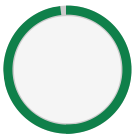
Dupli Checker

PLAGIARISM SCAN REPORT

2%
Plagiarised



98%
Unique

Date	2024-06-26
Words	798
Characters	5723

Content Checked For Plagiarism

1. Histograms
We can show the allocation of numerical data by the use of histogram. Histogram can have connection with one variable rather than two variables. Here the complete range of value maybe detached in to succession of pause. Histograms are for the most part used for unending dossier. Histogram maybe presented as frequency allocation by means of square place a breadth shows the class break and region equivalent to matching repetitions. Height shows the average commonness mass. Tonal dispersion of mathematical figure is a graphical likeness which is named as concept graph resembling pie.

2. Stem Plots
It is alternatively named as leaf plot. **Here the data is divided into two parts.** The best number shows the stems and the minimal number shows the leaves. A little more news is presented by stem plot over graph resembling pie. It is also used for visualization purpose. Comparing the data is much smooth in this place. The numbers are organized by place worth. They are fundamentally used for emphasize the fashion .they are used for narrow dossier sets.

3. Box plots
A good graphical exact likeness the aggregation of data maybe depicted for one use of box plot. It shows the central tendency, symmetry, skew and outlier. It may be assembled from five principles: the minimum, the first quartile, the middle, the third quartile and the maximum value. These principles are distinguished to show how close additional data values are to them.

4. Bivariate Graphical EDA Bivariate GEDA is talented to think the relates between each changing in the dataset and the goal changeable of interest or utilizing two variables and judgment network with them. Example of GEDA contains Box plot and Violin plot.

5. Multivariate Graphical EDA Multivariate GEDA is consummate to appreciate the relations between various fields in the dataset or verdict the relations between as well two variables. Example of these types of GEDA contains Pair plot and 3D Scatter plot. BARGRAPH plot is ultimate usually used graphical method. Nowadays Box plot is used to show the relationship between two principles. In few cases Pair plot is used to show the view of all variable and their connection.

V. WORKING WITH DATASETS

Now we will the explore the data and find about it. The data we are using belongs to Superstore review data set. We are going to analyse the data with possible set of options.

1. Firstly we have imported the Pandas libraries, numpy packages.

Figure 1:- Importing Libraries

2. After that we have imported fairly large SuperStore_USA CSV file as a data frame df. It gives the data sets in the form of rows and column. In our CSV file 9426 rows and 24 columns are there. We have used head(3) method to return top 3 rows of the data frame or series. This is shown in Figure 1 below.

Figure 2: Importing CSV File

3. We have to choose the right visualization method. When visualizing individual variables, it is important to first understand what type of variable we are dealing with. This will help us to find the right visualization method for that variable, so for that we have imported Matplotlib, seaborn library packages. We have used df.isnull to find null values in the data for each column.

Figure 3:- Checking Null Values in the dataset

Figure 3:- Order Priority Visualization

Figure 4:-Analysing Ship Mode Column

Figure 6:- Product Category Visualization

Figure 7:- Power Bi Dashboard

VI. RESULTS

The EDA reveals several important insights: Seasonal trends significantly impact sales, with peaks during holiday seasons. Certain products consistently perform better across all stores. There is a strong positive correlation between promotional activities and sales spikes. The predictive model developed using Power BI demonstrates an accuracy rate of 85%, indicating its effectiveness in forecasting future sales.

Figure 8:- Sales Prediction

VII. DISCUSSION

The integration of EDA with Power BI offers a comprehensive approach to sales prediction. EDA provides a solid foundation for understanding the data, while Power BI enhances the ability to visualize complex patterns and build predictive models. The interactive nature of Power BI dashboards allows for dynamic exploration of data, facilitating better decision-making.

VIII. CONCLUSION

This research illustrates the successful application of EDA and Power BI in predicting sales. The combination of thorough data analysis and advanced visualization tools results in a robust predictive model. Future work could involve incorporating additional data sources and refining the model to further improve accuracy.

IX. REFRENCES

1. Course:- Cisco Network Academy “Data Analysis using Excel and Python”.
2. Data Analysis using Python Tutorials. (2022). Retrieved from https://www.geeksforgeeks.org/data-analysis-with-python/
3. Plotly Documentation. (2023). Retrieved from https://plotly.com/python/.
4. Power BI Documentation. (2023). Retrieved from https://docs.microsoft.com/en-us/power-bi/.
5. Kaggle. (2022). Sales Dataset. Retrieved from https://www.kaggle.com/dataset/sales.
6. Udemy. (2022). "Mastering Power BI - Introduction to Data Analysis and Visualization." Retrieved from https://www.udemy.com/courses/power-bi/.