Consultation Report

Table of Contents

1. Introduction

This consultation report provides an in-depth analysis of user data for a telecommunications project aimed at understanding user behavior across multiple dimensions such as device usage, geographic distribution, and demographics. The insights gathered from this analysis will help inform decisions on enhancing customer experience, marketing strategies, and operational efficiency.

2. Project Description

The project aims to conduct an analysis on the distribution of users based on various factors including their geographic location, device type, gender, and age segments. The analysis will further extend to visualizing hourly phone call trends and user mapping to identify significant usage patterns and address customer needs.

3. Problem Statement

The primary problem addressed in this project is a lack of detailed knowledge regarding the distribution and behavior of users within a telecom network. Specifically, the client requires a clear understanding of:

- User distribution across states, age segments, and gender.

- Device preference and its correlation with demographic variables.

- Hourly trends in phone usage.

4. Problem Analysis (Strategy for the Problem Statement Analysis)

The strategy for solving the problem involved multiple steps of data analysis, from handling null values and duplicates to producing detailed visualizations. The key objectives were:

- Clean the dataset to ensure no duplication or missing values interfere with accurate analysis.

- Derive insights by analyzing the distribution of users across different segments (state, age, gender, and phone brand).

- Use interactive maps and charts to understand the geographic distribution of users and their activity trends.

5. Sources of Data

The data used in this project was sourced from the client-provided CSV files. The following columns were included:

- device_id: Unique identifier for each user.

- timestamp: Timestamps of user activities for hourly analysis.

- latitude and longitude: Used for mapping users geographically.

- state: Represents the geographic location of each user.

- phone_brand: Indicates the brand of the user's phone.

- gender: User's gender for demographic analysis.

- age: User's age, used to create different age segments.

6. Summary of Data Mining

During data mining, a few challenges were encountered:

- Null Values: Columns like `latitude`, `longitude`, and `state` contained missing values. Missing values we handled using 'Geopy' python library. Device_id is handled through the interpolate method to forward fill.

- Duplicate Entries: Multiple records for the same `device_id` were present. We have taken the unique Values user device_Id, during plotting the graphs.

- Merging files : "events_data, Phone_brand, and age_gender" Merging all the three files was one of the big challenges faces during the processing of data. We merged using left outer join having device_id in common. After merging phone_brands, device_model and age, gender and age_group were having missing values which is handled by taking the mode. Further the remaing null values was handled by dropping the null records using "dropna"

- Inconsistent data: Chinese characters were present in phone_brand and device_model, we translated from Chinese to English and we handled using the map function.

-Wrong Values : wrong values were available in 'longitude' and 'latitude' that we handled using Geopy library comparing the expected and then actual values. Expected is the data coming from Geopy and actual is the existing data in longitude and latitude.

-Inconsistent Format : device_id was in exponential notation that we converted to common format, that is the string datatype.

In summary, the data analysis involved deriving insights on user distribution across different states, age segments, and device types, alongside mapping and temporal analysis of phone call trends.

7. Proposed Solution for Customers (Describe Your Analysis in Detail)

User Distribution Across States:

- Maharashtra, Karnataka, and West Bengal have the highest concentration of users, with Maharashtra leading at 19%. These states represent a large portion of the user base, suggesting higher demand for mobile services.

- Recommendation: To address this high demand, consider allocating additional network resources to these regions. Furthermore, running targeted marketing campaigns could strengthen user engagement in these high-demand areas and improve brand loyalty.

Phone Brand Usage:

- Xiaomi is the most used brand (63% of users), followed by Huawei (25%) and Samsung (10%). This heavy concentration on specific brands indicates strong brand loyalty and preference.

- Recommendation: Leverage this insight by partnering with top brands (Xiaomi, Huawei, Samsung) to create exclusive service bundles, device-specific promotions, or discounts. This could enhance customer loyalty and potentially increase the market share through these brand partnerships.

User Demographics:

- Gender Distribution: Male users dominate the user base at 84.4%, with female users making up only 15.6%. This significant gender disparity suggests potential for engagement growth among female users.

- Recommendation: Create gender-specific campaigns targeting female users. Examples include women-oriented promotions, campaigns focused on mobile safety features, or initiatives that resonate more with female audiences to encourage higher engagement.

- Age Segmentation: Users between 19-30 years represent the largest age segment, comprising 40% of the user base. Younger users are more likely to switch to services that meet their specific interests, such as high-speed data for streaming or gaming.

- **Recommendation:** Develop special plans for this segment, such as gaming data packs or exclusive OTT content partnerships, to appeal to the entertainment and connectivity needs of younger users.

Correlation Analysis:

- **Device Preference by Age:** Younger users (19-30) predominantly use mid-tier brands like Xiaomi and Samsung, while older users (45+) show a preference for premium brands like Apple.

- **Recommendation:** Create segmented marketing campaigns tailored to each age group, such as promoting affordable data packs for younger users and premium service bundles for older users who may prioritize reliability and premium customer support.

- **Device Preference by Gender:** Male users have a higher preference for Xiaomi devices, while female users are more inclined towards Samsung.

- **Recommendation:** Tailored campaigns could emphasize Xiaomi's cost-effective and performance features to male users, while Samsung's camera quality and design could be highlighted in campaigns targeted toward female users.

Hourly Phone Call Trends:

- **Mobile** usage **drops** gradually **after 12 AM**. **Least** usage during **3-4 AM**. Mobile usage **sharply increase** after **4 AM** and reaches the **peak** at **10 AM**. **Steady** usage in the **Noon** time and **drops** after **10 PM**.

- Recommendation: Adjust network capacity to accommodate peak hours and potentially introduce limited-time offers or promotional events during these high-traffic periods to boost engagement.

Geographic Mapping:

- The geographic distribution map highlights high and low user concentrations across various states, with Maharashtra and neighboring states showing the densest user clusters.

- **Recommendation:** Use this geographic data to expand network infrastructure in high-density areas and improve service in underrepresented regions. In low-density regions, explore partnerships with local entities and promotional events to boost user acquisition and service penetration.

Customized Plans Based on Analysis:

- **Family Plans:** In states with high user concentrations, family plans that bundle services under one billing account could encourage higher retention and customer loyalty.

- **Youth Plans:** Create targeted youth plans with high-speed data, entertainment-focused content, and partnerships with gaming platforms to meet the preferences of the 19-30 age group.

- **Senior Packages:** For older users, consider bundles that include premium customer service and health or wellness app access to cater to their preferences and increase satisfaction.

8. Tools

DS Tools

- **pandas** - for data manipulation and analysis

- **numpy** - for numerical operations

- **matplotlib.pyplot** - for data visualization

- **seaborn** - for enhanced data visualizations

- **geopandas** - for working with geospatial data

- **geopy** - specifically, Nominatim and ArcGIS geocoders for converting addresses into geographic coordinates and vice versa

-time - for handling time-related functions

- warnings - to manage and suppress runtime warnings

Web UI Tools

- PHP

- JavaScript

9. Conclusion

The insights gained from this analysis provide valuable information on user behaviour, device preferences, and geographic trends. The findings support targeted marketing, improved service delivery, and better resource allocation in high-demand regions. Recommendations include:

-Targeted Marketing Campaigns: Focus on underrepresented demographics, particularly female users and older age segments, where there is growth potential.

-Network Optimization: Prioritize network improvements during peak hours (6 PM - 9 PM) to ensure service quality.

-State-Specific Strategies: Utilize state-level data to craft state-specific strategies in high-demand regions such as Maharashtra, Karnataka, and Tamil Nadu.

-Device-Specific Offers: Tailor promotions to suit younger users' preference for mid-tier Android brands and older users' preference for premium devices.

-Resource Allocation: Leverage geographic mapping data to optimize resource allocation and enhance service in low-density regions through local partnerships.

10. Key Takeaways for Stakeholders #(Added)

- **Marketing Focus**:
  - Gender and Age Group Targeting: The gender imbalance (84.4% male) and the concentration of younger users (19-30) highlight potential growth opportunities. Campaigns should target underrepresented groups, specifically female users and older age segments, to maximize untapped potential. This approach not only broadens the user base but also enhances brand inclusivity.
  - State-Specific Strategies: The high concentration of users in Maharashtra, Karnataka, and Tamil Nadu suggests that state-focused marketing efforts and resources will be more effective. Investing in campaigns and network improvements in these regions will likely yield better returns and increased user engagement.
- **Operational Efficiency:**
  - Network Optimization During Peak Hours: With peak call times, optimizing network resources during these hours will improve user satisfaction and reduce service complaints, contributing to long-term customer loyalty and retention. This investment will ultimately impact user retention rates positively and support brand reputation.
  - Resource Allocation Based on Geographic Insights: Geographic analysis of user distribution highlights where demand is highest and lowest. Expanding infrastructure in high-density regions will support current users effectively, while in lower-density areas, targeted partnerships and regional campaigns could boost market penetration.
- **Revenue-Driven Custom Plans:**
  - Custom plans such as family, youth, and senior packages based on demographic insights can help capture specific market segments and increase customer lifetime value. These plans can boost subscription rates, encourage cross-generational engagement, and improve user retention across segments, contributing to stable revenue growth.
- Brand Partnerships and Product Customization:

- The concentration of users in brands like Xiaomi, Huawei, and Samsung provides a clear opportunity for partnerships. Device-specific promotions or exclusive service bundles with these brands could enhance brand value and user loyalty, benefiting overall customer satisfaction. Device-based product customization can also help attract users based on their phone brand preference, potentially capturing the high-turnover younger segment.