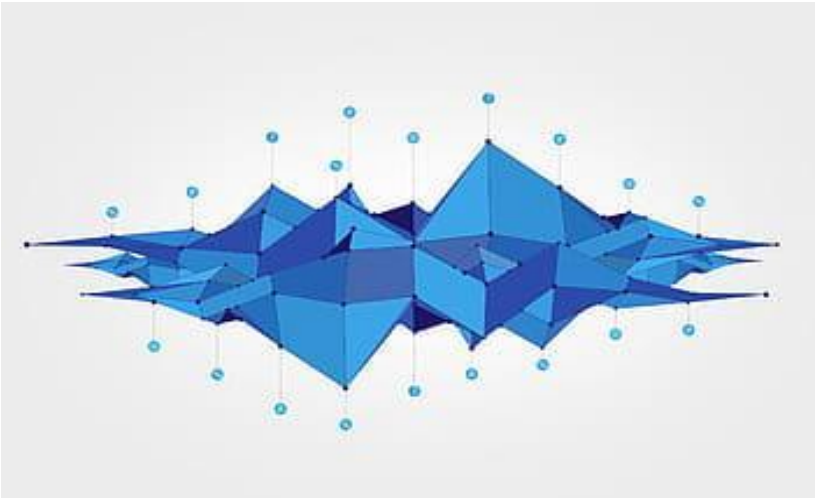
Leveraging Big Data for Targeted Travel Planning Insights

Group 6 :

Anisha Mediratta
Harshit Aggarwal
Ravneet Singh
Sagar Kapoor
Sakshi Arora
Yajur Khurana

Introduction

Welcome to our presentation on "Leveraging Big Data for Targeted Travel Planning Insights." In this session, we explore how data analytics transforms the travel industry, using Zeppelin and HDFS as our guiding tools. Our mission is to extract actionable insights from travel agency data, enabling precision in travel planning and enhancing customer experiences.



Data Collection and Ingestion

- **Data Sources:** Our project compiles diverse data sources, including booking records, customer profiles, and transaction histories, to form a comprehensive dataset.
- **Ingestion Process:** Leveraging the capabilities of Hadoop Distributed File System (HDFS), we efficiently store and manage this vast volume of data. HDFS ensures fault tolerance and scalability, enabling seamless data ingestion.

Data Preprocessing

Importance of Data Preprocessing:

- Data quality is crucial for accurate analysis.
- Preprocessing involves cleaning, transforming, and structuring the data.

Data Cleaning:

- Identify and handle missing values.
- Remove duplicate entries.
- Address outliers and inconsistencies.

Data Transformation:

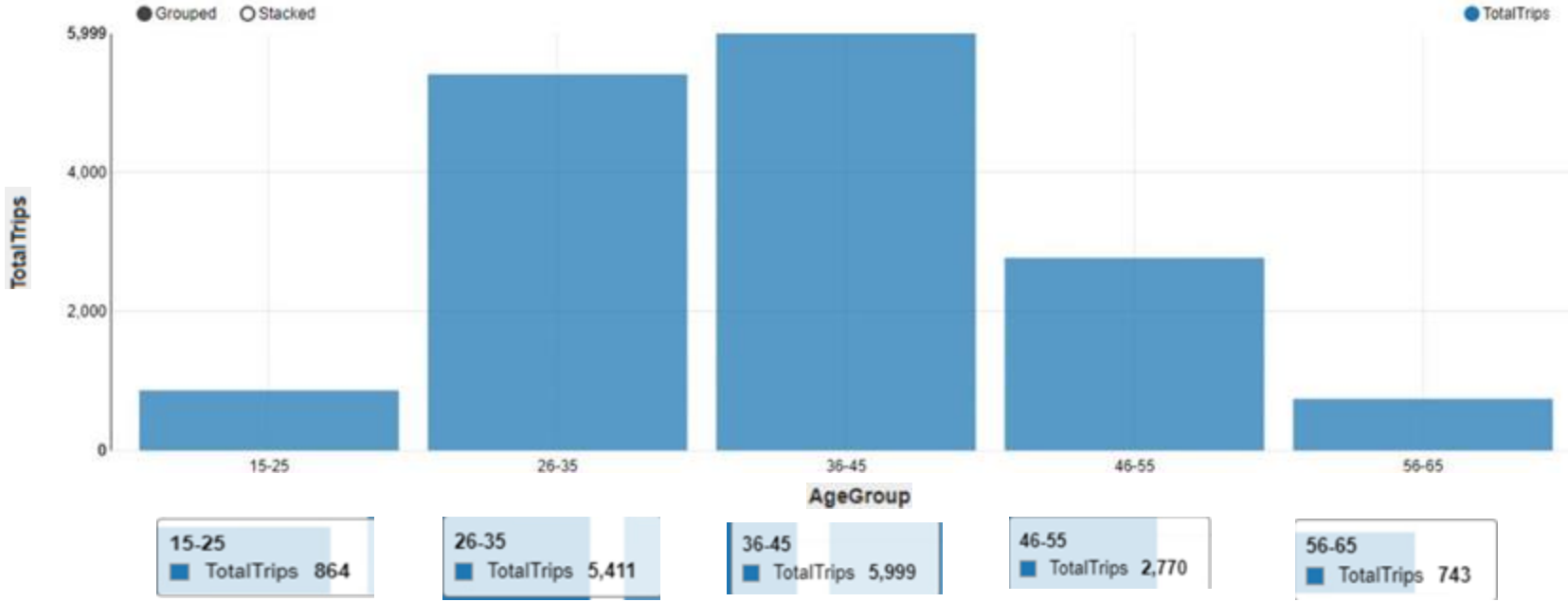
- Convert data into a consistent format.
- Normalize numerical features for meaningful comparisons.
- Encode categorical variables for analysis.

Data Structuring:

- Aggregate data as needed (e.g., grouping by time periods).
- Create new features for enhanced analysis (e.g., travel season, trip type)

Insights and Recommendations

1. **How does the age of customers correlate with the number of trips taken?**

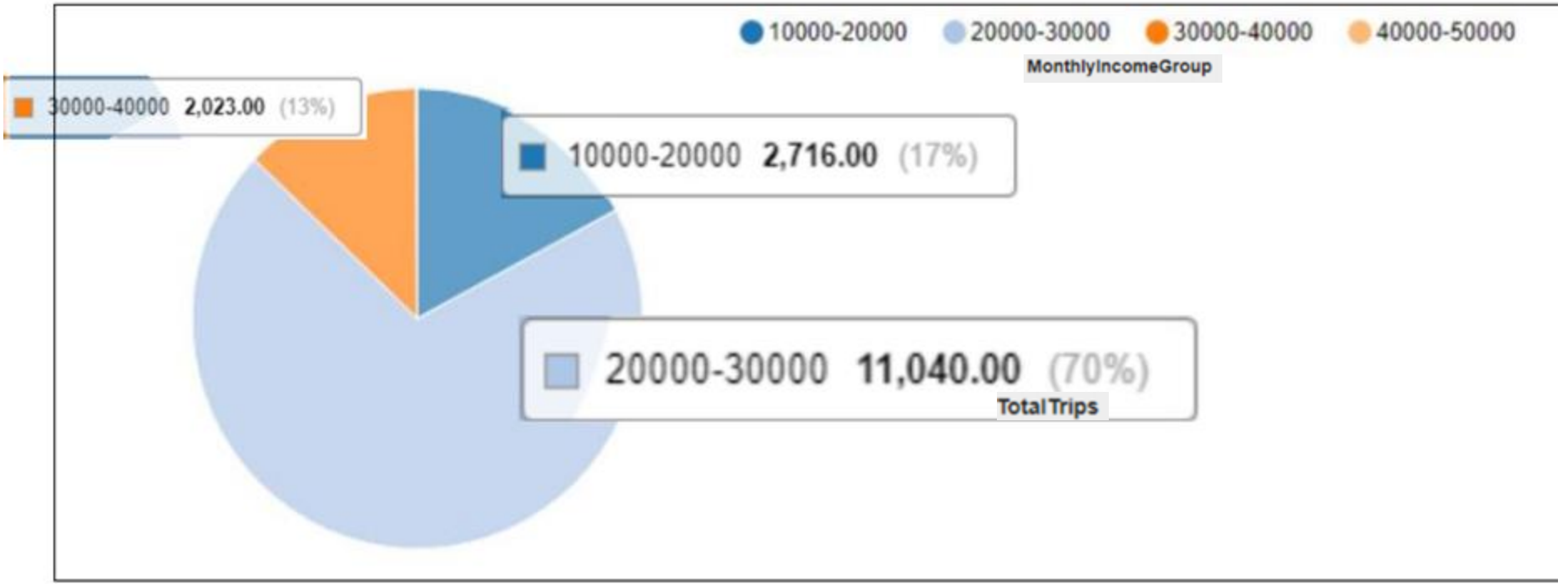


Conclusion :

- The data shows a clear correlation between age and the number of trips taken. The age group of 26 to 45 tends to take the highest number of trips, likely due to a combination of factors including career stability, disposable income, and work life balance. As age increases beyond this range, the number of trips generally starts to decrease, potentially due to retirement, changing priorities, and health considerations.

- This insight could be valuable for targeting marketing efforts, designing travel packages, and tailoring travel experiences to the preferences and lifestyles of different age groups.

2. **How does the salary range correlate with the number of trips taken?**



Conclusion :

- The income group with a monthly salary between 20000 and 30000 takes the highest number of trips, suggesting that a comfortable level of disposable income positively influences travel frequency. As income decreases below this range, the number of trips tends to decrease, likely due to financial constraints. Interestingly, in the highest income range (40000-50000), the number of trips is extremely low, possibly indicating that individuals in this income bracket have other commitments or preferences that limit their travel opportunities.

- This insight can guide travel agencies and businesses to tailor their offerings and marketing strategies based on different income segments.

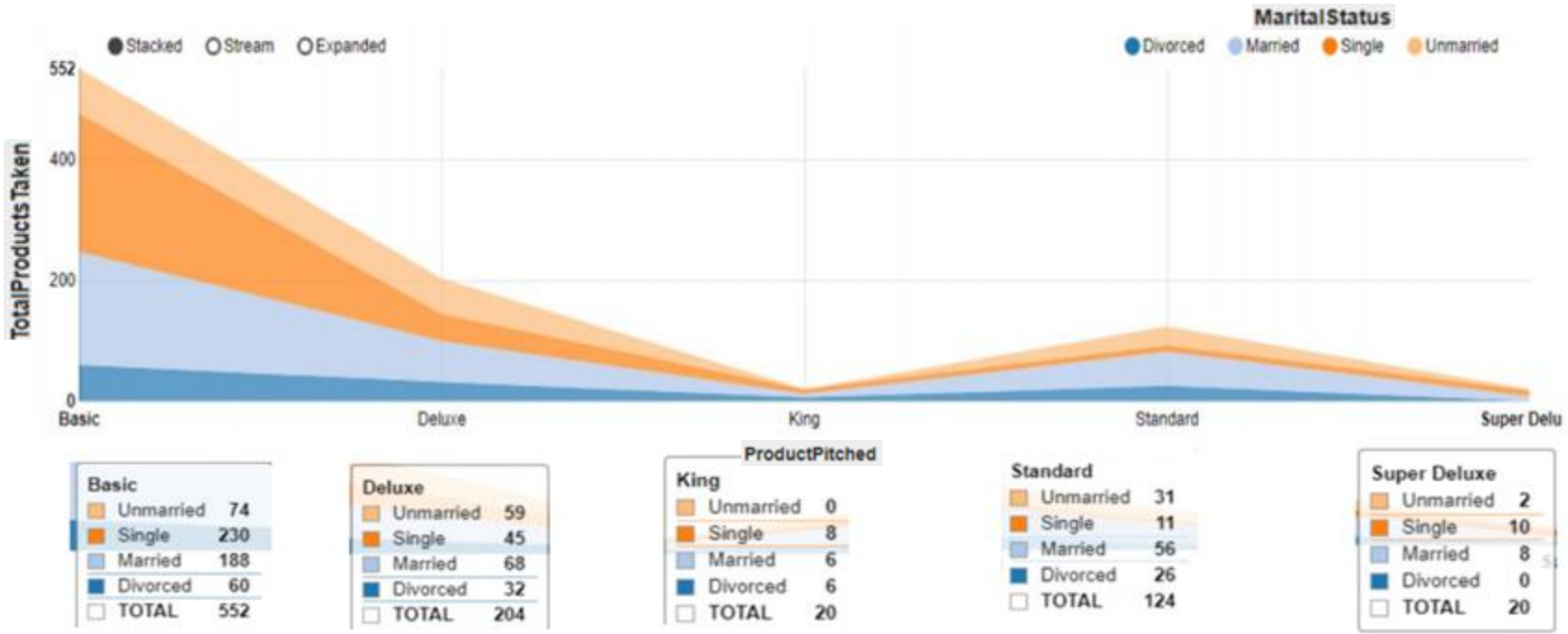
3. **How does the selection of products taken by customers vary based on different city tiers and their preferred property star ratings?**

Conclusion :

The data reveals distinct preferences for property star ratings across different city tiers:

- In City Tier 1, customers are inclined to choose properties with 3-star ratings, followed by 5-star properties. The popularity of 4-star properties is also notable but falls between the other two categories.
- In City Tier 2, customers have a preference for 5-star and 4-star properties, but there's a limited uptake of 3-star properties.
- In City Tier 3, 3-star properties are preferred, similar to City Tier 1. Customers also show a preference for 5-star properties, and 4-star properties are relatively popular as well.

4. What is the distribution of products taken by customers based on their marital status and the type of product pitched to them ?



Conclusion :

- From the data, we can observe that both divorced and married customers show a preference for the "Basic" and "Deluxe" products. These products seem to be attractive to both groups, indicating that they might offer a balance between affordability and desirable features.

- Interestingly, the "King" product seems to have limited uptake, regardless of marital status, with only a small number of customers from both groups selecting it.

- The absence of "Super Deluxe" products taken by divorced customers suggests that this premium option might not align with the preferences or needs of this specific demographic.

SUMMING UP:-

Travel and Age Frequency: There is a definite relationship between age and the number of trips taken. Because of variables such as work stability and disposable cash, the age group of 26 to 45 takes the most travels. Trips diminish with age due to retirement and shifting priorities.

Income and Travel Frequency: The 20000-30000 income bracket travels the most, whereas lower and higher income categories travel less. Those with the highest incomes (40000-50000) travel less, possibly due to other obligations.

Divorced and married customers both choose "Basic" and "Deluxe" travel products, which provide a good combination of pricing and functionality. "King" products are less popular across marital statuses, and divorced clients avoid "Super Deluxe" products.

City Tier Property Star Ratings: Customers favour 3- and 5-star properties in the city, with 4-star properties being somewhat popular for city tier 1 and customers prefer 5-star and 4-star properties, with less interest in 3-star properties for city tier 2.Similar to City Tier 1, 3-star properties are preferred in City Tier 3. 5-star and 4-star hotels are very popular.

These information help travel companies target certain age and income groups, customise product offerings to client preferences, and develop vacation packages that correspond to star rating preferences in different city tiers.

Thank You !

Appendix:-

https://www.kaggle.com/datasets/susant4learning/holiday-package-purchase-prediction