**Covid-19 Infection ETL Project: Analyzing Global Data**

**Abstract:**

The Covid-19 Infection ETL project explores the impact of the Covid-19 pandemic during January-March 2020, using data compiled by John Hopkins University, Kaggle. The data was sourced from reputable global organizations, including the World Health Organization and the European Centre for Disease Prevention and Control. The project's primary objective is to analyze the number of confirmed cases, deaths, and recoveries across various countries.

The process involves several key stages. First, the raw data was extracted from CSV files and cleaned using Python, particularly focusing on filtering the data and transforming date information for consistency. The cleaned data was then migrated to a local PostgreSQL database, where it was made query able for further analysis.

Through SQL queries, the project seeks to identify the top 5 countries with the most and least confirmed cases, deaths, and recoveries, along with the specific dates in 2020. That saw the highest numbers in these categories. This analysis is further enhanced by visualizations created using Power BI, providing an intuitive and interactive interface for the results. This project serves as an insightful exploration of the global impact of Covid-19 during its early stages and demonstrates the application of data engineering and visualization techniques.

**Members:**

1. Mohammad Aas Khan (22052561)
2. Naser Ahmad (22052741)
3. Jaish Nawed (22052731)
4. Snighdha (22052680)
5. Ishu Kant (22052730)
6. Mohammad Kaif (22051261)