

MCB Case Study – Engineering Lead (Data)

Nirantar Seewooruttun - 25th Nov 2021

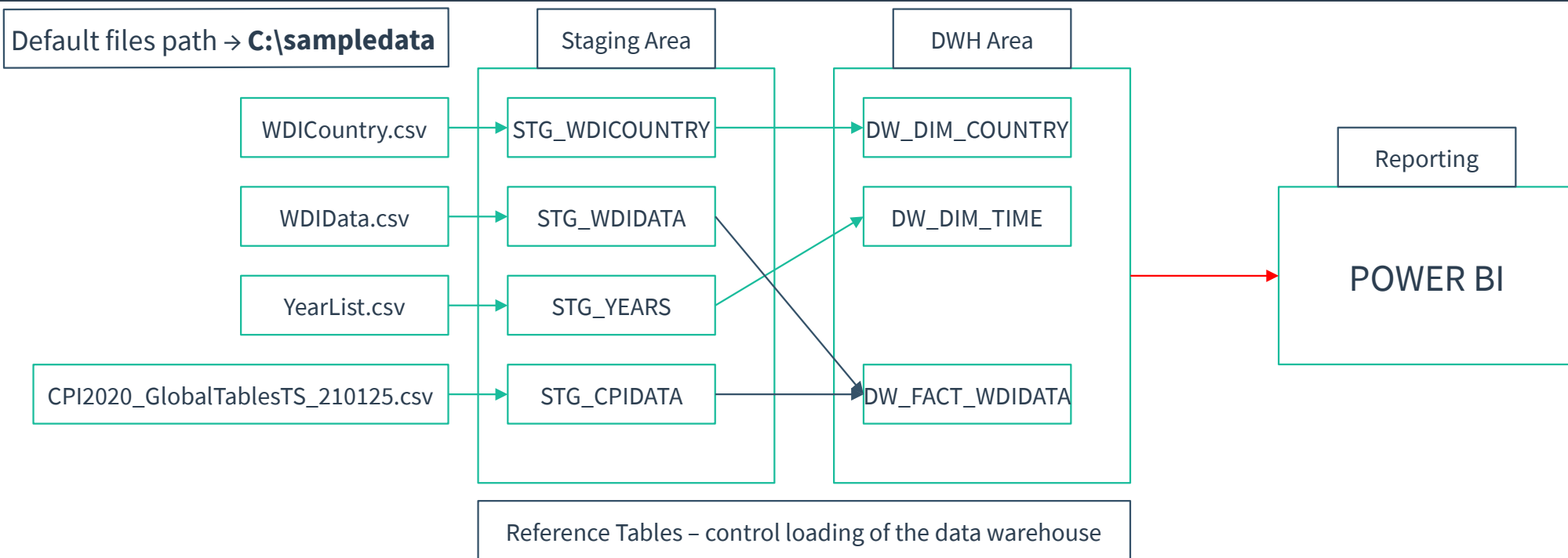
Contents

- BI Architecture Diagram
- ETL Process
- Power BI Report Layout
- Q&A
- Codebase access links
- Appendix 1 to 7

Assumptions:

1. Reporting will be conducted on WDI Data from 2010 onwards.

BI Architecture Diagram

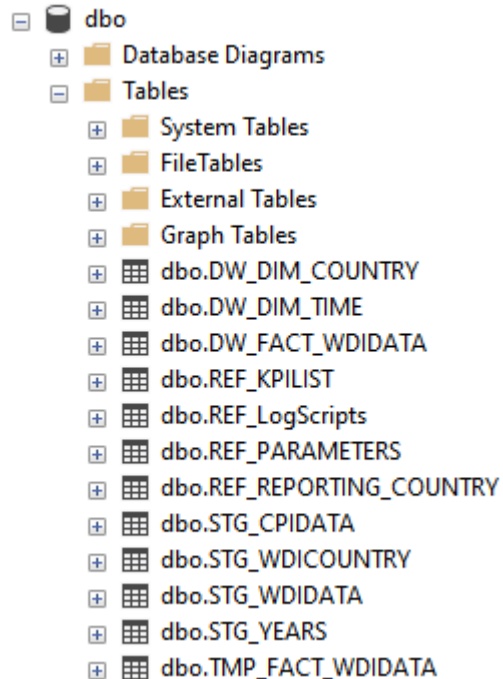
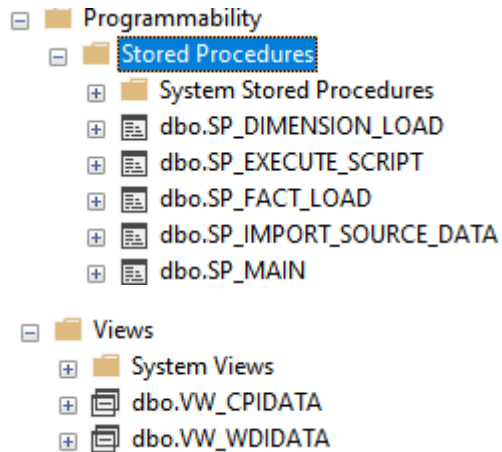


Notes:

- Developed locally on SQLServer Express edition. Local schema name : **DBO**
- 2 views : **VW_CPIDATA** and **VW_WDIDATA** are used to extract data from staging area for DWH. **WDIDATA.csv** was first unpivoted on years then pivoted on KPIs for reporting.

ETL Process

Default files path → C:\sampledata



EXECUTE **SP_MAIN** steps

→ SP_IMPORT_SOURCE_DATA :

Pulls source files from C:\sampledata\ folder
And loads data to staging tables.

→ SP_DIMENSION_LOAD : Loads data from staging tables to dimension tables

→ SP_FACT_LOAD : Loads data from staging fact tables in final fact table.

→ REF_LogScripts table logs all scripts executed on the data warehouse

Power BI Report Layout

Pick a Year

2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020

Country

All

Hold **Ctrl** to select multiple countries

Indicator (Single select)

Employment in agriculture (% of total employment) (modeled ILO estimate)

Country	Country Code	Region	Employment in agriculture (% of total employment) (modeled ILO estimate) Yr. 2019	Employment in agriculture (% of total employment) (modeled ILO estimate) Yr. 2018	Change in Scores	Employment in agriculture (% of total employment) (modeled ILO estimate) Rank Yr. 2019	Employment in agriculture (% of total employment) (modeled ILO estimate) Rank Yr. 2018	Changes in Rank
Australia	AUS	East Asia & Pacific	2.56	2.62	↓	7	7	0
Côte d'Ivoire	CIV	Sub-Saharan Africa	40.15	41.01	↓	3	3	0
Denmark	DNK	Europe & Central Asia	2.22	2.19	↑	8	8	0
Finland	Fin	Northern Europe	3.78	3.71	↑	6	6	0
India	IND	South Asia	42.60	43.33	↓	2	2	0
Kenya	KEN	Sub-Saharan Africa	54.34	55.08	↓	1	1	0
Maldives	MDV	South Asia	8.32	8.61	↓	5	5	0
Seychelles	SYC	Sub-Saharan Africa	0.00	0.00	→	10	10	0
Singapore	SGP	East Asia & Pacific	0.03	0.06	↓	9	9	0
Sri Lanka	LKA	South Asia	24.98	25.50	↓	4	4	0

Fields

Search

- > DW_DIM_COUNTRY
- > DW_DIM_TIME
- > DW_FACT_WDIDATA
- > REF_KPILIST



Questions

Codebase Access links

Backup Sample data

https://drive.google.com/file/d/1VWOya_LqnpVCqKHusGZXmGQAcJOYCniG/view?usp=sharing

Backup SQL Server Database link:

https://drive.google.com/file/d/14p87g9kVZGMRZiS7Y_DVbpvLkcV4dj47/view?usp=sharing

Codebase repository on GITHUB

<https://github.com/hanuseew/MCBAssignment.git>

Appendix 1 – Test cases

- List of countries in countries source data. → Finland manually added in SP_DIMENSION_IMPORT for the countries dimension : DW_DIM_COUNTRY

- ```
9 SELECT DISTINCT CountryCode, ShortName
10 FROM STG_WDICOUNTRY
11 WHERE CountryCode in (select CountryCode from REF_REPORTING_COUNTRY);
```

100 %

Results Messages

|   | CountryCode | ShortName     |
|---|-------------|---------------|
| 1 | AUS         | Australia     |
| 2 | CIV         | Côte d'Ivoire |
| 3 | DNK         | Denmark       |
| 4 | IND         | India         |
| 5 | KEN         | Kenya         |
| 6 | LKA         | Sri Lanka     |
| 7 | MDV         | Maldives      |
| 8 | SGP         | Singapore     |
| 9 | SYC         | Seychelles    |



# Appendix 2 – Source Data check for KPIs for one country - Australia

- Data extract as per the source file:

- `SELECT CountryCode, CountryName, IndicatorCode, IndicatorName, YR2010, YR2011, YR2012, YR2013, YR2014, YR2015, YR2016, YR2017, YR2018, YR2019, YR2020`

`FROM STG_WDIDATA WHERE CountryCode = 'AUS'`

and `IndicatorCode IN (`

`'ER.FSH.PROD.MT',  
'AG.LND.AGRI.K2',  
'IC.REG.DURS',  
'IC.BUS.NREG',  
'SL.AGR.EMPL.ZS',  
'SL.EMP.SELF.ZS');`

|                  |                  | Results          |                  | Messages                                            |                  |                  |                  |
|------------------|------------------|------------------|------------------|-----------------------------------------------------|------------------|------------------|------------------|
|                  | CountryCode      | CountryName      | IndicatorCode    | IndicatorName                                       | YR2010           | YR2011           | YR2012           |
| 1                | AUS              | Australia        | AG.LND.AGRI.K2   | Agricultural land (sq. km)                          | 3763720          | 3961210          | 3870760          |
| 2                | AUS              | Australia        | SL.AGR.EMPL.ZS   | Employment in agriculture (% of total employmen...  | 3.23000001907349 | 2.8199999332428  | 2.79999995231628 |
| 3                | AUS              | Australia        | IC.BUS.NREG      | New businesses registered (number)                  | 160615           | 168170           | 185009           |
| 4                | AUS              | Australia        | SL.EMP.SELF.ZS   | Self-employed, total (% of total employment) (mo... | 18.3700008392334 | 17.7900009155273 | 17.1800003051758 |
| 5                | AUS              | Australia        | IC.REG.DURS      | Time required to start a business (days)            | 2.5              | 2.5              | 2.5              |
| 6                | AUS              | Australia        | ER.FSH.PROD.MT   | Total fisheries production (metric tons)            | 292486.78125     | 290846.625       | 295670.15625     |
| YR2013           | YR2014           | YR2015           | YR2016           | YR2017                                              | YR2018           | YR2019           | YR2020           |
| 3717750          | 3744230          | 3481190          | 3426020          | 3718370                                             | 3588950          | NULL             |                  |
| 2.61999988555908 | 2.79999995231628 | 2.64000010490417 | 2.61999988555908 | 2.59999990463257                                    | 2.61999988555908 | 2.55999994277954 |                  |
| 226209           | 231920           | 242347           | 246623           | 250946                                              | 235654           | NULL             |                  |
| 16.8400001525879 | 17.1399993896484 | 17.0599994659424 | 17.0499992370605 | 16.8999996185303                                    | 16.4599990844727 | 16.5799999237061 |                  |
| 2.5              | 2.5              | 2.5              | 2.5              | 2.5                                                 | 2.5              | 2                |                  |
| 278254.375       | 272000.8125      | 286995.9375      | 313409.28125     | 319470.375                                          | 322235.21875     | NULL             |                  |

# Appendix 3 – DWH Data check for KPIs for one country – Australia

- Data extract as per data warehouse : (Data warehouse initialized as from Yr. 2010)

```
SELECT * FROM DW_FACT_WDIDATA where countrycode = 'AUS';
```

100 %

Results Messages

|    | WDIDataID | CountryID | TimeID | CountryName | CountryCode | Year | ER_FSH_PROD_MT | AG_LND_AGR_K2 | IC_REG_DURS | IC_BUS_NREG | SL_AGR_EMPL_ZS | SL_EMP_SELF_ZS | CPI     |
|----|-----------|-----------|--------|-------------|-------------|------|----------------|---------------|-------------|-------------|----------------|----------------|---------|
| 1  | 1         | 1         | 51     | Australia   | AUS         | 2010 | 292486.7813    | 3763720.0000  | 2.5000      | 160615.0000 | 3.2300         | 18.3700        | 0.0000  |
| 2  | 2         | 1         | 52     | Australia   | AUS         | 2011 | 290846.6250    | 3961210.0000  | 2.5000      | 168170.0000 | 2.8200         | 17.7900        | 0.0000  |
| 3  | 3         | 1         | 53     | Australia   | AUS         | 2012 | 295670.1563    | 3870760.0000  | 2.5000      | 185009.0000 | 2.8000         | 17.1800        | 85.0000 |
| 4  | 4         | 1         | 54     | Australia   | AUS         | 2013 | 278254.3750    | 3717750.0000  | 2.5000      | 226209.0000 | 2.6200         | 16.8400        | 81.0000 |
| 5  | 5         | 1         | 55     | Australia   | AUS         | 2014 | 272000.8125    | 3744230.0000  | 2.5000      | 231920.0000 | 2.8000         | 17.1400        | 80.0000 |
| 6  | 6         | 1         | 56     | Australia   | AUS         | 2015 | 286995.9375    | 3481190.0000  | 2.5000      | 242347.0000 | 2.6400         | 17.0600        | 79.0000 |
| 7  | 7         | 1         | 57     | Australia   | AUS         | 2016 | 313409.2813    | 3426020.0000  | 2.5000      | 246623.0000 | 2.6200         | 17.0500        | 79.0000 |
| 8  | 8         | 1         | 58     | Australia   | AUS         | 2017 | 319470.3750    | 3718370.0000  | 2.5000      | 250946.0000 | 2.6000         | 16.9000        | 77.0000 |
| 9  | 9         | 1         | 59     | Australia   | AUS         | 2018 | 322235.2188    | 3588950.0000  | 2.5000      | 235654.0000 | 2.6200         | 16.4600        | 77.0000 |
| 10 | 10        | 1         | 60     | Australia   | AUS         | 2019 | 0.0000         | 0.0000        | 2.0000      | 0.0000      | 2.5600         | 16.5800        | 77.0000 |
| 11 | 11        | 1         | 61     | Australia   | AUS         | 2020 | 0.0000         | 0.0000        | 0.0000      | 0.0000      | 0.0000         | 0.0000         | 77.0000 |

# Appendix 4 – Power BI reporting for Australia on : “Agricultural Land (sq. km)” KPI for current year = 2012

Pick a Year

2010 2011 **2012** 2013 2014 2015 2016 2017 2018 2019 2020

Country

All

Hold **Ctrl** to select multiple countries

Indicator (Single select)

Agricultural land (sq. km)

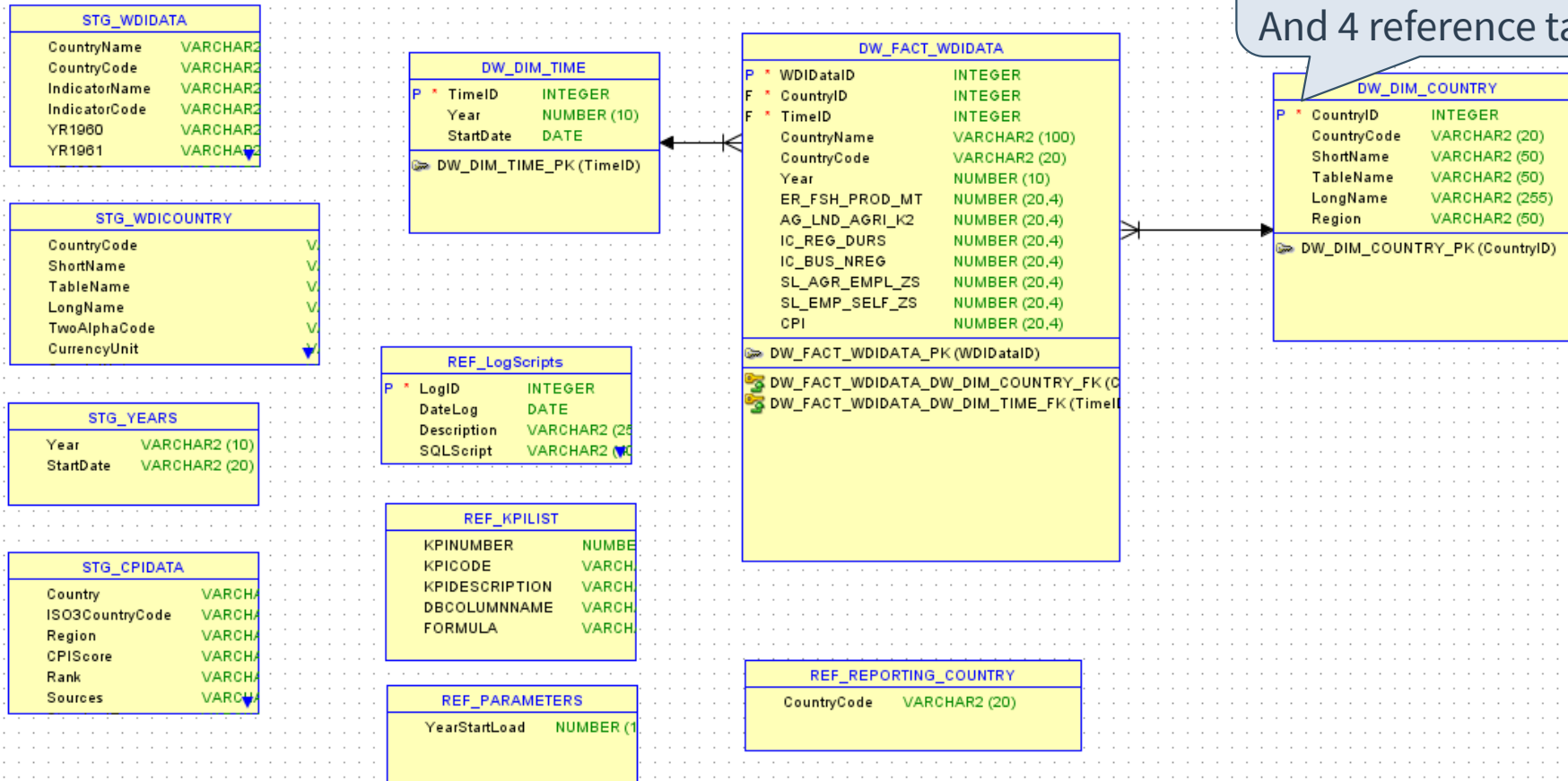
| Country       | Country Code | Region                | Agricultural land (sq. km) Yr. 2011 | Agricultural land (sq. km) Yr. 2010 | Change in Scores | Agricultural land (sq. km) Rank Yr. 2011 | Agricultural land (sq. km) Rank Yr. 2010 | Changes in Rank |
|---------------|--------------|-----------------------|-------------------------------------|-------------------------------------|------------------|------------------------------------------|------------------------------------------|-----------------|
| Australia     | AUS          | East Asia & Pacific   | 3,961,210.00                        | 3,763,720.00                        | ↑                | 1                                        | 1                                        | 0               |
| Côte d'Ivoire | CIV          | Sub-Saharan Africa    | 206,000.00                          | 206,000.00                          | →                | 4                                        | 4                                        | 0               |
| Denmark       | DNK          | Europe & Central Asia | 26,900.00                           | 26,260.00                           | ↑                | 6                                        | 5                                        | 1               |
| Finland       | Fin          | Northern Europe       | 22,866.00                           | 22,919.00                           | ↓                | 7                                        | 7                                        | 0               |
| India         | IND          | South Asia            | 1,796,700.00                        | 1,795,730.00                        | ↑                | 2                                        | 2                                        | 0               |
| Kenya         | KEN          | Sub-Saharan Africa    | 276,300.00                          | 273,200.00                          | ↑                | 3                                        | 3                                        | 0               |
| Maldives      | MDV          | South Asia            | 79.00                               | 79.00                               | →                | 8                                        | 8                                        | 0               |
| Seychelles    | SYC          | Sub-Saharan Africa    | 15.00                               | 24.00                               | ↓                | 9                                        | 9                                        | 0               |
| Singapore     | SGP          | East Asia & Pacific   | 7.30                                | 7.40                                | ↓                | 10                                       | 10                                       | 0               |
| Sri Lanka     | LKA          | South Asia            | 27,300.00                           | 26,200.00                           | ↑                | 5                                        | 6                                        | -1              |

## Appendix 5 – Reference Tables

- REF\_KPILIST → Configuration table for list of reporting KPIs. Only this list of KPI will goto reporting
- REF\_LogScripts → Scripts of all DB procedures are logged in this log table.
- REF\_PARAMETERS → Initializes the year from which to pull fact data on the data warehouse.
- REF\_REPORTING\_COUNTRIES → Initializes the list of countries for reporting. Can be modified to include more countries.

# Appendix 6 - Data warehouse tables layout

2 dimensions and  
1 fact table, 4 staging  
And 4 reference tables



# Appendix 7 – Log Table

```
22
23 SELECT * FROM REF_LogScripts order by LogID desc;
24
```

| 100 %            |       |                         |                                                          |                                                           |
|------------------|-------|-------------------------|----------------------------------------------------------|-----------------------------------------------------------|
| Results Messages |       |                         |                                                          |                                                           |
|                  | LogID | DateLog                 | Description                                              | SQLScript                                                 |
| 1                | 60    | 2021-11-25 20:18:48.250 | WDI Fact : Step 5 - Load final fact table for reporting  | INSERT INTO DW_FACT_WDIDATA (CountryID, TimeID, Countr... |
| 2                | 59    | 2021-11-25 20:18:48.247 | WDI Fact : Step 4 - Truncate fact table                  | TRUNCATE TABLE DW_FACT_WDIDATA;                           |
| 3                | 58    | 2021-11-25 20:18:35.697 | WDI Fact : Step 3 - Load source facts into temp fac...   | SELECT CountryName, CountryCode, Year, ...                |
| 4                | 57    | 2021-11-25 20:18:35.690 | WDI Fact : Step 2 - Drop temporary fact table if exists  | DROP TABLE IF EXISTS TMP_FACT_WDIDATA;                    |
| 5                | 56    | 2021-11-25 20:18:35.690 | Time dimension : Step 3 - Load dimension data from ...   | MERGE DW_DIM_TIME T USING STG_YEARS STG ...               |
| 6                | 55    | 2021-11-25 20:18:35.687 | Country dimension : Step 6 - Manual country import       | MERGE DW_DIM_COUNTRY C USING (SELECT 'Fin' ...            |
| 7                | 54    | 2021-11-25 20:18:35.667 | Country dimension : Step 5 - Load dimension data fr...   | MERGE DW_DIM_COUNTRY C USING STG_WDICO...                 |
| 8                | 53    | 2021-11-25 20:18:31.343 | WDIDATA Fact : Step 4 - Remove trailing characters       | UPDATE STG_WDIDATA SET YR2020 = REPLACE(REPLACE(...       |
| 9                | 52    | 2021-11-25 20:18:26.753 | WDIDATA Fact : Step 3 - Remove leading characters        | UPDATE STG_WDIDATA SET CountryName = REPLACE(REPL...      |
| 10               | 51    | 2021-11-25 20:18:17.177 | WDIDATA Fact : Step 2 - source data import               | BULK INSERT STG_WDIDATA FROM 'C:\sampledat...             |
| 11               | 50    | 2021-11-25 20:18:17.170 | WDIDATA Fact : Step 1 - Truncate staging table for...    | TRUNCATE TABLE STG_WDIDATA;                               |
| 12               | 49    | 2021-11-25 20:18:17.117 | CPI Fact : Step 2 - source data import                   | BULK INSERT STG_CPIDATA FROM 'C:\sampledat...             |
| 13               | 48    | 2021-11-25 20:18:17.117 | CPI Fact : Step 1 - Truncate staging table for source... | TRUNCATE TABLE STG_CPIDATA;                               |
| 14               | 47    | 2021-11-25 20:18:17.093 | Time Dimension : Step 2 - source data import             | BULK INSERT STG_YEARS FROM 'C:\sampledata\...             |
| 15               | 46    | 2021-11-25 20:18:17.093 | Time Dimension : Step 1 - Truncate staging table for...  | TRUNCATE TABLE STG_YEARS;                                 |
| 16               | 45    | 2021-11-25 20:18:17.087 | Country dimension : Step 4 - Remove trailing charac...   | UPDATE STG_WDICOUNTRY SET LatestTradeData = REPLAC...     |
| 17               | 44    | 2021-11-25 20:18:17.080 | Country dimension : Step 3 - Remove leading chara...     | UPDATE STG_WDICOUNTRY SET CountryCode = REPLACE(R...      |
| 18               | 43    | 2021-11-25 20:18:17.057 | Country dimension : Step 2 - source data import          | BULK INSERT STG_WDICOUNTRY FROM 'C:\sam...                |
| 19               | 42    | 2021-11-25 20:18:17.017 | Country dimension : Step 1 - Truncate staging table      | TRUNCATE TABLE STG_WDICOUNTRY;                            |