Cognifyz
Where Data Meets Intelligence

Task 1: Data Overview Internship at Cognifyz Technologies

Kamesh Bhardwaj
Business Analytics Intern

Cognifyz
Where Data Meets Intelligence

Objective: Understand the dataset structure

The objective of this task was to load a given dataset into a data analysis tool (such as Python with pandas) and perform an initial analysis to understand its structure. This involved importing the dataset, checking for basic information like the number of entries, columns, and data types, and using descriptive statistics to ensure the dataset is well-organized. By performing this analysis, I gained insights into the dataset's composition, which is crucial for further data processing and analysis.

Cognifyz
Where Data Meets Intelligence

Load the Dataset

- Imported the pandas library to handle and analyze the dataset.
- Used pd.read_csv() to load the dataset.
- Stored the loaded dataset in a pandas DataFrame named df.
- Previewed the first 5 rows of the dataset using df.head() to check its structure and ensure the data loaded correctly.

Cognifyz
Where Data Meets Intelligence

```python
#pandas library
import pandas as pd

#dataset
df = pd.read_csv(r'C:/Users/bhard/Desktop/DATA ANALYST OFFLINE/Python/internship/Dataset.csv')

#first 5 rows of the dataset
print(df.head())
```

```
   gender  age Investment_Avenues  Mutual_Funds  Equity_Market  Debentures  \
0  Female   34                Yes             1              2           5
1  Female   23                Yes             4              3           2
2    Male   30                Yes             3              6           4
3    Male   22                Yes             2              1           3
4  Female   24                 No             2              1           3

   Government_Bonds  Fixed_Deposits  PPF  Gold  ...           Duration  \
0                 3               7    6     4  ...          1-3 years
1                 1               5    6     7  ...  More than 5 years
2                 2               5    1     7  ...          3-5 years
3                 7               6    4     5  ...   Less than 1 year
4                 6               4    5     7  ...   Less than 1 year

  Invest_Monitor   Expect       Avenue What are your savings objectives?  \
0        Monthly  20%-30%  Mutual Fund                  Retirement Plan
1         Weekly  20%-30%  Mutual Fund                      Health Care
2          Daily  20%-30%       Equity                  Retirement Plan
3          Daily  10%-20%       Equity                  Retirement Plan
```

Cognifyz
Where Data Meets Intelligence

Descriptive Statistics

- Imported the pandas library to analyze the dataset's structure and statistics.
- Used df.info() to gather details about the dataset, such as the number of entries, columns, and data types.
- Analyzed the dataset's structure, ensuring all 40 entries had valid data across 24 columns with no missing values.
- Used df.describe() to calculate basic statistical measures (e.g., mean, standard deviation, min, max) for the numeric columns.
- The descriptive statistics helped in understanding the distribution and range of values for numerical attributes in the dataset.

Cognifyz
Where Data Meets Intelligence

```python
df_info = df.info()

# basic statistical details of the dataset (e.g., mean, std, min, max, etc.)
df_description = df.describe()

# Show the output
print("Dataset Information:")
print(df_info)

print("\nDataset Statistical Summary:")
print(df_description)
```

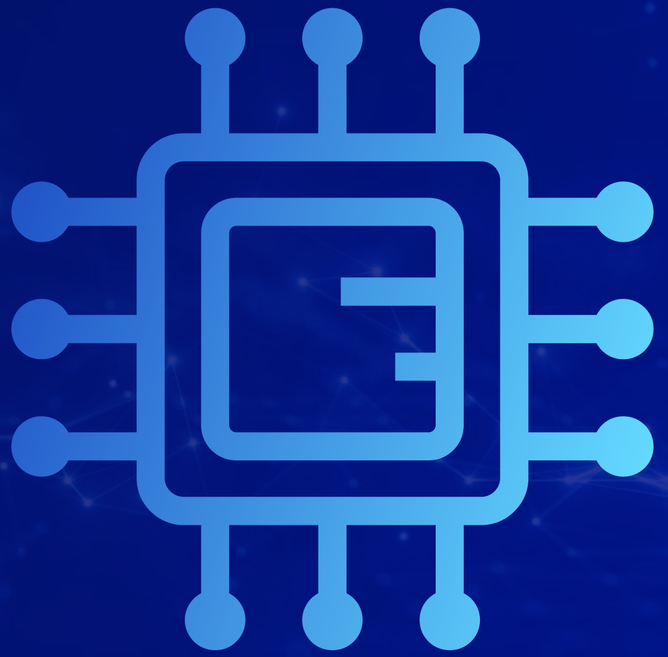
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 40 entries, 0 to 39
Data columns (total 24 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   gender                40 non-null     object
 1   age                   40 non-null     int64
 2   Investment_Avenues    40 non-null     object
 3   Mutual_Funds          40 non-null     int64
 4   Equity_Market         40 non-null     int64
 5   Debentures            40 non-null     int64
 6   Government_Bonds      40 non-null     int64
 7   Fixed_Deposits        40 non-null     int64
 8   PPF                   40 non-null     int64
 9   Gold                  40 non-null     int64
 10  Stock_Marktet         40 non-null     object
 11  Factor                40 non-null     object
 12  Objective             40 non-null     object
```

Cognifyz
Where Data Meets Intelligence

Conclusion & Next Steps

Conclusion:

- Task 1: Data Overview completed successfully (Dataset loaded and analyzed using descriptive statistics).
- Task 2: Gender distribution analysis will be the next step, which will include visualization of the dataset's gender distribution.



Next Steps:

- Task 2: Create and visualize gender distribution (bar chart/pie chart).

Cognifyz
Where Data Meets Intelligence

Thank You!

Feel free to reach out for any questions or feedback.

https://github.com/KameshBhardwaj

https://www.linkedin.com/in/kamesh-bhardwaj-270005321/

-