**pl\_mysql\_ingestion** -> This pipeline ingests the data(.parquet files) from external mysql server to azure cloud(landing->bronze->silver->gold).

Implementation steps:

1.Get Metadata details: Extract the relevant parameters from databricks metadata tables required to run adf pipeline.

2 Trigger pipeline inprogress email alert and store pipeline start time in a variable.

3.Copy Source Files: Copy table records(.parquet files) to landing container in azure ADLS gen2. If copy source files activity fails then go to step 3.1 otherwise go to step 3.2

3.1 Trigger copy failed email alert and collect logs.

3.2 Compare rowsread and rowscopied record count. If validation fails go to step 3.2.1 otherwise go to step 4

3.2.1 Trigger record mismatch email alert and collect logs.

4.Trigger mysql\_landing\_to\_bronze pipeline.

5.On completion of step 4, move files to processed folder in landing container.

6. On Successful completion of step 4,Trigger bronze\_to\_silver pipeline.

7. On successful completion of step 6, trigger silver\_to\_gold pipeline.

8. On completion of step 7, trigger pipeline completion email alert(Success/Failure).  
  
**pl\_sftp\_ingestion** -> This pipeline ingests the data(.csv files) from external SFTP location to azure cloud(landing->bronze->silver->gold).

Implementation steps:

1.Get Metadata details: Extract the relevant parameters from databricks metadata tables required to run adf pipeline.

1.1 Trigger pipeline inprogress email alert and store pipeline start time in a variable.

2.Get Source File details: Collects File names, file format from the SFTP source location.

3.Check Source File count: Validate whether files are available in SFTP source location or not. If Source File count = 0, then go to step 4, otherwise go the step 5.

4.Collect logs and trigger email alert and end the pipeline.

5.Copy Source Files: Copy source files to landing container in azure ADLS gen2.

6. Trigger sftp\_landing\_to\_bronze pipeline.

7.On completion of step 6, move files to processed folder in landing container.

8. On Successful completion of step 7,Trigger bronze\_to\_silver pipeline.

9. On successful completion of step 8, trigger silver\_to\_gold pipeline.

10. On completion of step 9, trigger pipeline completion email alert(Success/Failure).

**pl\_sftp\_landing\_to\_bronze** -> This pipeline loads data(.csv files) into bronze layer table.

Implementation steps:

1.loading\_to\_bronze: Trigger the databricks notebook which contains script to extract records from source file(in landing container) and insert those records in bronze layer table. If loading\_to\_bronze activity fails then go to step 2 otherwise go to step 3.

2. Trigger bronze ingestion failure email alert and collect logs.

3.Records Validation: Validate whether source files and records inserted in bronze layer tables are same or not. If validation fails go the step2, otherwise go to step 4.

4. Collect logs for bronze layer ingestion success.

**pl\_mysql\_landing\_to\_bronze** -> This pipeline loads data(.parquet files) into bronze layer table.

Implementation steps:

1.loading\_to\_bronze: Trigger the databricks notebook which contains script to extract records from source file(in landing container) and insert those records in bronze layer table. If loading\_to\_bronze activity fails then go to step 2 otherwise go to step 3.

2. Trigger bronze ingestion failure email alert and collect logs.

3.Records Validation: Validate whether source files and records inserted in bronze layer tables are same or not. If validation fails go the step2, otherwise go to step 4.

4. Collect logs for bronze layer ingestion success.

**pl\_bronze\_to\_silver** -> This pipeline loads data from bronze layer tables to silver layer tables.

Implementation steps:

1.loading\_to\_silver: Trigger the databricks notebook which contains script to insert records from bronze layer table to silver layer table. If loading\_to\_silver activity fails then go to step 2 otherwise go to step 3.

2. Trigger silver ingestion failure email alert and collect logs.

3. Collect logs for silver layer ingestion success.

**pl\_silver\_to\_gold** -> This pipeline loads data from silver layer tables to gold layer tables.

Implementation steps:

1.loading\_to\_gold: Trigger the databricks notebook which contains script to insert records from silver layer table to gold layer table. If loading\_to\_gold activity fails then go to step 2 otherwise go to step 3.

2. Trigger gold ingestion failure email alert and collect logs.

3. Collect logs for gold layer ingestion success.

**pl\_bronze\_to\_silver** -> This pipeline loads data from bronze layer tables to silver layer tables.

Implementation steps:

1.loading\_to\_silver: Trigger the databricks notebook which contains script to insert records from bronze layer table to silver layer table. If loading\_to\_silver activity fails then go to step 2 otherwise go to step 3.

2. Trigger silver ingestion failure email alert and collect logs.

3. Collect logs for silver layer ingestion success.