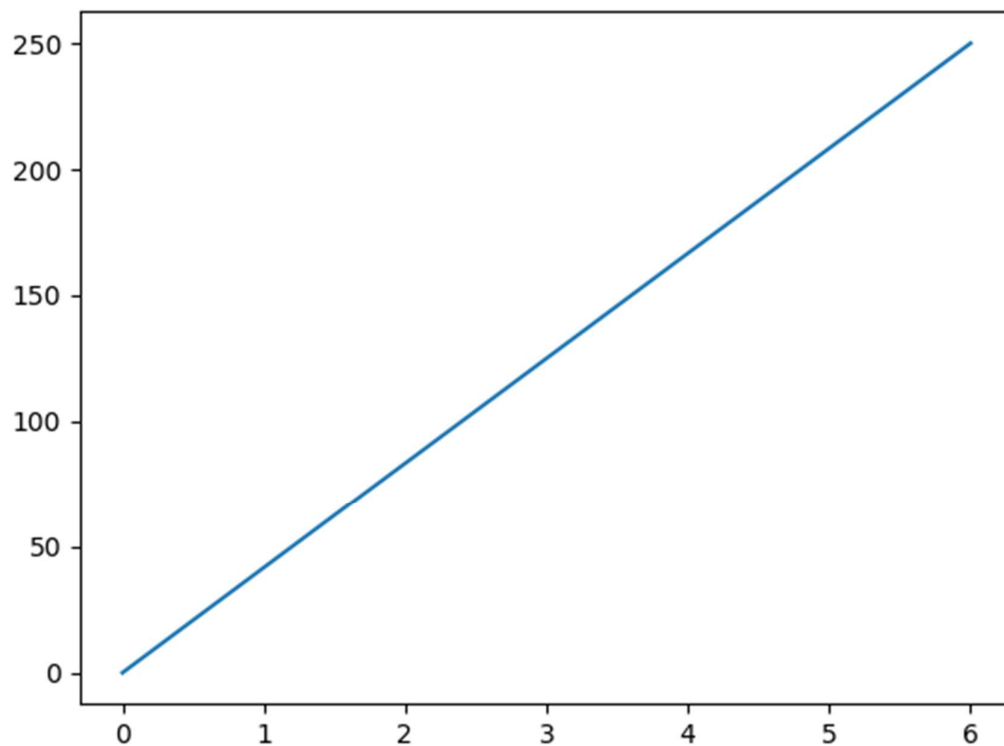
Draw a line in a diagram from position (0,0) to position (6,250):

```
import matplotlib.pyplot as plt
import numpy as np

xpoints = np.array([0, 6])
ypoints = np.array([0, 250])

plt.plot(xpoints, ypoints)
plt.show()
```

Result:

Matplotlib Plotting

Plotting x and y points

The `plot()` function is used to draw points (markers) in a diagram.

By default, the `plot()` function draws a line from point to point.

The function takes parameters for specifying points in the diagram.

Parameter 1 is an array containing the points on the **x-axis**.

Parameter 2 is an array containing the points on the **y-axis**.

If we need to plot a line from (1, 3) to (8, 10), we have to pass two arrays [1, 8] and [3, 10] to the plot function.

Example Get your own Python Server

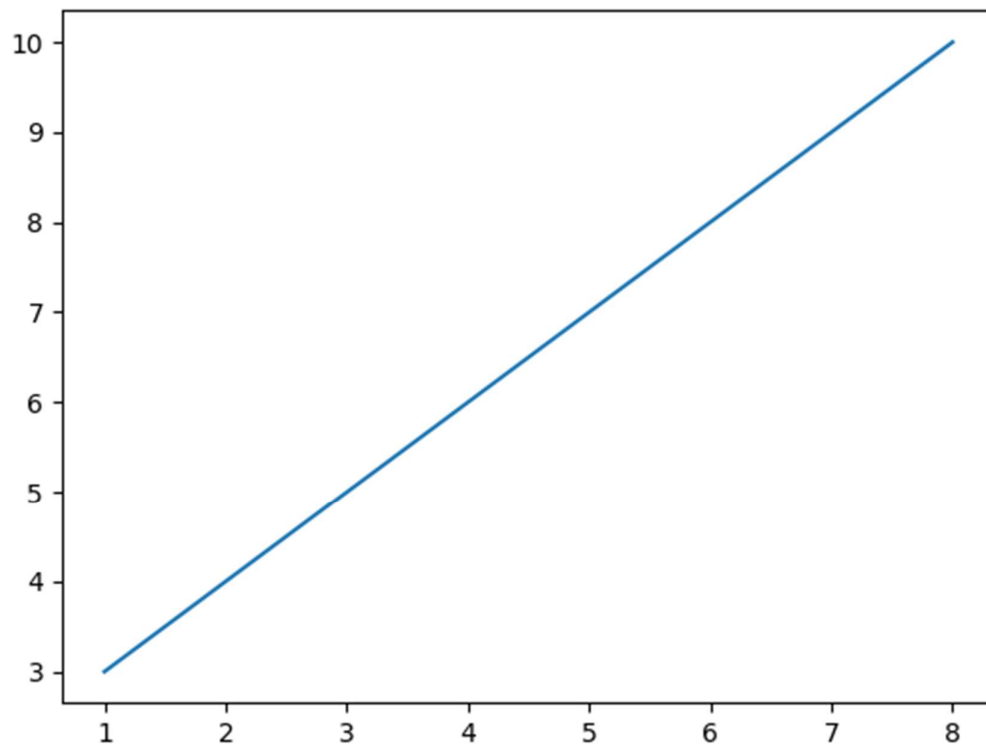Draw a line in a diagram from position (1, 3) to position (8, 10):

```python
import matplotlib.pyplot as plt
import numpy as np

xpoints = np.array([1, 8])
ypoints = np.array([3, 10])

plt.plot(xpoints, ypoints)
plt.show()
```

Result:

The **x-axis** is the horizontal axis.

The **y-axis** is the vertical axis.

Plotting Without Line

To plot only the markers, you can use *shortcut string notation* parameter 'o', which means 'rings'.
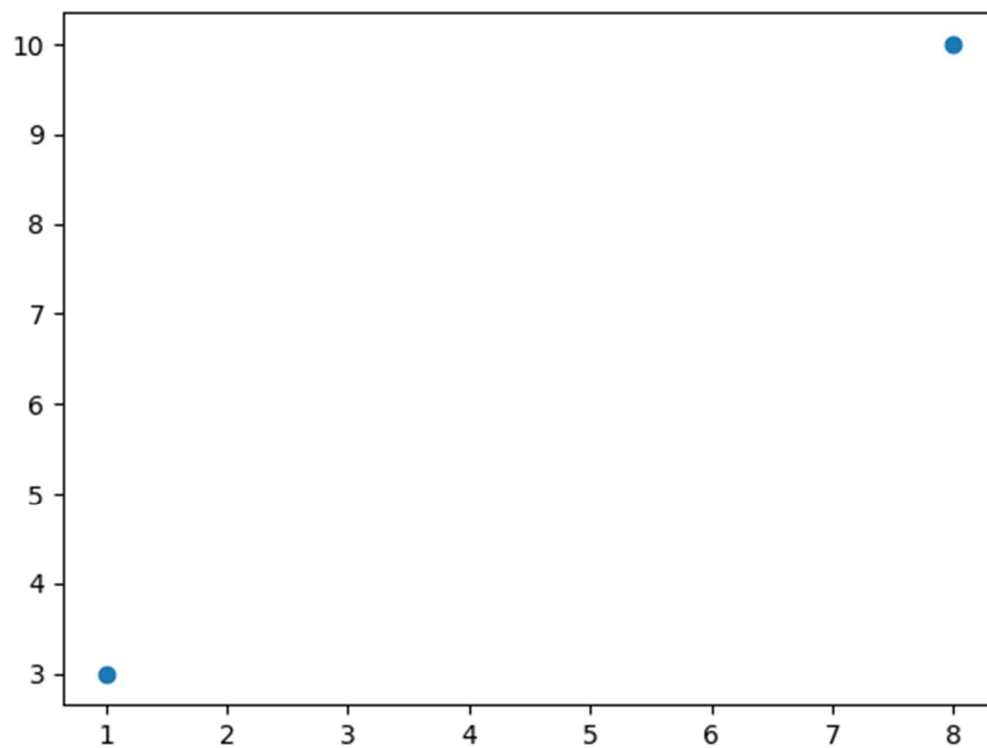
Example

Draw two points in the diagram, one at position (1, 3) and one in position (8, 10):

```python
import matplotlib.pyplot as plt
import numpy as np

xpoints = np.array([1, 8])
ypoints = np.array([3, 10])

plt.plot(xpoints, ypoints, 'o')
plt.show()
```

Result:

Try it Yourself »

Features	Matplotlib	Seaborn
Functionality	It is utilized for making basic graphs. Datasets are visualised with the help of bargraphs, histograms, piecharts, scatter plots, lines and so on.	Seaborn contains a number of patterns and plots for data visualization. It uses fascinating themes. It helps in compiling whole data into a single plot. It also provides distribution of data.
Syntax	It uses comparatively complex and lengthy syntax. Example: Syntax for bargraph-matplotlib.pyplot.bar(x_axis, y_axis).	It uses comparatively simple syntax which is easier to learn and understand. Example: Syntax for bargraph-seaborn.barplot(x_axis, y_axis).
Dealing Multiple Figures	We can open and use multiple figures simultaneously. However they are closed distinctly. Syntax to close one figure at a time: matplotlib.pyplot.close(). Syntax to close all the figures: matplotlib.pyplot.close("all")	Seaborn sets time for the creation of each figure. However, it may lead to (OOM) out of memory issues
Visualization	Matplotlib is well connected with Numpy and Pandas and acts as a graphics package for data visualization in python. Pyplot provides similar features and syntax as in	Seaborn is more comfortable in handling Pandas data frames. It uses basic sets of methods to provide beautiful graphics in python.

Features	Matplotlib	Seaborn
	MATLAB. Therefore, MATLAB users can easily study it.	
Pliability	Matplotlib is a highly customized and robust	Seaborn avoids overlapping of plots with the help of its default themes
Data Frames and Arrays	Matplotlib works efficiently with data frames and arrays.It treats figures and axes as objects. It contains various stateful APIs for plotting. Therefore plot() like methods can work without parameters.	Seaborn is much more functional and organized than Matplotlib and treats the whole dataset as a single unit. Seaborn is not so stateful and therefore, parameters are required while calling methods like plot()
Use Cases	Matplotlib plots various graphs using Pandas and Numpy	Seaborn is the extended version of Matplotlib which uses Matplotlib along with Numpy and Pandas for plotting graphs

Conclusion:

The aim of this project is to facilitate the retailers by detecting the demand accruals and setbacks. By the sincere guidance of our respected faculties of LNB and our esteemed GIT Faculties we completed this project. We are heartly thankful to our respected teachers of GIT and LNB.

Thanks with Regards.