

# Global SuperMarket Data Analysis

---

POOJA DAS

# OBJECTIVE

---

- Extract various information such as Sales, Profit, and variance.
- Comparing Sales, Profit, and Quantity with various attributes.
- Extract necessary information about Products and Category.
- Do ETL : Extract, transform and Load the data
- Use various visualization and features and make the best dashboard.
- Find key metrics and factors and show the meaningful relationships between attributes.

# **BENEFITS**

---

- Analyzing the Sales and Profit of various items.
- Determining the Sales in the year 2016
- Extracting the flows and achievements
- Detecting the trends of Profit in each Fiscal year

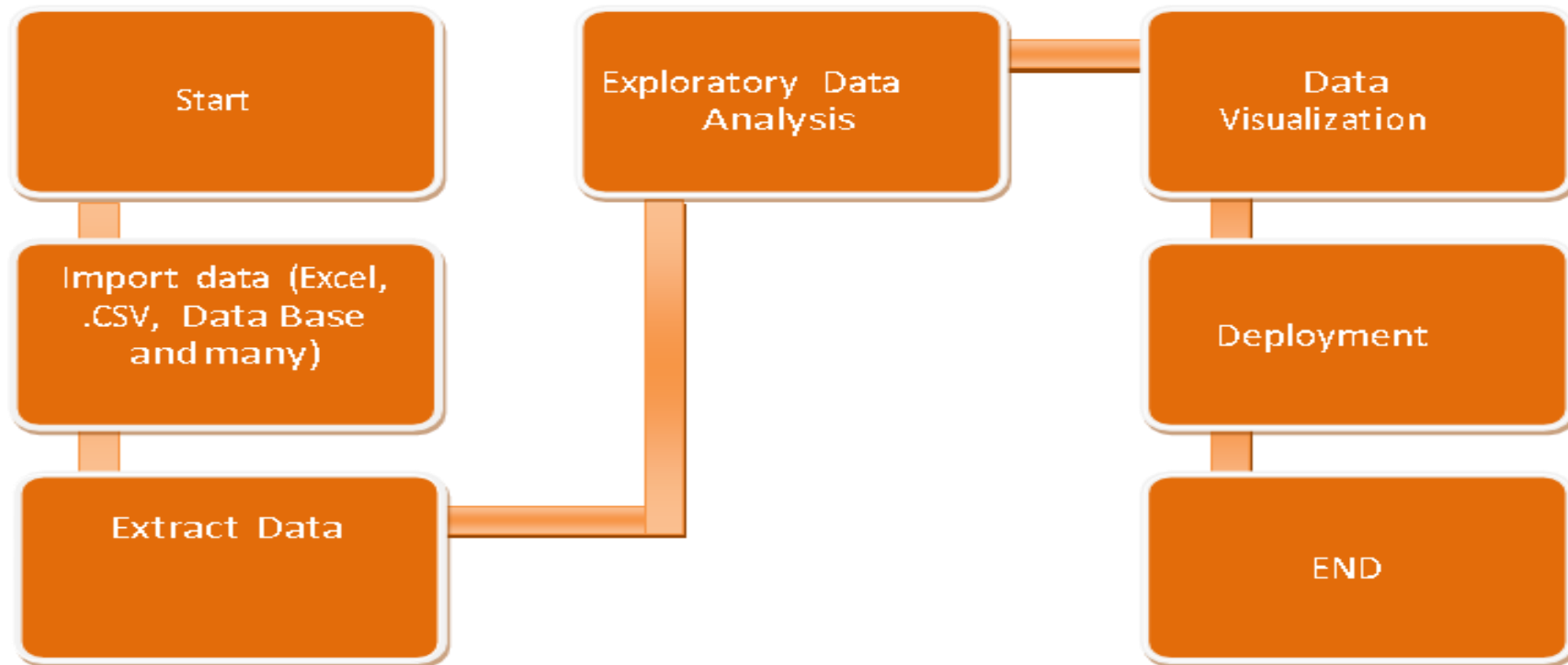
# DATA SHARING AGREEMENT

---

- Sample file name
- Length of date stamp(8 digits)
- Length of time stamp(6 digits)
- Number of Columns
- Column names
- Column data type

# Architecture

---



# Q & A

---

- Q1) What's the source of data?
- The data for training is provided by the client in multiple batches and each batch contain multiple files
- Q 2) What was the type of data?
- The data was a combination of numerical and Categorical values.
- Q 3) After the File validation what do you do with incompatible files or files which didn't pass the validation?
- Files like these are moved to the Achieve Folder and a list of these files has been shared with the client and we removed the bad data folder.

# Q & A

---

- Q 4) What techniques were you using for data pre-processing?
- Removing unwanted attributes
- Visualizing the relation of independent variables with each other and output variables
- Checking and changing the Distribution of continuous values
- Removing outliers
- cleaning data and imputing if null values are present.
- Converting categorical data into numeric values.
- Scaling the data
- Q 5) Where Dashboard deployed?
- Dashbaord was deployed in OneDrive cloud
- [global\\_superstore\\_Dashboard.xlsb](#) (Dashboard link)

# Q & A

---

- Q1) What's the source of data?
- The data for training is provided by the client in multiple batches and each batch contain multiple files
- Q 2) What was the type of data?
- The data was a combination of numerical and Categorical values.
- Q 3) After the File validation what do you do with incompatible files or files which didn't pass the validation?
- Files like these are moved to the Achieve Folder and a list of these files has been shared with the client and we removed the bad data folder.