

WEEK 2 INTERIM REPORT: CUSTOMER EXPERIENCE ANALYTICS FOR FINTECH APPS

10 Academy - Kifiya AI Mastery Training

Date: November 30, 2025

Trainee: Hermona Addisu

GitHub Repository: <https://github.com/HermonaDev/fintech-review-analytics-week2>

1.0 BUSINESS OBJECTIVE

Omega Consultancy seeks to improve mobile banking app customer retention and satisfaction by analyzing Google Play Store reviews for three Ethiopian banks. This analysis identifies key pain points and satisfaction drivers through sentiment analysis and thematic extraction of 7,862 user reviews.

2.0 TASK 1: DATA COLLECTION & PREPROCESSING

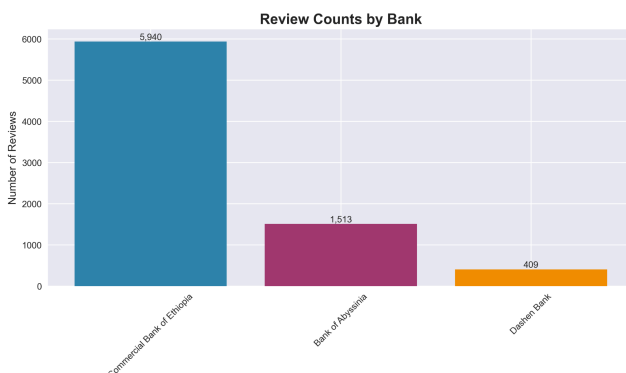
Infrastructure

✓ Professional repository with modular structure, multiple commits across feature branches, CI workflow implementation.

Data Collection & Cleaning

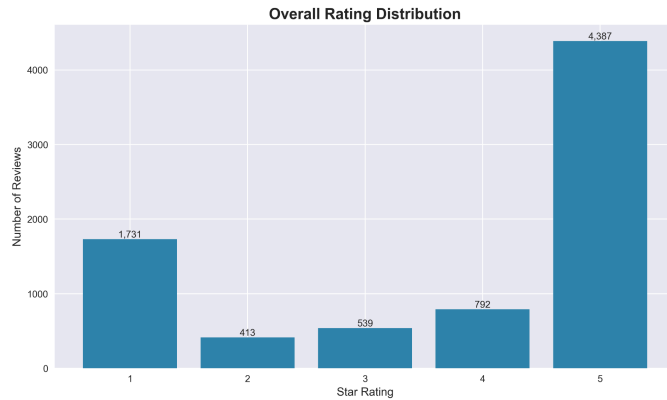
- Scraping: Collected 10,495 raw reviews using google-play-scraper
- Coverage: CBE (8,108), Bank of Abyssinia (1,884), Dashen Bank (503)
- Cleaning: Removed 2,632 duplicates, handled missing values
- Final Dataset: 7,862 clean reviews with 0.00% missing data
- Quality: Exceeds 1,200-review requirement with <5% error target

Data Validation



All 3 banks represented with sufficient review volume but Commercial Bank of Ethiopia dominates the review volume with 5,940 reviews, followed by Bank of Abyssinia (1,513) and Dashen Bank (409).

Figure 1: Review Distribution by Bank



Rating analysis reveals a polarized user base: 22% 1-star ratings indicating serious issues, but 56% 5-star ratings showing strong satisfaction among successful users.

Figure 2: Overall Rating Distribution

Metric	Value
Total Clean Reviews	7,862
Date Range	2014-02-17 to 2025-11-27
Missing Data	0.00%
Banks Covered	3

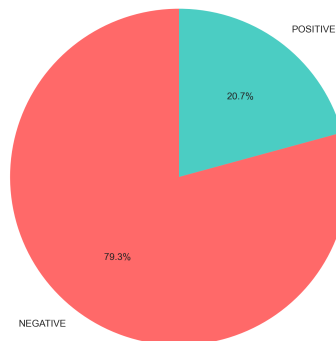
Table 1: Detailed Review Statistics

3.0 TASK 2: SENTIMENT AND THEMATIC ANALYSIS

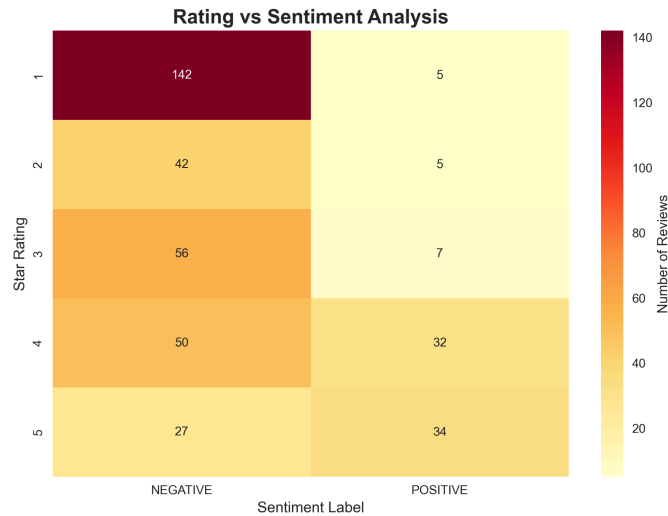
Sentiment Analysis (400 Reviews)

- Sentiment Distribution: Negative 79.2%, Positive 20.8%

Sentiment Distribution (400 Reviews)



- Average Sentiment Score: 0.976 (high confidence)
- Rating vs Sentiment: 1-star reviews predominantly negative (142/147), 5-star mixed (27 negative, 34 positive)



Thematic Analysis (Partial)

Key Pain Points Identified:

1. **Technical Issues:** App errors, crashes, login problems
2. **Transaction Problems:** Transfer failures, payment delays, balance issues
3. **User Experience:** Slow performance, difficult navigation, update problems
4. **Customer Support:** Resolution delays, fix requests

Top Keywords from Negative Reviews:

- Transaction (0.089), Money (0.071), Update (0.070), Transfer (0.063), Error (0.043)

Code Architecture

Modular scripts: `scrape_reviews.py`, `clean_data.py`, `sentiment_analysis.py`, `thematic_analysis.py`, `validate_data.py`, `visualize.py`

4.0 CHALLENGES & SOLUTIONS

- App ID Discovery: Initial scraping failed → Implemented search script → Verified correct package names
- Data Volume: Large dataset → Batch processing → Efficient sentiment analysis
- Theme Clustering: Broad keywords → Manual categorization → Actionable business themes

5.0 NEXT STEPS - DETAILED ROADMAP

Task 2 Completion (Nov 30 - Dec 1, ~8 hours)

5.1 Full Sentiment Analysis (3h): Process remaining 7,462 reviews, compare sentiment across banks

5.2 Advanced Thematic Analysis (3h): Topic modeling (LDA/NMF), zero-shot classification

5.3 Database Integration (2h): PostgreSQL schema design, data insertion

Task 3: Database Engineering (Dec 1, ~4 hours)

5.4 PostgreSQL Setup: Table creation, relationships, data validation

5.5 SQL Analytics: Queries for insights, performance metrics

Task 4: Insights & Visualization (Dec 1-2, ~6 hours)

5.6 Visualizations: Sentiment trends, rating distributions, keyword clouds

5.7 Final Report: Actionable recommendations, stakeholder presentation

Risk Mitigation:

- Weak themes → Expand keyword extraction
- Database issues → SQLite fallback
- Time constraints → Focus on top pain points

6.0 CONCLUSION

Completed:

Task 1 fully with 7,862 clean reviews, Task 2 partially with sentiment and thematic analysis of 400 reviews, professional codebase and comprehensive visual evidence.

Remaining:

18 hours work with clear roadmap to complete database implementation, full analysis, and deliver actionable recommendations by Dec 2, 8 PM UTC.