Virtual World User Engagement Pipeline Analytics Pipeline - Project Documentation

# 1. Introduction

Provide an overview of the project, its goals, and why it was chosen. Mention technologies used such as Apache Kafka, Spark Streaming, BigQuery/Redshift, dbt, and Google Data Studio.

# 2. Data Ingestion

Step-by-step guide to how the data ingestion process works, including code snippets, configuration, and screenshots for Kafka setup.

# 3. Data Processing

Explain the Spark Streaming setup and how the data is processed in real time. Include screenshots of logs or processing in action.

# 4. Data Storage & Transformation

Explain how data is stored in BigQuery/Redshift and transformed using dbt. Provide SQL queries or dbt models with screenshots of results.

# 5. Visualization

Explain how the data is visualized using Google Data Studio or Tableau Public. Include screenshots of dashboards and key insights.

# 6. Challenges & Solutions

Describe any challenges you faced during the project and how you overcame them.

# 7. Conclusion

Summarize the project results, key insights, and what could be improved in the future.