**Project Report**

Delivery time Prediction Based on Sorting Time

**1. Introduction**

**1.1 Objective**

The objective of the project is to predict delivery time based on sorting time using a simple linear regression model.

**1.2 Scope**

The scope includes data analysis, model development, and insights generation to inform human resource decisions.

**1.3 Project Duration**

The project was initiated on 06-11-2023 and completed on [12-11-2023]. The total time taken for development and analysis was 1-WEEK.

**2. Solution Architecture**

**2.1 Data Collection**

The dataset consists of delivery time and sorting time pairs.

Data collected through internal logistics systems.

**2.2 Data Preprocessing:**

Loaded and cleaned the dataset using Python and Pandas.

Split the data into training and testing sets.

**2.3 Exploratory Data Analysis (EDA)**

Utilized data visualization tools (Matplotlib, Seaborn) for EDA.

Conducted scatter plot and heatmap analysis to understand the relationship between sorting time and delivery time.

**2.4 Model Development**

Implemented a simple linear regression model using the scikit-learn library.

Split the data into training and testing sets.

Trained the model on the training set.

**2.5 Model Evaluation**

Evaluated the model using Mean Squared Error (MSE) and R-squared metrics.

Provided insights into the coefficients and intercept of the model.

**2.6 Real-Time Implementation:**

Considered opportunities for integration with existing logistics systems.

Explored continuous improvement strategies and feedback mechanisms.

**2.7 Challenges Faced:**

LimitedData**:** Limited dataset with a small number of observations.

Model Complexity: Balancing model complexity with interpretability.

Mitigation: Chose a simple linear regression model for transparency and ease of implementation.

Integration Challenges: Integrating the model with real-time logistics systems.

Mitigation: Collaborated with IT and logistics teams, implemented gradual integration steps.

**2.7 Complexity**

The project complexity was moderate, primarily involving a linear regression model.

**3. Methodology**

**3.1 Data Preparation**

Loaded and preprocessed the dataset, handling any missing or anomalous data points.

**3.2 Exploratory Data Analysis (EDA)**

Generated descriptive statistics and visualizations to gain insights into the data distribution.

**3.3 Model Building**

Developed a simple linear regression model using the scikit-learn library.

Split the data into training and testing sets for model evaluation.

**3.4 Model Evaluation**

Assessed model performance using Mean Squared Error (MSE) and R-squared metrics. Ensuring model convergence and performance.

**3.5 Insights Generation**

Provided business insights based on model coefficients and intercept. Ensuring seamless integration with existing systems .

**4. Conclusion**

**4.1 Results**

The model demonstrated a positive correlation between Sorting Time and Delivery.

Coefficients and intercept provided valuable insights into the relationship.

**4.2 Business Impact**

Challenges were addressed through thorough data preprocessing and model optimization. The project's methodology and architecture provide a foundation for future enhancements and integrations.

**4.3 Future Considerations**

Continuous improvement and adaptation to evolving business needs remain essential for sustained success.

**5. Acknowledgments**

**6. References**

Feel free to customize this template based on the specifics of your project and organization.