```
from google.colab import files
uploaded = files.upload()
```

Choose files bank.csv

- bank.csv(text/csv) - 461474 bytes, last modified: 06/07/2025 - 100% done

Saving bank.csv to bank.csv

```
# Step 2: Import Libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Step 3: Load CSV File with Correct Separator
df = pd.read_csv('bank.csv', sep=';')

# Step 4: Display First 5 Rows
print(df.head())

# Step 5: Dataset Information
print(df.info())

# Step 6: Show Available Columns
print(df.columns)

# Step 7: Summary Statistics
print(df.describe())

# Step 8: Plot Job Distribution
plt.figure(figsize=(8,5))
sns.countplot(data=df, x='job')
plt.show()

# Step 9: Plot Marital Status Distribution
plt.figure(figsize=(8,5))
sns.countplot(data=df, x='marital')
plt.show()

# Step 10: Grouped Summary by Education
print(df.groupby('education').size())
```

```
   age          job  marital  education default  balance housing loan  \
0   30   unemployed  married    primary      no     1787      no   no
1   33     services  married  secondary      no     4789     yes  yes
2   35   management   single   tertiary      no     1350     yes   no
3   30   management  married   tertiary      no     1476     yes  yes
4   59  blue-collar  married  secondary      no        0     yes   no

    contact  day month  duration  campaign  pdays  previous poutcome   y
0  cellular   19   oct        79         1     -1         0  unknown  no
1  cellular   11   may       220         1    339         4  failure  no
2  cellular   16   apr       185         1    330         1  failure  no
3   unknown    3   jun       199         4     -1         0  unknown  no
4   unknown    5   may       226         1     -1         0  unknown  no
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4521 entries, 0 to 4520
Data columns (total 17 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   age        4521 non-null   int64
 1   job        4521 non-null   object
 2   marital    4521 non-null   object
 3   education  4521 non-null   object
 4   default    4521 non-null   object
 5   balance    4521 non-null   int64
 6   housing    4521 non-null   object
 7   loan       4521 non-null   object
 8   contact    4521 non-null   object
 9   day        4521 non-null   int64
 10  month      4521 non-null   object
 11  duration   4521 non-null   int64
 12  campaign   4521 non-null   int64
 13  pdays      4521 non-null   int64
 14  previous   4521 non-null   int64
 15  poutcome   4521 non-null   object
 16  y          4521 non-null   object
dtypes: int64(7), object(10)
memory usage: 600.6+ KB
None
Index(['age', 'job', 'marital', 'education', 'default', 'balance', 'housing',
       'loan', 'contact', 'day', 'month', 'duration', 'campaign', 'pdays',
       'previous', 'poutcome', 'y'],
      dtype='object')
                age       balance          day     duration     campaign  \
count  4521.000000   4521.000000  4521.000000  4521.000000  4521.000000
mean     41.170095   1422.657819    15.915284   263.961292     2.793630
std      10.576211   3009.638142     8.247667   259.856633     3.109807
min      19.000000  -3313.000000     1.000000     4.000000     1.000000
25%      33.000000     69.000000     9.000000   104.000000     1.000000
50%      39.000000    444.000000    16.000000   185.000000     2.000000
75%      49.000000   1480.000000    21.000000   329.000000     3.000000
max      87.000000  71188.000000    31.000000  3025.000000    50.000000

             pdays     previous
count  4521.000000  4521.000000
mean     39.766645     0.542579
std     100.121124     1.693562
```

min	-1.000000	0.000000
25%	-1.000000	0.000000
50%	-1.000000	0.000000
75%	-1.000000	0.000000
max	871.000000	25.000000