Customer Dataset Segregation and Analysis

Building a Customer analysis and segregation using IBM Cognos for isualization involves several key steps, including defining analysis objectives, collecting campaign data, and processing and cleaning the data.

Here's a step-by-step guide on how to get started:

1. Define Analysis Objectives:

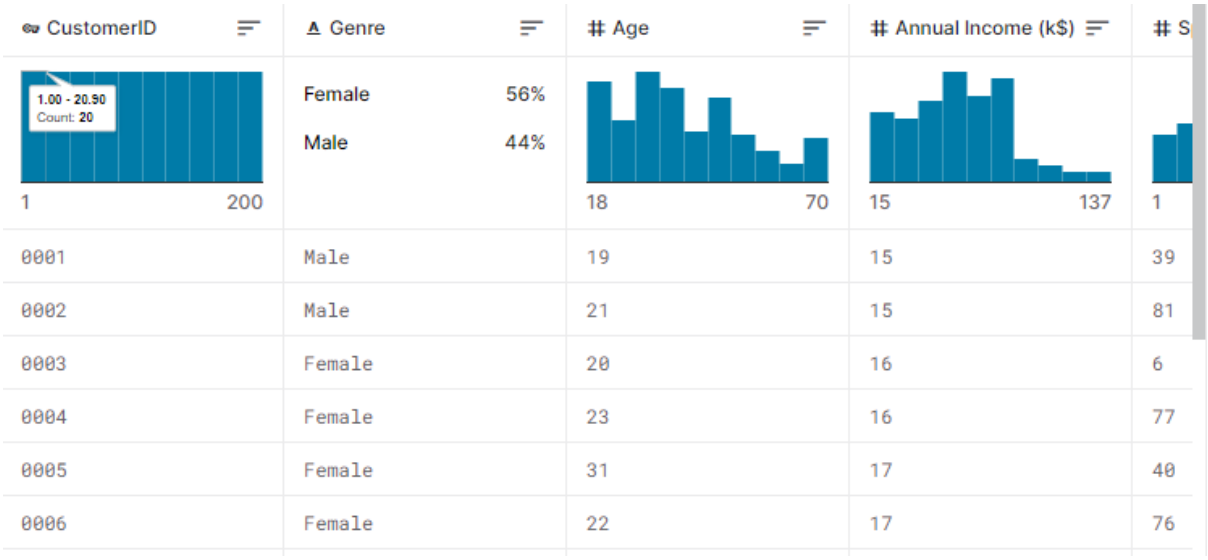
Start by clearly defining the objectives of your customer segregation and analysis. What specific insights are you trying to gain from the data?

The objective of the design analysis for "Customer Segregation and Analysis" is to assess the effectiveness, efficiency, and scalability of the proposed system in partitioning customers into distinct segments based on relevant criteria, and subsequently evaluating the analytical capabilities to derive actionable insights and tailored strategies for each segment. This analysis aims to ensure that the design optimally addresses the business objectives, accommodates diverse customer profiles, and provides a robust framework for ongoing analysis and adaptation.

2.Implementing a dataset

Kaggle dataset:

Dataset Link: https://www.kaggle.com/datasets/akram24/mall-customers



3. Collect Customer Data:

To collect data for your analysis, you'll need access to the data source. This may involve reaching out to relevant public, commodities, retail shops, or partners. Ensure you have permission to use and analyse the data.

In [1]:

```
#imports necessary libraries to do basic things on the dataset
import pandas as pd
import numpy as np

import seaborn as sns
import matplotlib.pyplot as plt


print('Successfully imported')
```

```
Successfully imported
```

4. Process and Clean Data:

Data processing and cleaning are critical to ensuring the quality and accuracy of your analysis. Here are some steps to follow:

a. Data Integration:
   - If your customer data is stored in different formats or sources, integrate them into a single dataset. This may involve using ETL (Extract, Transform, Load) tools.

b. Data Cleaning:

- Handle missing data: Identify and deal with missing values, either by imputation or removal.
- Remove duplicates: Eliminate duplicate records.
- Data format standardisation: Ensure that date formats, units of Measurement, and naming conventions are consistent.

In [3]:

```
#Check the dataset for missing data
if data.isnull().sum().sum() == 0 :
    print ('There is no missing data in our dataset')
else:
    print('There is {} missing data in our dataset '.format(data.isnull().sum().sum()))
```

```
There is 1892 missing data in our dataset
```

In [4]:

```
#Check our missing data from which columns and how many unique features they have.
frame = pd.concat([data.isnull().sum(), data.nunique(), data.dtypes], axis = 1, sort= False)
frame
```

c. Data Transformation:

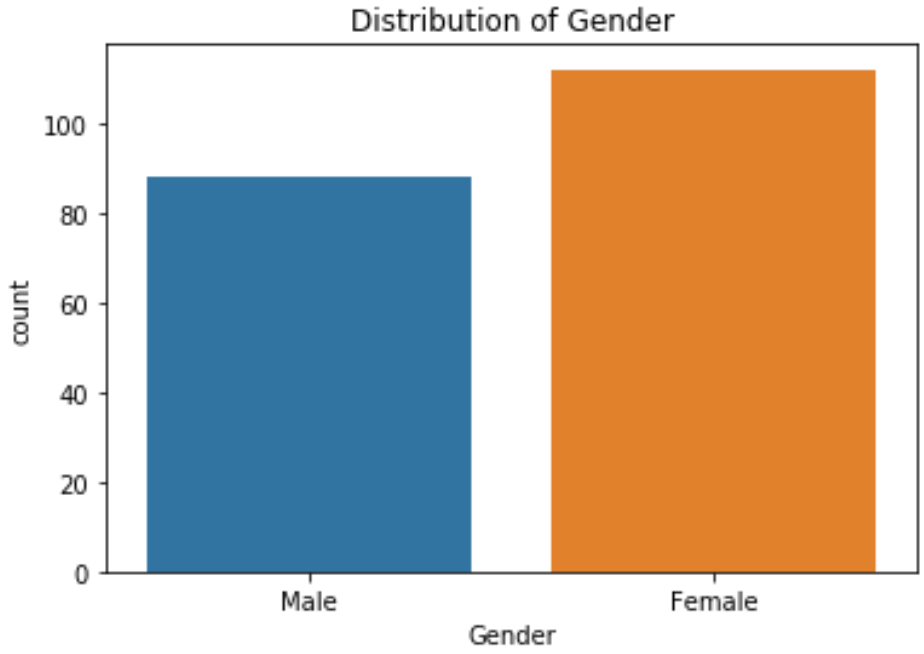
- Convert data types: Ensure that data types are appropriate for analysis (e.g., dates as date objects, numbers as numeric types).
- Create calculated fields: Generate new variables if needed.
- Aggregation: Summarise data as needed, e.g., daily, weekly, or monthly aggregates.

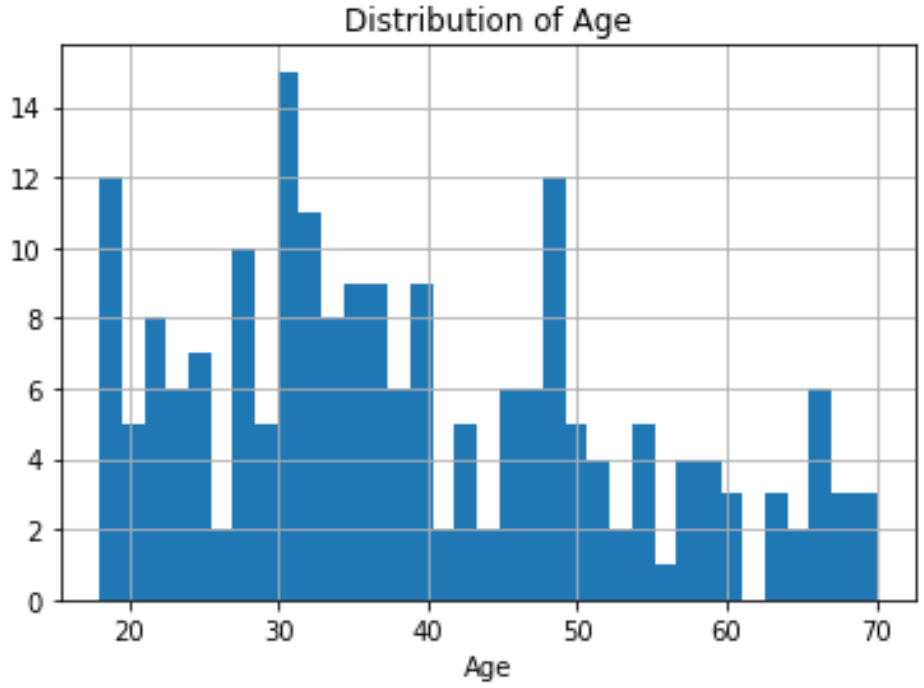
d. Data Quality Check:

- Check for outliers and anomalies that may affect analysis.
- Validate data against the defined analysis objectives.

4. Create Visualizations and Reports:

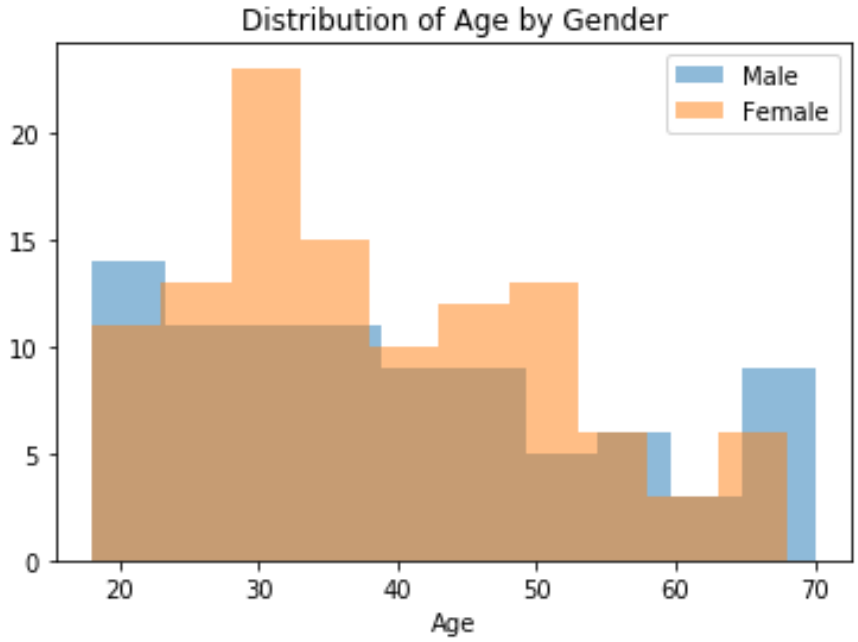
With your data in IBM Cognos, you can start building visualizations and reports to address your analysis objectives. You can use various chart types, tables, and graphs to present the data effectively.





5. Analyze the Data:

Analyze the data to draw insights and conclusions related to your dataset's effectiveness. Use features within IBM Cognos to perform statistical analysis or apply business intelligence techniques to discover patterns and trends.



6. Share Insights:

Finally, share the insights and findings from your analysis with relevant stakeholders, managers, product manager, executives, or other decision-makers. Use the reports and visualizations created in IBM Cognos for this purpose.

7. Iterate and Refine:

Data analysis is an iterative process. If your analysis reveals areas for

improvement in the analysis, work with stakeholders to refine strategies and potentially run new programs to improve the efficiency.

Remember to document your analysis process, including the data sources, data cleaning steps, and the rationale behind your analysis choices. This documentation will be valuable for future reference and for ensuring the transparency and reproducibility of your work.