

```
# IMPORTANT: SOME KAGGLE DATA SOURCES ARE PRIVATE
# RUN THIS CELL IN ORDER TO IMPORT YOUR KAGGLE DATA SOURCES.
import kagglehub
kagglehub.login()
```

```
# IMPORTANT: RUN THIS CELL IN ORDER TO IMPORT YOUR KAGGLE DATA SOURCES,
# THEN FEEL FREE TO DELETE THIS CELL.
# NOTE: THIS NOTEBOOK ENVIRONMENT DIFFERS FROM KAGGLE'S PYTHON
# ENVIRONMENT SO THERE MAY BE MISSING LIBRARIES USED BY YOUR
# NOTEBOOK.

madhvii0911_fuzzylogic_path = kagglehub.dataset_download('madhvii0911/fuzzyl
print('Data source import complete.')
```

```
# This Python 3 environment comes with many helpful analytics libraries inst
# It is defined by the kaggle/python Docker image: https://github.com/kaggle
# For example, here's several helpful packages to load
```

```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
```

```
# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will l
```

```
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

```
# You can write up to 20GB to the current directory (/kaggle/working/) that
# You can also write temporary files to /kaggle/temp/, but they won't be sav
```

```
/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv
/kaggle/input/fuzzylogic/IPL IMB381IPL2013 - IPL IMB381IPL2013.csv
```

fuzzylab - lab\_plotting

q1)plot the barchart

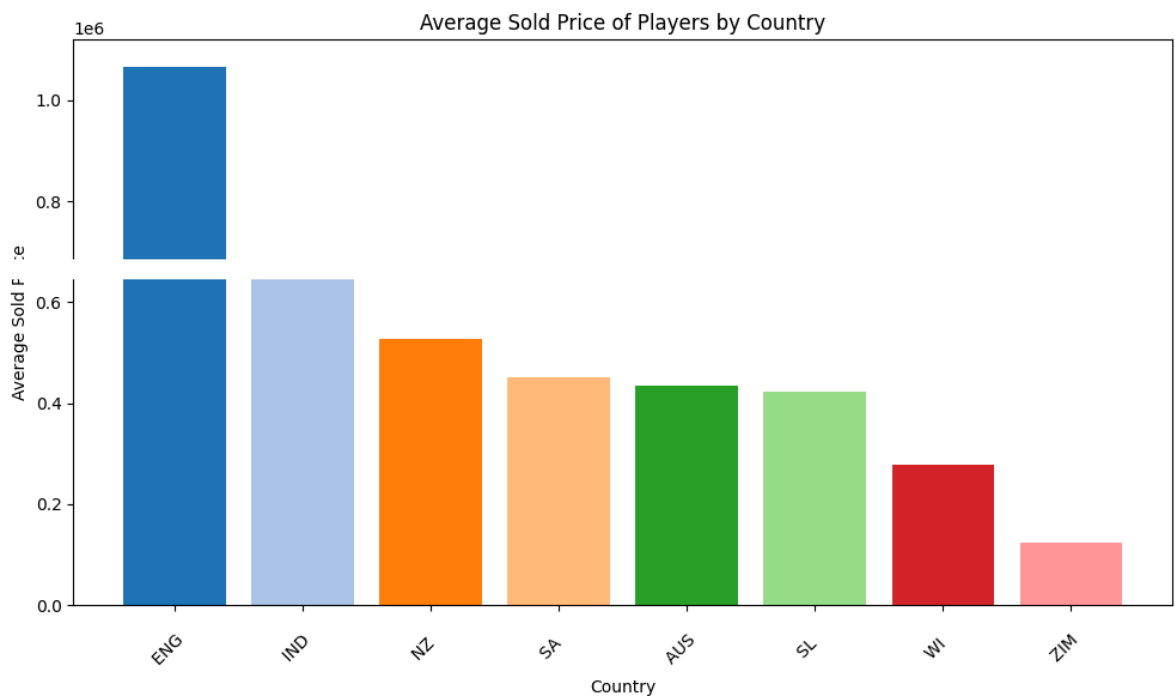
```
import pandas as pd
import matplotlib.pyplot as plt
file_path = "/kaggle/input/fuzzylogic/IPL IMB381IPL2013 - IPL IMB381IPL2013.
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
country_avg_price = (
    df.groupby("COUNTRY")["SOLD PRICE"]
```



```

        .mean()
        .sort_values(ascending=False)
    )
    colors = plt.cm.tab20.colors
    bar_colors = colors[:len(country_avg_price)]
    plt.figure(figsize=(10,6))
    plt.bar(country_avg_price.index, country_avg_price.values, color=bar_colors)
    plt.title("Average Sold Price of Players by Country")
    plt.xlabel("Country")
    plt.ylabel("Average Sold Price")
    plt.xticks(rotation=45)
    plt.tight_layout()
    plt.show()

```



q2)Plot the piechart

```

import pandas as pd
import matplotlib.pyplot as plt
file_path = "/kaggle/input/fuzzylogic/IPL IMB381IPL2013 - IPL IMB381IPL2013."
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
role_counts = df["PLAYING ROLE"].value_counts()
top_roles = role_counts.head(5)
others = role_counts[5:].sum()
role_counts_clean = top_roles.copy()
role_counts_clean["Others"] = others
role_percent = (role_counts_clean / role_counts_clean.sum()) * 100
plt.figure(figsize=(7,7))
wedges, texts, autotexts = plt.pie(
    role_percent,
    labels=role_percent.index,
    autopct='%1.1f%%',

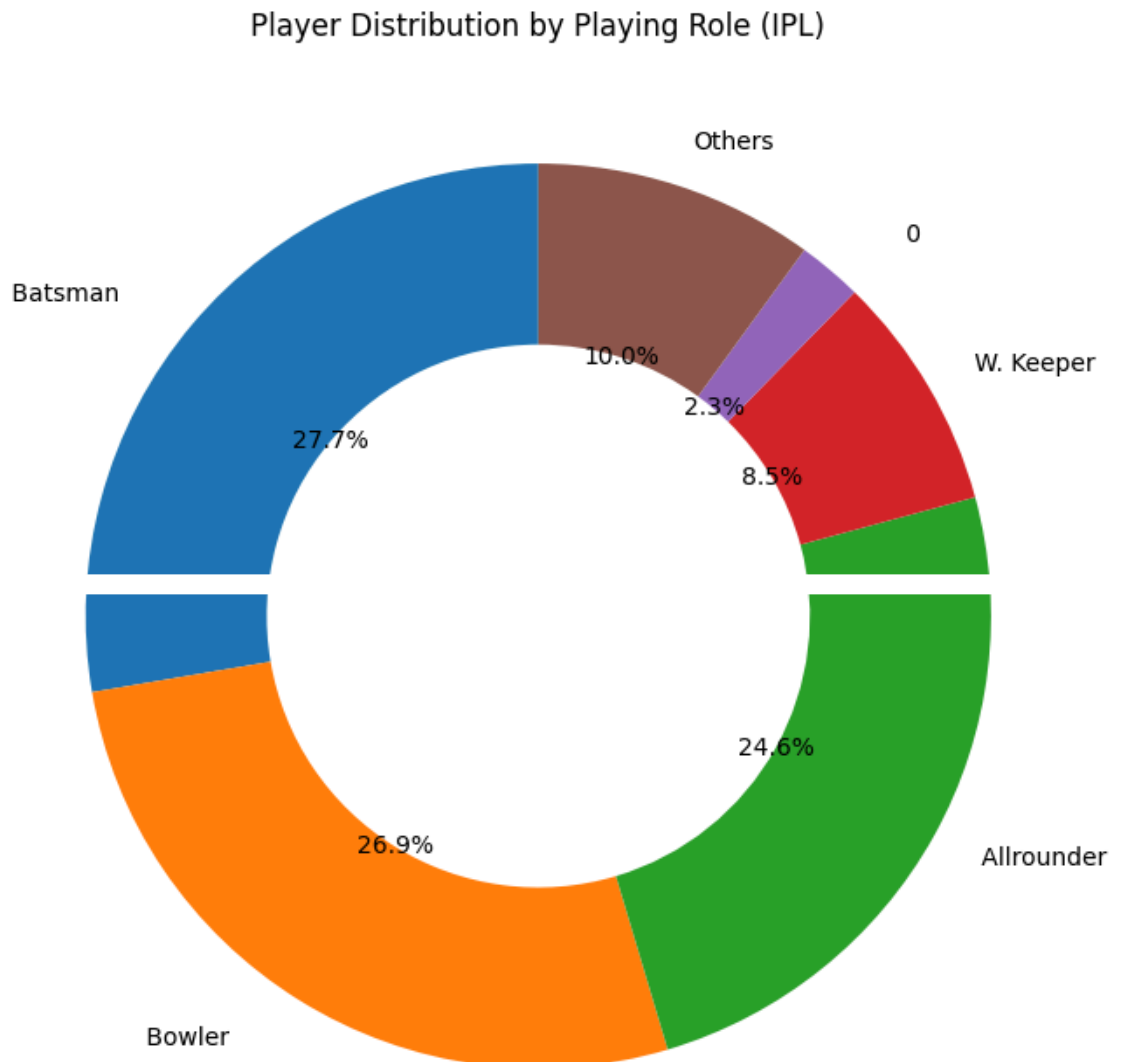
```



```

startangle=90,
wedgeprops={'width':0.4})
plt.title("Player Distribution by Playing Role (IPL)")
plt.tight_layout()
plt.show()

```



q3) Plot the multiseriess barchart

```

import pandas as pd
import matplotlib.pyplot as plt
file_path = "/kaggle/input/fuzzylogic/IPL IMB381IPL2013 - IPL IMB381IPL2013"
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
df["AGE"] = pd.to_numeric(df["AGE"], errors="coerce")
df = df.dropna(subset=["AGE"])
age_map = {
    1: "Under 25",

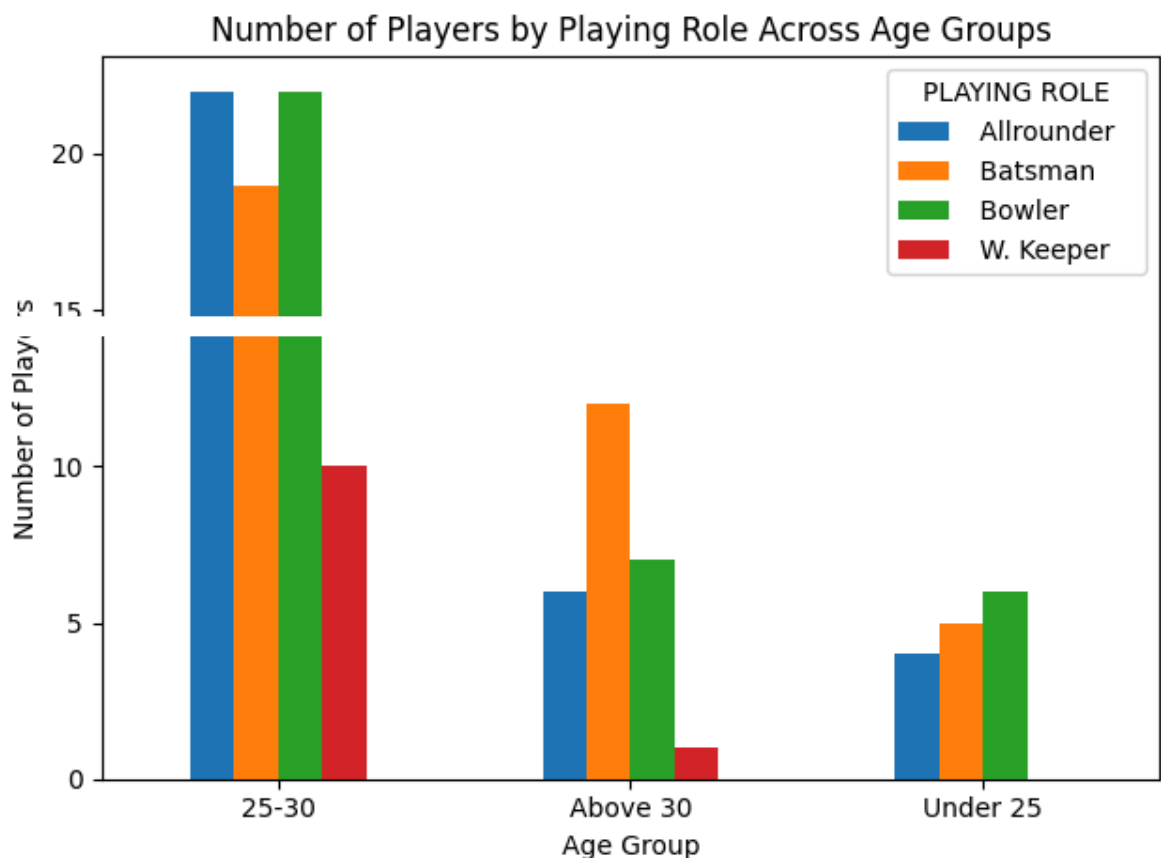
```



```

2: "25-30",
3: "Above 30"
}
df["AGE GROUP"] = df["AGE"].map(age_map)
grouped = df.groupby(["AGE GROUP", "PLAYING ROLE"]).size().unstack(fill_value=0)
grouped.plot(kind="bar")
plt.title("Number of Players by Playing Role Across Age Groups")
plt.xlabel("Age Group")
plt.ylabel("Number of Players")
plt.xticks(rotation=0)
plt.tight_layout()
plt.show()

```



#### q4) Scatterplot

```

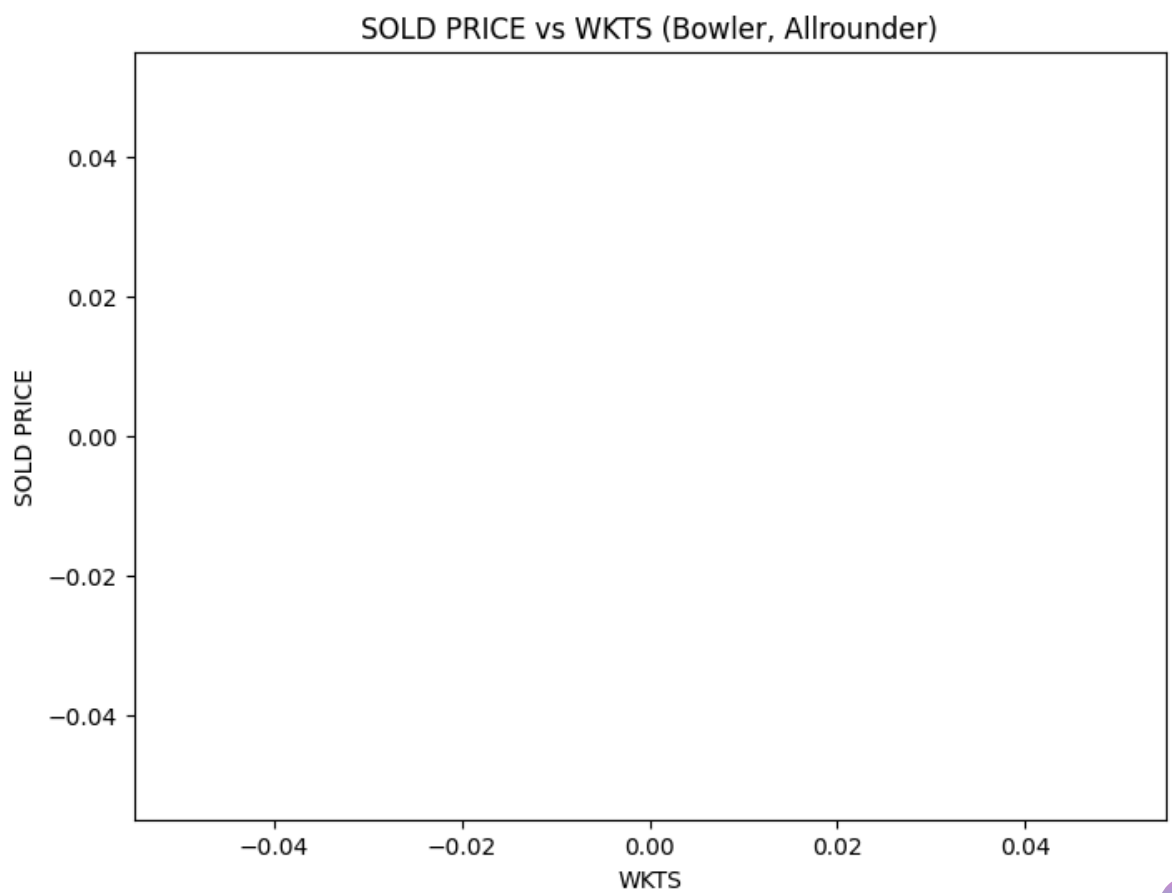
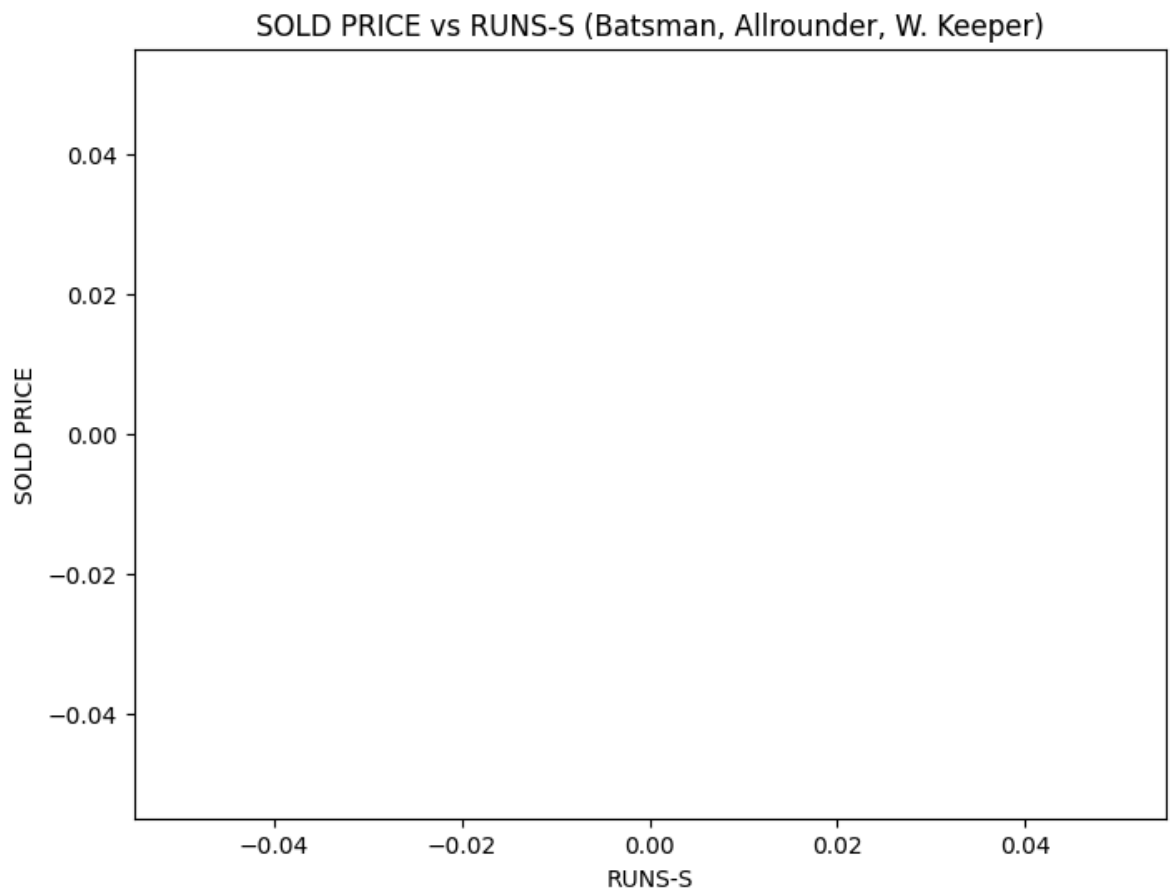
%matplotlib inline
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
file_path = "/kaggle/input/fuzzylogic/IPL IMB381IPL2013 - IPL IMB381IPL2013.csv"
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
df["SOLD PRICE"] = pd.to_numeric(df["SOLD PRICE"], errors="coerce")
df["RUNS-S"] = pd.to_numeric(df["RUNS-S"], errors="coerce")
df["WKTS"] = pd.to_numeric(df["WKTS"], errors="coerce")
df = df.dropna(subset=["SOLD PRICE", "RUNS-S", "WKTS", "PLAYING ROLE"])
bat_roles = ["Batsman", "Allrounder", "W. Keeper"]

```



```
bat_df = df[df["PLAYING ROLE"].isin(bat_roles)]
plt.figure(figsize=(8,6))
sns.regplot(data=bat_df, x="RUNS-S", y="SOLD PRICE", scatter_kws={"alpha":0.
plt.title("SOLD PRICE vs RUNS-S (Batsman, Allrounder, W. Keeper)")
plt.show()
bowl_roles = ["Bowler", "Allrounder"]
bowl_df = df[df["PLAYING ROLE"].isin(bowl_roles)]
plt.figure(figsize=(8,6))
sns.regplot(data=bowl_df, x="WKTS", y="SOLD PRICE", scatter_kws={"alpha":0.6
plt.title("SOLD PRICE vs WKTS (Bowler, Allrounder)")
plt.show()
```





## q1) printing dataset

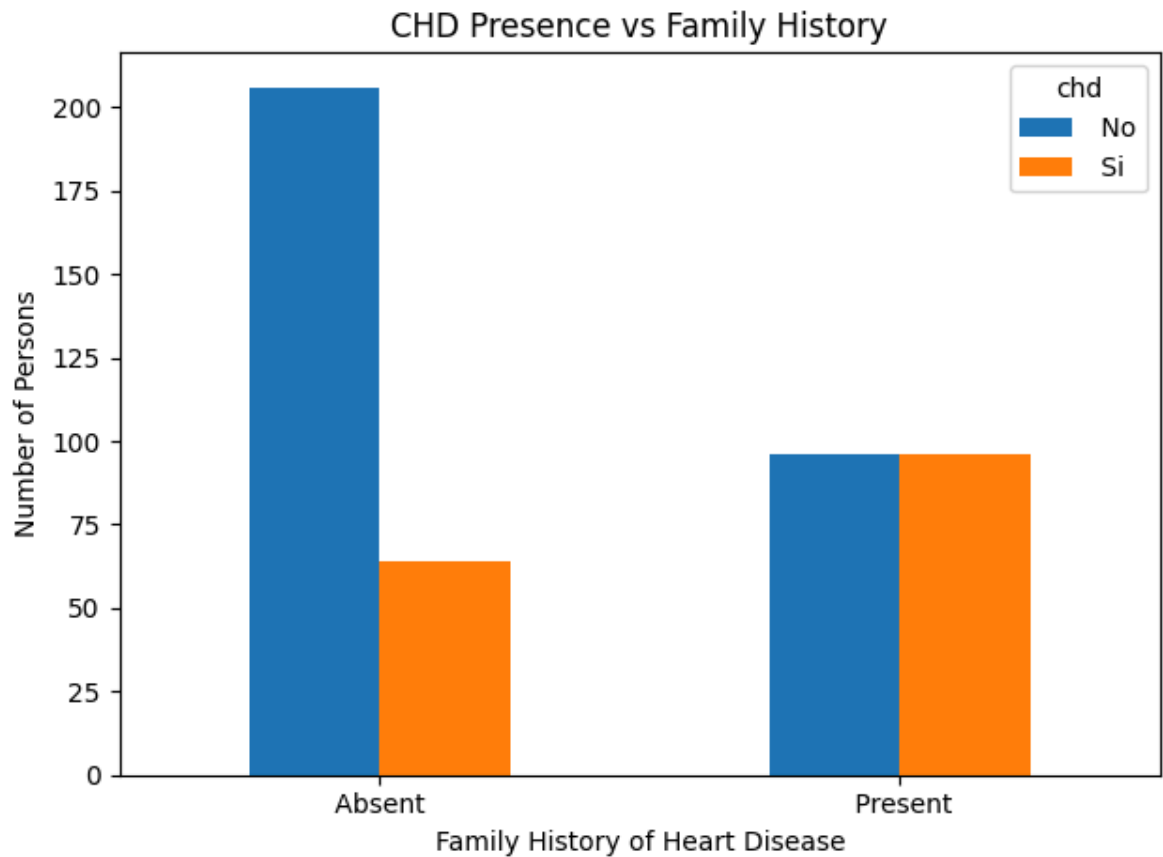
```
import pandas as pd
file_path = "/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv"
df = pd.read_csv(file_path)
print("Total number of records in dataset:",df.shape[0])
df.info()
```

```
Total number of records in dataset: 462
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 462 entries, 0 to 461
Data columns (total 10 columns):
#   Column      Non-Null Count  Dtype
---  -
0   sbp         462 non-null   int64
1   tobacco     462 non-null   float64
2   ldl         462 non-null   float64
3   adiposity   462 non-null   float64
4   famhist     462 non-null   object
5   typea       462 non-null   int64
6   obesity     462 non-null   float64
7   alcohol     462 non-null   float64
8   age         462 non-null   int64
9   chd         462 non-null   object
dtypes: float64(5), int64(3), object(2)
memory usage: 36.2+ KB
```

## q2)bar plot

```
import pandas as pd
import matplotlib.pyplot as plt
file_path = "/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv"
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
ct = pd.crosstab(df["famhist"],df["chd"])
ct.plot(kind = "bar")
plt.title("CHD Presence vs Family History")
plt.xlabel("Family History of Heart Disease")
plt.ylabel("Number of Persons")
plt.xticks(rotation = 0)
plt.tight_layout()
plt.show()
```





```
import pandas as pd
file_path = "/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv"
df = pd.read_csv(file_path)
print(df.head())
```

|   | sbp | tobacco | ldl  | adiposity | famhist | typea | obesity | alcohol |
|---|-----|---------|------|-----------|---------|-------|---------|---------|
| 0 | 160 | 12.00   | 5.73 | 23.11     | Present | 49    | 25.30   | 97.20   |
| 1 | 144 | 0.01    | 4.41 | 28.61     | Absent  | 55    | 28.87   | 2.06    |
| 2 | 118 | 0.08    | 3.48 | 32.28     | Present | 52    | 29.14   | 3.81    |
| 3 | 170 | 7.50    | 6.41 | 38.03     | Present | 51    | 31.99   | 24.26   |
| 4 | 134 | 13.60   | 3.50 | 27.78     | Present | 60    | 25.99   | 57.34   |

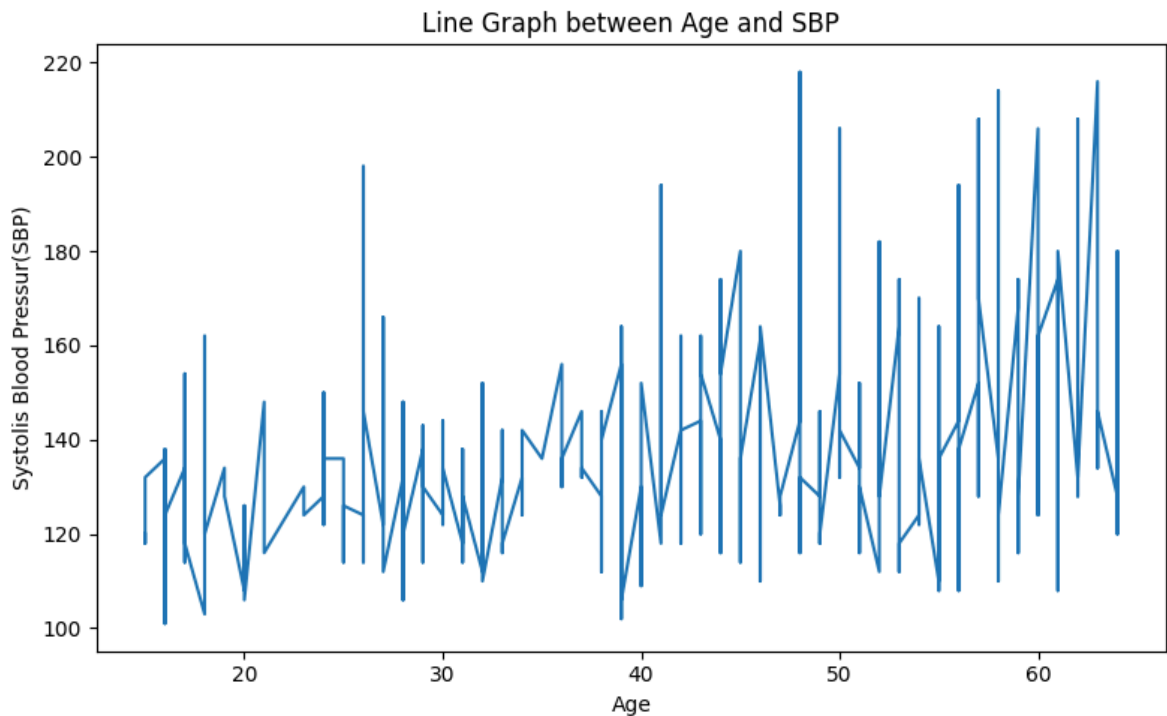
|   | age | chd |
|---|-----|-----|
| 0 | 52  | Si  |
| 1 | 63  | Si  |
| 2 | 46  | No  |
| 3 | 58  | Si  |
| 4 | 49  | Si  |

```
import pandas as pd
import matplotlib.pyplot as plt
file_path = "/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv"
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
df_sorted = df.sort_values(by = "age")
plt.figure(figsize = (8,5))
plt.plot(df_sorted["age"],df_sorted["sbp"])
plt.title("Line Graph between Age and SBP")
```





```
plt.xlabel("Age")
plt.ylabel("Systolis Blood Pressur(SBP)")
plt.tight_layout()
plt.show()
```



```
%matplotlib inline
import pandas as pd
import matplotlib.pyplot as plt

plt.close('all')

file_path = "/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv"
df = pd.read_csv(file_path)

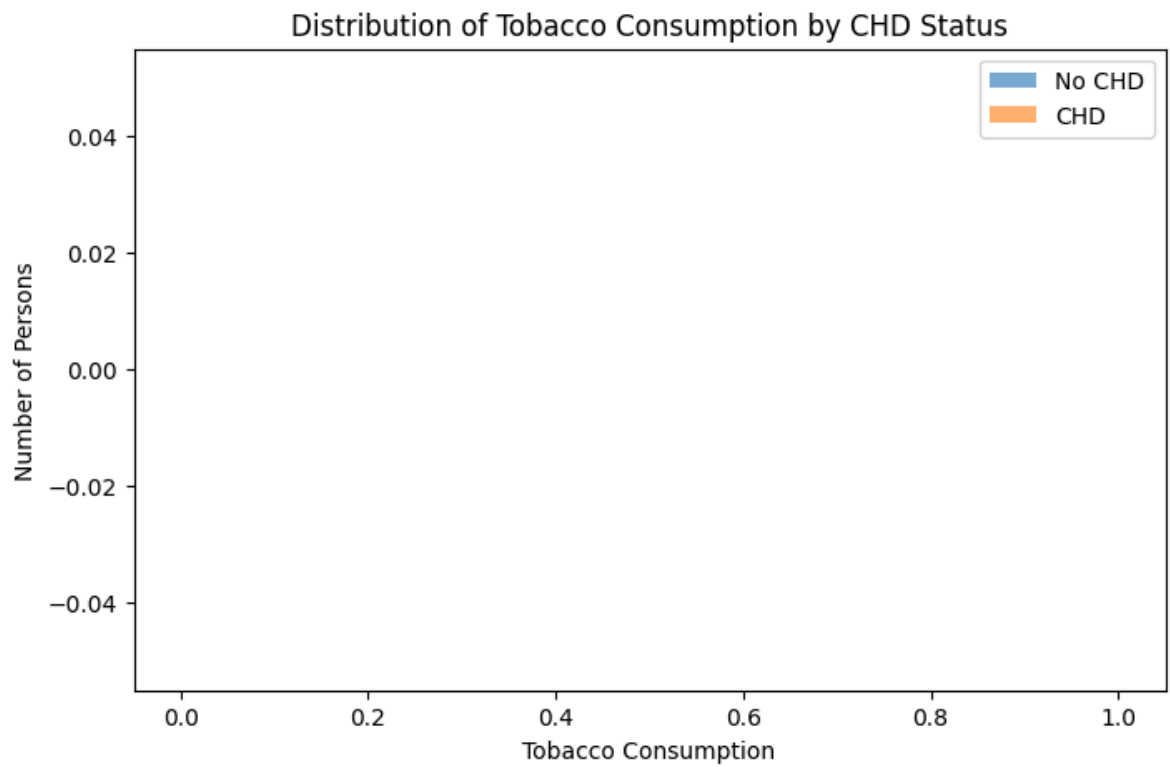
df.columns = df.columns.str.strip()

# Separate groups
chd_yes = df[df["chd"] == 1]["tobacco"]
chd_no = df[df["chd"] == 0]["tobacco"]

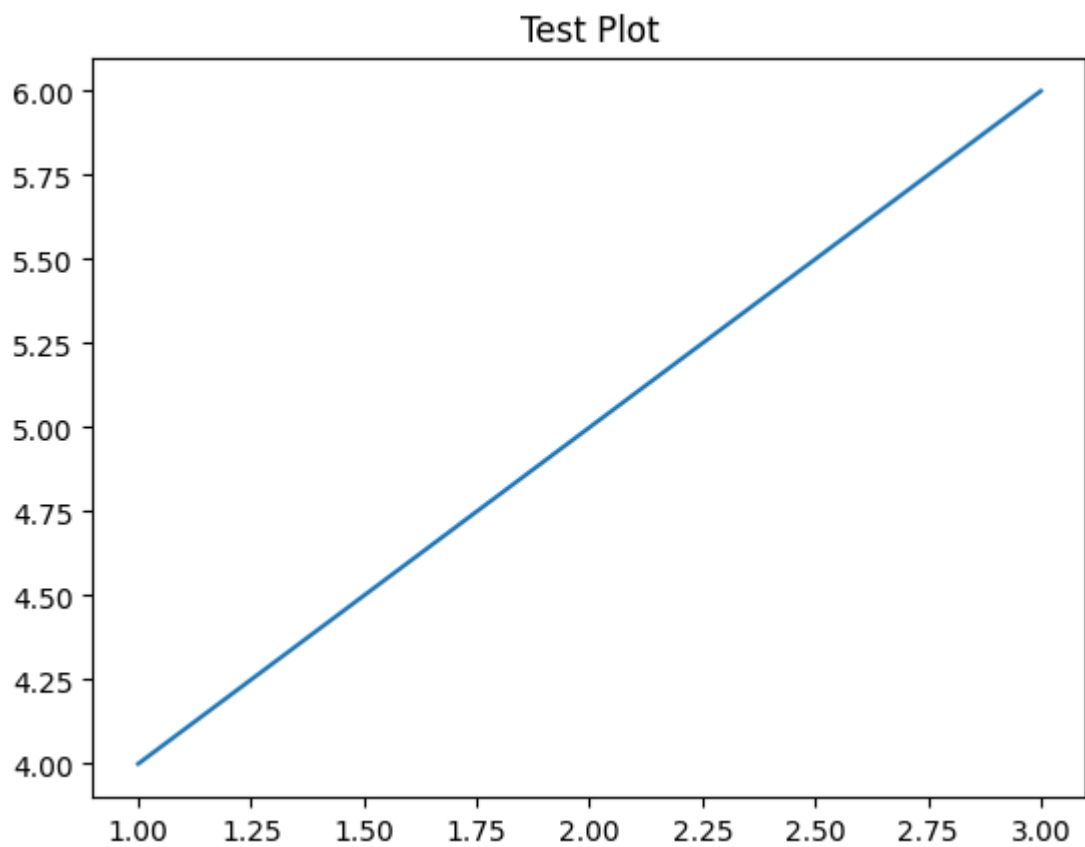
# Plot histograms
plt.figure(figsize=(8,5))
plt.hist(chd_no, bins=15, alpha=0.6, label="No CHD")
plt.hist(chd_yes, bins=15, alpha=0.6, label="CHD")

plt.title("Distribution of Tobacco Consumption by CHD Status")
plt.xlabel("Tobacco Consumption")
plt.ylabel("Number of Persons")
plt.legend()
plt.show()
```





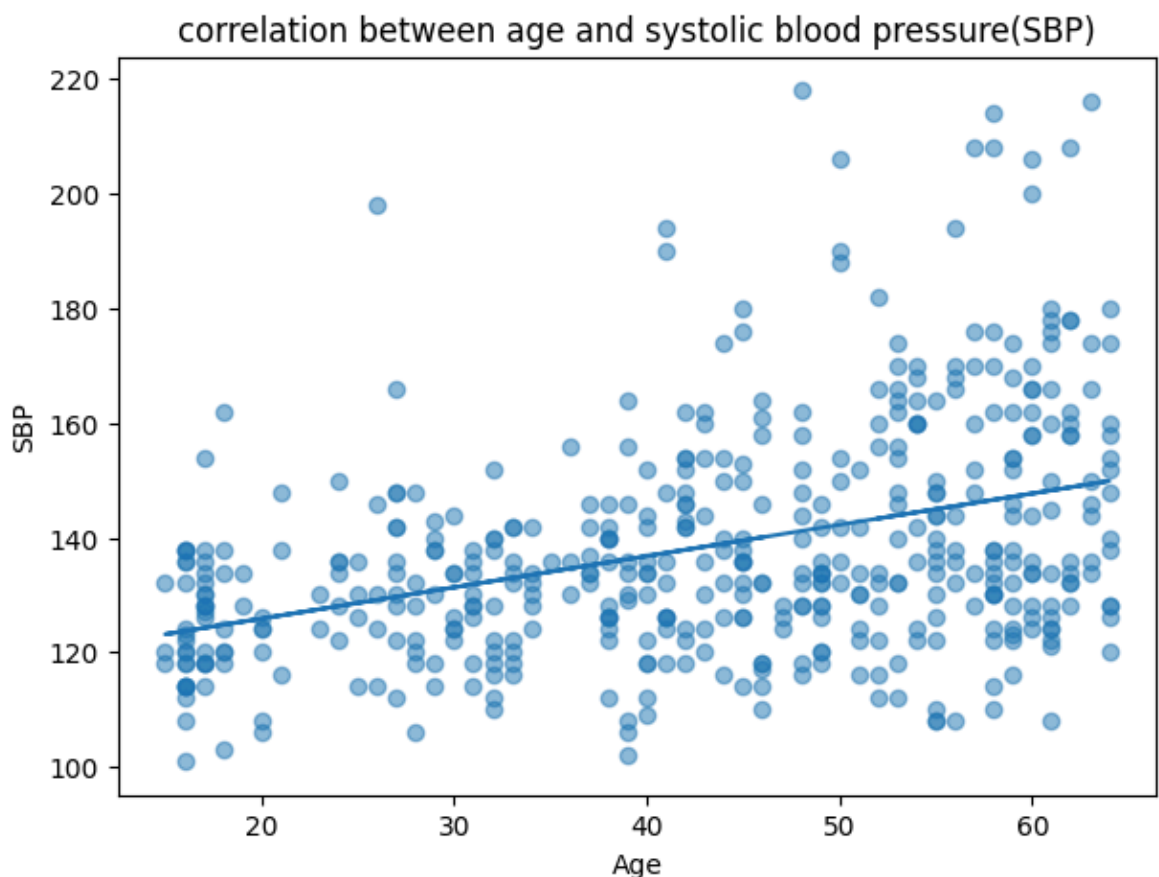
```
import matplotlib.pyplot as plt
plt.plot([1,2,3], [4,5,6])
plt.title("Test Plot")
plt.show()
```



## corelation of age with sbp

```
%matplotlib inline
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
file_path = "/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv"
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
corr = df["age"].corr(df["sbp"])
print("correlation between age and srp:",corr)
plt.figure(figsize = (7,5))
plt.scatter(df["age"],df["sbp"],alpha = 0.5)
m , b = np.polyfit(df["age"],df["sbp"],1)
plt.plot(df["age"],m*df["age"]+b)
plt.title("correlation between age and systolic blood pressure(SBP)")
plt.xlabel("Age")
plt.ylabel("SBP")
plt.show()
```

correlation between age and srp: 0.3887705972765735



```
%matplotlib inline
import pandas as pd
import matplotlib.pyplot as plt
file_path = "/kaggle/input/fuzzylogic/SAheart (1) - SAheart (1).csv"
df = pd.read_csv(file_path)
df.columns = df.columns.str.strip()
```



```
data = df[["sbp","obesity","age","ldl"]]  
pd.plotting.scatter_matrix(data,figsize = (8,8),diagonal = 'hist')  
plt.suptitle("Relationship between SBP,Obesity,Age, and LDL",y = 1.02)  
plt.show()
```

Relationship between SBP,Obesity,Age, and LDL

