**Project Name:** Public Transportation Analysis

**Date:** 30/09/2023

**Project Team Member:** Varsha, Abinaya, Tharani, Neha Shalini, Vinoth

**Table of Contents**

1. **Introduction**
   * Problem Statement
   * Objective
   * Scope
2. **Problem Statement Understanding**
   * Problem Overview
   * Key Challenges
   * Stakeholders
3. **Solution Design**
   * High-Level Architecture
   * Data Flow Diagram
   * Technology Stack
   * Algorithmic Approach
   * User Interface Design
   * Security Measures
   * Testing Strategy
4. **Project Timeline**
   * Milestones
   * Task Breakdown
5. **Conclusion**
   * Summary
   * Next Steps

**1. Introduction**

**Problem Statement**

The project aims to analyse public transportation data to assess service efficiency, on-time performance, and passenger feedback. The goal is to provide insights that support transportation improvement initiatives and enhance the overall public transportation experience.

**Objective**

* Define analysis objectives for improving public transportation.
* Collect and analyse transportation data.
* Design relevant visualizations using IBM Cognos.
* Utilize code for advanced data analysis.

**Scope**

* Data collection from various sources.
* Data analysis for on-time performance and service efficiency.
* Visualization design in IBM Cognos.
* Code-based data analysis.

**2. Problem Statement Understanding**

**Problem Overview**

Public transportation systems play a vital role in urban areas, but their efficiency and reliability can often be a challenge. This project addresses the need for data-driven insights to enhance public transportation services. It will involve collecting and analysing data related to routes, schedules, passenger feedback, and more.

**Key Challenges**

* Data Integration: Gathering data from diverse sources.
* Data Quality: Ensuring the accuracy and reliability of the data.
* Analysis Complexity: Developing algorithms for performance assessment.
* Visualization: Creating informative and accessible visualizations.
* Security: Protecting sensitive transportation data.

**Stakeholders**

* **Public Transport Authorities:** Interested in improving services.
* **Passengers:** Expecting better service and on-time performance.
* **Transportation Analysts:** Needing data for analysis.

**3. Solution Design**

**High-Level Architecture**

* Data Collection & Integration
* Data Preprocessing
* Data Analysis & Visualization
* User Interface
* Security Layer

**Technology Stack**

* Data Collection: APIs, databases
* Data Analysis: Python (Pandas, NumPy), IBM Cognos
* Data Visualization: IBM Cognos, Matplotlib, Seaborn
* Security: Encryption, authentication, access controls

**Algorithmic Approach**

* On-Time Performance: Analyse schedules vs. actual arrivals.
* Service Efficiency: Route optimization algorithms.
* Passenger Feedback: Sentiment analysis.

**User Interface Design**

* Dashboard in IBM Cognos for real-time insights.
* User-friendly web portal for stakeholders.

**Security Measures**

* Data encryption in transit and at rest.
* Role-based access control.
* Regular security audits.

**Testing Strategy**

* Unit testing for data processing components.
* Integration testing for the entire system.
* User acceptance testing with stakeholders.

**4. Project Timeline**

**Milestones**

1. Data Collection & Integration
2. Data Analysis & Visualization
3. User Interface Development
4. Security Implementation
5. Testing & Quality Assurance
6. Deployment

**5. Conclusion**

**Summary**

This document outlines our understanding of the problem and our proposed solution for the Public Transportation Data Analysis Project. By following this plan, we aim to provide valuable insights that will support transportation improvement initiatives and enhance the public transportation experience.

**Next Steps**

The next steps include initiating data collection, development, and testing phases according to the outlined timeline. Regular communication and collaboration among team members will be crucial to the success of this project.