Billionaires Data Analysis

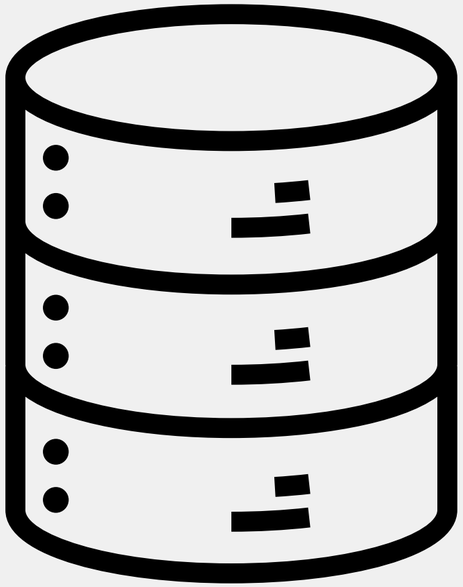
📊 Just wrapped up a deep dive into a billionaires dataset!

💰 Explored global wealth trends, top industries, and insights into billionaires.

🌍 Check out the visualizations and let's discuss the fascinating world of extreme wealth!

Loading Libraries and Dataset

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```



```
df = pd.read_csv("/content/drive/MyDrive/My projects/Billionaires Statistics Dataset.csv")
```

Dataset

```
df.columns
```

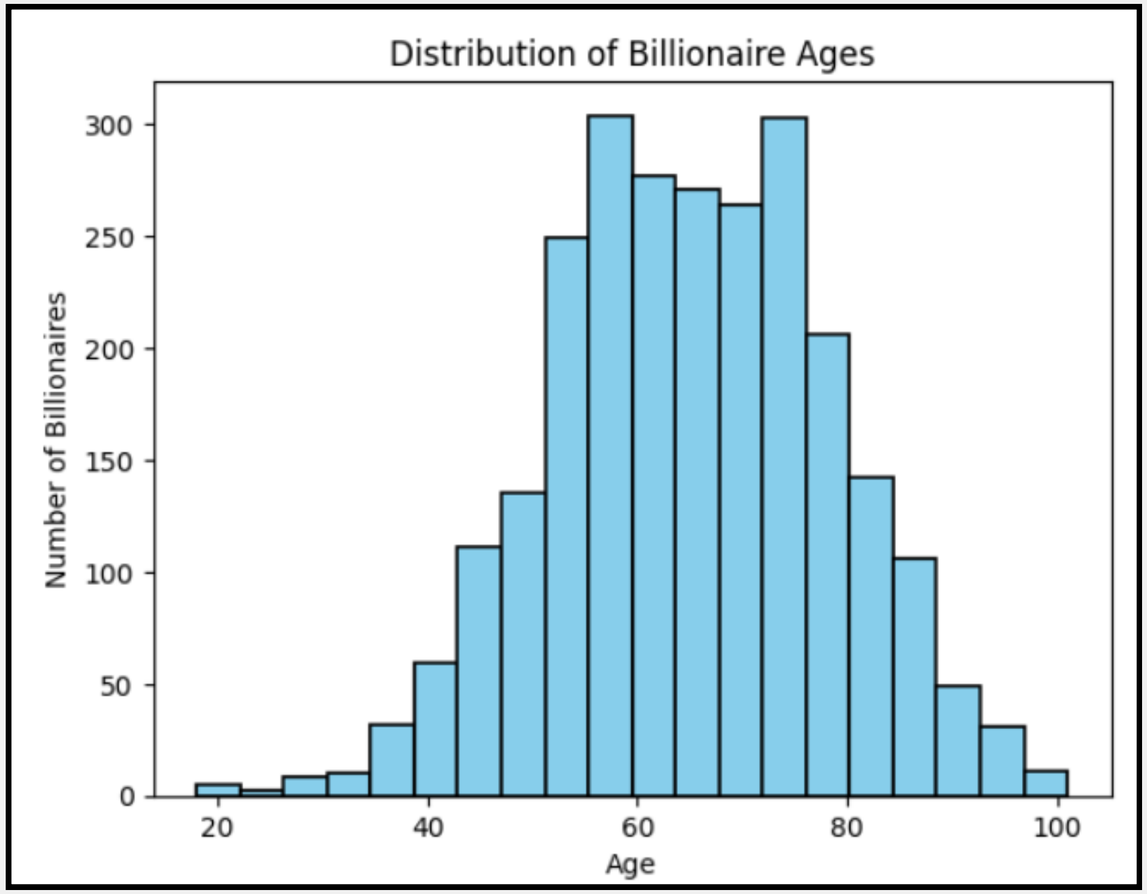
```
Index(['rank', 'finalWorth', 'category', 'personName', 'age', 'country',
       'city', 'industries', 'countryOfCitizenship', 'organization',
       'selfMade', 'status', 'gender', 'birthYear', 'birthMonth'],
      dtype='object')
```

```
df.head()
```

	rank	finalWorth	category	personName	age	country	city	industries	countryOfCitizenship	organization	selfMade	status	gender	birthYear	birthMonth
0	1	211.0	Fashion & Retail	Bernard Arnault & family	74.0	France	Paris	Fashion & Retail	France	LVMH Moët Hennessy Louis Vuitton	False	U	M	1949.0	3.0
1	2	180.0	Automotive	Elon Musk	51.0	United States	Austin	Automotive	United States	Tesla	True	D	M	1971.0	6.0
2	3	114.0	Technology	Jeff Bezos	59.0	United States	Medina	Technology	United States	Amazon	True	D	M	1964.0	1.0
3	4	107.0	Technology	Larry Ellison	78.0	United States	Lanai	Technology	United States	Oracle	True	U	M	1944.0	8.0
4	5	106.0	Finance & Investments	Warren Buffett	92.0	United States	Omaha	Finance & Investments	United States	Berkshire Hathaway Inc. (Cl A)	True	D	M	1930.0	8.0

What patterns do you observe in the age distribution of billionaires?

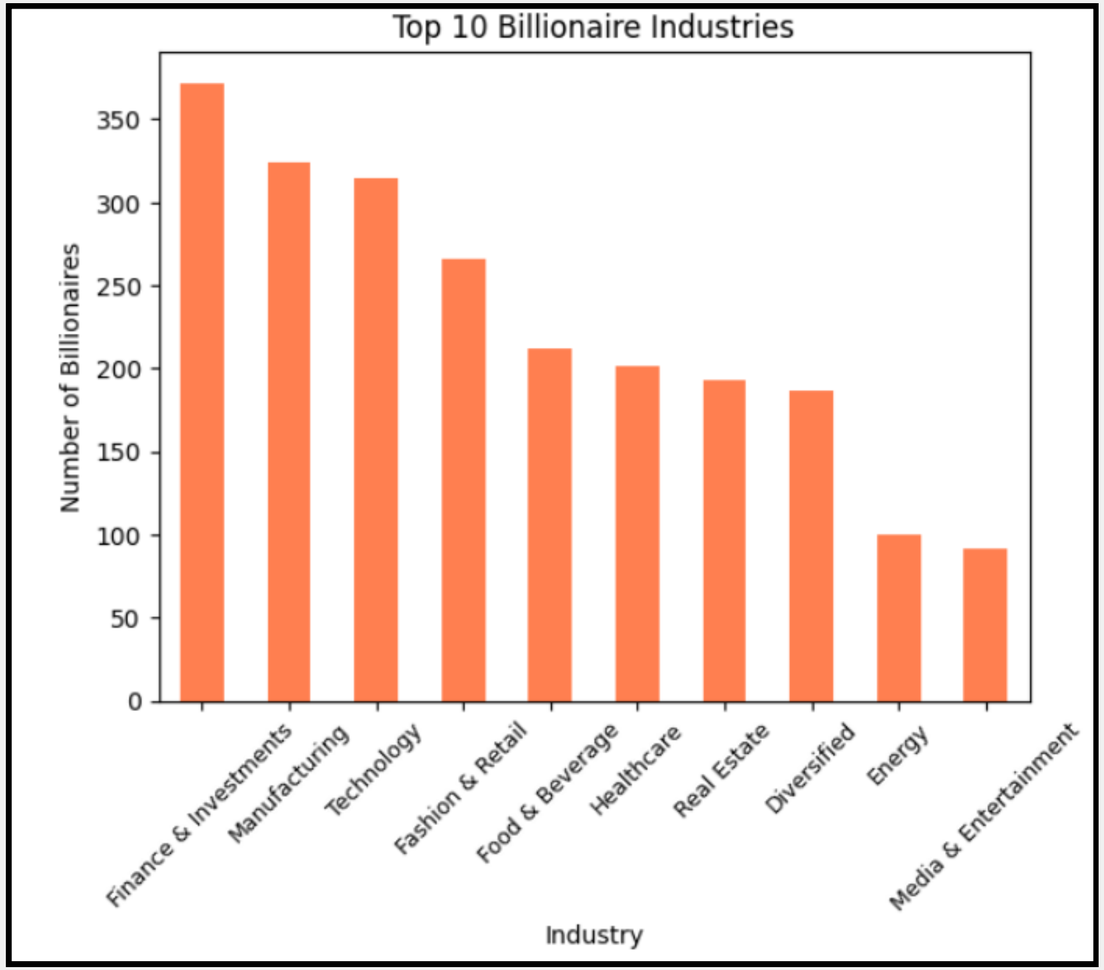
```
plt.hist(df['age'], bins=20, color='skyblue', edgecolor='black', linewidth=1.2)
plt.xlabel("Age")
plt.ylabel("Number of Billionaires")
plt.title("Distribution of Billionaire Ages")
plt.show()
```



- Most of the billionaires have age between 50Yr and 70Yr
- This suggests that success often comes with experience

Which industry dominates the billionaire landscape?

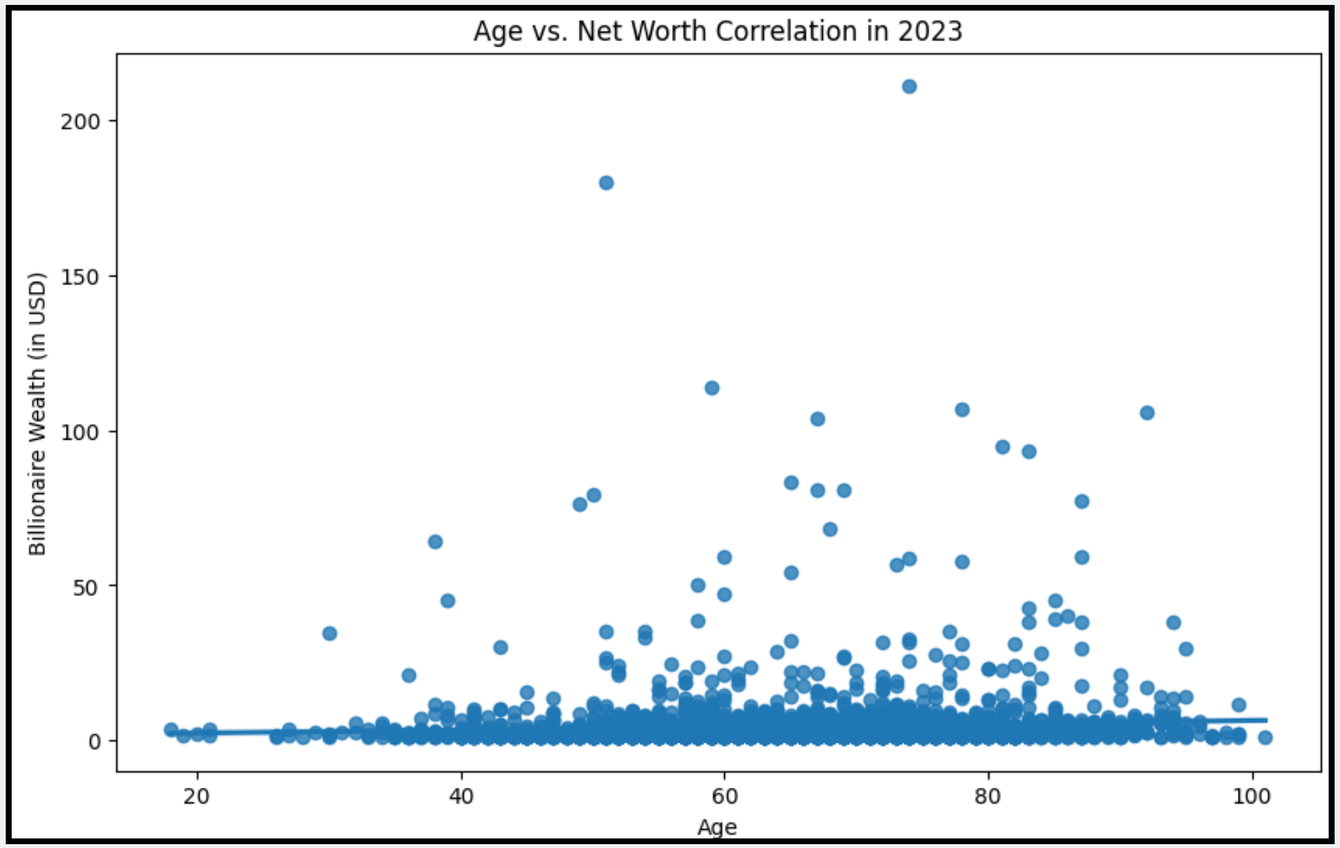
```
top_industries = df['category'].value_counts().head(10)
top_industries.plot(kind='bar', color='coral')
plt.xlabel("Industry")
plt.ylabel("Number of Billionaires")
plt.title("Top 10 Billionaire Industries")
plt.xticks(rotation=45,fontsize=9)
plt.show()
```



- Finance & Investing stands out as dominant sector among billionaires , more than 350+
- It also highlights the industry growth
- Manufacturing and Technology sector also plays significant role

What observations can you make about the relationship between age and net worth?

```
plt.figure(figsize=(10, 6))
sns.regplot(data=df, x='age', y='finalWorth')
plt.xlabel('Age')
plt.ylabel('Billionaire Wealth (in USD)')
plt.title('Age vs. Net Worth Correlation in 2023')
plt.show()
```

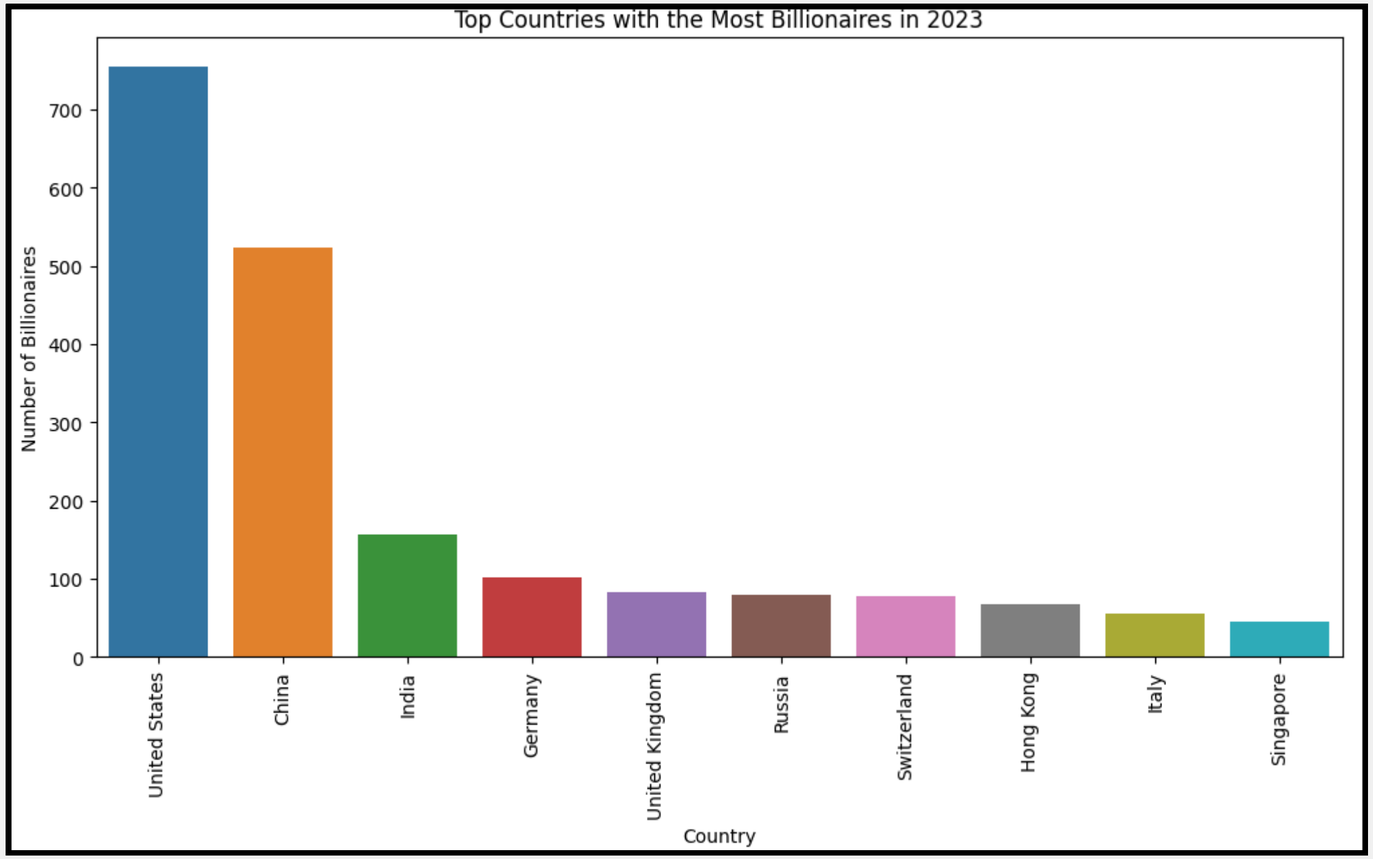


- A positive correlation between age and net worth, indicating that wealth tends to accumulate over time.
- Notable outliers suggest other factors also play a significant role in success.

Which countries have the highest concentration of billionaires?

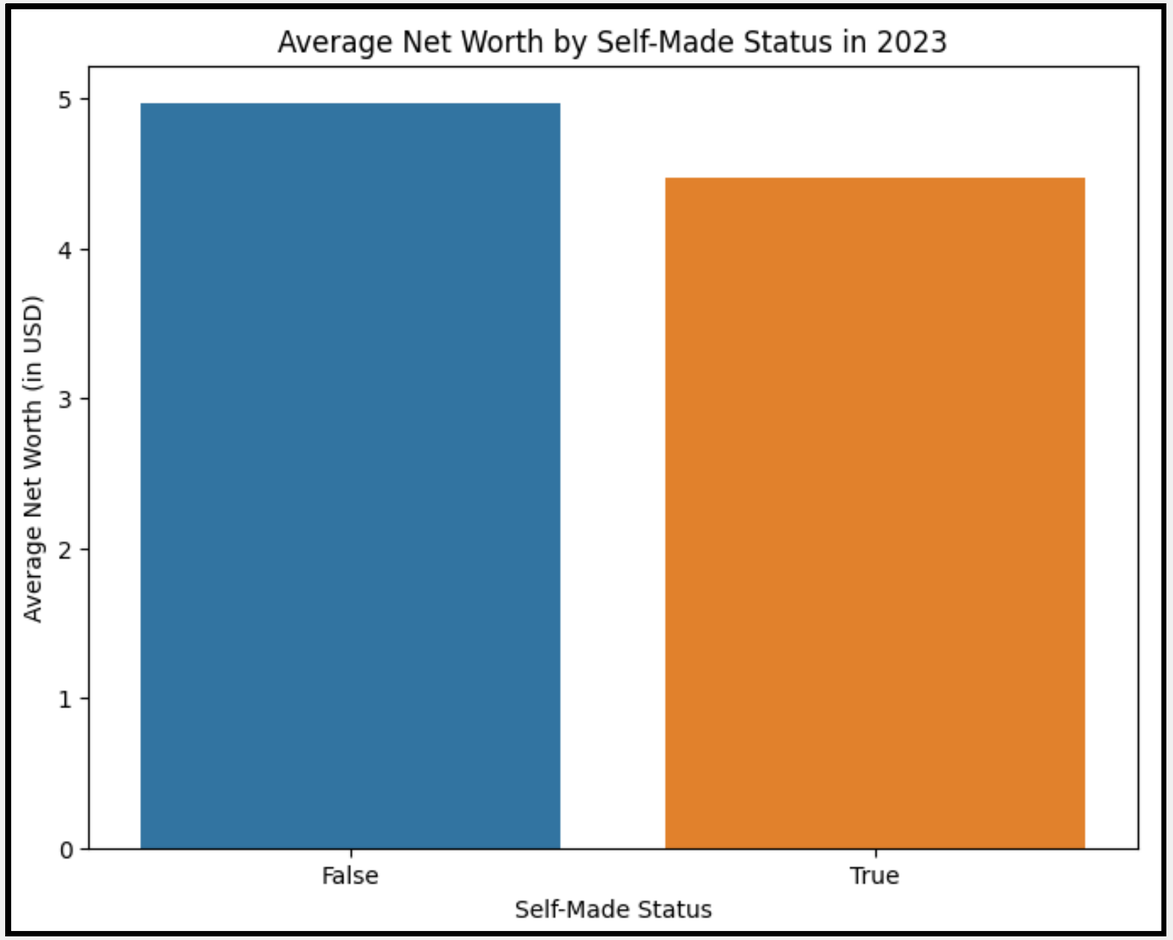
```
top_countries = df['country'].value_counts().head(10)
plt.figure(figsize=(12, 6))
sns.barplot(x=top_countries.index, y=top_countries.values)
plt.xticks(rotation=90)
plt.xlabel('Country')
plt.ylabel('Number of Billionaires')
plt.title('Top Countries with the Most Billionaires in 2023')
plt.show()
```

- The United States leads in billionaire concentration, reflecting its robust economy and other factors
- After US, China has significant number of billionaires



What is the difference in average net worth between self-made and inherited billionaires?

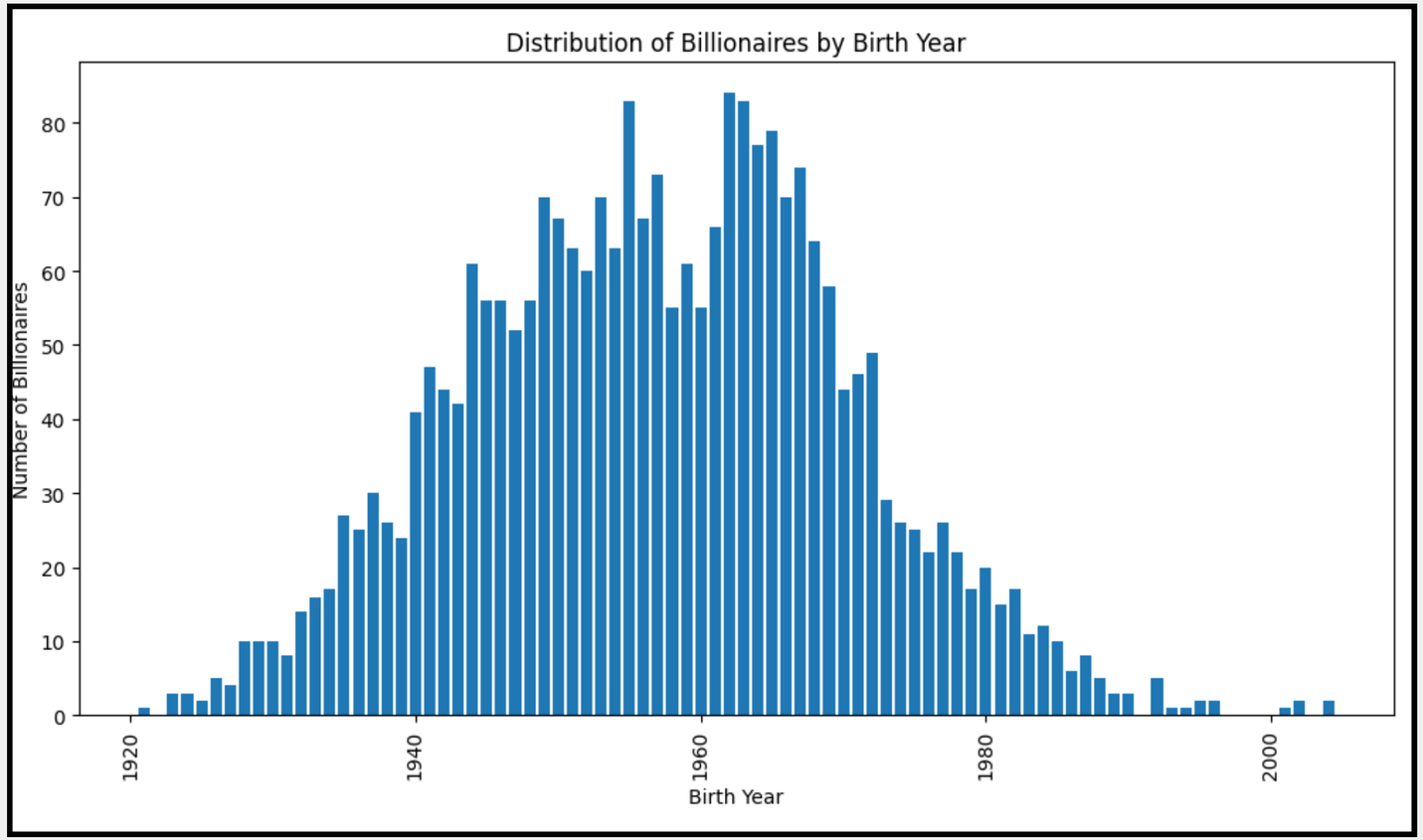
```
average_net_worth = df.groupby('selfMade')['finalWorth'].mean()
plt.figure(figsize=(8, 6))
sns.barplot(x=average_net_worth.index, y=average_net_worth.values)
plt.xlabel('Self-Made Status')
plt.ylabel('Average Net Worth (in USD)')
plt.title('Average Net Worth by Self-Made Status in 2023')
plt.show()
```



- Inherited billionaires typically have a higher net worth than self-made counterparts.
- Inherited billionaires benefit from substantial family assets, leading to elevated net worth
- Self-made billionaires face hurdles, resulting in a comparatively lower net worth. Factors include starting from scratch, and higher financial risks.

Are there any birth years that seem to produce more billionaires?

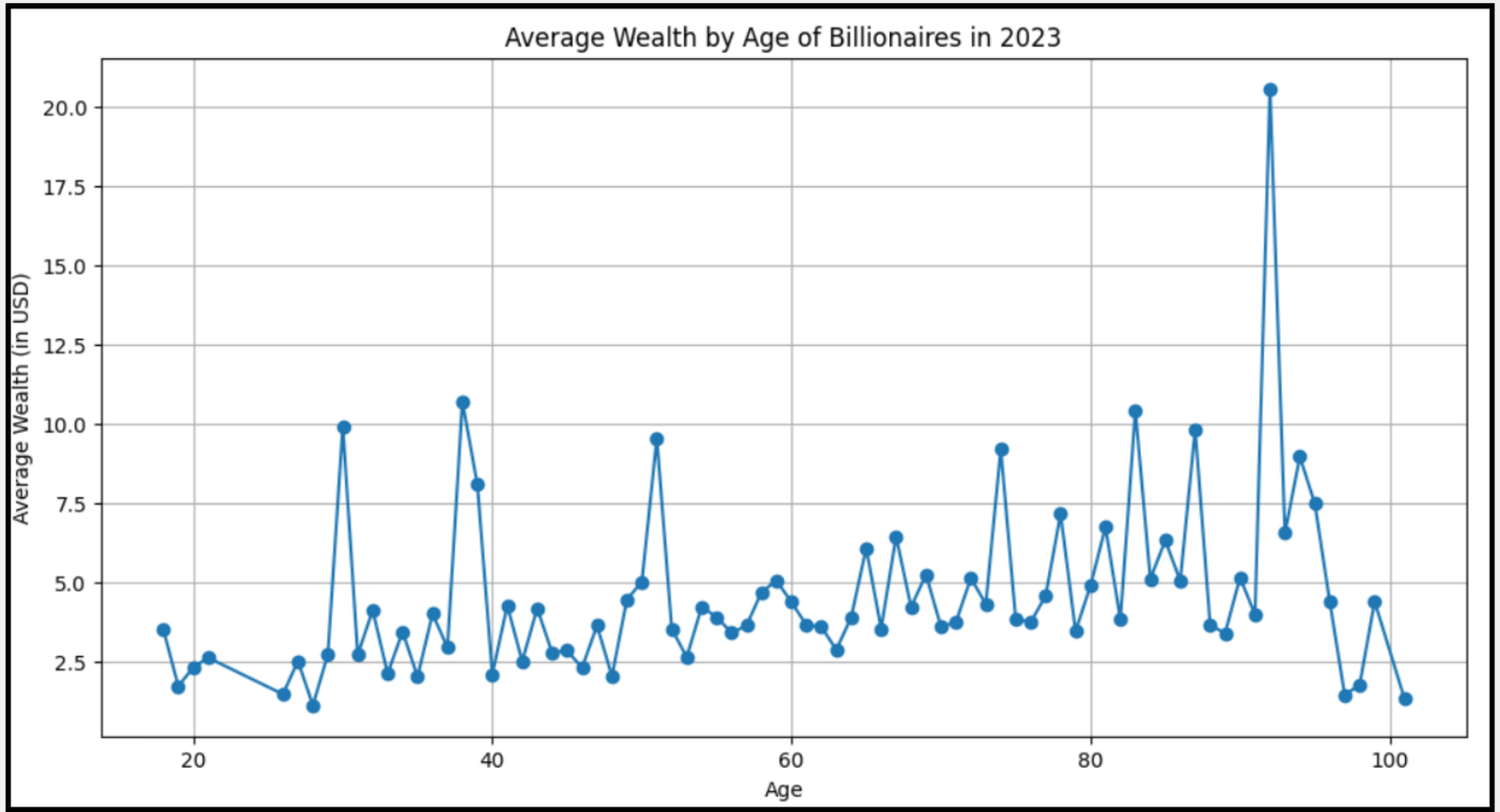
```
birth_year_counts = df['birthYear'].value_counts()
birth_year_counts = birth_year_counts.sort_index()
plt.figure(figsize=(12, 6))
plt.bar(birth_year_counts.index, birth_year_counts.values)
plt.xlabel('Birth Year')
plt.ylabel('Number of Billionaires')
plt.title('Distribution of Billionaires by Birth Year')
plt.xticks(rotation=90)
plt.show()
```



- Peaks in certain birth years prompt questions about the impact of historical events on wealth accumulation.
- Most of the Billionaires were born between 1950s - 1970s

How does the average wealth of billionaires change with age?

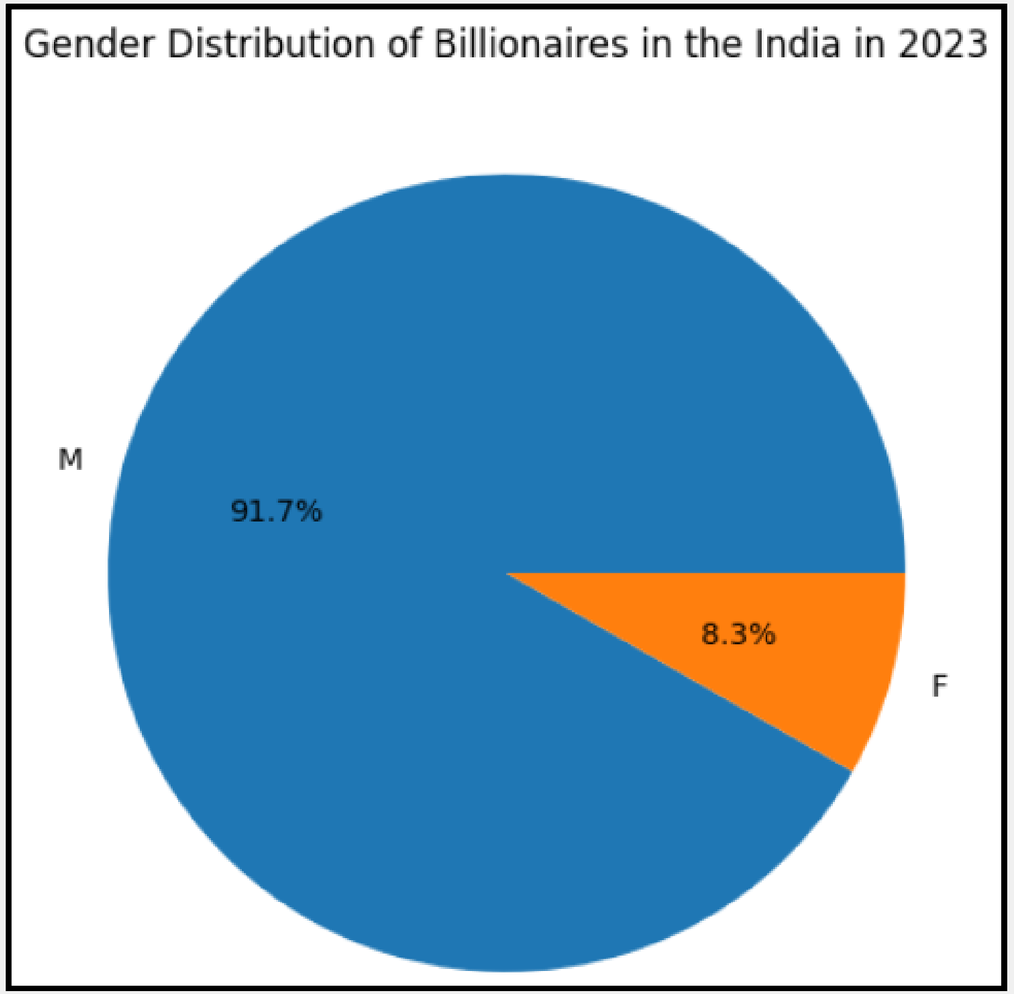
```
age_vs_wealth = df.groupby('age')['finalWorth'].mean().reset_index()

plt.figure(figsize=(12, 6))
plt.plot(age_vs_wealth['age'], age_vs_wealth['finalWorth'], marker='o', linestyle='-')
plt.xlabel('Age')
plt.ylabel('Average Wealth (in USD)')
plt.title('Average Wealth by Age of Billionaires in 2023')
plt.grid(True)
plt.show()
```



Analysis based on ‘INDIA’

What can you infer about the gender distribution of billionaires in India?

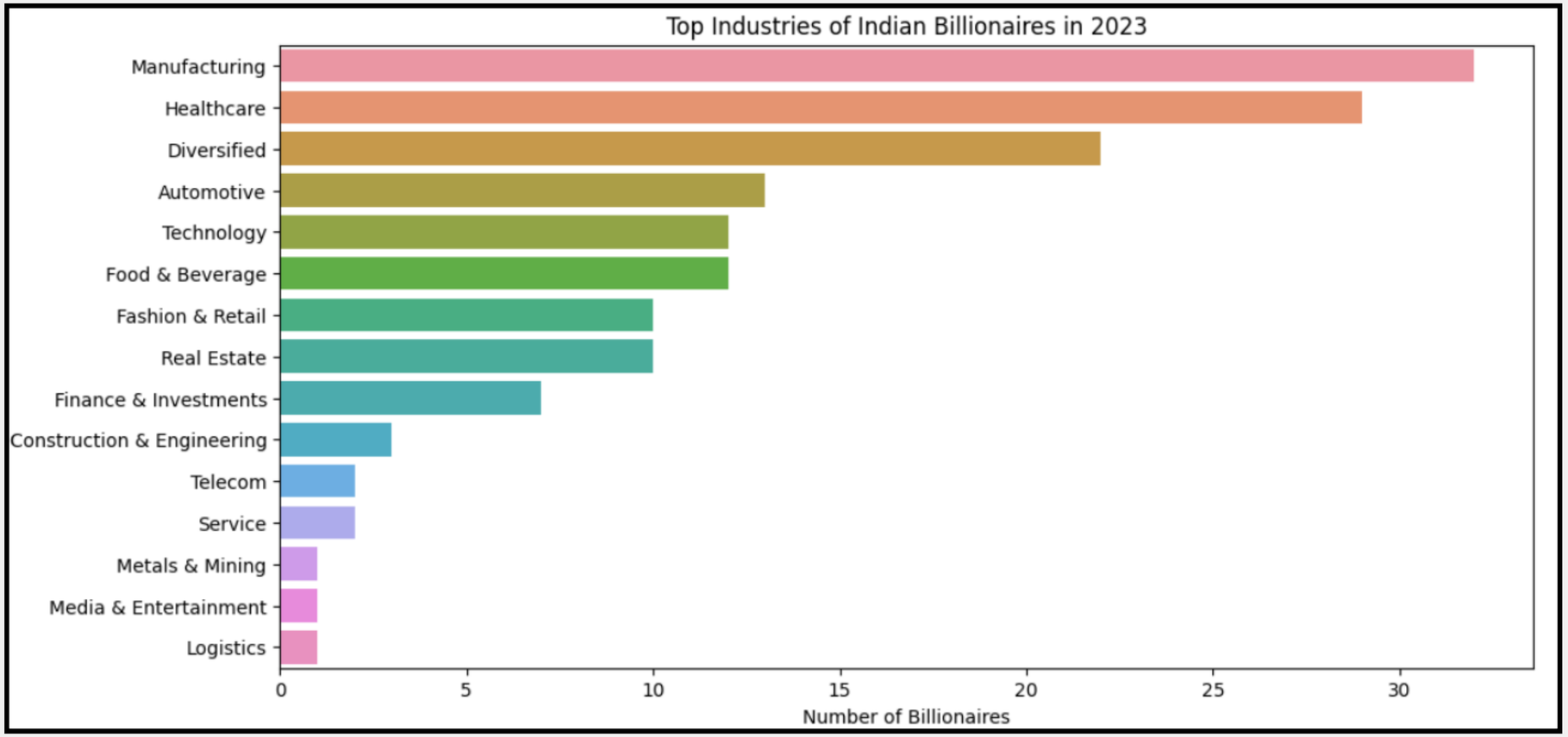
```
india_data = df[df['country'] == 'India']

gender_data = df[df['country'] == 'India']['gender'].value_counts()
plt.figure(figsize=(6, 6))
plt.pie(gender_data, labels=gender_data.index, autopct='%1.1f%%')
plt.title('Gender Distribution of Billionaires in the India in 2023')
plt.show()
```



- Gender distribution of Indian billionaires highlights a need for increased diversity.
- 91.3% of the billionaires are male

Which industries dominate the billionaire scene in India?

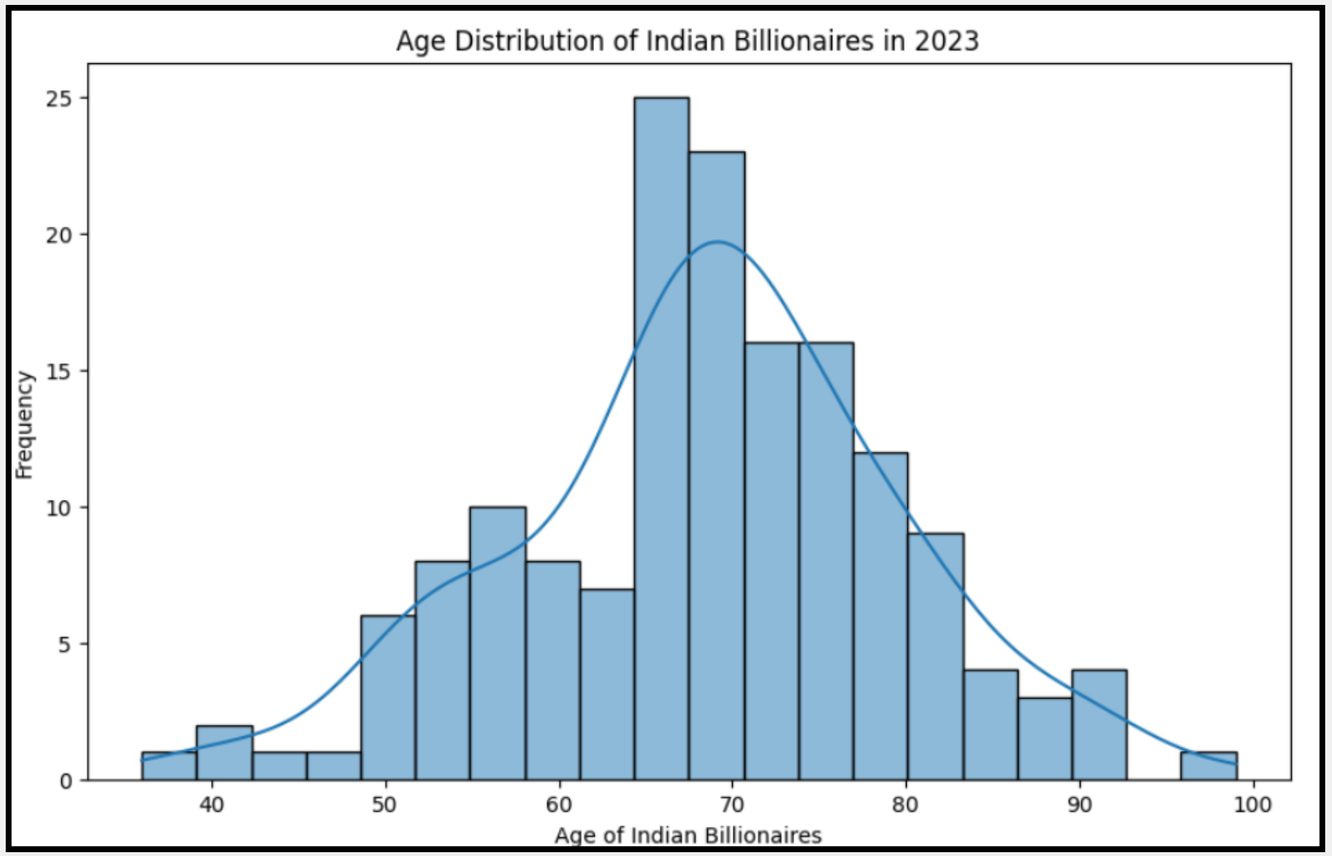
```
plt.figure(figsize=(12, 6))
sns.countplot(data=india_data, y='category', order=india_data['category'].value_counts().index)
plt.xlabel('Number of Billionaires')
plt.ylabel('Industry Category')
plt.title('Top Industries of Indian Billionaires in 2023')
plt.show()
```



- Manufacturing and Healthcare dominate in India, reflecting the nation's economic diversification.
- Followed by other diversified industries and the automotive sector

What is the age distribution of Indian billionaires?

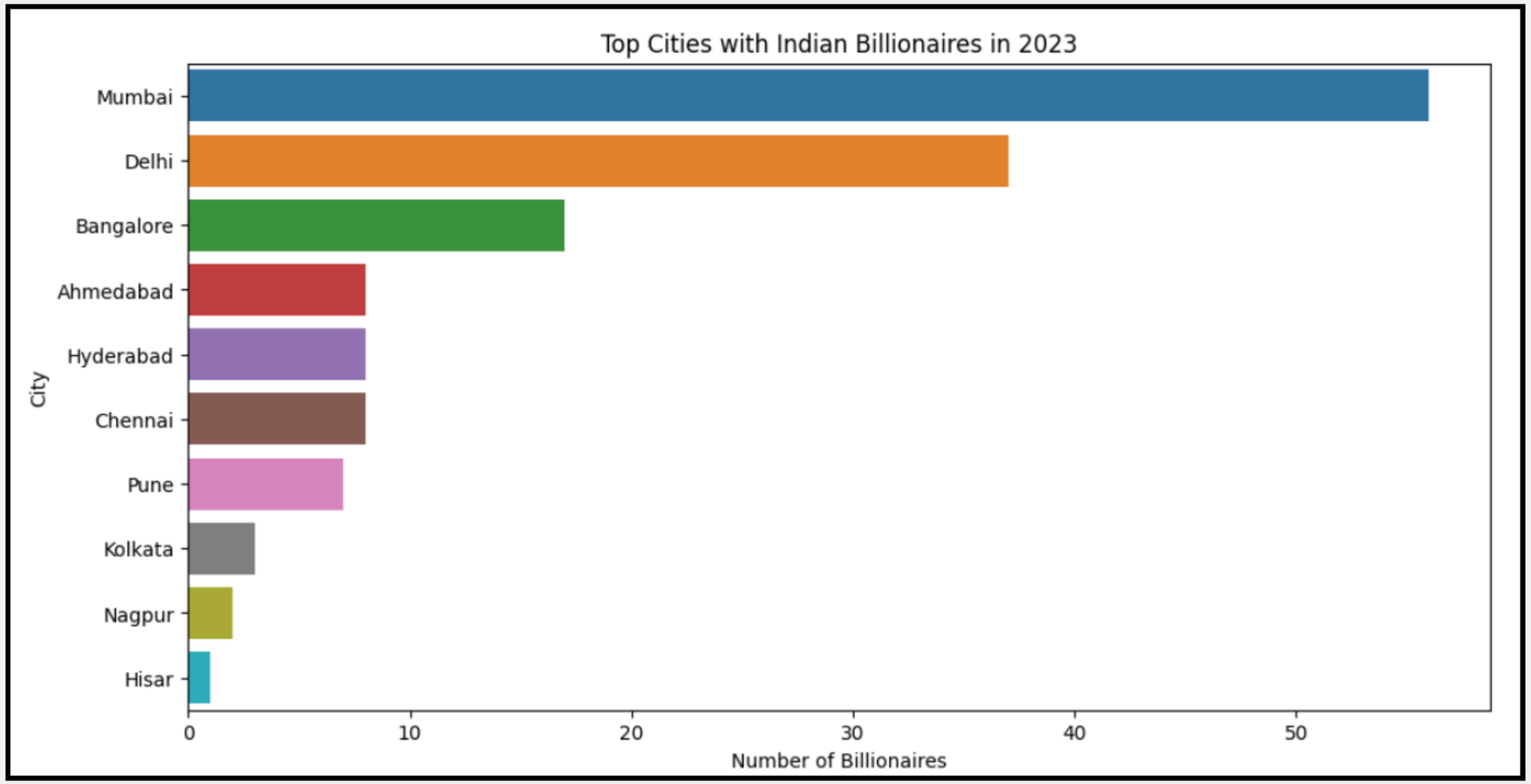
```
plt.figure(figsize=(10, 6))
sns.histplot(data=india_data, x='age', bins=20, kde=True)
plt.xlabel('Age of Indian Billionaires')
plt.ylabel('Frequency')
plt.title('Age Distribution of Indian Billionaires in 2023')
plt.show()
```



- Most of the Billionaires in India have age between 65yr and 70yr
- A steady decrease in number of Billionaires as age increases

Which cities in India have the highest concentration of billionaires?

```
top_cities = india_data['city'].value_counts().head(10)
plt.figure(figsize=(12, 6))
sns.barplot(x=top_cities.values, y=top_cities.index)
plt.xlabel('Number of Billionaires')
plt.ylabel('City')
plt.title('Top Cities with Indian Billionaires in 2023')
plt.show()
```



- Majority of Billionaires are from tier 1 cities
- With Mumbai has highest numbers followed by Delhi and Bangalore