

## ETL Bank Data Analysis

- Payal Kshirsagar

Dataset:

Rank	Bank_Name	Market cap
1	JPMorgan Chase	432.92
2	Bank of America	231.52
3	Industrial and Commercial Bank of China	194.56
4	Agricultural Bank of China	160.68
5	HDFC Bank	157.91
6	Wells Fargo	155.87
7	HSBC Holdings PLC	148.9
8	Morgan Stanley	140.83
9	China Construction Bank	139.82
10	Bank of China	136.81

Dataset schema:

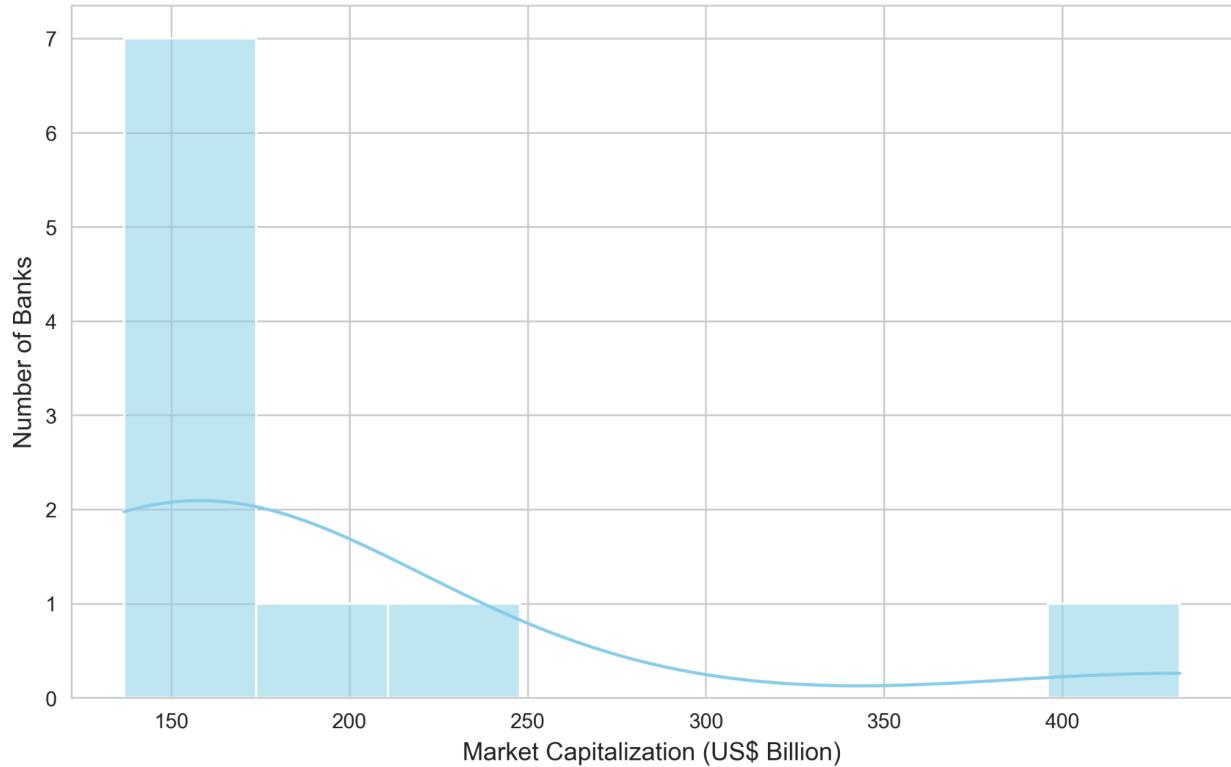
```
root
|--- Rank: long (nullable = true)
|--- Bank_Name: string (nullable = true)
|--- Market cap: double (nullable = true)
```

The dataset is clean and well-structured, containing 3 columns and 10 rows with no missing or duplicate values.

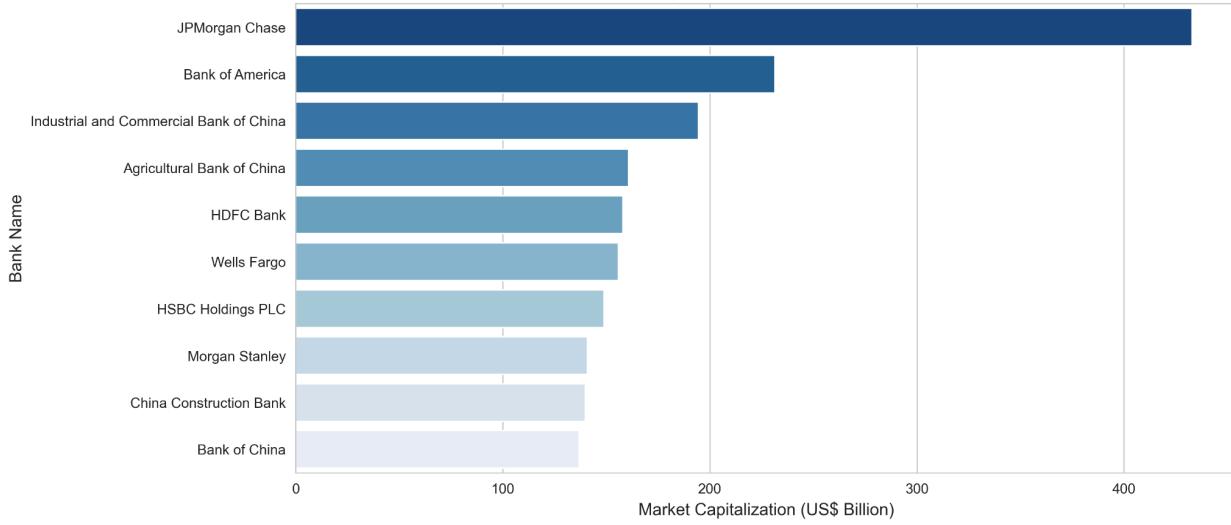
### Outlier analysis:

Outlier analysis reveals that JPMorgan Chase is a clear statistical outlier with a market capitalization of \$432.92 billion, significantly above the interquartile range of its peers. This indicates its dominant market position and exceptional investor valuation relative to other global banks, contributing to a right-skewed distribution in the dataset.

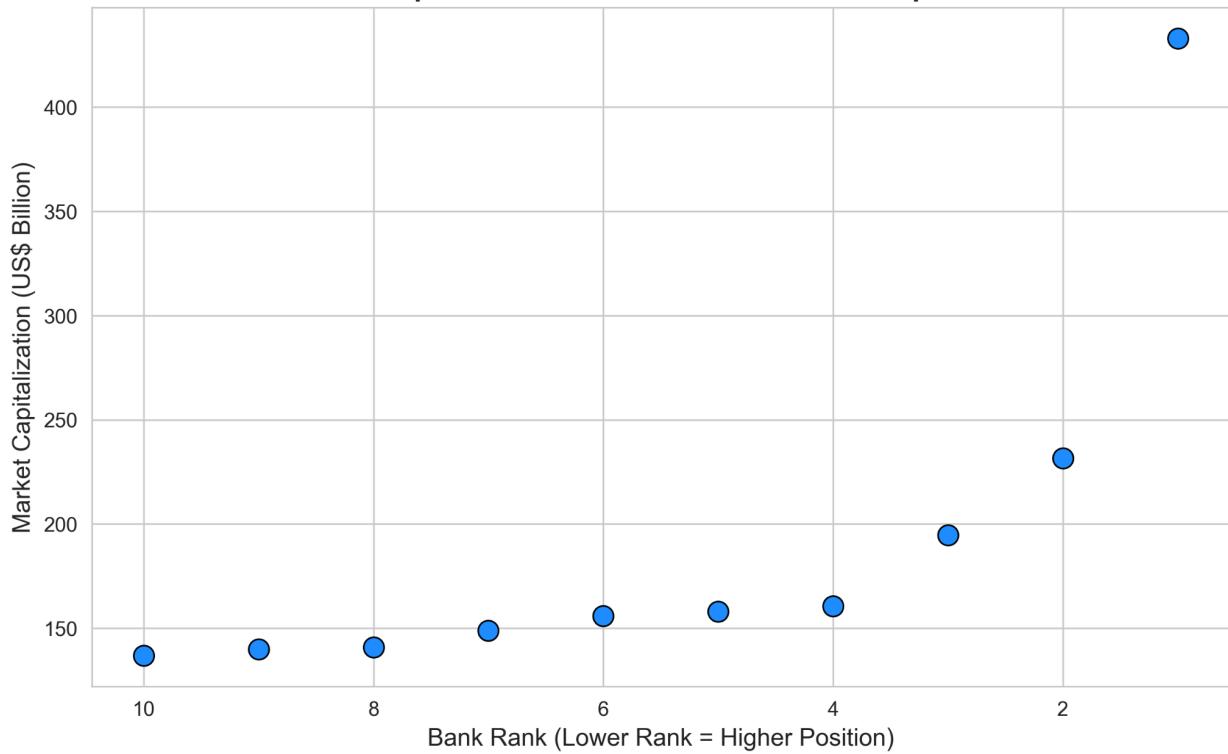
### Distribution of Market Capitalization (US\$ Billion)



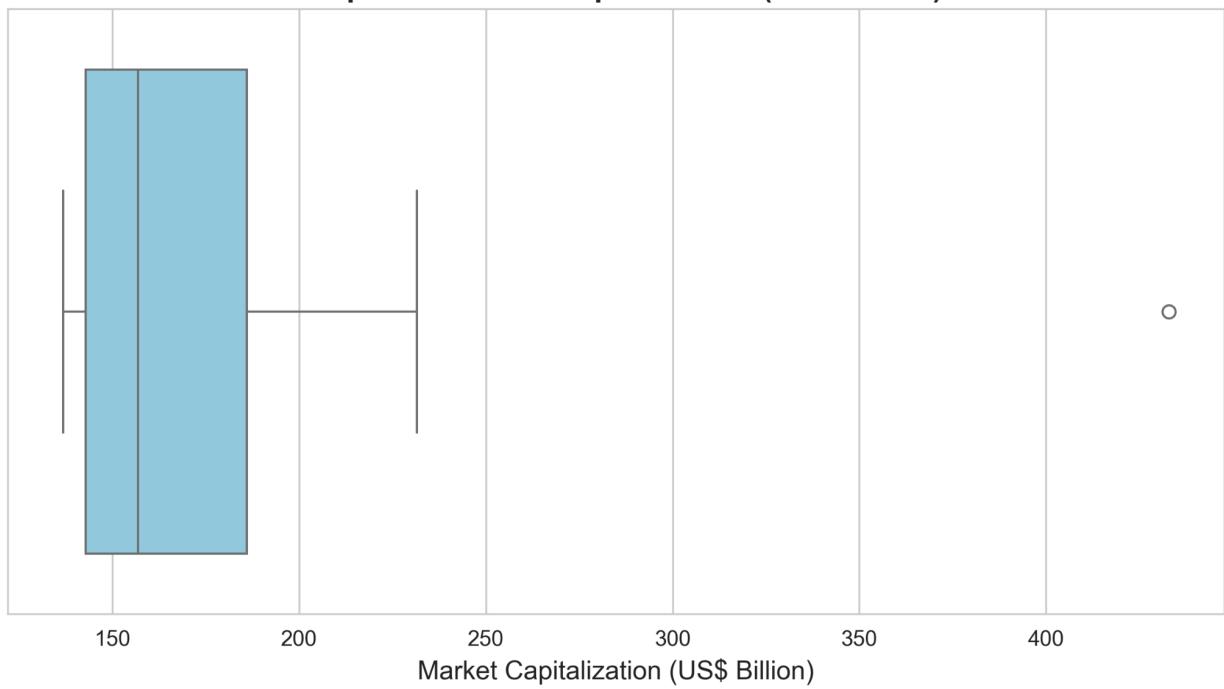
### Top 10 Banks by Market Capitalization (US\$ Billion)



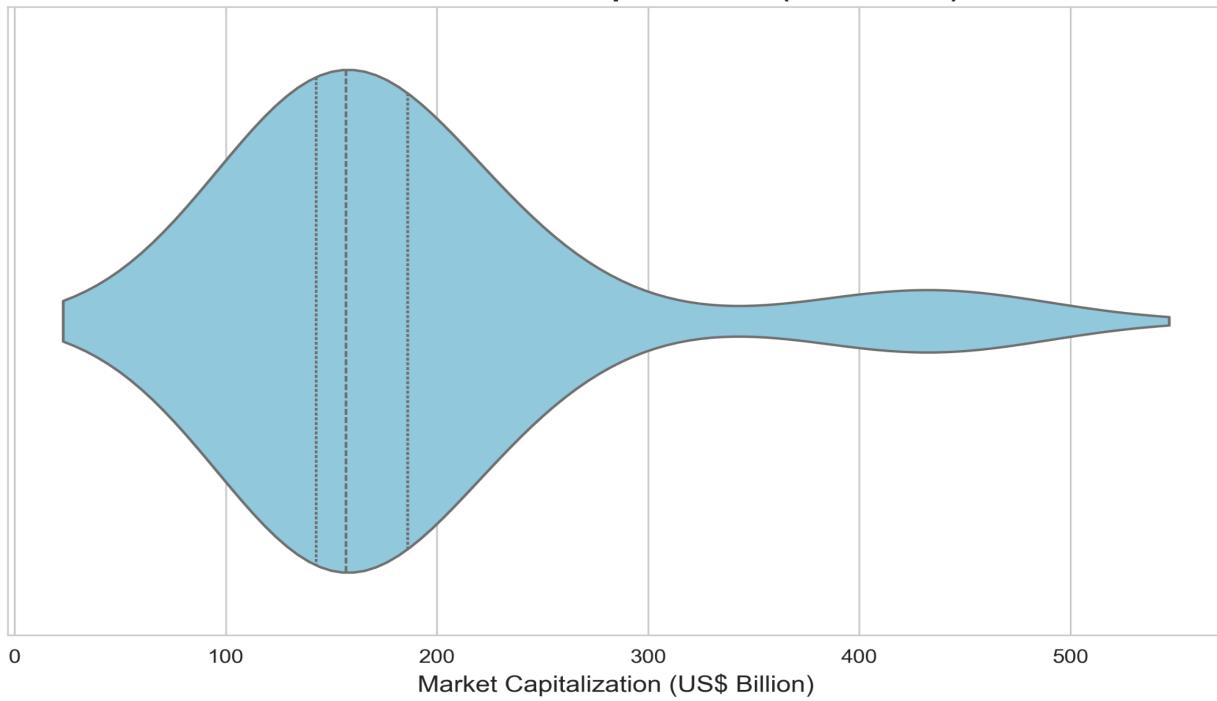
**Relationship between Bank Rank and Market Capitalization**



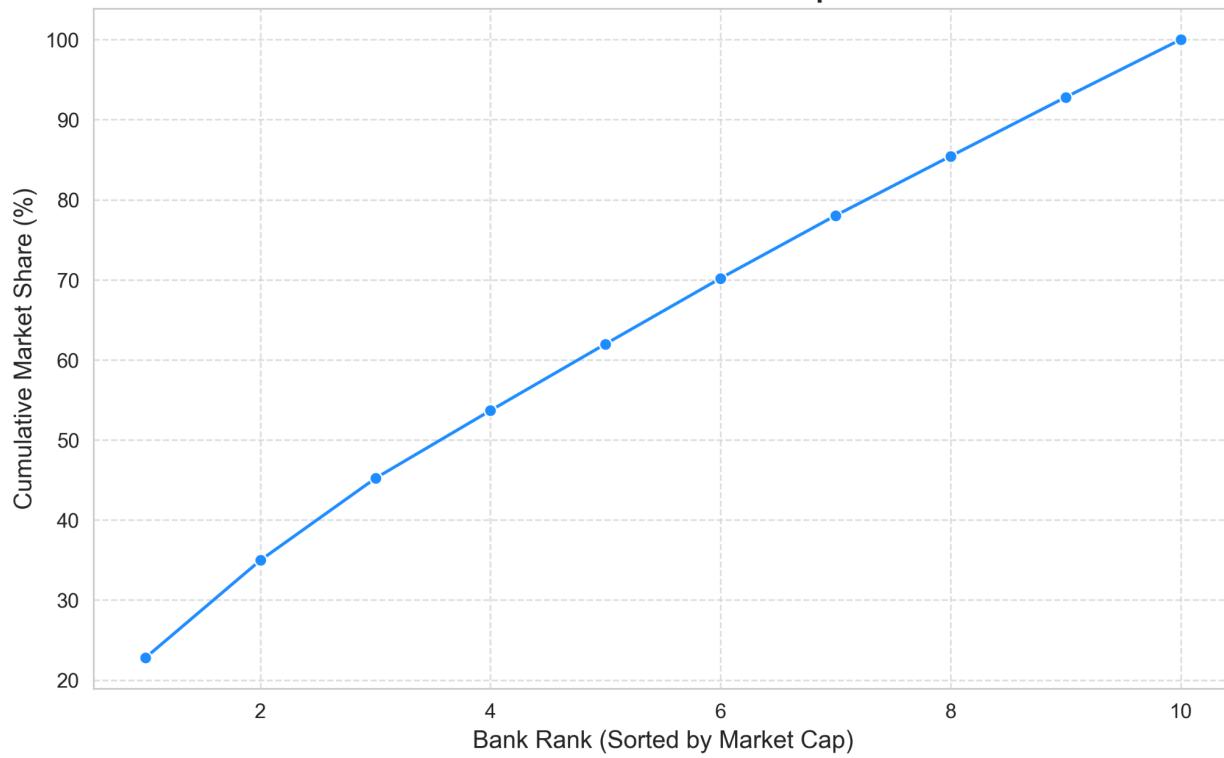
**Boxplot of Market Capitalization (US\$ Billion)**



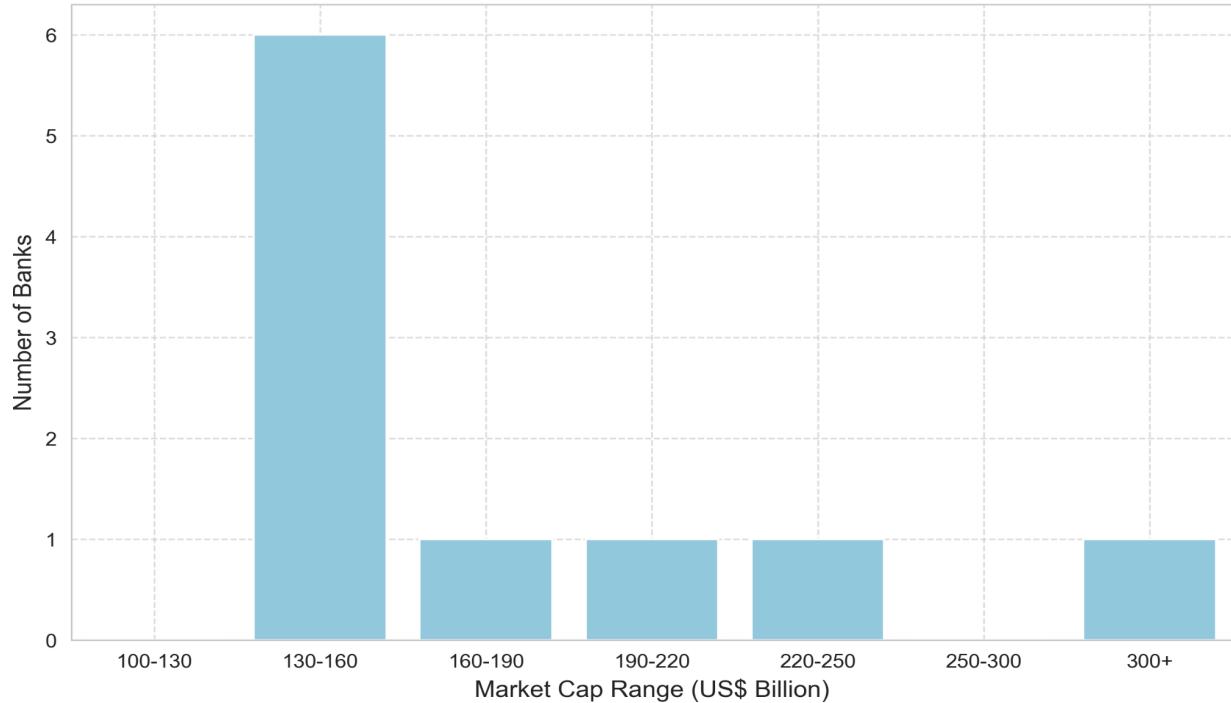
**Violin Plot of Market Capitalization (US\$ Billion)**



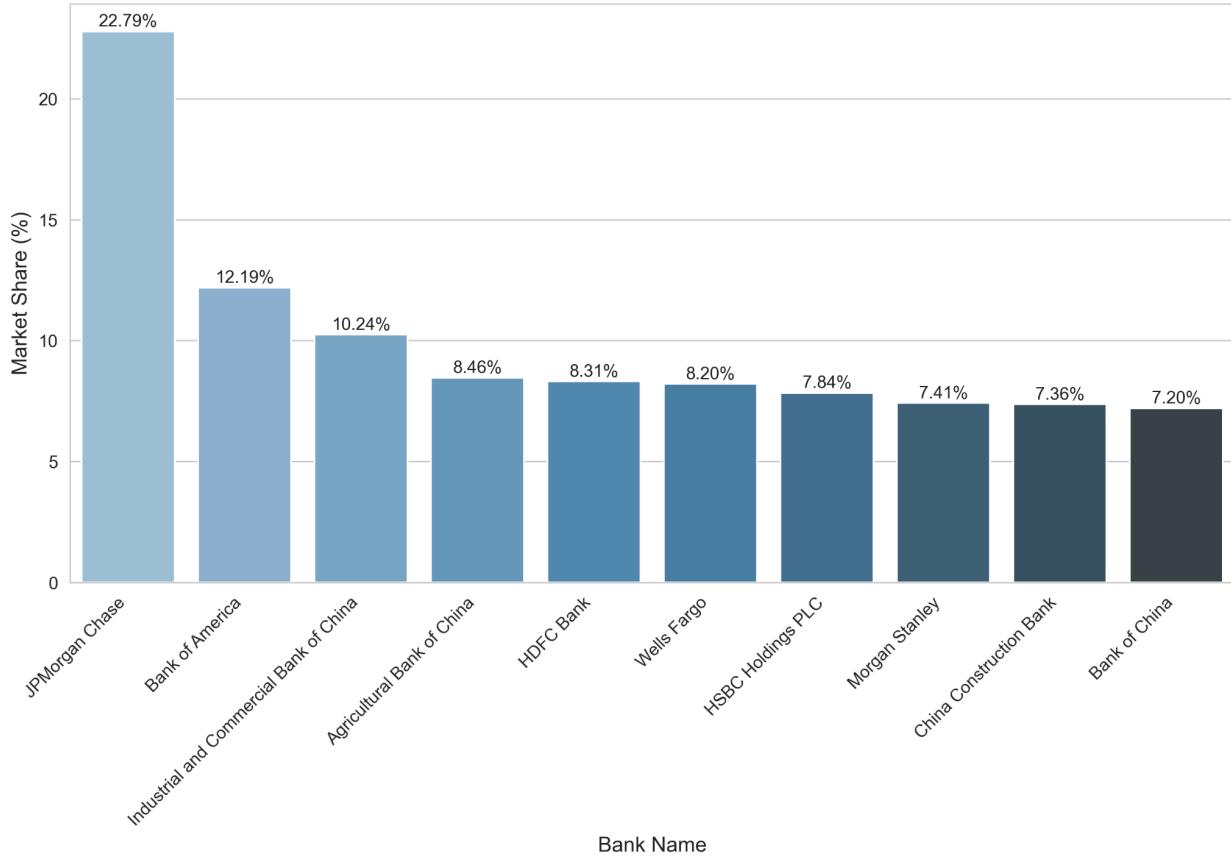
**Cumulative Market Share of Top Banks**



### Market Cap Range Distribution of Banks



### Top 10 Banks by Market Share (%)



## Advanced Market Cap Analysis with Growth Metrics

Rank_by_MarketCap	Bank_Name	Market_cap	Prev_MarketCap	MarketCap_Difference	MarketCap_Growth_Percent
1	JPMorgan Chase	432.92	0.0	0.0	0.0
2	Bank of America	231.52	432.92	-201.4	-46.52
3	Industrial and Commercial Bank of China	194.56	231.52	-36.96000000000001	-15.96
4	Agricultural Bank of China	160.68	194.56	-33.87999999999995	-17.41
5	HDFC Bank	157.91	160.68	-2.770000000000102	-1.72
6	Wells Fargo	155.87	157.91	-2.039999999999992	-1.29
7	HSBC Holdings PLC	148.9	155.87	-6.969999999999999	-4.47
8	Morgan Stanley	140.83	148.9	-8.069999999999993	-5.42
9	China Construction Bank	139.82	140.83	-1.010000000000193	-0.72
10	Bank of China	136.81	139.82	-3.009999999999991	-2.15

Rank\_by\_MarketCap – This ranks banks in descending order of their market capitalization, where Rank 1 represents the largest bank.

Prev\_MarketCap – This captures the market capitalization of the previous bank in the ranking (the one just above). It helps measure how much smaller each subsequent bank is compared to the one before it.

MarketCap\_Difference – This is the simple difference between each bank's market cap and the previous bank's market cap.

MarketCap\_Growth\_Percent – This represents how much each bank's market cap changes (in percentage) compared to the one above it.

### 1. Clear Market Leader

JPMorgan Chase stands far ahead of all competitors with a market capitalization of \$432.9 billion — almost double that of the second-largest bank, Bank of America (\$231.5 billion). This huge gap highlights JPMorgan's dominance and financial stability, making it the undisputed global leader in the banking sector.

### 2. Sharp Decline After the Top Rank

There's a dramatic 46% drop from Rank 1 to Rank 2, and another 16% drop from Rank 2 to Rank 3. After that, the decline becomes more gradual, suggesting that while the top few banks are far ahead, the middle ranks are more closely clustered together.

### 3. Concentration by Region

The list is dominated by banks from the United States and China, with four entries each. The remaining two — HDFC Bank (India) and HSBC Holdings (UK) — show the growing influence of emerging markets and the continued relevance of international banks.

#### 4. What the Growth Metrics Tell Us

The negative growth percentages don't mean the banks are losing value. Instead, they show how much smaller each bank is compared to the one above it.

The steep negative percentages at the top ranks show that market capitalization is heavily concentrated among the top few institutions.

#### 5. Overall Market Structure

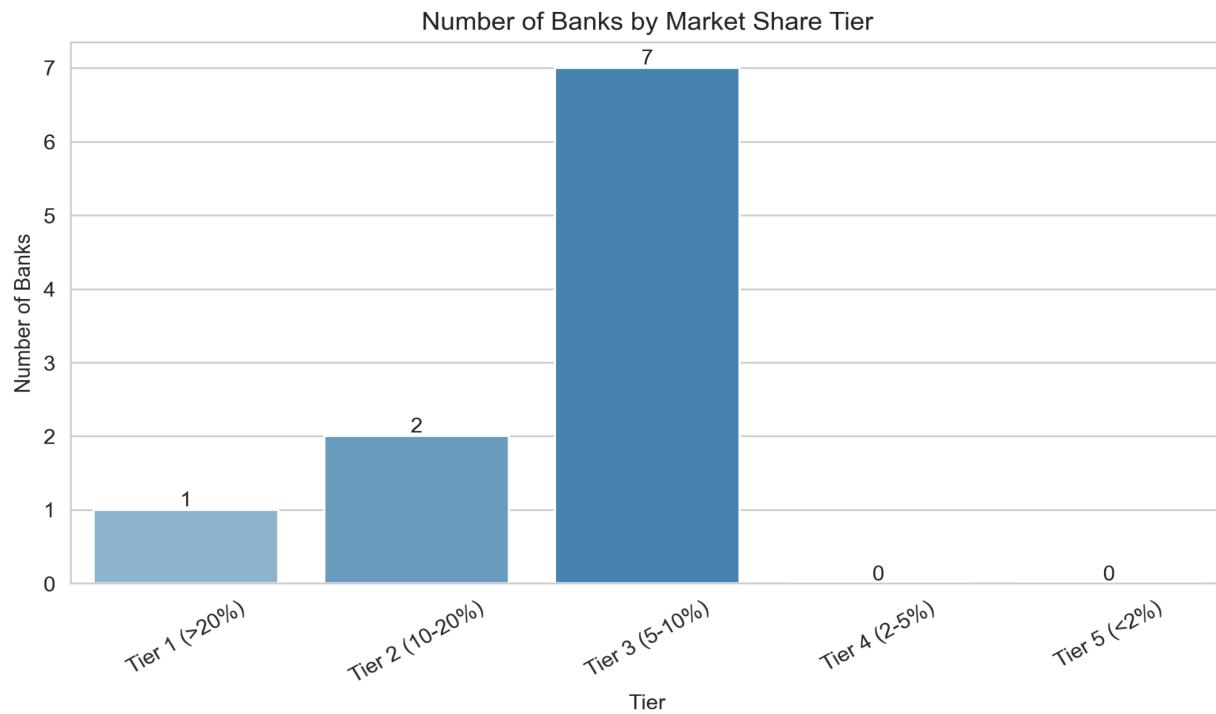
The top 3 banks together account for nearly half of the total market cap of the top 10 — a strong indicator of oligopolistic control in the global banking industry.

There's a visible "long tail" effect: a few giants dominate, while the rest hold significantly smaller shares.

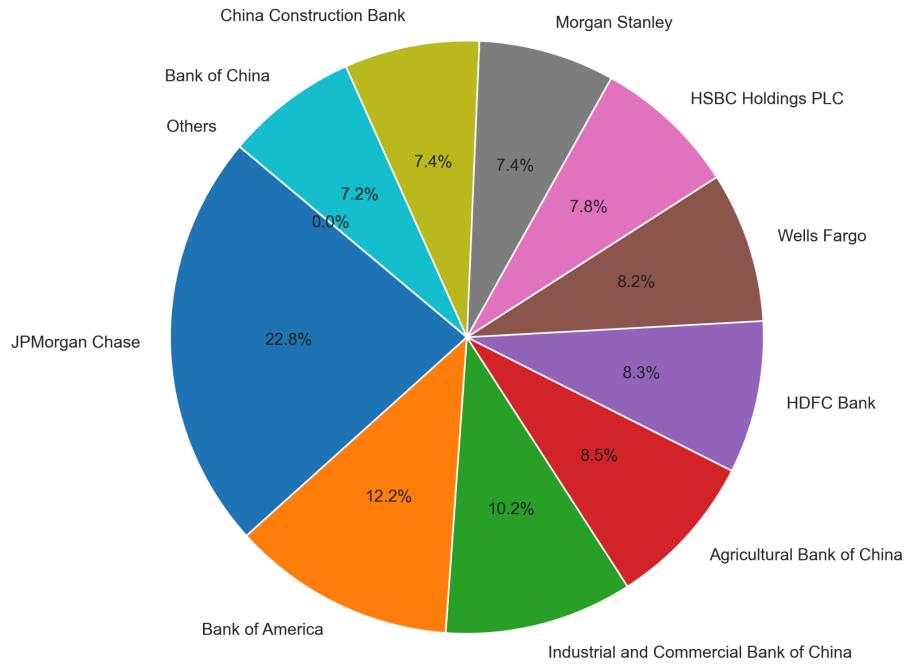
### Market Concentration Analysis

The global banking market, as represented by these top 10 institutions, exhibits a moderately concentrated yet balanced structure.

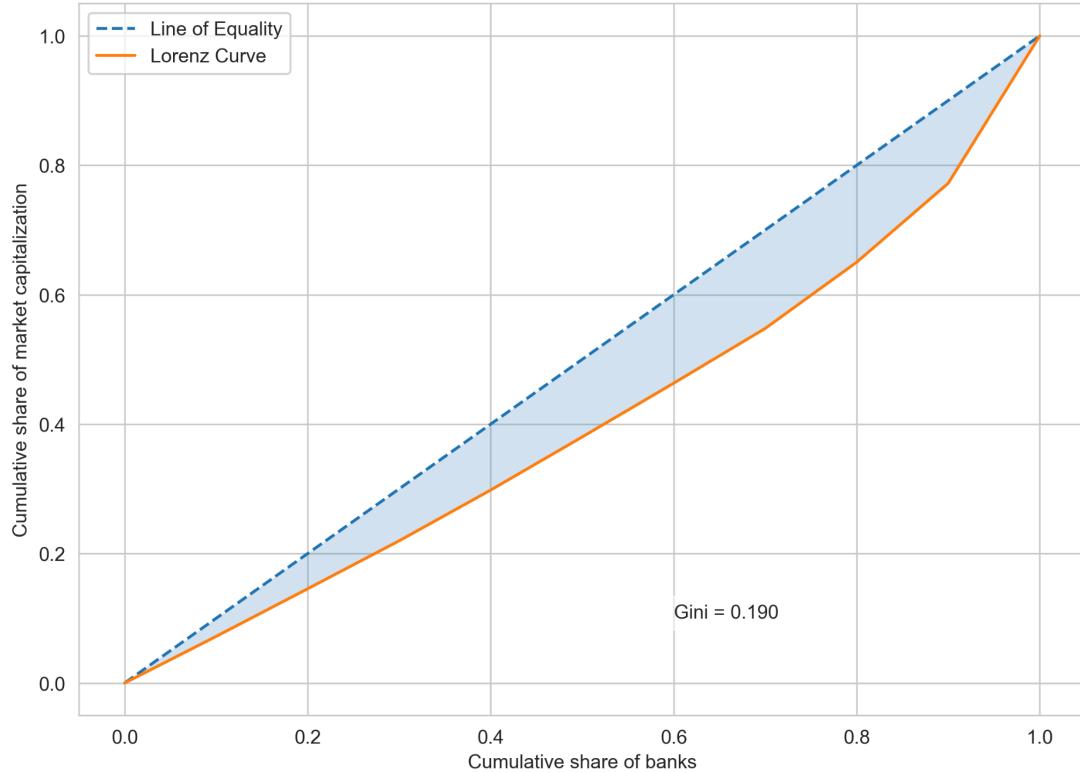
While a few American and Chinese banks command the largest individual market shares, the overall spread of capitalization suggests healthy competition with limited risk of market dominance or monopolistic tendencies.

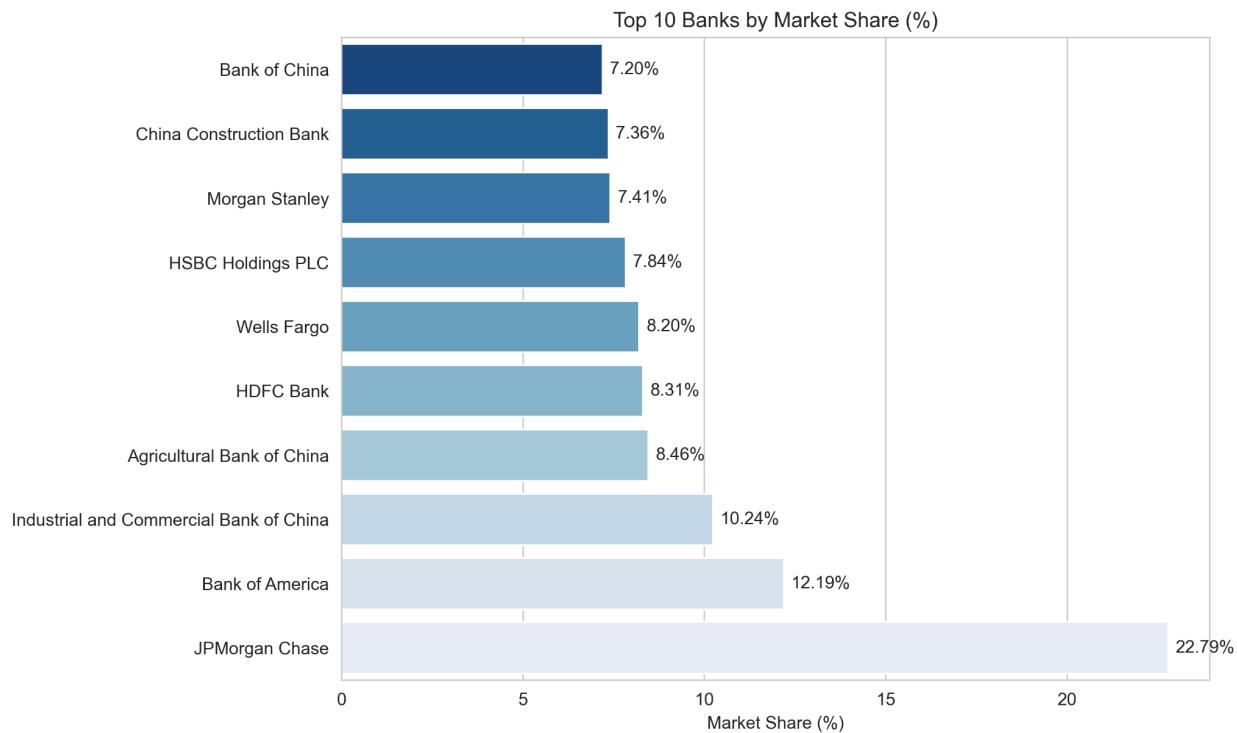


Market Share: Top 10 Banks vs Others



Lorenz Curve of Market Capitalization (Top Banks)



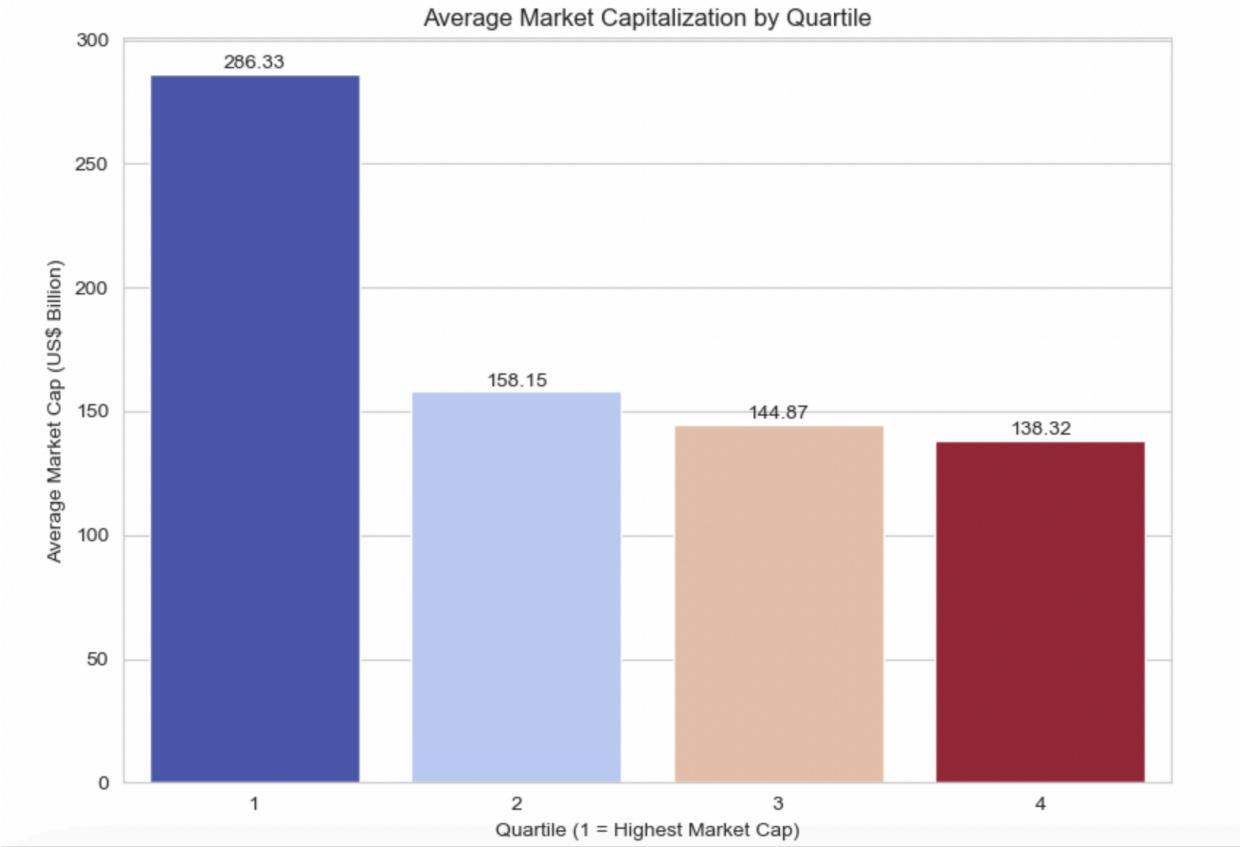


## Market Capitalization Distribution

The steep drop from Quartile 1 to Quartile 2 highlights that size inequality among the largest banks is significant — top institutions are nearly 2× larger than the next tier on average.

Beyond Q2, market caps become relatively uniform, implying that once outside the global mega-bank segment, the rest of the market is more balanced.

Investors and regulators can view Q1 banks as systemically critical, while Q2–Q4 represent a broad base of regional and diversified institutions.



## Comparative Size Analysis

### Large Cap Segment:

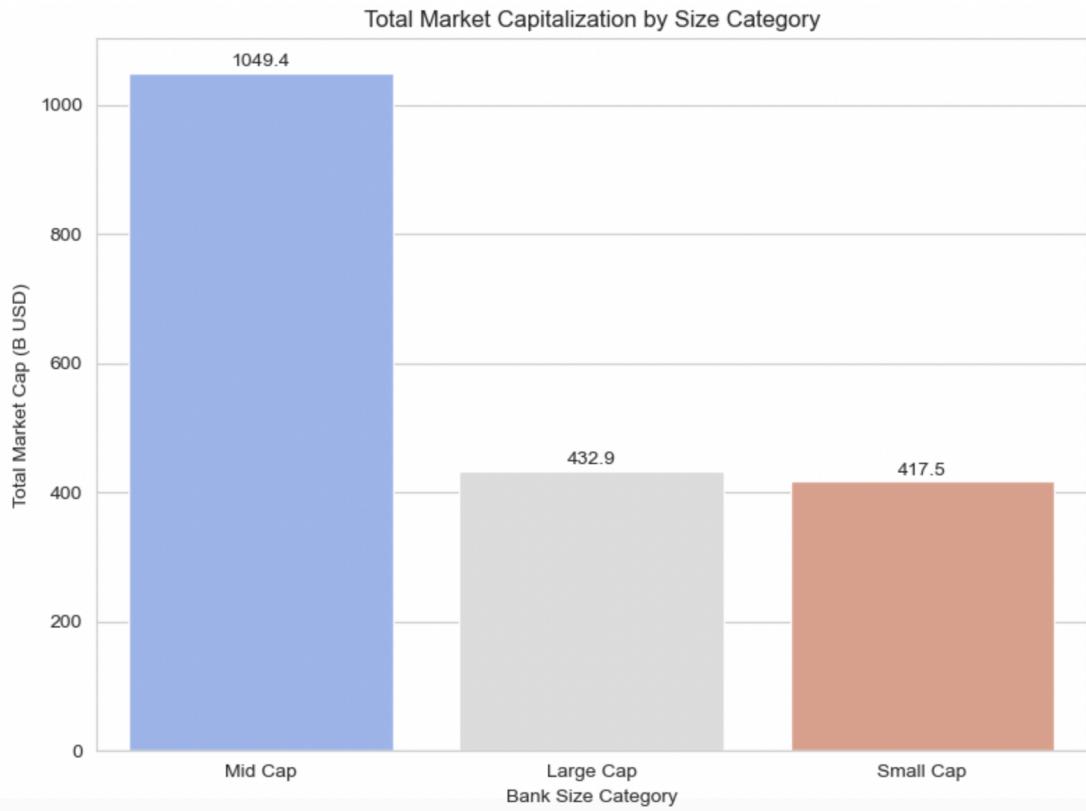
Only one bank, JPMorgan Chase, falls into the large-cap category, holding a market capitalization of \$432.92 billion. This single institution alone accounts for nearly 23% of the total market capitalization, underscoring its dominant position in the global banking landscape.

### Mid Cap Segment:

The mid-cap group, consisting of six banks, has an average market capitalization of \$174.91 billion. Collectively, they contribute around \$1,049 billion, representing the bulk of total market value. This segment contains major international players like Bank of America, ICBC, HDFC Bank, and Wells Fargo, signifying a strong cluster of competitive but diversified institutions.

### Small Cap Segment:

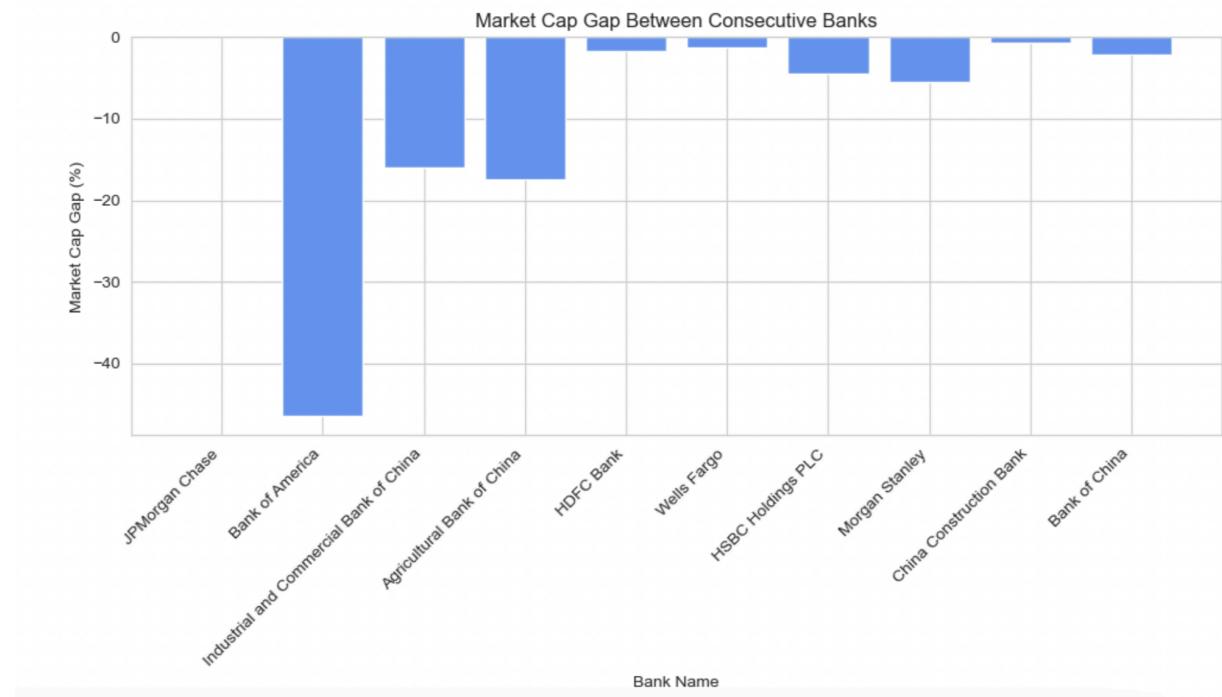
The small-cap category includes three banks, each with an average capitalization of approximately \$139 billion, summing to \$417 billion in total. These banks represent the smaller end of the top-tier institutions — still significant in scale but less dominant compared to the global leaders.



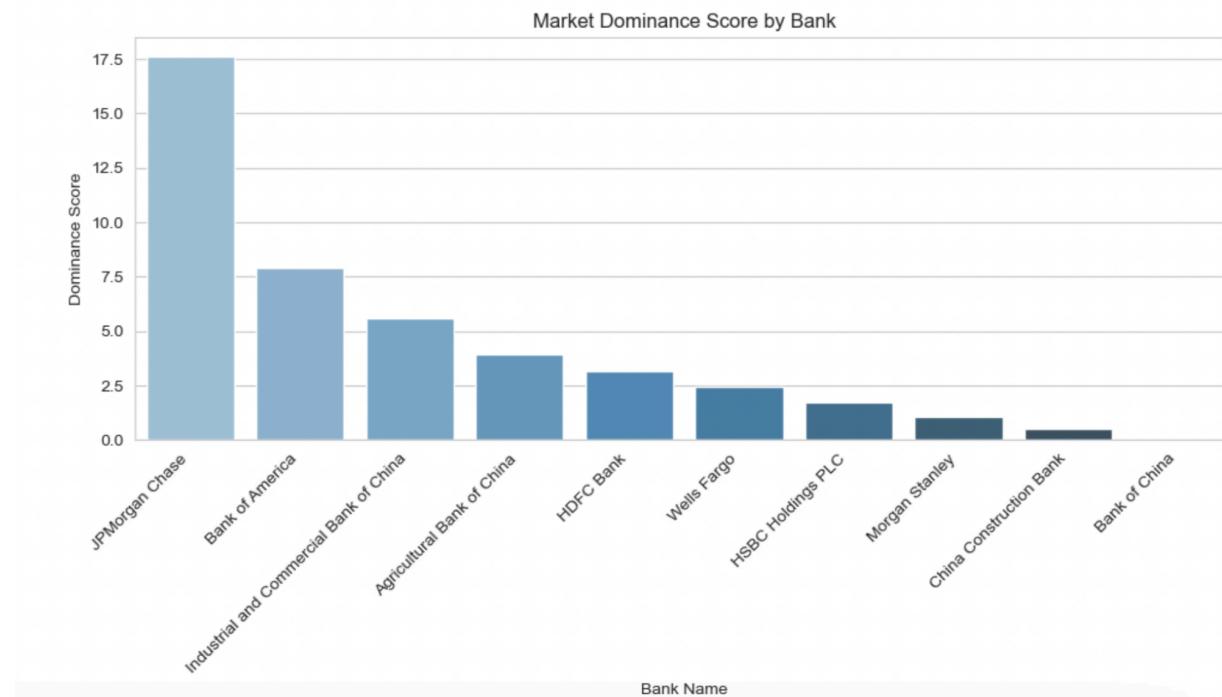
==== Comparative Size Analysis Report ====

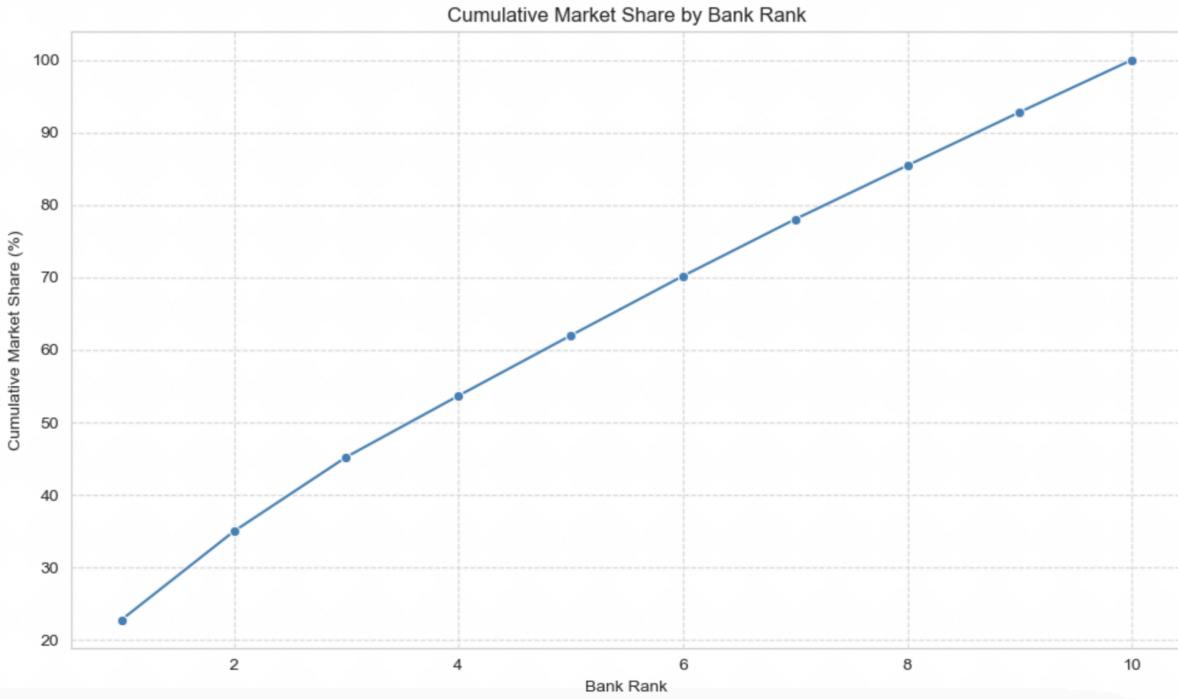
Size_Category	Bank_Count	Avg_MarketCap	Total_MarketCap
Mid Cap	6	174.906667	1049.44
Large Cap	1	432.920000	432.92
Small Cap	3	139.153333	417.46

## Growth and Gap Analysis



## Market Dominance Analysis





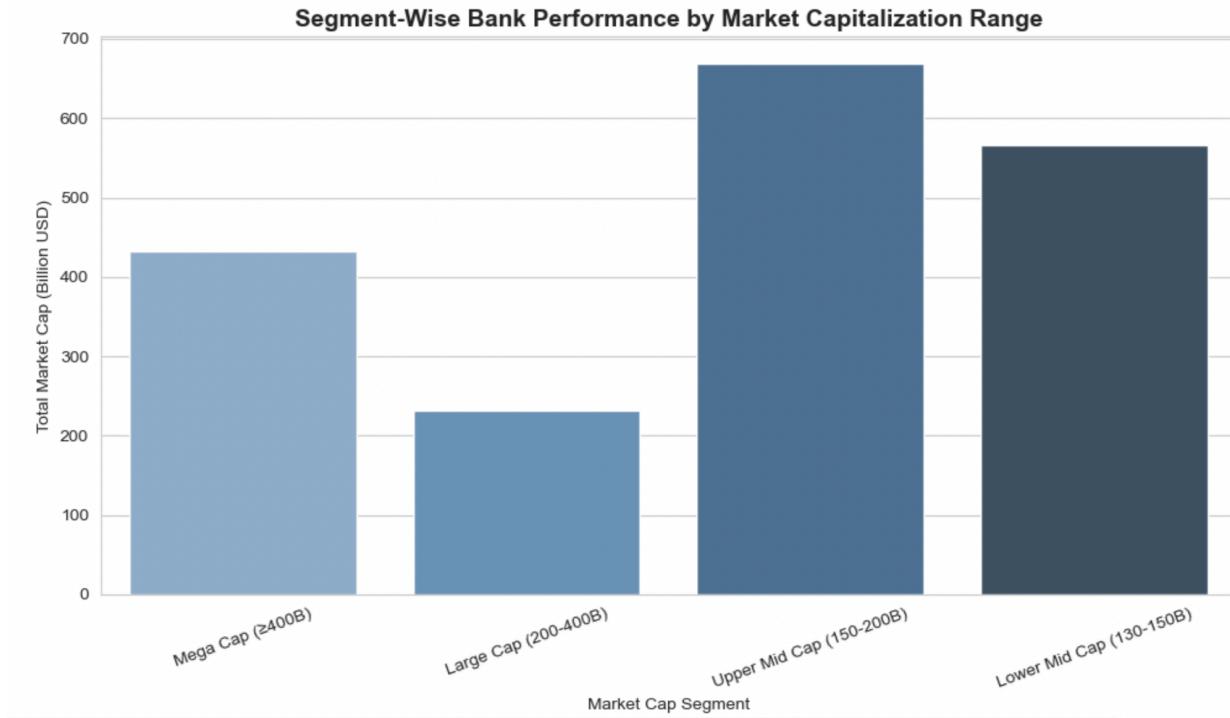
The global banking landscape is highly top-heavy, with JPMorgan Chase alone exerting dominant influence, followed by a closely packed cluster of mid-tier banks.

The top 5 banks together hold over 60% of total market capitalization, confirming that market dominance is concentrated among a few players, while smaller peers share the remaining slice of global value. This pattern reinforces the notion of a globally integrated but asymmetrically distributed financial system.

### Segment-Wise Bank Analysis

The market capitalization segmentation highlights a clearly tiered banking ecosystem, dominated by one Mega Cap leader and balanced by strong mid-tier groups.

MarketCap_Range	Bank_Count	Avg_MarketCap	Total_MarketCap	Min_MarketCap	Max_MarketCap
Mega Cap ( $\geq 400B$ )	1	432.92	432.92	432.92	432.92
Large Cap (200–400B)	1	231.52	231.52	231.52	231.52
Upper Mid Cap (15...)	4	167.26	669.02	155.87	194.56
Lower Mid Cap (13...)	4	141.59	566.36	136.81	148.9



The overall market capitalization analysis shows that the average market cap across all banks is \$189.98 billion, with a total of \$1.90 trillion, indicating a strong and sizable banking sector. The smallest bank has a market cap of \$136.81 billion, while the largest, JPMorgan Chase, stands at \$432.92 billion, more than three times larger. The average market share per bank is 10%, summing up to a total of 100%, suggesting proportional representation. However, the distribution is skewed, as a few top banks hold significantly higher valuations, creating a moderately concentrated market where dominant players control a large portion of total capitalization, while mid and lower-tier banks collectively contribute a smaller yet stable share.