**Title : COVID-19 Data Analysis Documentation with EDA and Machine Learning Approach**

**I. Introduction**

**Background**

This documentation presents a comprehensive analysis of COVID-19 data using Exploratory Data Analysis (EDA) and a Machine Learning (ML) approach. The ongoing global pandemic necessitates a thorough understanding of the virus's impact, and this project aims to contribute insights through data-driven methodologies.

**II. Objectives**

**A. Primary Objectives**

1. Conduct Exploratory Data Analysis (EDA) to unveil patterns and trends in COVID-19 datasets.
2. Implement a Machine Learning approach to predict and analyze key aspects of the pandemic.
3. Provide actionable insights for better public health decision-making.

**B. Scope**

The analysis covers regional COVID-19 test datasets, incorporating cough\_symptoms, fever, shortness\_of\_breath , known contact etc variables.

**III. Methodology**

**A. Exploratory Data Analysis (EDA)**

1. **Data Collection:** Gathered comprehensive COVID-19 test datasets from odinschool sources.
2. **Data Cleaning:** Addressed missing values, outliers, and inconsistencies to ensure data integrity.
3. **Exploration and Visualization:** Utilized statistical measures and visualizations to reveal patterns, correlations, and anomalies.

**B. Machine Learning Approach**

1. **Feature Engineering:** Selected relevant features for predictive modeling.
2. **Data Splitting:** Partitioned datasets into training and testing sets.
3. **Model Selection:** Explored various ML models, with a focus on algorithms suitable for time-series and classification tasks.
4. **Training and Evaluation:** Trained models on historical data and evaluated performance metrics.
5. **Predictions:** Applied logistic regression model to predict future trends and outcomes.

**IV. Results and Findings**

**A. EDA Insights**

1. **Temporal Trends:** Analyzed temporal patterns in shortness\_of\_breath, fever, and cough\_symptoms, soar throat.
2. **Variable Correlations:** Investigated relationships between gender, and fever rates.

**B. Machine Learning Predictions**

1. **Accuracy:** Achieved a 91% accuracy in predicting covid positive or not.
2. **Feature Importance:** Identified key features influencing predictions.
3. **Future Projections:** Utilized models to project potential trends and inform proactive measures.

**VI. Conclusion**

This documentation provides a thorough analysis of COVID-19 test data using EDA and a Machine Learning approach. The insights gained contribute to the collective understanding of the pandemic, supporting informed decision-making and paving the way for future research and interventions.