## **Phase-2**

**Student Name:** Naznin Fathima A   
**Register Number:** 611823243038  
**Institution:** PSV College of Engineering and Technology  
**Department:** B tech (AI&DS)  
**Date of Submission:** 10-05-2025  
**GitHub Repository Link:** https://github.com/Naznin-08/Survey-data-analysis.git

### **1. Problem Statement**

The problem being addressed is understanding the key factors that influence customer satisfaction, which is crucial for businesses to improve their products and services. By analyzing survey data, we aim to identify the most significant drivers of customer satisfaction and provide actionable insights for decision-making. This analysis falls under diagnostic analytics, as we seek to understand the underlying reasons for customer satisfaction or dissatisfaction.

### **2. Project Objectives**

* Identify the key drivers of customer satisfaction through survey data analysis
* Determine the most significant factors influencing customer satisfaction
* Provide recommendations for businesses to improve customer satisfaction
* Answer key questions:
* What are the most critical factors affecting customer satisfaction?
* How do different demographic groups perceive satisfaction differently?
* What are the areas where businesses can improve to increase customer satisfaction ?
* Expected deliverables include:
* Insights into the key drivers of customer satisfaction
* Trends and patterns in customer satisfaction data
* Recommendations for businesses to improve customer satisfaction

### **3. Flowchart of the Project Workflow**

Data Collection

↓

Cleaning

↓

EDA

↓

Insight Extraction

↓

Visualization

↓

Reporting

### **4. Data Description**

* Dataset name and source (e.g., Kaggle, UCI, company data)
* Data type: Structured / Unstructured
* Number of rows and columns
* Static or dynamic dataset
* Key fields or attributes relevant to the problem]

### **5. Data Preprocessing**

* Handling missing values
* Removing duplicates
* Formatting and parsing data
* Encoding categorical variables (if required)
* Identifying and optionally treating outliers
* Document all transformations and their reasons.]

### 

### 

### **6. Exploratory Data Analysis (EDA)**

* **Univariate Analysis**: Distribution of single variables using plots
* **Bivariate/Multivariate Analysis**: Heatmaps, pairplots, grouped bars, etc.
* Analysis of key metrics or KPIs (e.g., average revenue, churn rate, region-wise performance)
* Summary of insights and patterns identified

### 

### **7. Tools and Technologies Used**

* **Programming Language:** Python
* **Notebook/IDE:** Google Colab, Jupyter Notebook
* **Libraries:** pandas, numpy, matplotlib, seaborn, plotly
* **Optional Automation Tools:** pandas-profiling]

### **8. Team Members and Contributions**

| **Name** | **Contribution** |
| --- | --- |
| Naznin Fathima A | **Data Collection :** Gathering and measuring data from various sources to obtain a comprehensive dataset.  **Data cleaning :** Data preprocessing or data scrubbing, is the process of transforming and correcting raw data into a format suitable for analysis. |
| Puvisadhana R | **Exploratory Data Analysis:** visualizing and statistically exploring the data to understand its underlying structure, patterns, and relationships.  **Insight Extraction:** Identifying and interpreting meaningful patterns, trends, and relationships in the data. |
| Rajeshwari J | **Visualization :** Creating visual representations of complex data to communicate project progress, trends, and insights effectively.  **Reporting:** Define project objectives, deliverables, and timelines. |