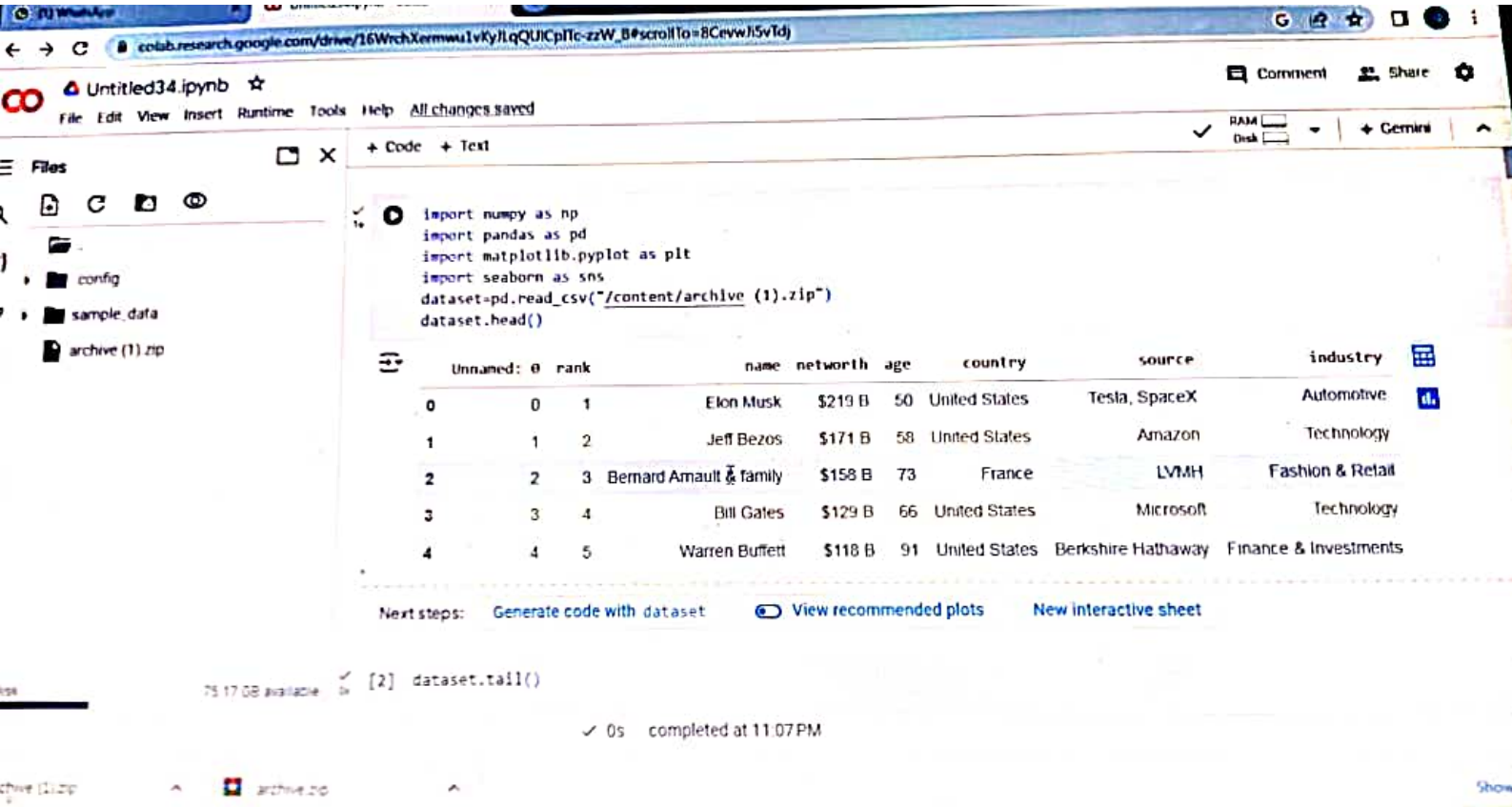
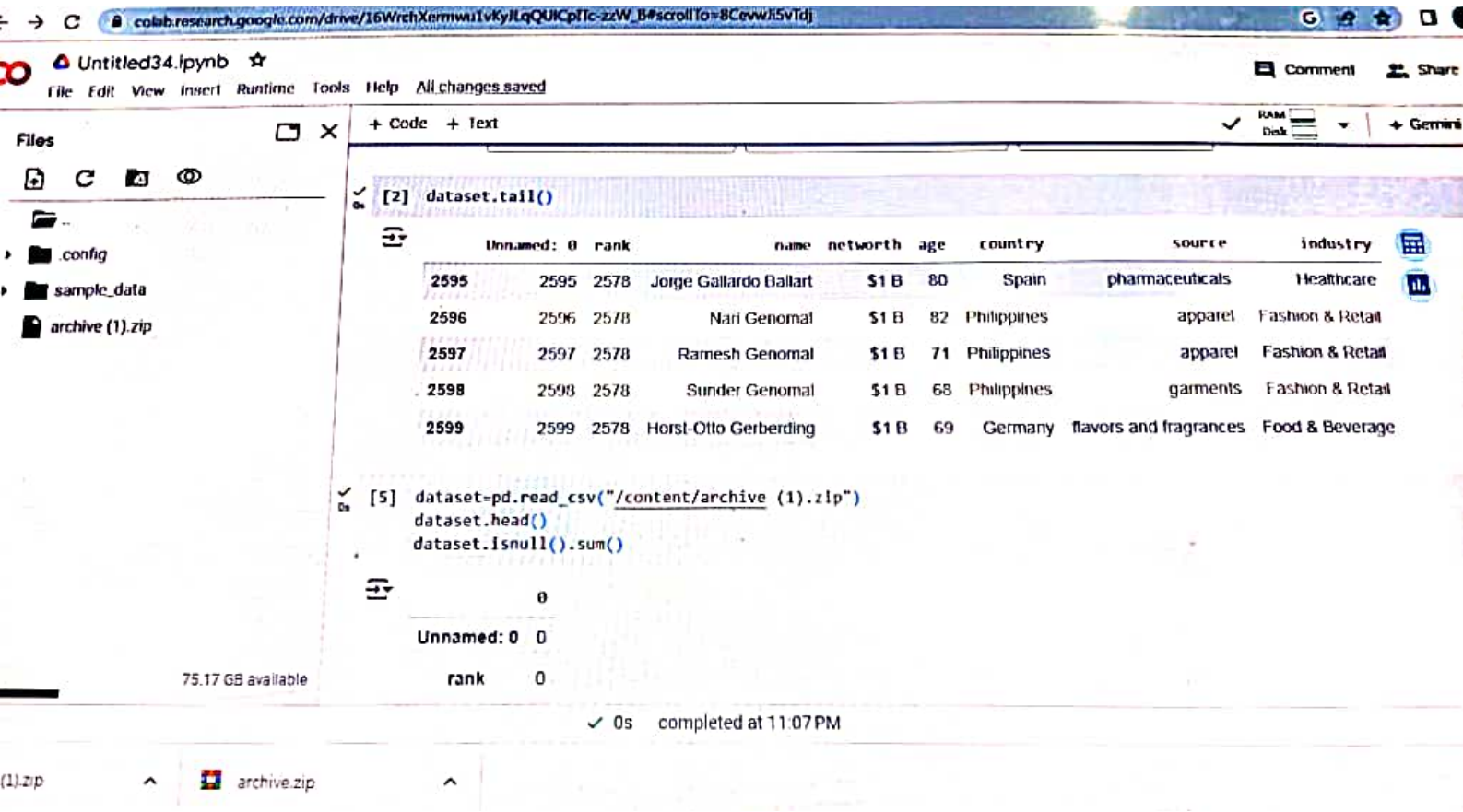
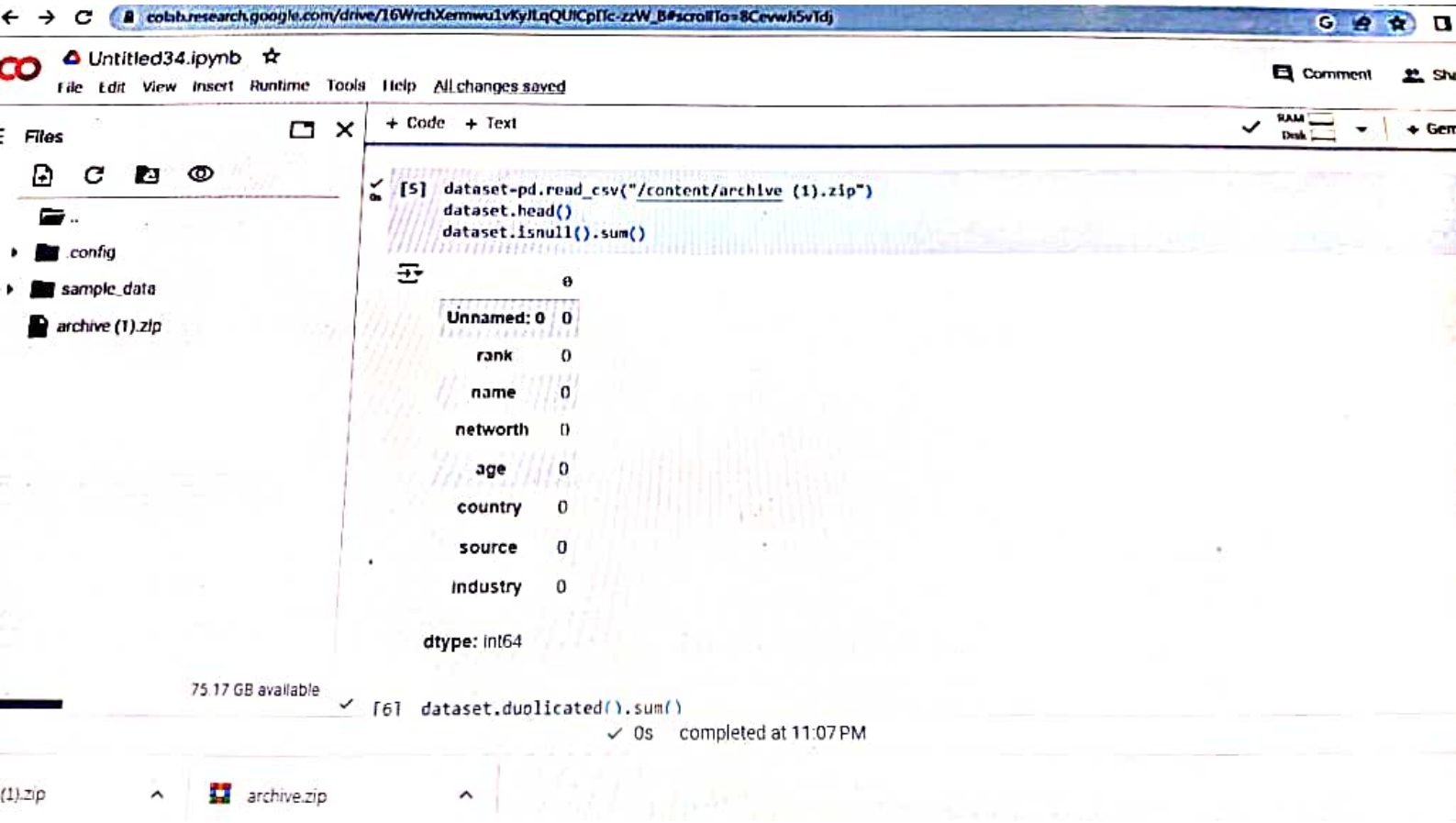
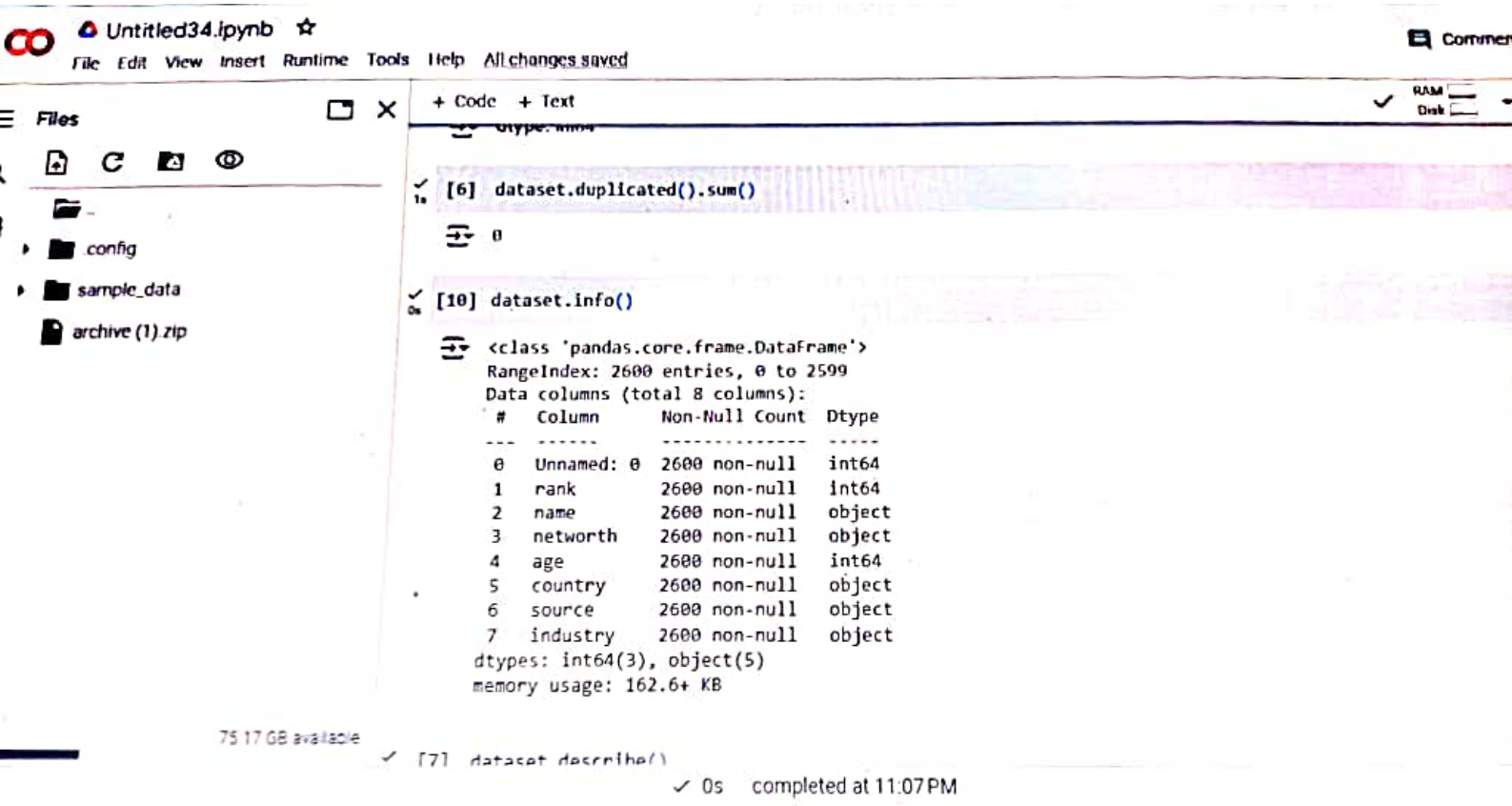
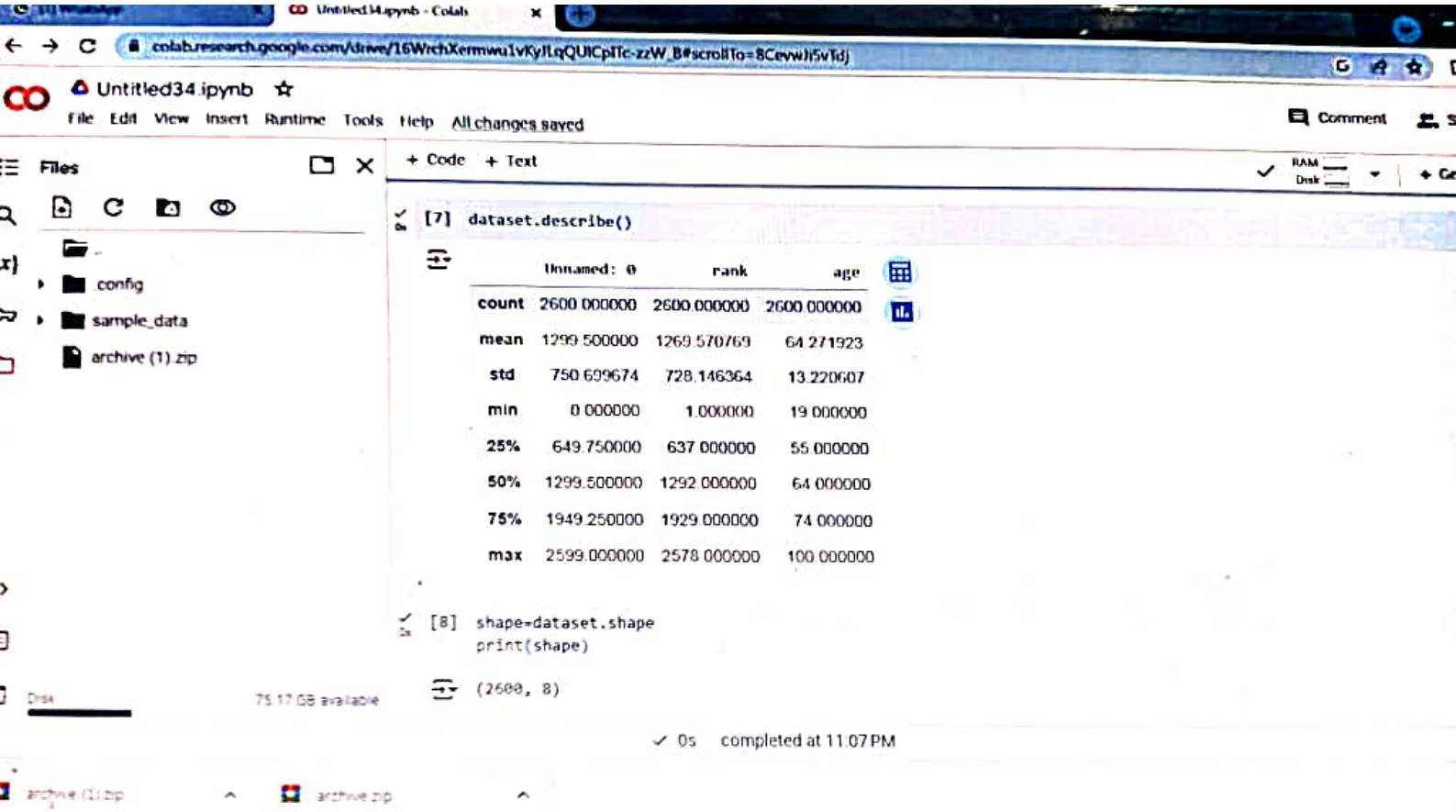

colab.research.google.com/drive/16WrchXermwu1vKyJLqQUICpITc-zzW_8#scrollTo=8CevwJi5vTdj
Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Gemini
Files
+ Code + Text
config
sample_data
archive (1).zip
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
dataset=pd.read_csv("/content/archive (1).zip")
dataset.head()
Unnamed: 0 rank name networth age country source industry
0 0 1 Elon Musk $219 B 50 United States Tesla, SpaceX Automotive
1 1 2 Jeff Bezos $171 B 58 United States Amazon Technology
2 2 3 Bernard Arnault & family $158 B 73 France LVMH Fashion & Retail
3 3 4 Bill Gates $129 B 66 United States Microsoft Technology
4 4 5 Warren Buffett $118 B 91 United States Berkshire Hathaway Finance & Investments
Next steps: Generate code with dataset View recommended plots New interactive sheet
[2] dataset.tail()
✓ 0s completed at 11:07 PM



	Unnamed: 0	rank	name	networth	age	country	source	industry
2595	2595	2578	Jorge Gallardo Ballart	$1 B	80	Spain	pharmaceuticals	Healthcare
2596	2596	2578	Nari Genomal	$1 B	82	Philippines	apparel	Fashion & Retail
2597	2597	2578	Ramesh Genomal	$1 B	71	Philippines	apparel	Fashion & Retail
2598	2598	2578	Sunder Genomal	$1 B	68	Philippines	garments	Fashion & Retail
2599	2599	2578	Horst-Otto Gerberding	$1 B	69	Germany	flavors and fragrances	Food & Beverage


Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Files
+ Code + Text
.config
sample_data
archive (1).zip
[5] dataset=pd.read_csv("/content/archive (1).zip")
dataset.head()
dataset.isnull().sum()
0
Unnamed: 0 0
rank 0
name 0
networth 0
age 0
country 0
source 0
industry 0
dtype: int64
75.17 GB available
[6] dataset.duplicated().sum()
0s completed at 11:07 PM
(1).zip
archive.zip


Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
config
sample_data
archive (1).zip
75.17 GB available
+ Code + Text
[6] dataset.duplicated().sum()
0
[10] dataset.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2600 entries, 0 to 2599
Data columns (total 8 columns):
Column Non-Null Count Dtype
0 Unnamed: 0 2600 non-null int64
1 rank 2600 non-null int64
2 name 2600 non-null object
3 networth 2600 non-null object
4 age 2600 non-null int64
5 country 2600 non-null object
6 source 2600 non-null object
7 industry 2600 non-null object
dtypes: int64(3), object(5)
memory usage: 162.6+ KB
[7] dataset.describe()
0s completed at 11:07 PM
archive.zip


Untitled34.ipynb - Colab
colab.research.google.com/drive/16WrchXermwu1vKyILqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj
Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Files
+ Code + Text
RAM
Disk
config
sample_data
archive (1).zip
[7] dataset.describe()
Unnamed: 0 rank age
count 2600.000000 2600.000000 2600.000000
mean 1299.500000 1269.570769 64.271923
std 750.699674 728.146364 13.220607
min 0.000000 1.000000 19.000000
25% 649.750000 637.000000 55.000000
50% 1299.500000 1292.000000 64.000000
75% 1949.250000 1929.000000 74.000000
max 2599.000000 2578.000000 100.000000
[8] shape=dataset.shape
print(shape)
(2600, 8)
Disk 75.17 GB available
0s completed at 11:07 PM
archive (1).zip
archive.zip

colab.research.google.com/drive/16WrchXermwu1vKyJLqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

.config

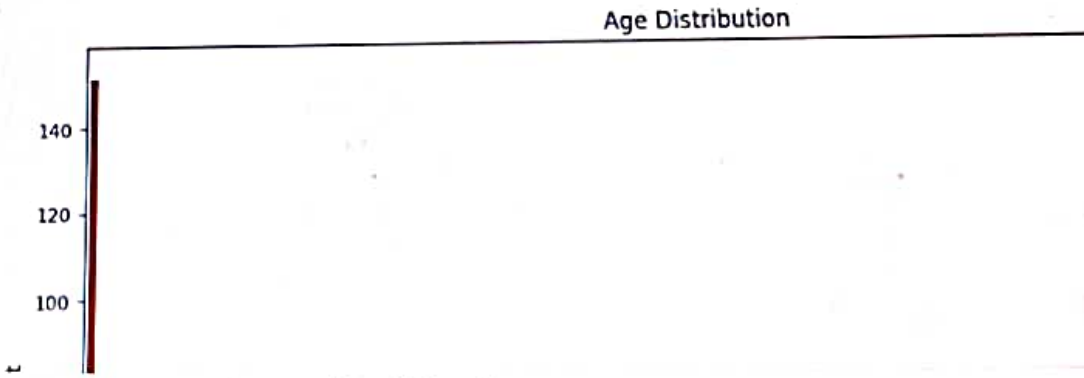
sample_data

archive (1).zip

75.17 GB available

+ Code + Text

```
[39] #Show the Age distribution among the data using bar plot
     age_distribution = dataset['age'].value_counts()
     plt.figure(figsize=(10,6))
     age_distribution.plot(kind='bar', color='maroon')
     plt.title('Age Distribution')
     plt.xlabel('Age')
     plt.ylabel('Count')
     plt.xticks(rotation=45)
     plt.tight_layout()
     plt.show()
```



✓ 0s completed at 11:07 PM

archive (1).zip

archive.zip

colab.research.google.com/drive/16WrchXermwu1vKyJLqQUlCpITc-zzW_B#scrollTo=8CevwJi5vTdj

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

+ Gemini

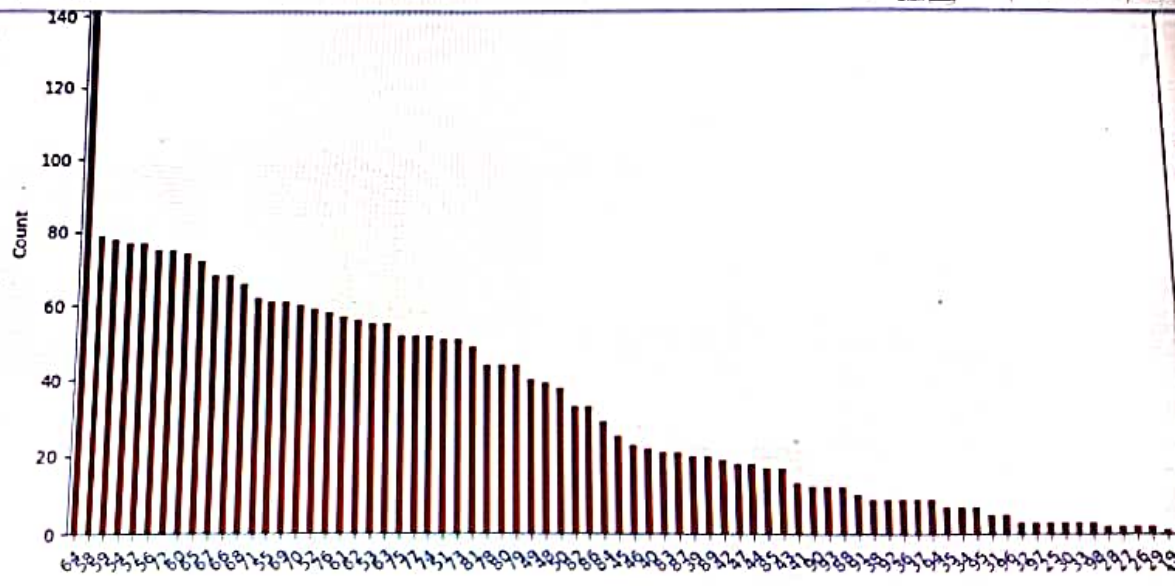
config

sample_data

archive (1).zip

75.17 GB available

[39]



✓ 0s completed at 11:07 PM

archive.zip

Show all

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

config

sample_data

archive (1).zip

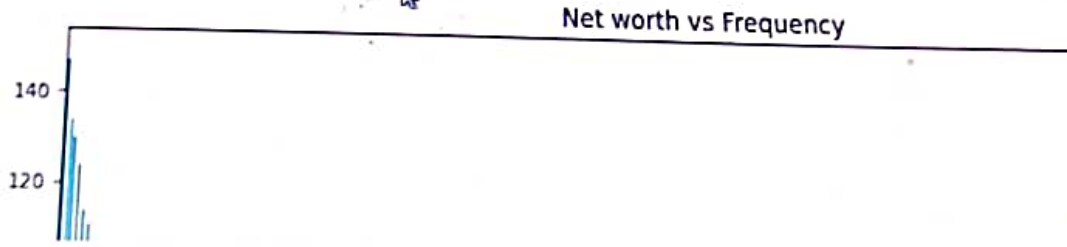
+ Code + Text

RAM Disk

[39]

Age

```
#Show the Net worth vs Frequency using bar plot
net_worth_frequency = dataset['networth'].value_counts()
plt.figure(figsize=(10,6))
net_worth_frequency.plot(kind='bar', color='teal')
plt.title("Net worth vs Frequency")
plt.xlabel('Net worth')
plt.ylabel('Frequency')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



75.17 GB available

✓ 0s completed at 11:07 PM

archive.zip

Untitled34.ipynb - Colab

colab.research.google.com/drive/16WrchXermwu1vKyJLqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment

Files

..

.config

sample_data

archive (1).zip

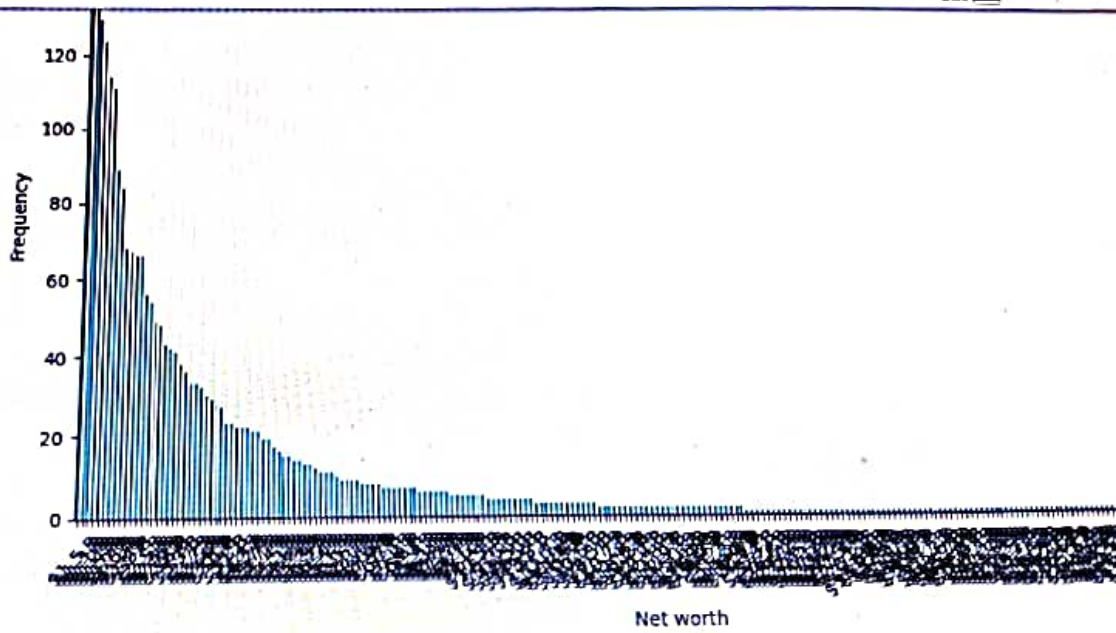
75.17 GB available

+ Code + Text

RAM

Disk

[40]



✓ 0s completed at 11:07 PM

archive.zip

lenovo

colab.research.google.com/drive/16WrchXermwu1vKyJLqQUlCpITc-zzW_B#scrollTo=8CevwJi5vTdj

Untitled34.ipynb

Edit View Insert Runtime Tools Help All changes saved

Comment

+ Code + Text

RAM Disk

config

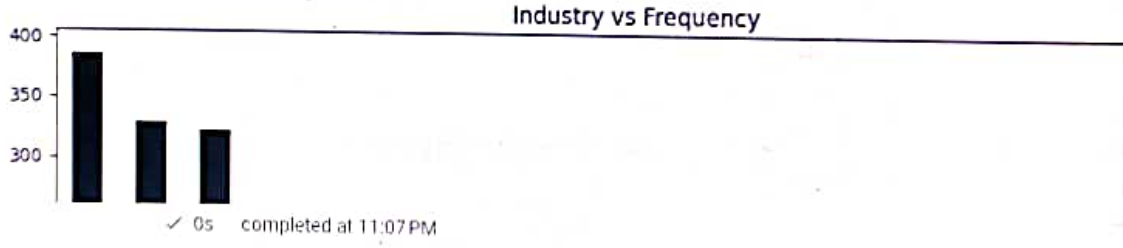
sample_data

archive (1).zip

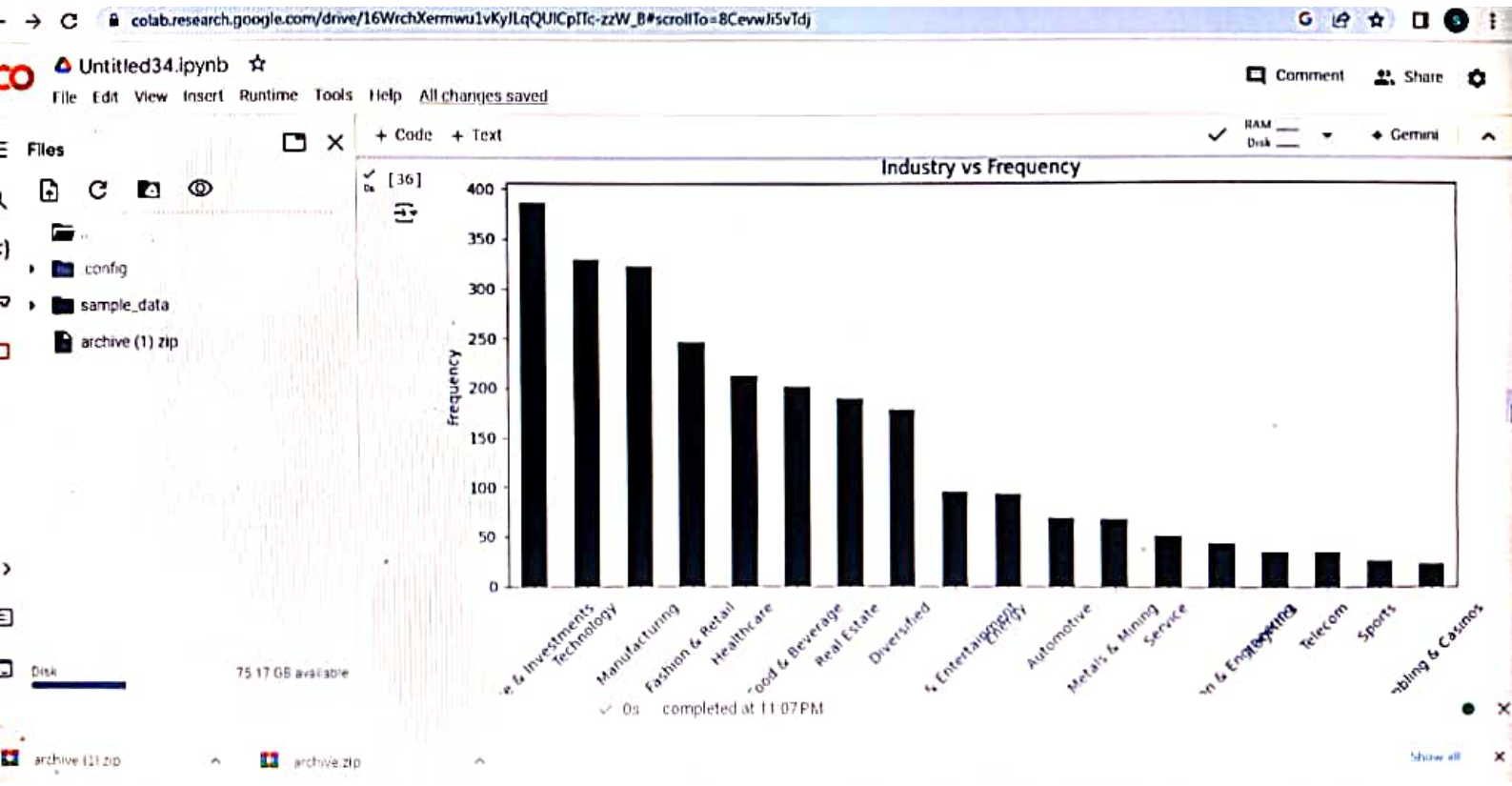
[40]

Net worth

[36]
```
#Show Industry vs Frequency using bar plot
industry_frequency = dataset['industry'].value_counts()
plt.figure(figsize=(10,6))
industry_frequency.plot(kind='bar', color='black')
plt.title("Industry vs Frequency")
plt.xlabel('Industry')
plt.ylabel('Frequency')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



✓ 0s completed at 11:07 PM


colab.research.google.com/drive/16WrchXermwu1vKyILqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj
Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code
+ Text
Gemini
config
sample_data
archive (1).zip
[36]
Industry vs Frequency
Frequency
400
350
300
250
200
150
100
50
0
Technology
Manufacturing
Fashion & Retail
Healthcare
Real Estate
Diversified
Automotive
Service
Telecom
Sports
0s completed at 11:07 PM
Disk
75.17 GB available
archive (1).zip
archive.zip

titled34.ipynb ☆

lit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

✦ Gemini

nfig

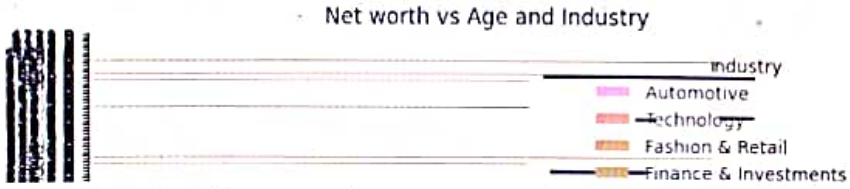
nple_data

hive (1).zip

75.17 GB available

[18] #Show how does net worth change with age and Imdustry using cat plot
plt.figure(figsize=(12,8))
sns.catplot(x='age',y='networth',hue='Industry',data=dataset,kind='bar')
plt.title("Net worth vs Age and Industry")
plt.xlabel('Age')
plt.ylabel('Net worth')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
print(dataset.columns)

<Figure size 1200x800 with 0 Axes>



✓ 0s completed at 11:07 PM

archive.zip

Show all

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

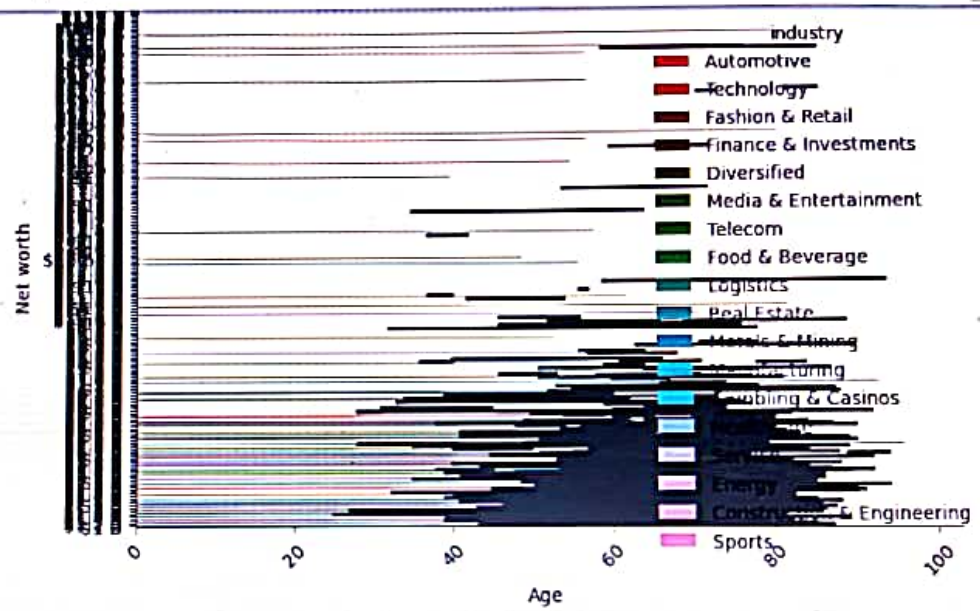
Comment Share

+ Code + Text

RAM Disk

Gemini

[18]




```
Index(['Unnamed: 0', 'rank', 'name', 'networth', 'age', 'country', 'source',
```

✓ 0s completed at 11:07PM

.config

sample_data

archive (1).zip

75.17 GB available

archive.zip

colab.research.google.com/drive/16WrchXermwu1vKyILqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj

tled34.ipynb ☆

View Insert Runtime Tools Help All changes saved

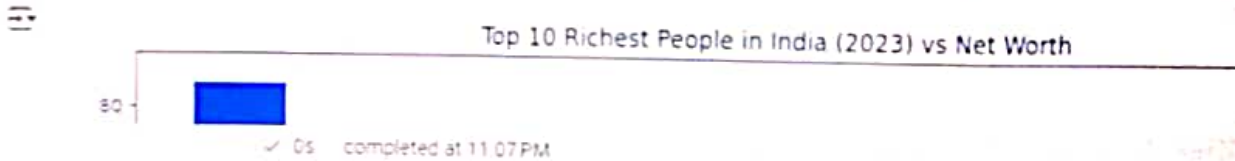
Comm

+ Code + Text

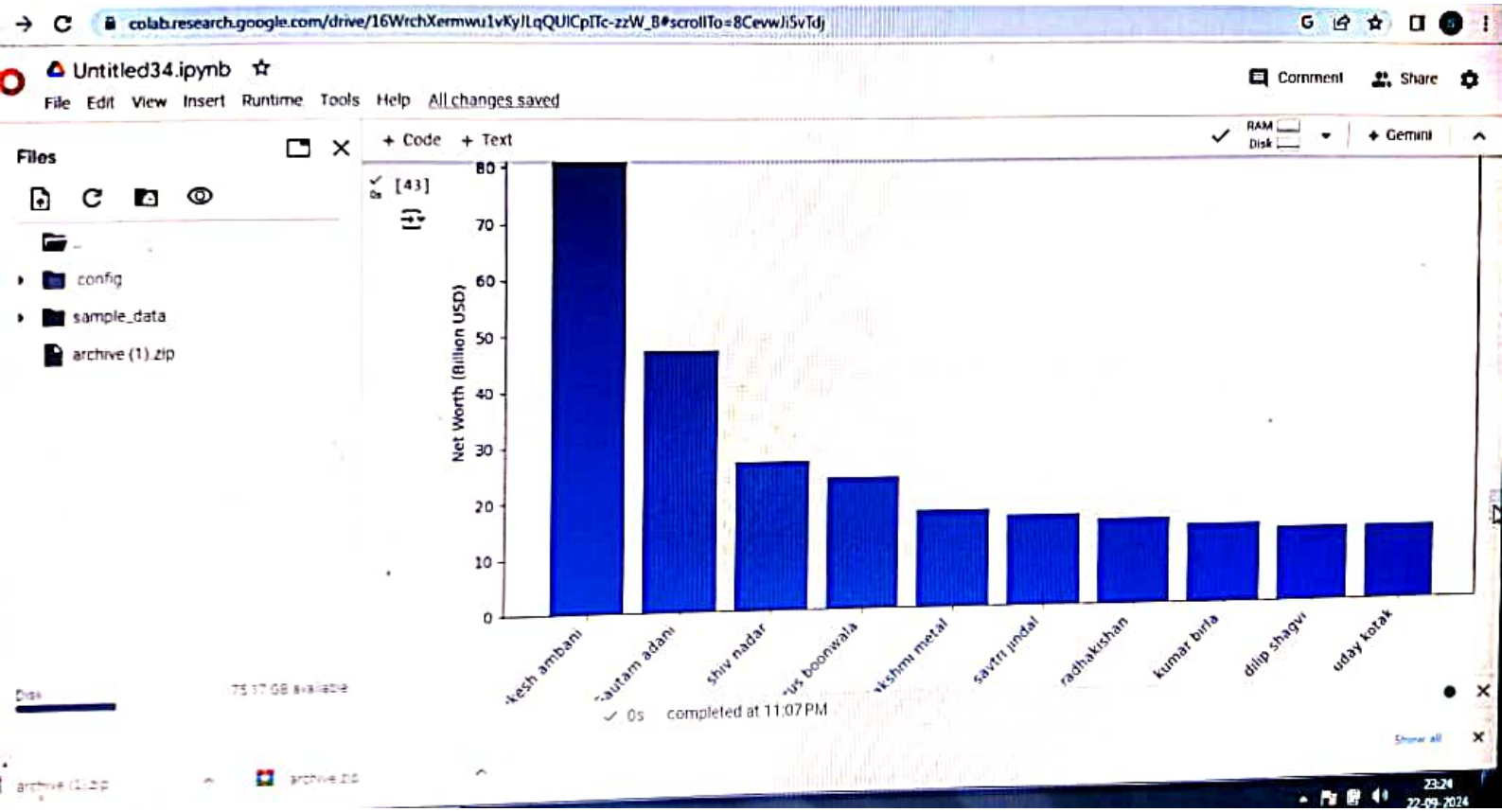
RAM
Disk

```
        dtype='object')
```

```
[43] #Show the top 10 richest people vs net worth
     import matplotlib.pyplot as plt


     wealth = [83.4, 47.2, 26.8, 23.7, 17.6, 16.4, 15.3, 14.2, 13.5, 13.4]


     plt.figure(figsize=(10,6))
     plt.bar(names, wealth, color='navy')
     plt.xticks(rotation=45, ha='right')  # Rotate x-axis labels for better readability
     plt.ylabel('Net Worth (Billion USD)')
     plt.title('Top 10 Richest People in India (2023) vs Net Worth')


     plt.tight_layout()
     plt.show()
```



✓ 0s completed at 11:07 PM


colab.research.google.com/drive/16WrchXermwu1vKyJLqQUlCpITc-zzW_B#scrollTo=8CevwJi5vTdj
Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
config
sample_data
archive (1).zip
+ Code + Text
RAM
Disk
Gemini
[43]
Net Worth (Billion USD)
0
10
20
30
40
50
60
70
80
kesh ambani
autam adani
shiv nadar
ys boonwala
kshmi metal
savitri jindal
radhakishan
kumar birla
dilip shagvi
uday korak
0s completed at 11:07 PM
23:24
22-09-2024

ive/16WrchXermwu1vKyJLqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj

ols Help All changes saved

Comment Share

+ Code + Text

RAM Disk

Gemini

[48] #Show the richest people from India with the names in any plot
```
import matplotlib.pyplot as plt
# data: names of the richest people and their net worth (in billion USD)
wealth = [83.4,47.2,26.8,23.7,17.6,16.4,15.3,14.2,13.5,13.4]
names=["mukesh ambani","Gautam adani","shiv nadar","cyrus boonwala","lakshmi metal","savtri jindal","radhakishan","kumar
plt.figure(figsize=(10,6))
plt.barh(names,wealth,color='purple')
plt.xlabel('Net Worth (Billion USD)')
plt.title('Richest People from India(2023)')
plt.tight_layout()
plt.show()
```

✓ 0s completed at 11:07 PM

zip

Show all


Untitled34.ipynb - Colab
colab.research.google.com/drive/16WrchXermwu1vKyJLqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj
Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
RAM
Disk
Gemini
+ Code
+ Text
[48]
config
sample_data
archive (1).zip
dilip shagvi
kumar birla
radhakishan
savtri jindal
lakshmi metal
cyrus boonwala
shiv nadar
Gautam adani
mukesh ambani
0
10
20
30
40
50
60
70
80
Net Worth (Billion USD)
0s completed at 11:07 PM

pynb ☆

Insert Runtime Tools Help All changes saved

Comment

+ Code + Text

RAM
Disk

[26]
```
#Show the minimum age billionaire<=50 with name and industry
import pandas as pd

# Load the Forbes Billionaires CSV file
df = pd.read_csv('/content/archive (1).zip')

# Filter the data for billionaires aged 50 or less
young_billionaires=df[df['age']<=50]

# Find the minimum age
min_age= young_billionaires['age'].min()

# Get the name and industry of the minimum age billionaire
min_age_billionaire = young_billionaires.loc[young_billionaires['age'] == min_age]

# Print the result
print("name:", min_age_billionaire['name'].values[0])
print("Industry:", min_age_billionaire['industry'].values[0])
print("age:", min_age)
```

```
name: Kevin David Lehmann
Industry: Fashion & Retail
age: 19
```

75.17 GB available

0s completed at 11:07PM

archive.zip

b.research.google.com/drive/16WrchXermwu1vKyJLqQUICpITc-zzW_B#scrollTo=8CevwJi5vTdj

4.ipynb ☆

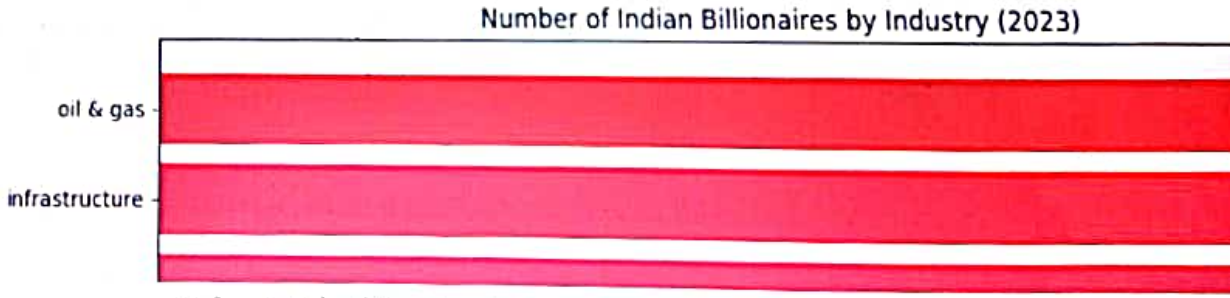
Insert Runtime Tools Help All changes saved

Comment

+ Code + Text

RAM Disk

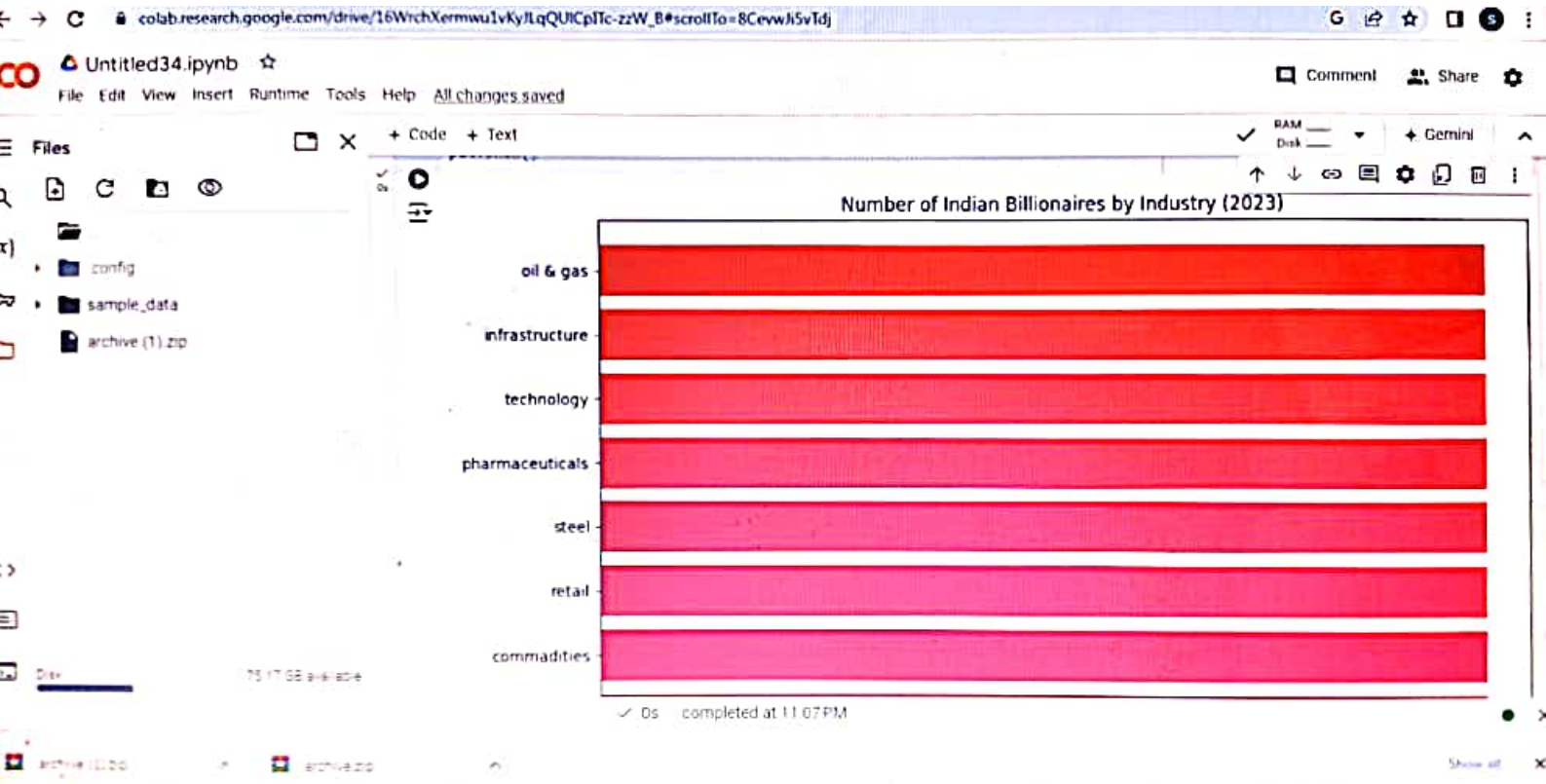
```
import matplotlib.pyplot as plt
from collections import Counter
industries=("Banking","commadities","retail","steel","pharmaceuticals","technology","infrastructure","oil &
no_of_billionaires=(83.4,47.2,26.8,23.7,17.6,16.4,15.3,14.2)
industry_count=Counter(industries)
industry_names = list(industry_count.keys())
industry_values = list(industry_count.values())
plt.figure(figsize=(10,6))
plt.barh(industry_names, industry_values, color='crimson')
plt.xlabel('Number of Billionaires')
plt.title('Number of Indian Billionaires by Industry (2023)')
plt.tight_layout()
plt.show()
```



75.17 GB available

✓ 0s completed at 11:07 PM

archive.zip


colab.research.google.com/drive/16WrchXermwu1vKyJLqQUlCpITc-zzW_B#scrollTo=8CevwJi5vTdj
Untitled34.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code
+ Text
Gemini
config
sample_data
archive (1).zip
Number of Indian Billionaires by Industry (2023)
oil & gas
infrastructure
technology
pharmaceuticals
steel
retail
commadities
0s completed at 11:07PM