EXPLORATORY DATA ANALYSIS ON CHASE BANK BRANCH PERFORMANCE

BY

JIBRIL YAHAYA JIBRIL

Table of Contents

1. Introduction

1.1 Objective
The objective of this project is to analyze the performance of Chase Bank branches across the United States, identifying trends, opportunities, and areas for improvement. The analysis aims to provide insights that will inform business strategies and optimize branch operations.

1.2 Dataset Description
The dataset used for this analysis contains records for every Chase Bank branch in the United States, with the following information:

1. Institution Name: The name of the bank or financial institution.
2. Main Office: Indicates if the branch is the main office (likely a headquarters or central location).
3. Branch Name: The name of the branch.
4. Branch Number: A unique identifier for the branch.
5. Established Date: The date the branch was established.
6. Acquired Date: The date the branch was acquired (if applicable).
7. Street Address: The physical street address of the branch.
8. City: The city where the branch is located.
9. County: The county where the branch is located.
10. State: The state where the branch is located.
11. Zipcode: The zip code of the branch's location.
12. Latitude and Longitude: The geographic coordinates (location) of the branch.
13. 2010 Deposits to 2016 Deposits: The total deposits (amount of money) for each year from 2010 to 2016.

1.3 Scope and Limitations
The scope of this analysis is limited to branch-level data, excluding individual customer information. The dataset only includes branches in the United States and does not account for global operations. The analysis focuses on the specified time period and does not consider prior or subsequent data.

2. Data Cleaning and Preprocessing

2.1 Handling missing values

The following is the missing values summary:

Columns	Missing Values
Institution Name	0
Main Office	0
Branch Name	0
Branch Number	0
Established Date	0
Acquired Date	1615
Street Address	0
City	0
County	0
State	0
Zipcode	0
Latitude	66
Longitude	66
2010 Deposits	740
2011 Deposits	578
2012 Deposits	329
2013 Deposits	175
2014 Deposits	56
2015 Deposits	19
2016 Deposits	0

The following issues (and the respective methodologies used to resolve them) are as follows:

✓ **Issue 1:**
Out of the total number of branches in the dataset, **1615** branches do not have an 'Acquired Date' recorded.

This could be because:
- These branches were not acquired (they were newly established)
- The acquisition date is not applicable (e.g., the branch was never acquired)
- The data is missing or not available

For this issue, i assume that these branches were newly established and as such, i proceed to impute the missing values with 'NA'.

✓ **Issue 2:**
There are **1,897** missing values across six columns ranging for deposits for year 2010 to 2015. The table thus:

2010 Deposits	740
2011 Deposits	578
2012 Deposits	329
2013 Deposits	175
2014 Deposits	56
2015 Deposits	19
2016 Deposits	0

I assume the following as to why there are missing values in these columns:

i. Data collection issues
ii. Changes in data sources
iii. Mergers and acquisitions:
iv. Data entry errors

For this issue, I proceed to fill the missing values with the median value [since it is a play project].

2.2 Data type verification and normalization

To ensure the quality and consistency of the data and verify it is in the correct format and meets the expected standards, I performed data profiling where I checked for the data type, converted all 'Deposits' columns to int64, checked for data range and uniqueness.

2.3 Date format standardization

For this, I typecast the 'Established Date' column and the 'Acquired Date' column from object datatype to the datetime datatype.

3. Exploratory Data Analysis (EDA)

3.1 Summary Statistics

Below is the tabular representation of the summary statistics:

	Main Office	Branch Number	Established Date	Acquired Date	Zipcode	2010 Deposits	2011 Deposits
count	5413	5413	5413	5413	5413	5413	5413
mean	0.000185	4303.761685	1978-05-08 15:12:12.237206752	2005-12-01 03:24:18.488823040	58800.016072	236271.3	276063.3
min	0	0	1800-01-01 00:00:00	1824-01-01 00:00:00	2110	0	0
25%	0	2548	1962-03-23 00:00:00	2004-11-13 00:00:00	33334	25453	25668
50%	0	4274	1989-04-30 00:00:00	2007-09-12 00:00:00	60654	43072	44426
75%	0	6172	2006-10-26 00:00:00	2008-09-25 00:00:00	89102	74298	78720
max	1	7989	2016-05-24 00:00:00	2016-05-24 00:00:00	99362	633131000	743268000
std	0.013592	2146.930304	NaN	NaN	30863.015686	8890869	10498670

3.2 Initial insights and observations

The summary statistics suggest that:

- The average deposits have consistently increased over the years, with a significant jump from 2015 to 2016.

- The standard deviation of deposits is relatively high, indicating significant variation in deposits across branches.

- The median deposits are consistently lower than the mean deposits, indicating that a few branches have exceptionally high deposits, skewing the mean.

- The minimum deposits are zero, indicating that some branches have no deposits or have not reported deposits.

- The maximum deposits are extremely high (over 1.15 billion), indicating potential outliers or errors in data entry.

These observations suggest that:

- There may be a need to investigate the data quality and handle potential outliers or errors.

- The bank has expanded its branches over time, with a significant increase in deposits in recent years.

3.3 Annual Total Deposits

This is to answer the question '*What is the overall performance of the branches in terms of deposits, and how has it changed over the years (2010-2016)?*'

3.4 Ranking and identification of top-performing branches

Below are the top-10 most performing branches (by deposits):

2010

Branch Name	Branch Number	2010 Deposits
JPMorgan Chase Bank Main Office	0	633131000
Madison and 48th St Branch	3586	153577328
Madison Ave Abd 48th St Branch	544	45099076
Houston Main Office	1315	26356388
Chicago's Main Office Branch	1838	22101086
Dallas Main Banking Center Branch	4311	15209318
Detroit Main Branch	2616	6953356
Broad St Downtown Branch	3481	4916291
Baton Rouge Main Branch	2036	4301756
Chase Tower Branch	2414	4246556

2011

Branch Name	2011 Deposits
JPMorgan Chase Bank Main Office	743268000
Madison and 48th St Branch	194388000
Madison Ave Abd 48th St Branch	62630984
Houston Main Office	36991358
Chicago's Main Office Branch	27521204
Dallas Main Banking Center Branch	17537788
Detroit Main Branch	10480927
Broad St Downtown Branch	5986462
Fifth Avenue & 44th Street Banking Center Branch	4792944
One Utah Branch	4549988

2012

Branch Name	2012 Deposits
JPMorgan Chase Bank Main Office	832455000
Madison and 48th St Branch	207367988
Madison Ave Abd 48th St Branch	81960379
Houston Main Office	38794794
Chicago's Main Office Branch	36219931
Dallas Main Banking Center Branch	21265872
Detroit Main Branch	12544817
Fifth Avenue & 44th Street Banking Center Branch	6278005
Broad St Downtown Branch	6114913
Baton Rouge Main Branch	5014857

2013

Branch Name	2013 Deposits
JPMorgan Chase Bank Main Office	916543000
Madison and 48th St Branch	223395353
Madison Ave Abd 48th St Branch	85177807
Houston Main Office	52284845
Chicago's Main Office Branch	39152650
Dallas Main Banking Center Branch	22966580
Detroit Main Branch	14030471
Fifth Avenue & 44th Street Banking Center Branch	8334885
Broad St Downtown Branch	6447273
Chase Tower Branch	5582684

2014

Branch Name	2014 Deposits
JPMorgan Chase Bank Main Office	1032549000
Madison and 48th St Branch	241540582
Madison Ave Abd 48th St Branch	98322162
Houston Main Office	82408236
Chicago's Main Office Branch	47103774
Dallas Main Banking Center Branch	26173793
Detroit Main Branch	18121683
Broad St Downtown Branch	7142562
Fifth Avenue & 44th Street Banking Center Branch	6822099
Chase Tower Branch	5857726

2015

Branch Name	2015 Deposits
JPMorgan Chase Bank Main Office	1069425000
Madison and 48th St Branch	261292750
Madison Ave Abd 48th St Branch	84569038
Houston Main Office	60702170
Chicago's Main Office Branch	43326953
Dallas Main Banking Center Branch	26944549
Detroit Main Branch	18855647
One Utah Branch	9078323
Empire State Branch	7099781
Broad St Downtown Branch	6915884

2016

Branch Name	2016 Deposits

JPMorgan Chase Bank Main Office	1155185000
Madison and 48th St Branch	313686851
Madison Ave Abd 48th St Branch	69044786
Houston Main Office	57820585
Chicago's Main Office Branch	36539028
Dallas Main Banking Center Branch	25414792
Detroit Main Branch	19499502
Empire State Branch	15580897
One Utah Branch	9138734
Broad St Downtown Branch	8142300

Overall Top performing Branches

Branch Name	Total Deposits
JPMorgan Chase Bank Main Office	6382556000
Madison and 48th St Branch	1595248852
Madison Ave Abd 48th St Branch	526804232
Houston Main Office	355358376
Chicago's Main Office Branch	251964626
Dallas Main Banking Center Branch	155512692
Detroit Main Branch	100486403
Broad St Downtown Branch	45665685
Fifth Avenue & 44th Street Banking Center Branch	39783307
Chase Tower Branch	39524482

Least Performing Branches (136)

Branch Name	Total Deposits
Bay Area Boulevard Motor Bank Branch	0
Westminister Thurber Ret Ctr Branch	0
Youree South Drive-Up Branch	0
South Meridian Drive-Up Branch	0
Detached Drive-In Facility Branch	0
...	...
Park Central Motor Bank Branch	0
Upper Montclair Drive-In Branch	0
Cs First Boston Corp Cashier Branch	0
Wheaton Main Drive-Up Branch	0
Quarum Motor Bank Branch	0

There are upon deep anaysis, I found out that there are 136 branches that have a $0 deposit over all the years; however, since I do not have any concrete information about the cause I assume the following:
  - The branch may have closed or been merged with another branch.
 - The branch may not have had any deposit activity during that time period.

- The data may be incomplete or missing for that branch.
- The branch may be a new location that hasn't had any deposits yet.

3.5 Branch establishment and acquisition trends

From the data, we can see:

- The earliest establishment date is 01/01/1824 (which appears to be a default or placeholder date)
- The latest establishment date is 01/01/2016
- Some branches have no acquisition date (see 2.1)
- The first acquisition was on 01/18/1975
- The last acquisition was on 10/14/2011
- There is a total of 3798 acquisitions all together
- Most acquisitions happened in the year 2004 with a total of 1602 acquisitions

3.6 Visualizations

Branches Acquired by Year
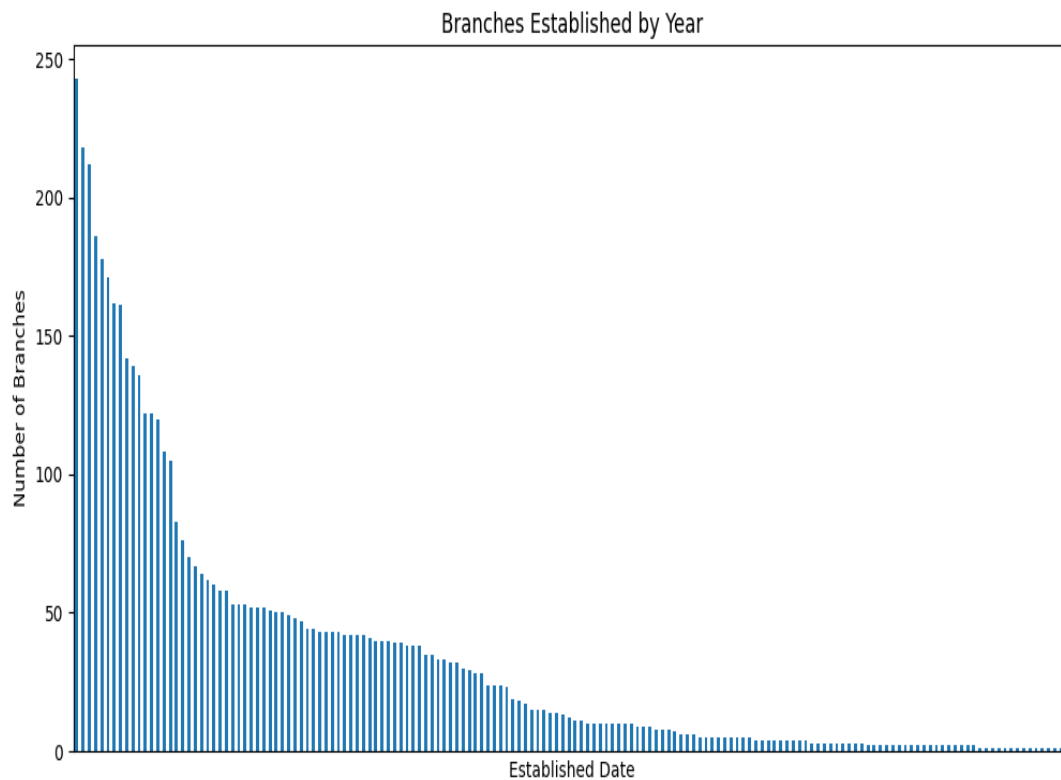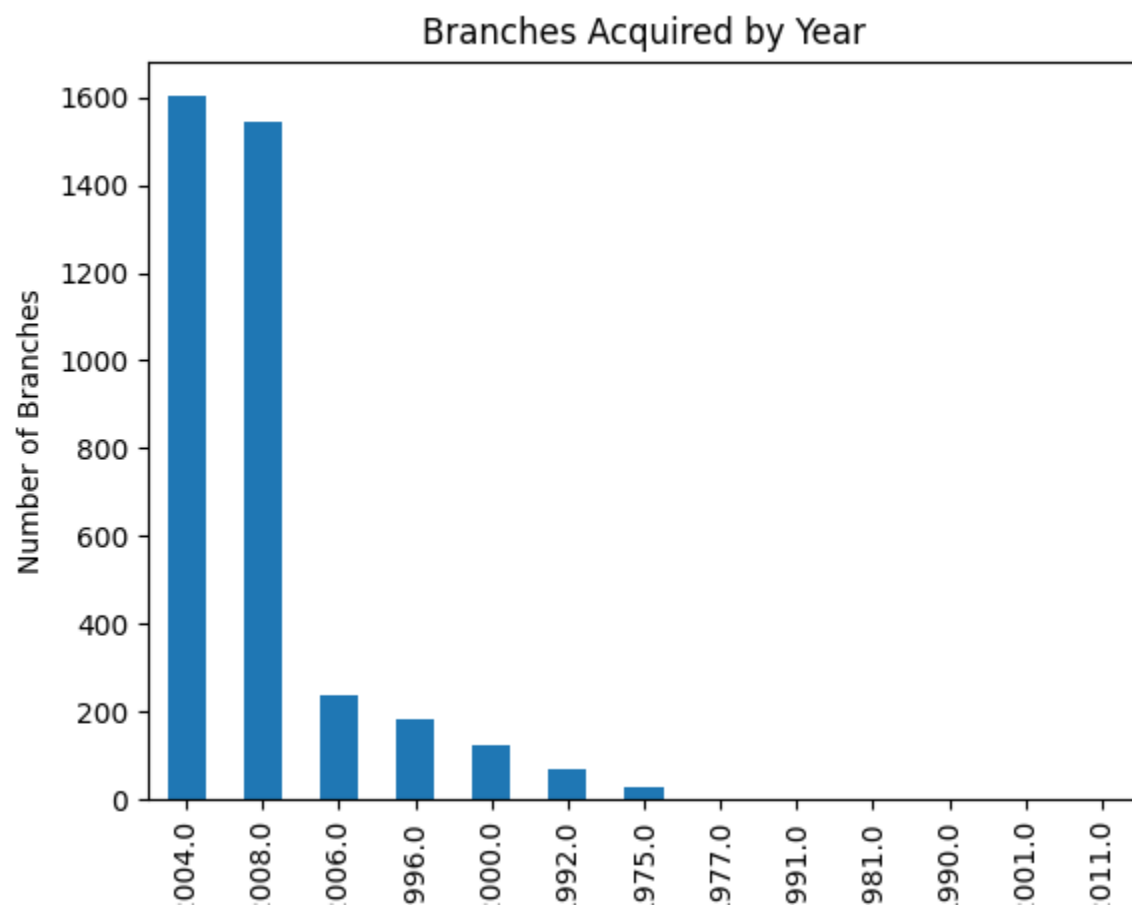
4. Conclusion

4.1 Key Insights:

1. Significant growth in deposits from 2015 to 2016
2. Top 10 branches account for a large portion of total deposits
3. 136 branches have zero deposits across all years
4. Most acquisitions occurred in 2004

4.2 Recommendations

1. Investigate and address potential data quality issues
2. Analyze and replicate best practices from top-performing branches
3. Consider consolidating or closing underperforming branches
4. Develop a strategic expansion plan based on acquisition trends and market analysis

4.3 Summary of findings

This analysis of Chase Bank's branch performance revealed several key trends and insights. Firstly, total deposits have consistently increased over the years, with a significant jump from 2015 to 2016. The top 10 performing branches account for a substantial portion of total deposits, with the main office being the highest performer. However, 136 branches have zero deposits across all years, indicating potential closure, merger, or incomplete data.

The analysis also revealed trends in branch establishment and acquisition. The earliest establishment date is 1824, and the latest is 2016. Most acquisitions occurred in 2004, with a total of 3798 acquisitions.

4.4 Implications for business strategies

1. Optimize Branch Operations: Identify best practices from top-performing branches and replicate them across other branches to improve overall performance.
2. Strategic Expansion: Develop a targeted expansion plan based on acquisition trends and market analysis to increase market share and revenue.
3. Data-Driven Decision Making: Improve data quality and leverage data insights to inform business strategies and optimize branch operations.
4. Performance Management: Establish clear performance metrics and benchmarks to monitor and improve branch performance.
5. Consolidation and Closure: Consider consolidating or closing underperforming branches to optimize resources and improve overall efficiency.
6. Growth through Acquisitions: Continue to pursue strategic acquisitions to expand market share and increase deposits.
7. Market Analysis: Conduct regular market analysis to identify trends and opportunities for growth.

4.5 Future Research Directions

1. Branch Typology: Develop a typology of branches based on performance, demographics, and market characteristics to inform targeted strategies.
2. Market Segmentation: Conduct market segmentation analysis to identify high-value customer segments and tailor strategies to meet their needs.
3. Acquisition Strategy Evaluation: Evaluate the effectiveness of past acquisitions and develop a framework for future acquisitions.
4. Data Analytics: Explore advanced data analytics techniques, such as predictive modeling and machine learning, to improve branch performance forecasting.

5. Customer Behavior Analysis: Analyze customer behavior and preferences to inform product development and marketing strategies.

6. Competitive Analysis: Conduct a comprehensive competitive analysis to identify market gaps and opportunities.

7. Geographic Expansion: Explore opportunities for geographic expansion into new markets and regions.

8. Digital Transformation: Investigate the impact of digital transformation on branch operations and develop strategies to leverage technology for improved performance.