ADT Group-25
Final Project Proposal

AIRLINES CUSTOMER SATISFACTION DASHBOARD

Team Members

- Ayantika Nandi (*Team Lead*)
- Megha Nagabhushana Reddy
- Shahrukh Quraishi

Project Summary

This project showcases the practical implementation of an Extract, Transform, Load (ETL) pipeline, illustrating its pivotal role in generating downstream applications like Tableau. Tableau serves as a powerful dashboard tool, crucial for business stakeholders seeking insights into and tracking key performance indicators (KPIs).

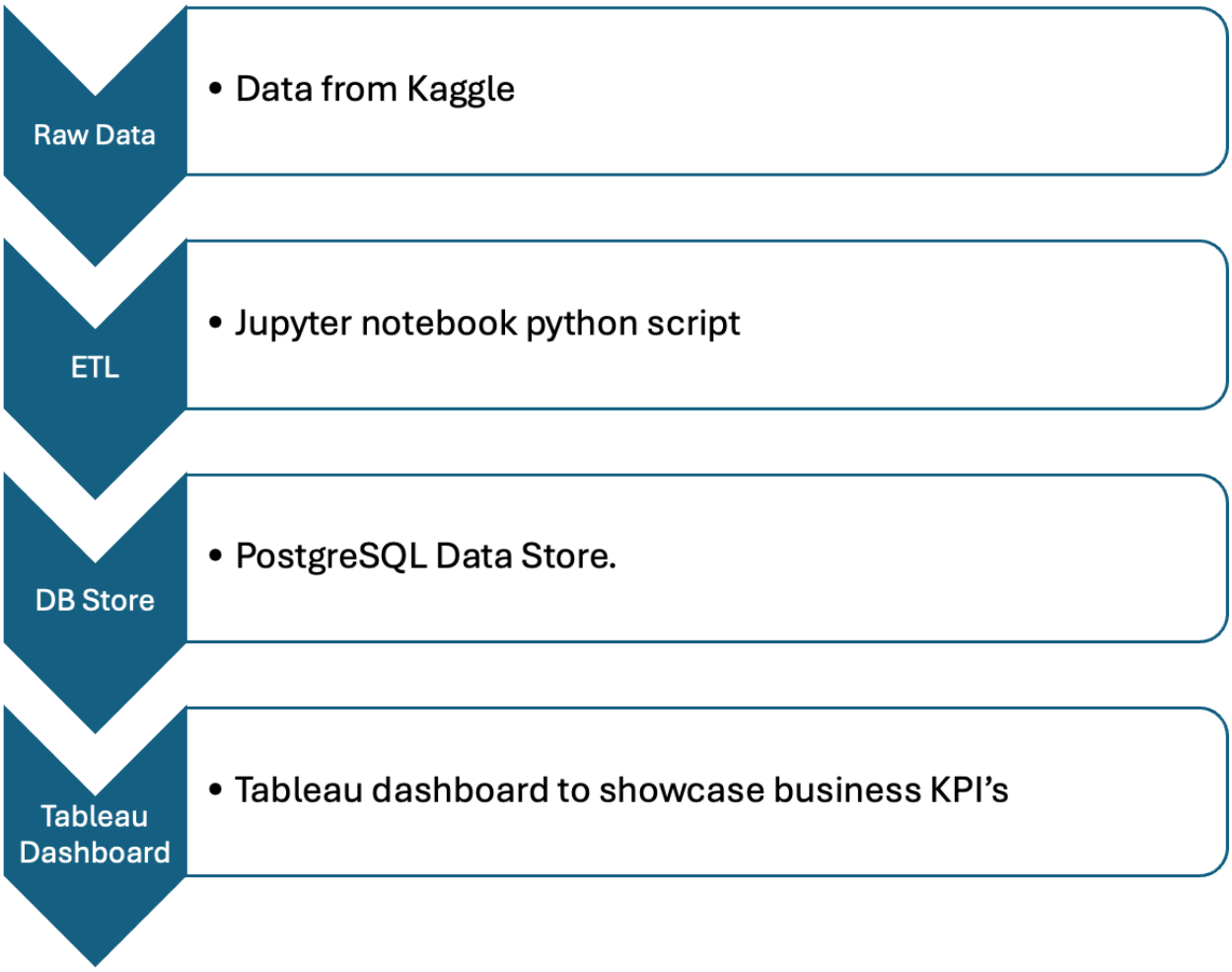
Our focus lies in leveraging a dataset on customer satisfaction sourced from Kaggle, utilizing it to monitor and analyze various business metrics and KPIs. Through the ETL pipeline, we streamline the process of extracting relevant data, transforming it into actionable insights, and loading it into Tableau for intuitive visualization and comprehensive analysis.

By harnessing the synergy between ETL pipelines and Tableau, our project aims to empower businesses with the ability to make informed decisions, optimize processes, and enhance customer satisfaction based on insightful data-driven strategies.

Project Description

- Objectives:

  In this project, we delineate our workflow into four distinct stages, each contributing significantly to the seamless transformation of raw data into actionable insights. Illustrated below is an overview of these stages:



1. **Raw Data (Acquisition)**: Initially, we acquire the customer satisfaction data from Kaggle, a reputable source, in the CSV format. This raw data serves as the foundation for our analysis and insights generation.

2. **ETL Processing**: Next, we implement a Python script tailored for ETL processes. This script efficiently loads the raw CSV data and executes a series of crucial tasks including data cleaning, transformation, and aggregation. Subsequently, the refined data is seamlessly inserted into a PostgreSQL database. This pivotal step ensures that the data is prepared and structured optimally for downstream analysis.

3. **Database Schema Design and Storage**: In this stage, we meticulously design the schema for our PostgreSQL database. This schema is thoughtfully crafted to accommodate the aggregated KPI data, organized at various levels such as customer,

demographic, or date. By structuring the data in this manner, we facilitate efficient querying and retrieval of insights during subsequent analysis.

4. **Tableau Dashboard Creation:** The culmination of our solution lies in the creation of a dynamic and visually engaging Tableau dashboard. This meticulously crafted dashboard acts as the interface for showcasing our business KPIs, with a primary focus on customer satisfaction metrics. By leveraging Tableau's intuitive visualization capabilities, stakeholders gain access to meaningful insights and trends, enabling them to make informed decisions and drive strategic initiatives effectively.

- Usefulness:

Utility of the Database

The dataset "Invistico_Airline.csv" covering airline operations, customer feedback, and flight details, becomes an important tool when combined with Tableau for visualization and ETL pipelines. This setup will be particularly useful due to:

- *Data-Driven Decisions*: By transforming raw data into visually appealing and easily understandable charts and graphs, stakeholders can make informed decisions quickly, enhancing operational efficiency and customer satisfaction.
- *Interactive Visual Analytics*: Tableau's powerful visual analytics allow users to interact with the data in real-time, enabling them to uncover hidden insights through a highly intuitive interface.
- *Trend Identification*: Visualization helps in identifying trends and patterns over time, which may not be apparent through raw data analysis. This is crucial for strategic planning.

Comparative Analysis

There are several databases and visualization tools available in the market such as Microsoft Power BI, Google Data Studio, and industry-specific solutions like FlightStats for flight data analytics, OAG for flight schedules and AirNav RadarBox for real-time flight tracking. However, our project would stand out by:

- ***Customization and Flexibility***: Tailoring the dashboard to meet the specific needs of the airline industry, with customized visualizations that highlight the most relevant metrics for this sector.
- ***Integration with Tableau***: Utilizing Tableau's advanced visualization capabilities can provide a more sophisticated analysis platform compared to basic dashboards. This

includes dynamic filtering, drill-down capabilities, and the ability to handle complex datasets.

- ***Comprehensive Data Integration***: Our project stands out by integrating a wide range of data types, from customer feedback and satisfaction scores to detailed flight performance and operational efficiency metrics. This kind of comprehensive approach provides a holistic view of the airline's operations, which is often lacking in more specialized tools such as OAG for flight schedules or FlightStats for flight data analytics.
- ***User-Centric Design***: Focusing on ease of use and accessibility for non-technical users, making complex data analysis approachable for a broader audience.

Target User Group

The dashboard will be designed to cater to a wide audience, focusing on roles that benefit most from data-driven insights and visual data exploration:

- ***Airline Management and Operations Teams:*** For monitoring flight performance, customer satisfaction, and operational efficiency.
- ***Aviation Authorities***: To oversee compliance and performance standards within the airline industry.
- ***Customer Experience Managers:*** To deeply understand customer feedback and areas requiring improvement, directly impacting service quality.
- ***Data Analysts and Business Intelligence Professionals:*** Who require deep dives into data for comprehensive analysis, trend spotting, and generating actionable insights.

- Dataset:

The dataset is sourced from Kaggle. The dataset was gathered by an airline group "Invistico Airlines". The company's true name is not divulged, most likely owing to privacy or confidentiality concerns. The dataset includes information about customers who have previously traveled with the airline, as well as input on various areas of the service.

The primary goal of the dataset appears to be predicting whether future passengers would be satisfied with the airline's service using a variety of characteristics and to identify the areas where the airline's services might be improved to increase customer satisfaction.

The dataset was most likely compiled over time and includes a significant number of records (about 129K) with 23 attributes. The particular time frame for data collection is

not specified, however it appears to be recent, considering that it is being used in a tech competition. It was distributed to competitors for analysis and modeling tasks.

Dataset Source: https://www.kaggle.com/datasets/sjleshrac/airlines-customer-satisfaction/data

- Sharing: Link to GitHub repo - https://github.com/megha1823/ADT-Team25-FinalProjectSpring2024

Group Contribution

Name	***Tasks***	***Average Time Spent (per milestone)***
Ayantika Nandi	• Dataset selection and data exploration • Data extraction • Dataset cleaning and pre-processing • Data normalization and conversion • Data Loading into database • Optimization and maintenance • Documentation	~10 hours
Shahrukh Quraishi	• Dataset selection and data exploration • Using a data processing framework like Apache Airflow to build the pipeline stages • Loading data into database • Optimization and maintenance • Documentation	~10 hours

Megha Nagabhushana Reddy	● Dataset selection and data exploration ● Data visualization using Tableau ● Optimization and maintenance ● Documentation	~10 hours