India Banks Reviews

An analysis of customer experiences and feedback on Indian Bank's services.

Monitor customer reviews and ratings to enhance customer service, identify areas for improvement, and ensure a better banking experience.

Data Set Used for Analysis

Columns Names:

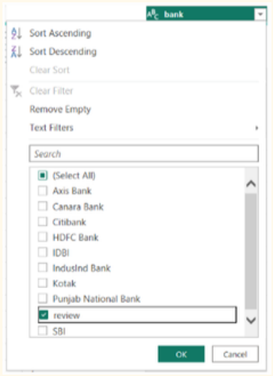
Author, Date, Address, Bank, Rating, Review title by user, Review, Bank image, Rating title by user, Useful count

COUNT 10 COLUMNS	
7 text	3 numerical
3000 Rows	

Data Issues Encountered & Recommendations

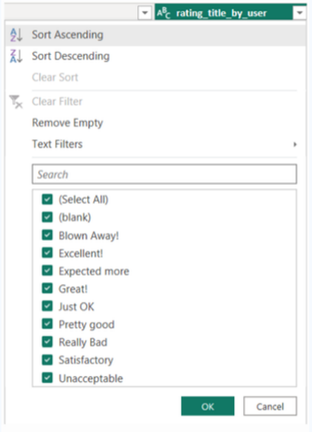
1- Handling of Invalid Entries

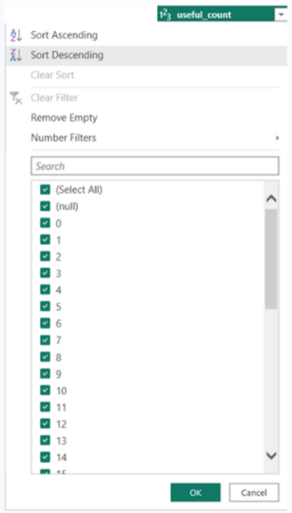
Entries in the **bank** column that contained the word "**review**" were considered illogical and were replaced with "**Unknown Bank**", to avoid misclassification and maintain data integrity.



Blank entries in **Rating Title by User** were replaced with "**Untitled Review**", to ensure every review has a descriptive title for analysis and visualization consistency.

Null values in **Useful Count** were replaced with **0**, since a missing usefulness score reasonably implies that the review has not received any usefulness votes.



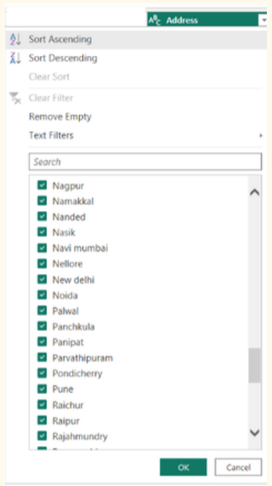


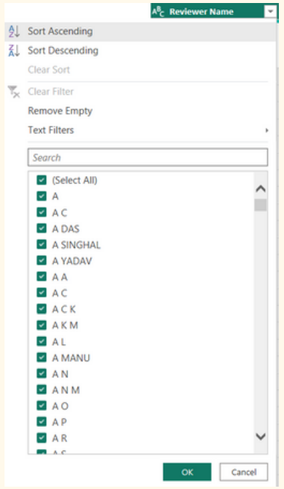
Data Issues Encountered & Recommendations

2- Author & Address Columns Standardization

Trimmed extra white spaces

Capitalized each word





3- Dataset Column Standardization

The dataset headers were standardized by capitalizing all column names and replacing underscores (_) with spaces for improved readability.

Additionally, the Author column was renamed to Reviewer to more accurately reflect the role of the person providing feedback in the dataset, and the Date column is now labeled Review Date to clearly indicate that it represents the date the review was submitted.

Data Issues Encountered & Recommendations

4- Date Format Transformation

The date column was reformatted from **text-based format** (e.g., 21-Mar-20) to a consistent **numeric format** (e.g.,21/03/2020) ensuring uniformity and enabling accurate sorting, filtering, and time-series analysis.

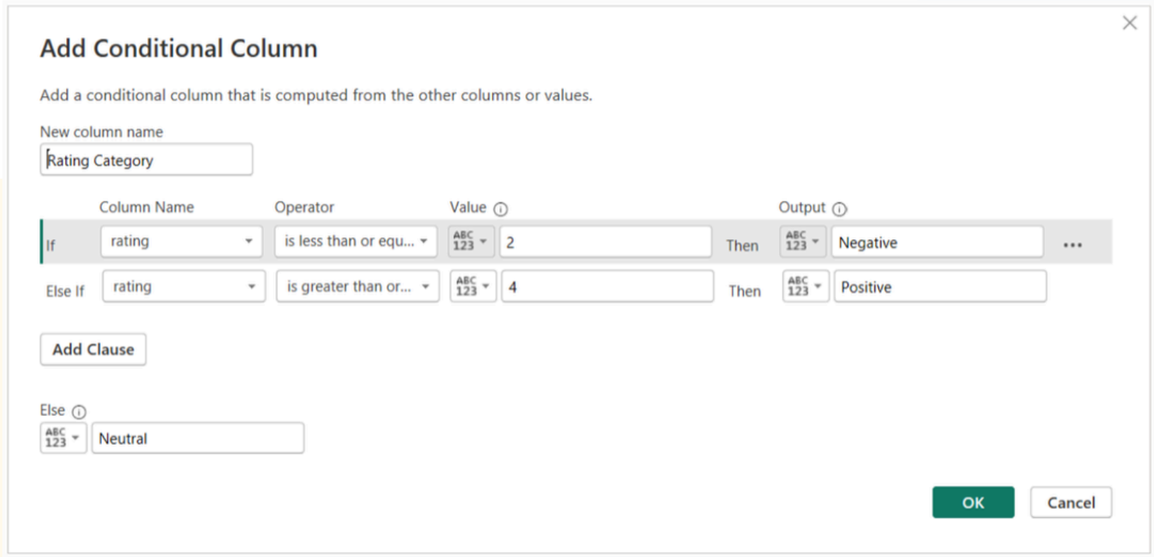
date	
Valid	100%
Error	0%
Empty	0%
Mar 21, 2020	
Mar 20, 2020	
Mar 20, 2020	
Mar 20, 2020	
Mar 20, 2020	
Mar 20, 2020	
Mar 20, 2020	
Mar 20, 2020	
Mar 20, 2020	
Mar 19, 2020	
Mar 19, 2020	

Created Columns for Analysis

1- Rating Category Based on Rate Column

Groups customer feedback into predefined categories for easier analysis.

Positive (4–5),
Neutral (3),
Negative (1–2)

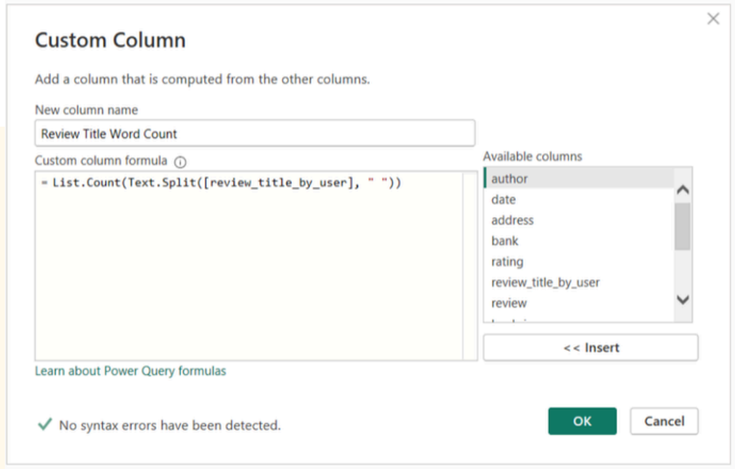


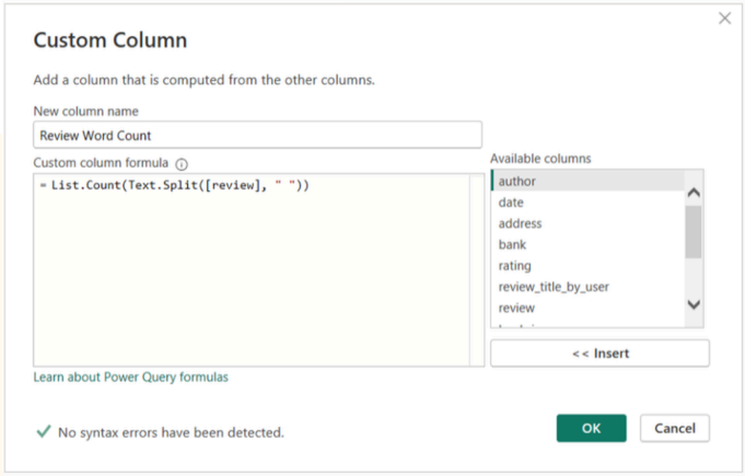
Created Columns for Analysis

2- Review Word Count & Review Title Word Count

Measures the total number of words in the review text and the review title. This provides insight into how detailed or concise customer feedback is.

Useful in Customer Review Insights to analyze patterns — for example, longer reviews may indicate more detailed experiences.



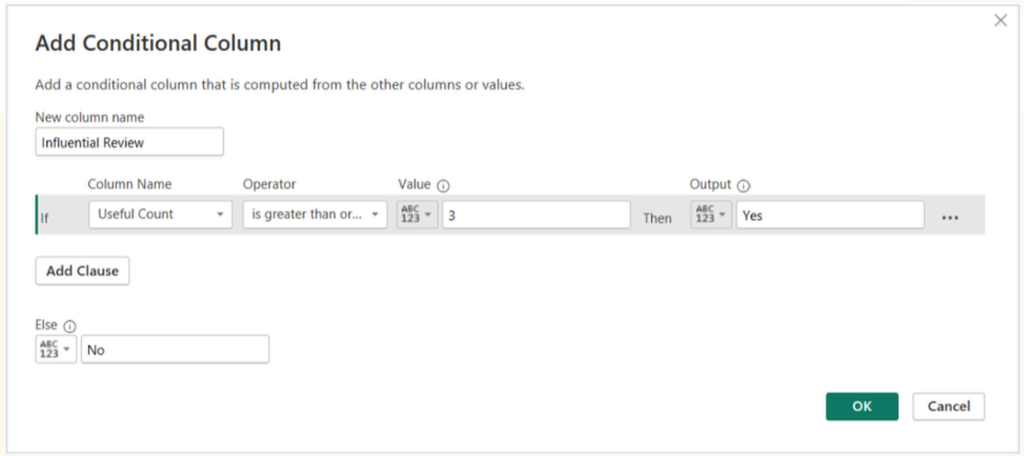


Created Columns for Analysis

3- Influential Review Based on Useful Count

An Influential Review is a binary indicator based on the Useful Count column.

- Yes → the review's Useful Count is above the 75th percentile of all reviews.
- No → the review's Useful Count is at or below this threshold.



Created Columns for Analysis

4- City Column Based on Address

With the purpose of unifying all entries as valid cities in India. To ensure accuracy, an AI tool was used to check all unique entries from the Address column

The tool helped categorize entries into: Valid Cities, Districts, States & Variants

Category	Recommended Action (Unify as Indian Cities)
Valid Cities	Keep all as valid city entries.
Districts / Localities / Suburbs (not cities, but administrative regions)	Replace with nearest major city
States (not cities, but Union Territories)	Replace with the capital
Variants (Cities Changed its name or duplicates)	Keep only the official/modern city name for consistency

Created Columns for Analysis

4- City Column Based on Address

Non-city entries were mapped to their respective cities, and variant names were unified.

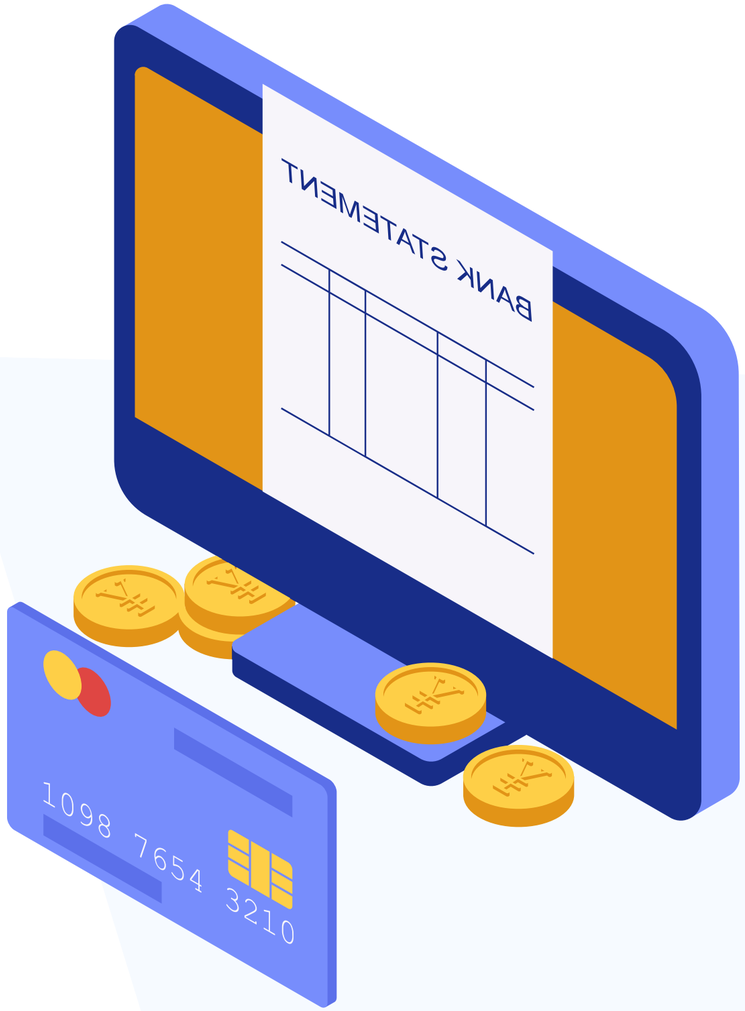
This process ensures consistency and reliability in the City column, enabling accurate location-based analysis and insights.

Replace with nearest major city	Replace with the capital	Keep only the official/modern city name for consistency
1. Hooghly → Kolkata 2. Mahabubnagar → Hyderabad 3. Mahbubnagar → Hyderabad 4. Medak → Hyderabad 5. Rangareddy → Hyderabad 6. Ernakulam → Kochi 7. Bhalswa Jahangir Pur → Delhi 8. Karawal Nagar → Delhi 9. Sultan Pur Majra → Delhi 10. Nangloi Jat → Delhi 11. North Dumdum → Kolkata 12. South Dumdum → Kolkata 13. Rajpur Sonarpur → Kolkata 14. Raurkela Industrial Township → Rourkela	1. Goa -> Panaji	1. New Delhi → Delhi 2. Allahabad → Prayagraj 3. Pondicherry → Puducherry 4. Gurgaon → Gurugram 5. Tuticorin → Thoothukudi 6. Gulbarga → Kalaburagi 7. Alleppey → Alappuzha 8. Hubli → Hubballi 9. Mangalore → Mangaluru 10. Nasik → Nashik 11. Vijayanagaram → Vizianagaram 12. Hubli–Dharwad → Hubli-Dharwad 13. Hubliâ€“Dharwad → Hubli-Dharwad 14. Trivandrum → Thiruvananthapuram 15. Disabled Keonjhar → Keonjhar

Data Model Overview (Star Schema)

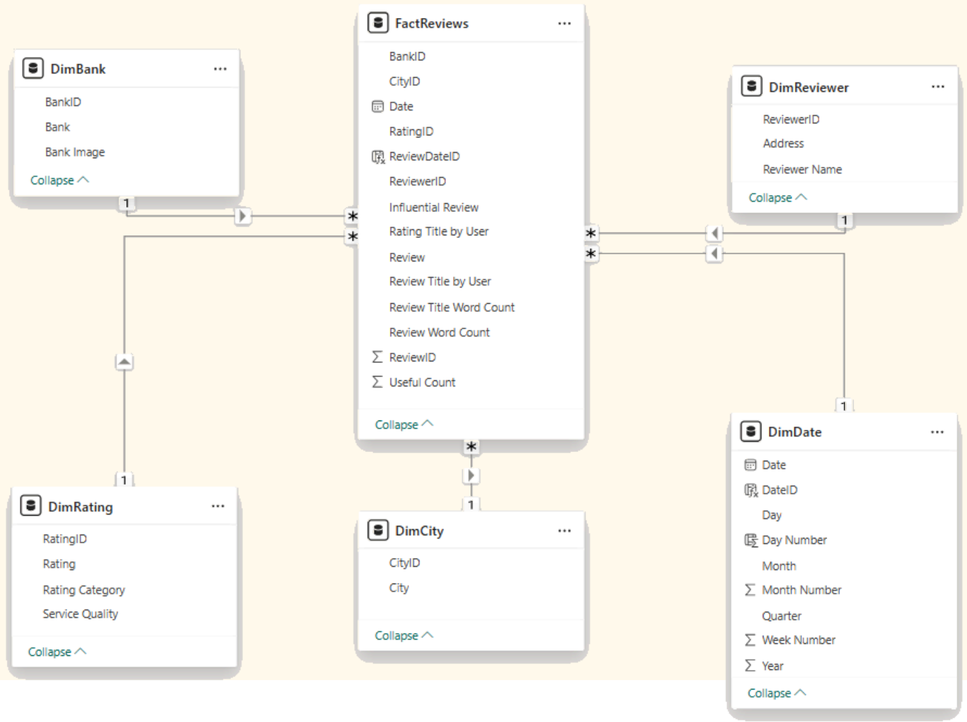
The star schema design for the dataset

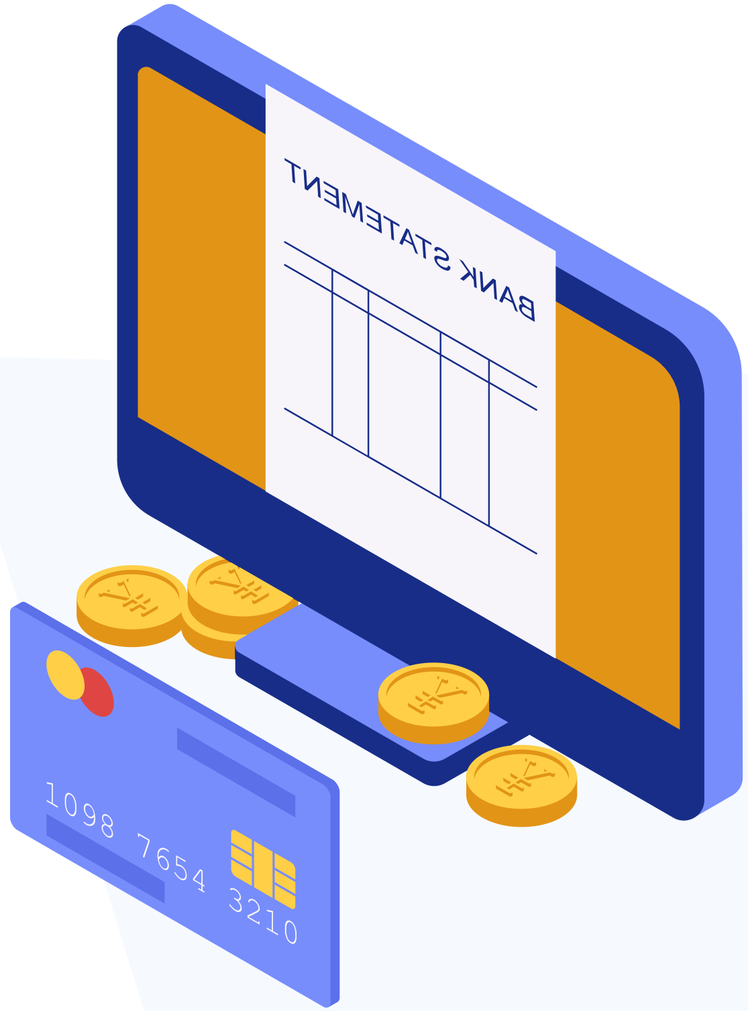
Table Name	Columns	Type	Details
DimDate	DateID Date	Dim	Fact Create
DimBank	BankID Bank Bank Image	Dim	Fact Foreign Key: BankID
DimReviewer	ReviewerID Reviewer Address	Dim	Fact Foreign Key: ReviewerID
DimCity	CityID City	Dim	Fact Foreign Key: CityID
DimRating	RatingID Rating Rating Categories Service Quality	Dim	Fact Foreign Key: RatingID
FactReviews	ReviewID Rating Title by User Review Title by User Review ReviewDate ReviewDateID Useful Count Review Word Count Review Title Word Count Influential Review	Fact	Foreign Keys: BankID, ReviewerID, CityID, RatingID



Data Model Overview (Star Schema)

The relationship between the fact table and dimension tables.





Dashboard Plan

The dashboard will have five pages to give a complete view of bank performance and customer feedback:

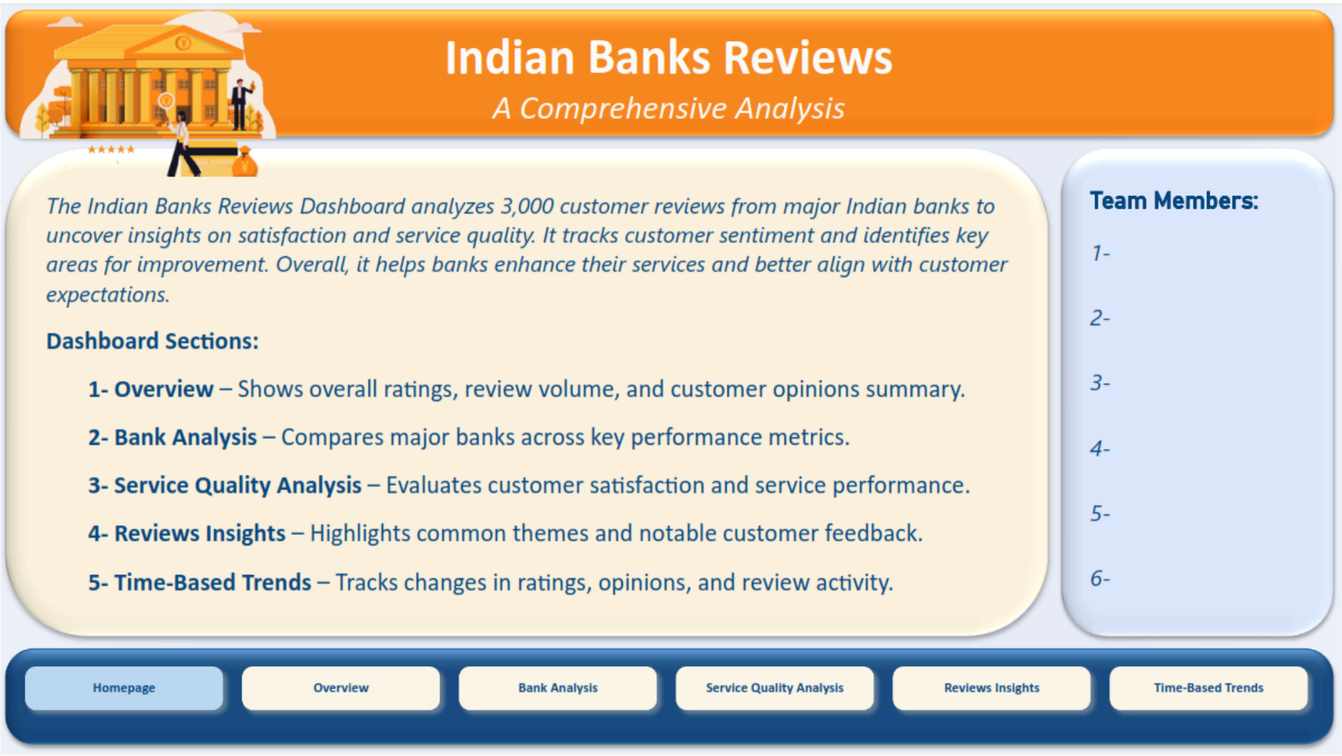
- **Overview:** Quick snapshot of overall metrics and KPIs.
- **Bank Analysis:** Compare banks across key performance indicators.
- **Service Quality Analysis:** Segment banks by service excellence and satisfaction.
- **Reviews Insights:** Analyze customer feedback and influential reviewers.
- **Time-Based Trends:** Track changes in ratings, service quality, and reviews over time.

This layout ensures clarity, actionable insights, and both high-level and detailed analysis in one dashboard.

Dashboard Pages

0- Home Page

Objective: Offer a concise overview of the dashboard, highlighting its sections, a brief summary, a user-friendly page navigator, and the contributing team members.



Dashboard Pages Overview

1. Overview (Summary & Performance)

Objective: Provide a high-level snapshot of overall bank performance, satisfaction, and engagement across the entire dataset.

Key Metrics:

- Average Ratings,
- Total reviews,
- Average Words Count,
- Number of Cities
- Customer Feedback Breakdown

Visuals:

- KPI cards,
- Top 10 Banks by Average Rating,
- Reviews Split by Service Quality per Bank,
- Ratings vs. Reviews by Bank,
- Ratings vs. Reviews by Bank,
- Reviews Density Map

Slicers: City, Date Range.

Dashboard Pages Overview

1. Overview (Summary & Performance)

Thank You