

colab.research.google.com/drive/16WrchXermwu1vKyJlqQUICpTc-zzW_B#scrollTo=8CevwJi5vTdj

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

config
sample_data
archive (1).zip

+ Code + Text

```
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](#)

75.17 GB available

[2] dataset.tail()

0s completed at 11:07 PM

archive (1).zip

archive.zip

Files

- config
- sample_data
- archive (1).zip

+ Code + Text

```
[2] dataset.tail()
```

	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

```
[5] dataset=pd.read_csv("/content/archive (1).zip")
dataset.head()
dataset.isnull().sum()
```

Unnamed: 0	0
rank	0

75.17 GB available

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

config

sample_data

archive (1).zip

+ Code + Text

[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

colabresearch.google.com/drive/16WrchXermwu1vKyJlqQUICpIle-zzW_B#scrollTo=8Ceywh5vIdj

Untitled34.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

config
sample_data
archive (1).zip

+ Code + Text

[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)

RAM
Disk

75.17 GB available

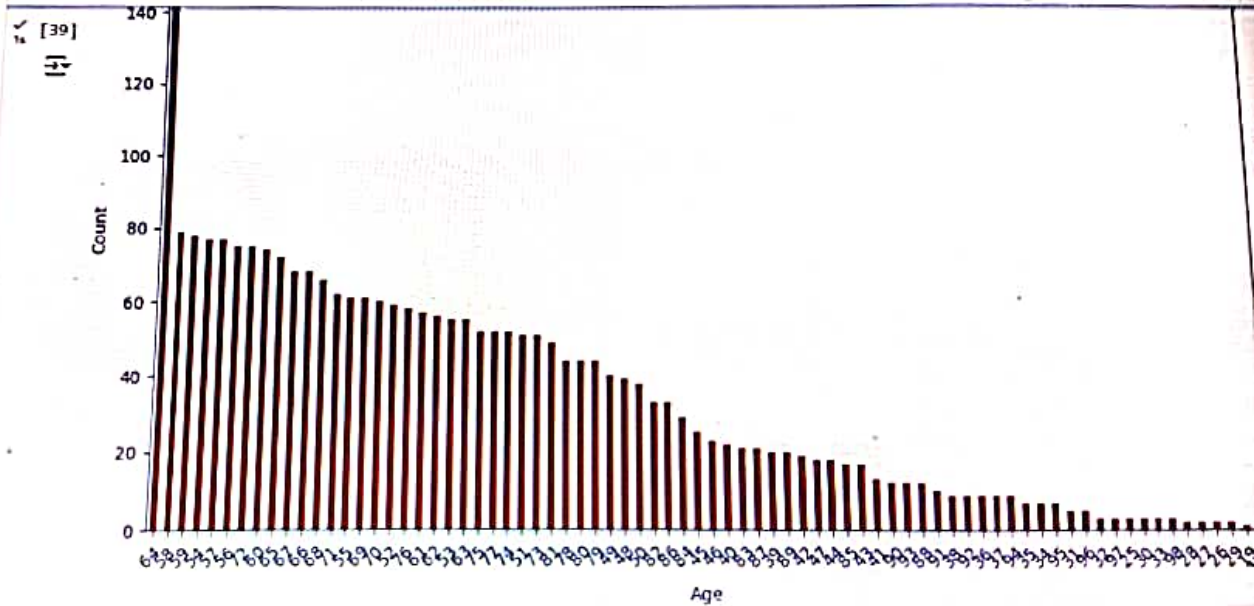
0s completed at 11:07 PM

archive (1).zip archive.zip

Files

- config
- sample_data
- archive (1).zip

75.17 GB available



✓ 0s completed at 11:07 PM

← → C colab.research.google.com/drive/16WrchXermwu1vKy/LqQUICpITc-zzW_B#scrollTo=8Cevwli5vIdj

Untitled34.ipynb ☆
File Edit View Insert Runtime Tools Help All changes saved

Files

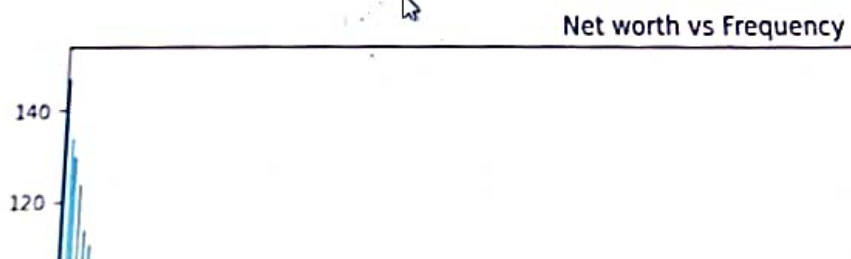
📁 -

- config
- sample_data
- archive (1).zip

+ Code + Text

```
[39]
#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()
```

🔄



✓ 0s completed at 11:07 PM

archive.zip

Untitled34.ipynb ☆

Edit View Insert Runtime Tools Help All changes saved

Comment

+ Code + Text

✓ RAM
Disk

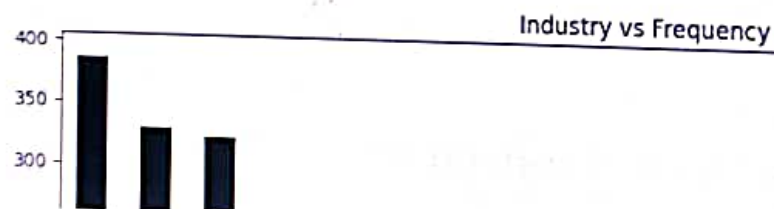
[40]



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()
```

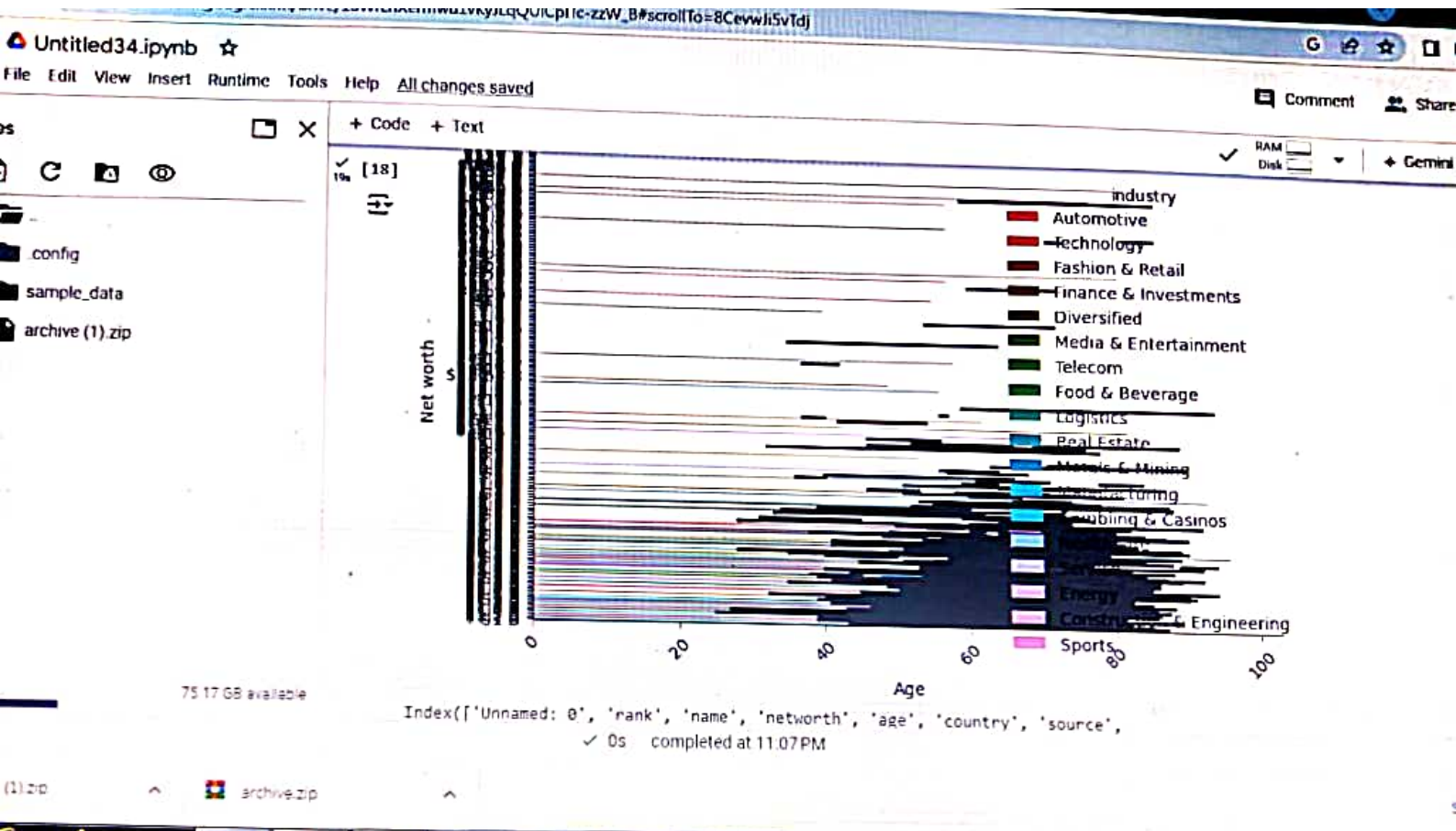


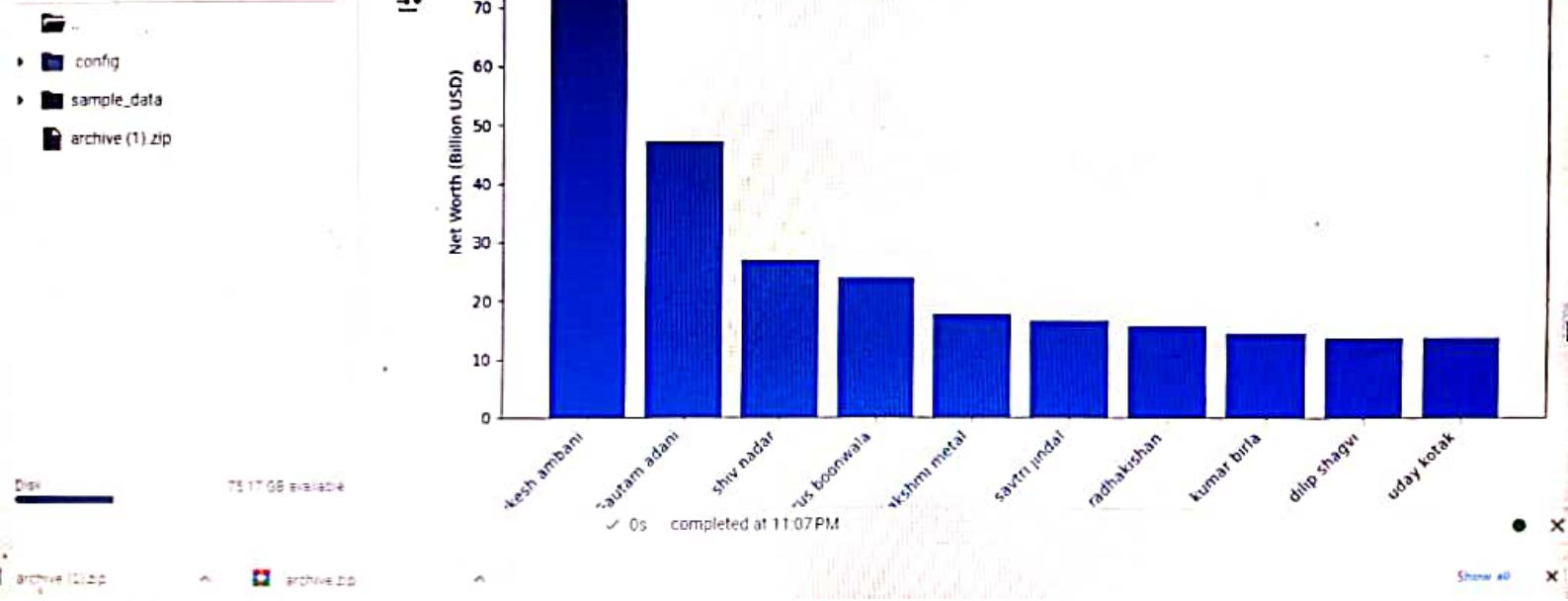
✓ 0s completed at 11:07 PM

75.17 GB available

Archive.zip







ve/16WrchXermwu1vKyJLqQUICpITc-zzW_B#scrollTo=8CewWj5vTdJ



ols Help All changes saved

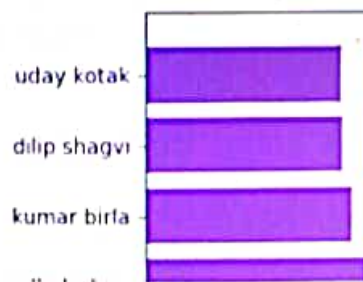
Comment Share

+ Code + Text

✓ RAM Disk Gemini

```
✓ [48] #Show the richest people from India with the names in any plot
0s 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 boomwala","lakshmi metal","savitri 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()
```

Richest People from India(2023)



✓ 0s completed at 11:07 PM

Show all

pynb ☆

Insert Runtime Tools Help [All changes saved](#)

Comment

+ Code + Text

✓ RAM
Disk

```
✓ [26] #Show the minimum age billionaire<=50 with name and industry
0s 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 11:07 PM

archive.zip

research.google.com/drive/16WrchXermwu1vKy/LqQUICpITc-zzW_B#scrollTo=8Cewwli5vTdj

.ipynb ☆

Comment

Insert Runtime Tools Help All changes saved

✓ RAM
Disk

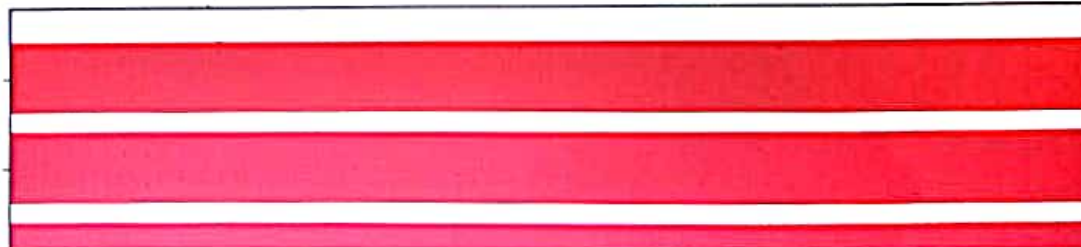
+ Code + Text

↑ ↓ ↻ ⌂ ⚙

```
import matplotlib.pyplot as plt
from collections import Counter
industries=("Banking","commodities","retail","steel","pharmaceuticals","technology","infrastructure","oil & gas")
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()
```

Number of Indian Billionaires by Industry (2023)

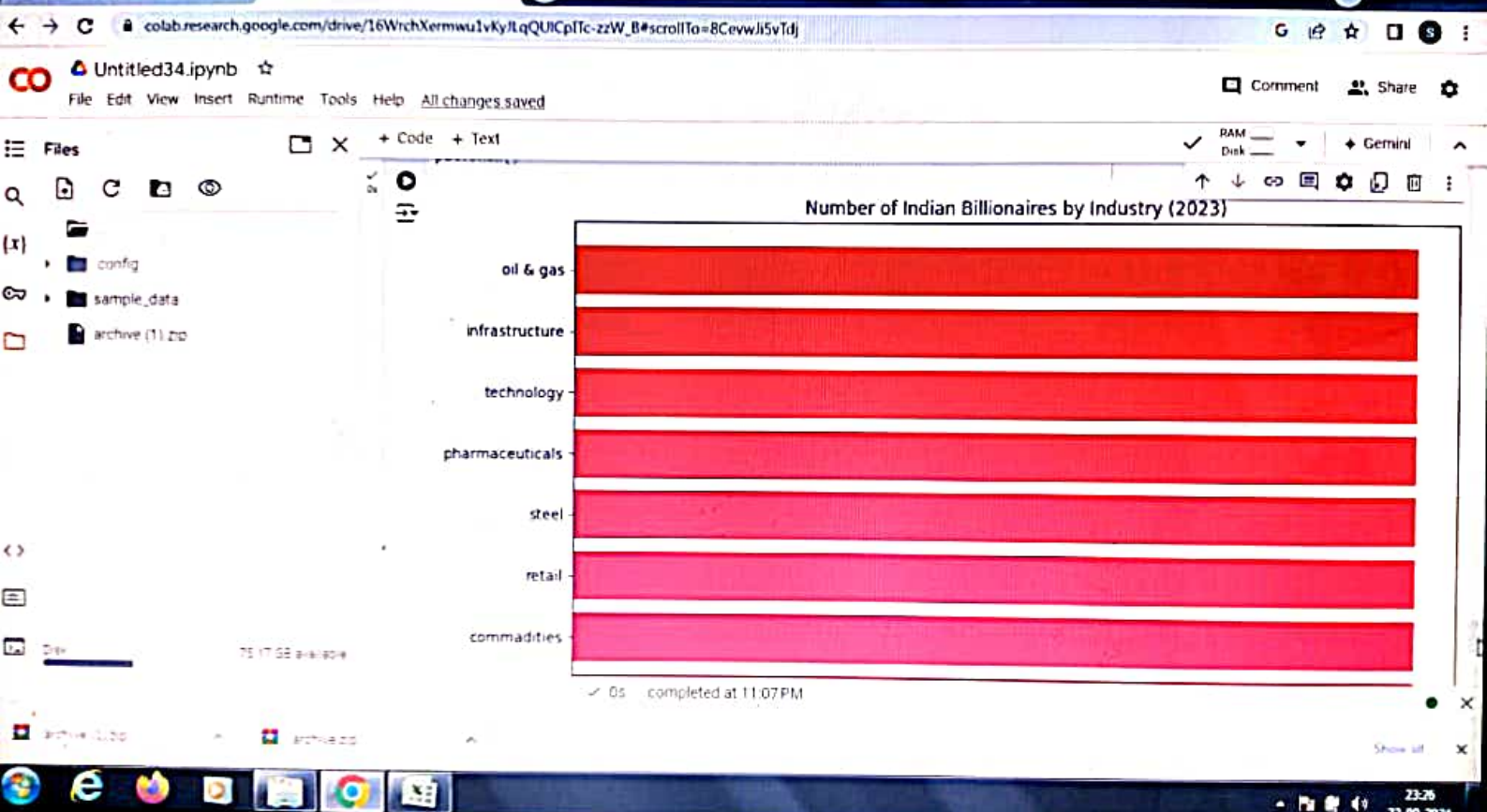
oil & gas
infrastructure



✓ 0s completed at 11:07 PM

75.17 GB available

archive.zip





titled34.ipynb ☆

File View Insert Runtime Tools Help All changes saved

Comm

Code Text
dtype='object')

RAM
Disk

```
[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()
```

Top 10 Richest People in India (2023) vs Net Worth

80

✓ Os completed at 11:07 PM



Untitled34.ipynb ☆

File View Insert Runtime Tools Help All changes saved

Comment Share



+ Code + Text



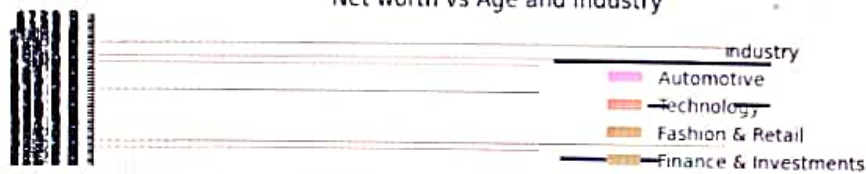
RAM
Disk

+ Gemini

```
[18]: #Show how does net worth change with age and industry 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>

Net worth vs Age and Industry



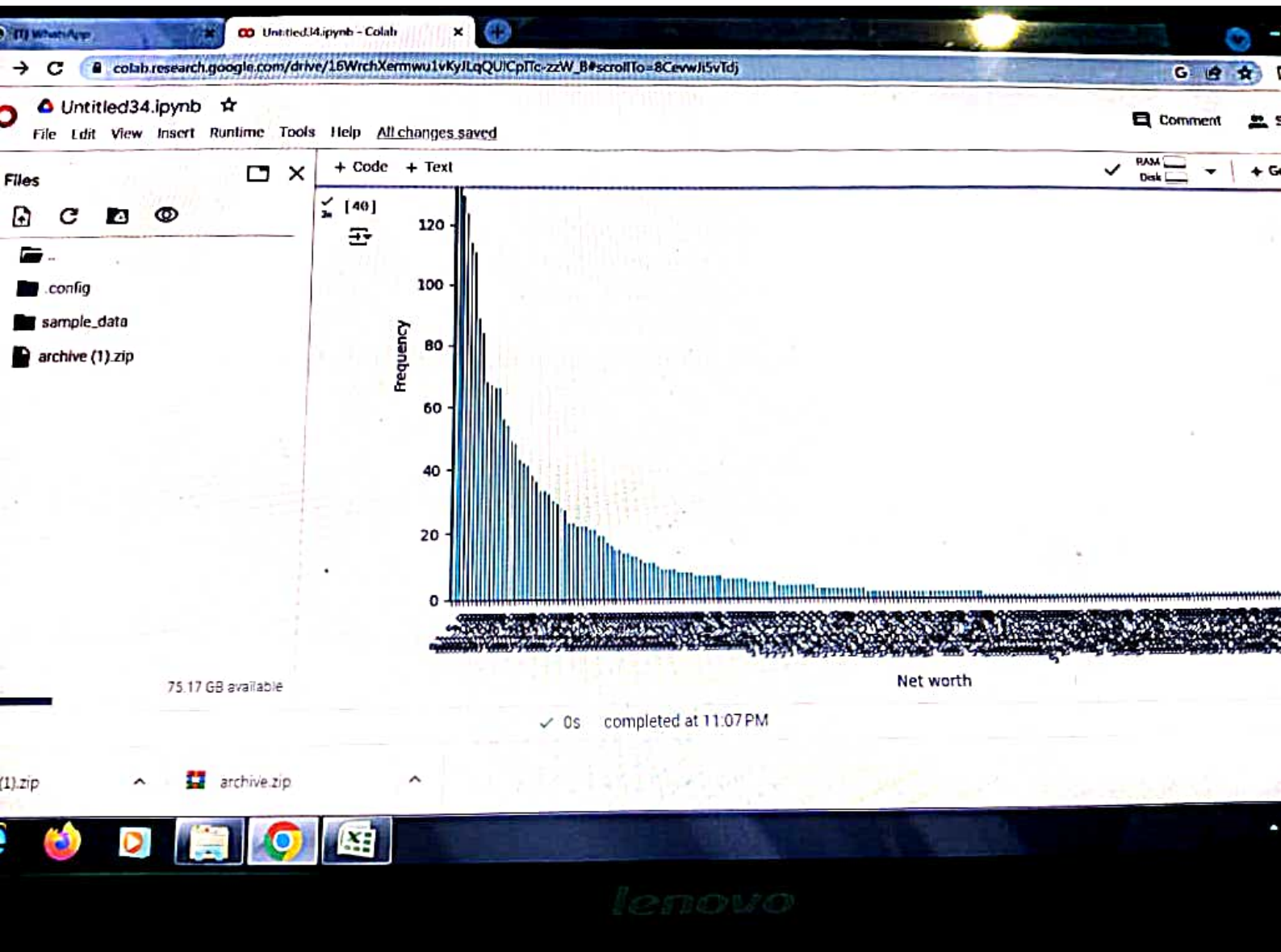
✓ 0s completed at 11:07 PM

Show all

zip

archive.zip





Files

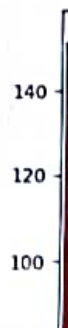
- ..
- .config
- sample_data
- archive (1).zip

+ 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()
```



Age Distribution



✓ 0s completed at 11:07 PM

archive (1).zip

archive.zip



colab.research.google.com/drive/16WrchXermwu1vkyJlqQUICpTc-zzW_B#scrollTo=8CevwJi5vIdj

Untitled34.ipynb ☆

File Edit View Insert Runtime Tools Help All changes saved

Files

- config
- sample_data
- archive (1).zip

+ 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
```

75.17 GB available

✓ [7] dataset.describe()

0s completed at 11:07 PM