Data analyst project  
**Diwali\_Festival\_Sales\_Analysis**

Imp Words:-

1. Numpy :- NumPy is a Python library that stands for "Numerical Python." It provides support for large, multi-dimensional arrays and matrices, along with mathematical functions to operate on these arrays. In simple terms, NumPy makes it easier to perform numerical operations on large sets of data.
2. Pandas:- Pandas is like a super-tool for working with this kind of data in Python. It provides you with a bunch of functions and tools to easily manipulate and analyze this tabular data.

Each column might represent a different type of information (like names, ages, or scores), and each row is a set of values for a particular instance (like a person or a measurement).

* With Pandas, you can do things like:
* Read Data: Import data from various file formats, like Excel, CSV, or databases.
* View Data: Look at parts of your data easily to understand what's in it.
* Filter and Select: Choose specific rows or columns based on certain conditions.
* Modify Data: Add or change values in your table.
* Group and Aggregate: Group your data based on certain criteria and calculate summaries.
* Merge and Join: Combine multiple tables based on common columns.
* Handle Missing Data: Deal with empty or missing values in a smart way.
* Plotting: Visualize your data using plots and charts.

1. Matplotlib:- Matplotlib is a popular Python library used for creating static, animated, and interactive visualizations in a simple and effective way. It provides a wide range of functions for producing various types of plots, charts, and graphs, such as line plots, scatter plots, bar charts, histograms, and more.

In simpler terms, if you have data that you want to visualize in the form of a graph or chart, Matplotlib helps you do that easily with Python code. It's like a tool that allows you to turn your data into visually understandable representations, making it easier to analyze and interpret.

1. Seaborn:- Seaborn is a Python data visualization library that makes it easy to create informative and attractive statistical graphics. It is built on top of Matplotlib, another popular visualization library, and provides a high-level interface for drawing attractive and informative statistical graphics.

In simpler terms, Seaborn helps you create visually appealing charts and plots for your data, making it easier to understand and interpret. It simplifies the process of creating common statistical visualizations, such as scatter plots, bar plots, and heatmaps, with a focus on aesthetics and ease of use. Seaborn is particularly useful for exploring and presenting patterns and relationships in your data.

Project steps:-

1. We install all the above library .
2. Then we import the csv file in jupyter notebook
3. And the we apply all the data cleaning function(eg:- df. xyz)
4. Then we do EDA (Exploratory Data Analysis)

