

PRESENTATION

SESSION: 2025-26

PRESENTED BY:- (MEMBERS)

SANIDHYA PRAJAPATI – 202410101150105

VIVEK KUMAR – 202410101150106

B.TECH CS(DS+AI)

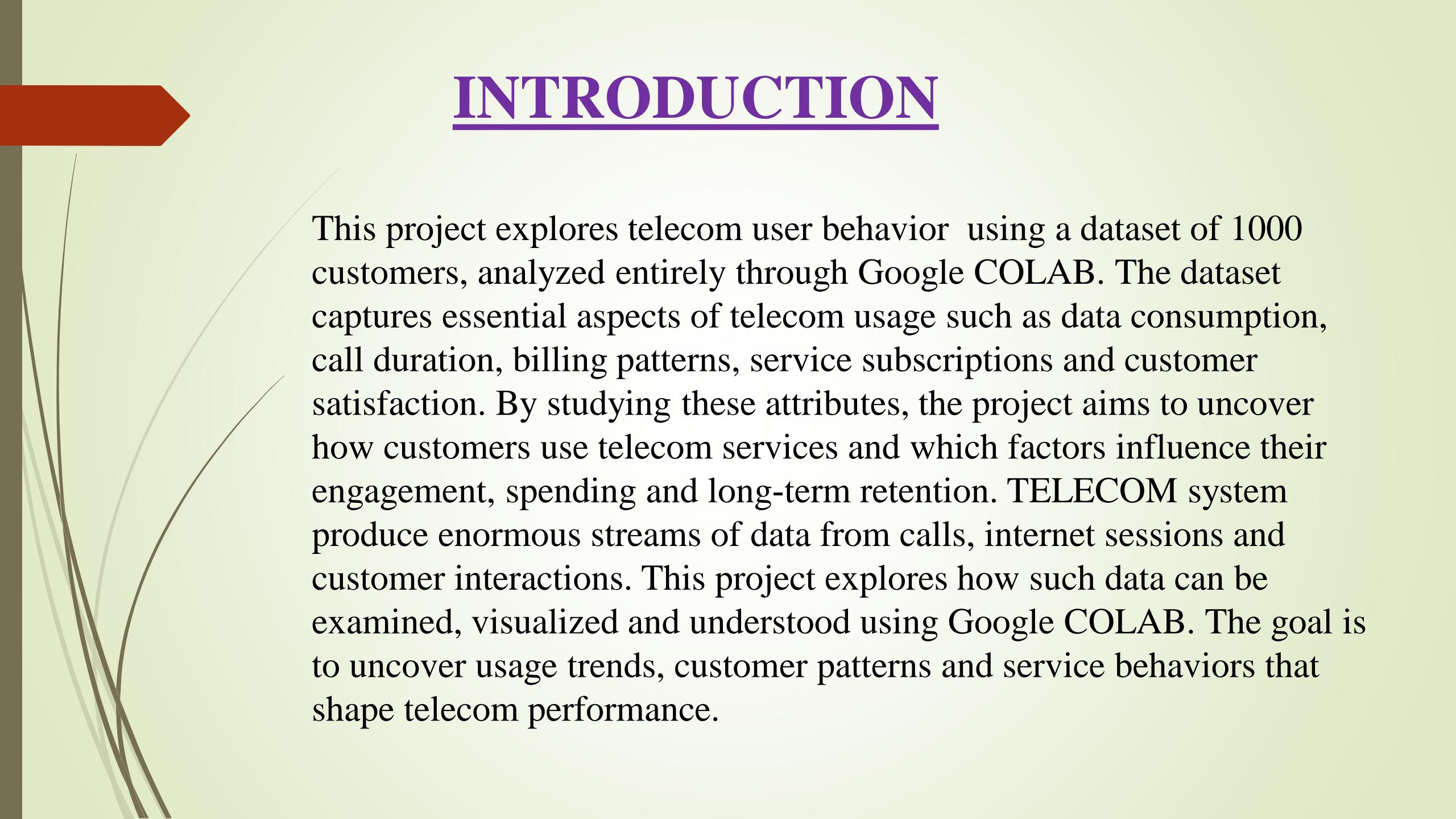
33,34

PRESENTED TO:-

MRS. DEEPIKA TIWARI

FACULTY OF IBM

DATA ANALYTICS AND REPORTING



INTRODUCTION

This project explores telecom user behavior using a dataset of 1000 customers, analyzed entirely through Google COLAB. The dataset captures essential aspects of telecom usage such as data consumption, call duration, billing patterns, service subscriptions and customer satisfaction. By studying these attributes, the project aims to uncover how customers use telecom services and which factors influence their engagement, spending and long-term retention. TELECOM system produce enormous streams of data from calls, internet sessions and customer interactions. This project explores how such data can be examined, visualized and understood using Google COLAB. The goal is to uncover usage trends, customer patterns and service behaviors that shape telecom performance.



PROBLEM STATEMENT

- Telecom companies often struggle to understand customer behavior hidden in massive datasets. Churn patterns remain unclear, high-value users are hard to identify and spending behavior shifts quietly. This project addresses these challenges by examining real telecom attributes and extracting insights that help predict churn, improve service quality and enhance customer experience.



DATASET OVERVIEW

The dataset includes customer demographics such as age, gender and region, along with key service attributes like internet type, plan category and multiple-line subscriptions. It also contains numerical metrics including data usage in gigabytes, call duration in minutes, monthly bill amounts, network ratings and churn status. Together, these attributes create a detailed map of how telecom services are consumed and how customers interact with the network.

- Contains 1000+ customer entries
- Includes demographic fields like gender, age and region
- Covers usage features: DATA_USAGE_GB and CALL_DURATION_MIN
- Billing information captured through MONTHLYCHARGES and TOTALCHARGES
- Service attributes like INTERNETSERVICE, MULTIPLELINES and PLANTYPE
- Satisfaction measured using NETWORK_RATING
- Churn column indicates whether a customer left the service



DATA CHARACTERISTICS

- Within the dataset, patterns emerge across different age groups, regions and service plans. Customers vary widely in their data usage, with some relying heavily on internet-based services while others show higher call engagement. Billing ranges from low monthly charges to premium-tier spending. Network rating values reflect user satisfaction, while churn labels reveal which customers are leaving the service. These factors collectively shape the foundation of the analysis.



LIBRARIES USED

The project is carried out in Google COLAB using Python libraries designed for large-scale analytics. Pandas and NumPy handle data preparation and transformation. Matplotlib create expressive visuals that reveal trends across data usage, call duration and churn. Google COLAB's cloud environment ties everything together with smooth, hardware-free execution.

- **Pandas:** cleaning and restructuring telecom data
- **NumPy:** numerical operations behind usage metrics
- **Matplotlib:** visualizing call and data trends
- **Seaborn:** correlation heatmaps and distribution plots
- **Google COLAB:** cloud workspace for fast execution



METHODOLOGY

- The analysis begins with loading and inspecting the dataset, followed by cleaning missing values in columns such as gender, tenure and payment method. Once the data becomes consistent, descriptive statistics and visual patterns are examined to understand usage behavior. Relationships between variables like monthly charges, data usage, network rating and churn are analyse. Predictive modeling is then applied to examine churn patterns and identify critical drivers behind customer exits.
- The project begins by loading the dataset into Google COLAB, followed by cleaning and formatting each attribute. Once the data becomes consistent, exploratory analysis is performed to uncover how customers differ in usage, spending and satisfaction. Various visualizations are created to highlight patterns, and a machine learning model is optionally added to evaluate churn probability. Insights are then interpreted and organized to support decision-making.

EXPLORATORY DATA OBSERVATION

- Early exploration shows that customers display significant variation in both data usage and call duration, reflecting different service priorities. Monthly billing reveals how spending grows with increased data demand, especially for fiber optic users. Regions show distinct usage patterns, while network ratings highlight satisfaction differences. Churn cases often align with lower ratings, higher bills or short tenure, revealing early clues about behavioral risks.
- high data usages often linked with higher monthly charges
- Fiber optic users show strong engagement but a higher risk of churn when bills rise
- Customers with lower network rating frequently appear in churn cases
- Short- tenure users tend to leave early, hinting at onboarding issues
- Urban regions show heavier data consumptions compared to other.



VISULATION FINDINGS

Visual maps of the dataset expose several interesting relationships. Scatter plots show how monthly charges increase with both network usage and added services. Distribution graphs reveal the concentration of users within mid-range billing categories. Heatmaps demonstrate connections among variables like age, data usage, tenure and churn. These visuals turn raw numbers into meaningful storylines that describe how customers behave across the telecom network.

- Scatter plots show clear relationships between bill amount and data use
- Line graphs reveal rising trends in average spending
- Heatmaps highlight strong correlations among network rating, tenure and churn
- Distribution curves show that most customers fall within mid-range billing



PROJECT OUTCOMES

The project delivers a clear picture of customer behavior across data consumption, call activity, billing and satisfaction. It identifies the core factors influencing churn and showcases how analytics can guide business decisions. By using Google COLAB, the analysis remains efficient, transparent and easy to reproduce. The findings create a strong foundation for telecom companies aiming to improve customer retention, optimize billing strategies and enhance service quality.

- ▶ Clear understanding of telecom user behavior
- ▶ Identification of high-risk churn customers
- ▶ Insightful visual dashboards for decision support
- ▶ Predictive model built for churn analysis
- ▶ Complete workflow executed using Google COLAB
- ▶ Actionable recommendations for telecom improvement