PROJECT REQUIREMENTS

DATA INGESTION REQUIREMENTS

1.INGEST ALL THE 8 FILES INTO THE DATA LAKE

2.INGESTED DATA MUST HAVE THE SCHEMA APPLIED

3.INGESTED DATA MUST HAVE THE AUDIT COLUMNS

4.INGESTED DATA MUST BE STORED IN COLUMNAR FORMAT(EG.PARQUET)

5.MUST BE ABLE TO ANALYZE THE INGESTED DATA VIA SQL

6.INGESTED LOGIC MUST BE ABLE TO HANDLE INCREMENTAL LOAD

DATA TRANSFORMATION REQUIREMENTS

1.JOIN THE KEY INFORMATION REQUIRED FOR REPORTING TO CREATE A NEW TABLE

2. JOIN THE KEY INFORMATION REQUIRED FOR ANALYSIS TO CREATE A NEW TABLE

3.TRANSFORMED TABLES MUST HAVE AUDIT COLUMNS

4.MUST BE ABLE TO ANALYSE TRANSFORMED DATA VIA SQL

5.TRANSFORMED DATA MUST BE STORED IN COLUMNAR FORMAT(PARQUET)

6.TRANSFORMATION LOGIC MUST BE ABLE TO HANDLE INCREMENTAL LOAD

REPORTING REQUIREMENTS

1.DRIVER STANDINGS

2.CONSTRUCTOR STANDINGS

ANALYTICAL REQUIREMENTS

1.DOMINANT DRIVERS

2.DOMINANT TEAMS

3.VISUALIZE THE OUTPUTS

4.CREATE DATABRICKS DASHBOARDS

SCHEDULING REQUIREMENTS

1.SCHEDULED TO RUN EVERY SUNDAY 10PM

2.ABILITY TO MONITOR PIPELINES

3.ABILITY TO RUN FAILED PIPELINES

4.ABILITY TO SET-UP ALERTS ON FAILURES