MCB Software Engineer
Data Practical Test Paper

Summary
An analysis of World Happiness Report has been conducted using an open source data-set. The objective of the project includes studying the various factors which lead up to the calculation of a “Happiness Index” for each country and understanding its distribution throughout the world.

The process involves ingesting the data-set in SQL, cleaning the dataset and then using it in R through RODBC server. Further, a regression model and clustering analysis has been done in R. Finally, the data has been visualized through Tableau in which the data has been imported from Microsoft SQL Server using the SQL server connection.

Data Overview

Source: The dataset is an open source dataset from a report published on world happiness, first in 2012 and thereafter every year. This report is an outcome of the survey results of Gallup World Poll which takes representative sample from each country and asks them questions in the form of Cantril ladder, which is asking respondents to think of a ladder with the best possible life for them being a 10 and the worst possible life being a 0 and to rate their own current lives on that scale.
The dataset is present here for 2015-2017 : https://www.kaggle.com/unsdsn/world-happiness/data

Data Description:

The datasets are identical except for the year they contain information of and have the following columns:

- **Country**: Name of the country
- **Region**: Region of the world, the country belongs to
- **Happiness Rank**: Rank of the country according to happiness score
- **Happiness Score**: Metric measured as a combination of various factors
- **Economy (GDP per capita):** The extent to which GDP contributes to happiness
- **Family**: The extent to which Family contributes to happiness
- **Health (Life Expectancy):** The extent to which Life Expectancy contributes to happiness
- **Freedom**: The extent to which Freedom contributes to happiness
- **Trust (Government Corruption):** The extent to which trust in government contributes to happiness
- **Generosity**: Generosity of the general public and its contribution to happiness
- **Dystopia Residual**: Contribution to Dystopia residual to happiness. Dystopia is an imaginary country that has the world's least happy people. The purpose of having this is to have a lower benchmark so that all countries do positively against it. This variable has no physical significance.

Task 1.

Data Loading:

Data is present in the form of 4 csv files, one for each year – 2016, 2017, 2018 and 2019. The files are imported in a Database called "Project_MCB" through the import task. A general schema of the imported file is here:

(i) Creating database:

CREATE DATABASE Project_BCM

(ii) Since we will be using this database for the creation of the tables:

USE Project_MCB

```
/* Creating Database "Project_BCM" */
CREATE DATABASE Project_BCM */

/* Using "Project_BCM" Database */
USE Project_BCM
```

(iii) Creating table "dbo.Country" for storing values from "Country_List.csv".

```sql
/* Creating table for storing Country data */
/* */
CREATE TABLE dbo.Country(Country VARCHAR(100) NULL,
                         Image_File VARCHAR(MAX) NULL,
                         Image_URL VARCHAR(MAX) NULL,
                         Alpha_2 VARCHAR(100) NULL,
                         Alpha_3 VARCHAR(100) NULL,
                         Country_Code VARCHAR(MAX) NULL,
                         iso_3166_2 VARCHAR(100) NULL,
                         Region VARCHAR(100) NULL,
                         Sub_Region VARCHAR(100) NULL,
                         Intermediate_Region VARCHAR(100) NULL,
                         Region_Code VARCHAR(MAX) NULL,
                         Sub_Region_Code VARCHAR(MAX) NULL,
                         Intermediate_Region_Code VARCHAR(MAX) NULL)

/* Note: "dbo.Country" will consume data from the file "Country_List.csv". */
```

(iv) Creating tables to import raw data from .csv files.

```sql
/* (a) - Creating table "Raw_WHR_2016" to store raw data from "HR_2016.csv". */

CREATE TABLE dbo.Raw_WHR_2016(Country VARCHAR(100) NULL,
                              Happiness_Score DECIMAL(6,3),
                              Lower_Confidence_Interval DECIMAL(7,5),
                              Upper_Confidence_Interval DECIMAL(7,5),
                              GDP_per_Capita DECIMAL(7,5),
                              Family DECIMAL(7,5),
                              Health DECIMAL(7,5),
                              Freedom DECIMAL(7,5),
                              Trust DECIMAL(7,5),
                              Generosity DECIMAL(7,5),
                              Dystopia DECIMAL(7,5))

/* (b) - Creating table "Raw_WHR_2017" to store raw data from "happiNess_report_2017.csv". */

CREATE TABLE dbo.Raw_WHR_2017(Country VARCHAR(100) NULL,
                              Happiness_Score DECIMAL(6,3),
                              Whisker_High DECIMAL(7,5),
                              Whisker_Low DECIMAL(7,5),
                              GDP_per_Capita DECIMAL(7,5),
                              Family DECIMAL(7,5),
                              Health DECIMAL(7,5),
                              Freedom DECIMAL(7,5),
                              Generosity DECIMAL(7,5),
                              Trust DECIMAL(7,5),
                              Dystopia DECIMAL(7,5))

/* (c) - Creating table "Raw_WHR_2018" to store raw data from "2018.csv". */

CREATE TABLE dbo.Raw_WHR_2018(Country VARCHAR(100) NULL,
                              Happiness_Score DECIMAL(6,3),
                              GDP_per_Capita DECIMAL(7,5),
                              Family DECIMAL(7,5),
                              Health DECIMAL(7,5),
                              Freedom DECIMAL(7,5),
                              Generosity DECIMAL(7,5),
                              Trust DECIMAL(7,5))
```

```
/* (d) - Creating table "Raw_WHR_2019" to store raw data from "report_2019.csv". */

CREATE TABLE dbo.Raw_WHR_2019(Country VARCHAR(100) NULL,
                              Happiness_Score DECIMAL(6,3),
                              GDP_per_Capita DECIMAL(7,5),
                              Family DECIMAL(7,5),
                              Health DECIMAL(7,5),
                              Freedom DECIMAL(7,5),
                              Generosity DECIMAL(7,5),
                              Trust DECIMAL(7,5))
```

Using “TRUNCATE” to remove any rows in the tables before importing data.

```
/* Using "TRUNCATE" to remove all rows (data) from a table and "BULK INSERT" to populate the tables. */

TRUNCATE TABLE dbo.Country
TRUNCATE TABLE dbo.Raw_WHR_2016
TRUNCATE TABLE dbo.Raw_WHR_2017
TRUNCATE TABLE dbo.Raw_WHR_2018
TRUNCATE TABLE dbo.Raw_WHR_2019
```

Task2.

Inserting data into table “dbo.Country”

```
BULK INSERT dbo.Country
FROM 'C:\Users\pkavi\Documents\MCB_Assignment\Country_List.csv'
WITH
(
    FIRSTROW=2, /* Import of data starts as from row 2, else header will be imported as well. */
    FORMAT='CSV'
)
```

```
SELECT *
FROM dbo.Country
```

100 %

Results | Messages

	Country	Image_File	Image_URL	Alpha_2	Alpha_3	Country_Code	iso_3166_2	Region	Sub_Region	Intermediate_Region	Region_Code	Sub_Region_Code	Intermedi
1	Afghanistan	Flag_of_Afghanistan.svg	https://upload.wikimedia.org/wikipedia/commons/9...	AF	AFG	4	ISO 3166-2:AF	Asia	Southern Asia	NULL	142	34	NULL
2	Albania	Flag_of_Albania.svg	https://upload.wikimedia.org/wikipedia/commons/3...	AL	ALB	8	ISO 3166-2:AL	Europe	Southern Europe	NULL	150	39	NULL
3	Algeria	Flag_of_Algeria.svg	https://upload.wikimedia.org/wikipedia/commons/7...	DZ	DZA	12	ISO 3166-2:DZ	Africa	Northern Africa	NULL	2	15	NULL
4	Andorra	Flag_of_Andorra.svg	https://upload.wikimedia.org/wikipedia/commons/1...	AD	AND	20	ISO 3166-2:AD	Europe	Southern Europe	NULL	150	39	NULL
5	Angola	Flag_of_Angola.svg	https://upload.wikimedia.org/wikipedia/commons/9...	AO	AGO	24	ISO 3166-2:AO	Africa	Sub-Saharan Africa	Middle Africa	2	202	17
6	Antigua and Barbuda	Flag_of_Antigua_and_Barbuda.svg	https://upload.wikimedia.org/wikipedia/commons/8...	AG	ATG	28	ISO 3166-2:AG	Americas	Latin America and the Caribbean	Caribbean	19	419	29
7	Argentina	Flag_of_Argentina.svg	https://upload.wikimedia.org/wikipedia/commons/1...	AR	ARG	32	ISO 3166-2:AR	Americas	Latin America and the Caribbean	South America	19	419	5
8	Armenia	Flag_of_Armenia.svg	https://upload.wikimedia.org/wikipedia/commons/2...	AM	ARM	51	ISO 3166-2:AM	Asia	Western Asia	NULL	142	145	NULL
9	Australia	Flag_of_Australia_%28converted%29.svg	https://upload.wikimedia.org/wikipedia/commons/8...	AU	AUS	36	ISO 3166-2:AU	Oceania	Australia and New Zealand	NULL	9	53	NULL
10	Austria	Flag_of_Austria.svg	https://upload.wikimedia.org/wikipedia/commons/4...	AT	AUT	40	ISO 3166-2:AT	Europe	Western Europe	NULL	150	155	NULL
11	Azerbaijan	Flag_of_Azerbaijan.svg	https://upload.wikimedia.org/wikipedia/commons/d...	AZ	AZE	31	ISO 3166-2:AZ	Asia	Western Asia	NULL	142	145	NULL

Query executed successfully. | DESKTOP-S2LJ53N\SQLEXPRESS ... | DESKTOP-S2LJ53N\pkavi ... | Project_BCM | 00:00:00 | 273 rows

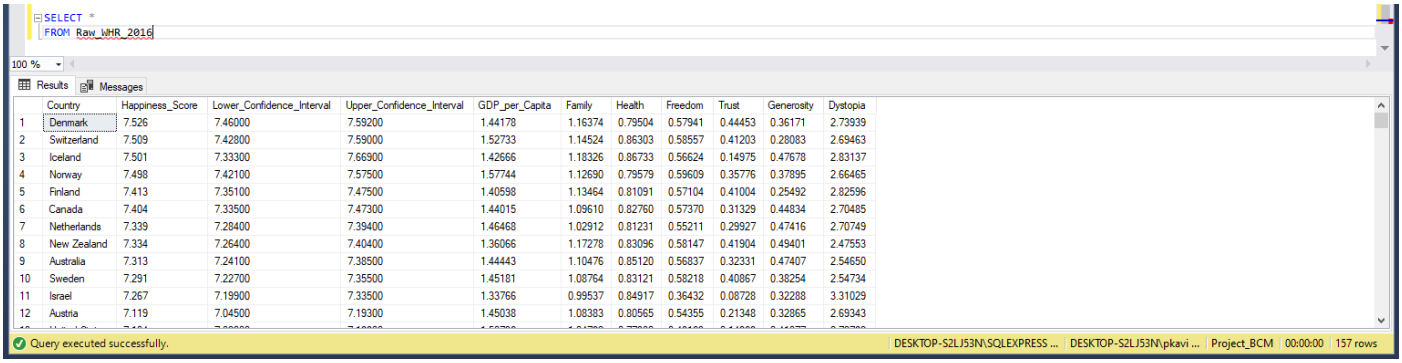
Inserting data into the report tables.

```
/* (a) - Inserting data into "dbo.Raw_WHR_2016" from "HR_2016.csv". */

BULK INSERT dbo.Raw_WHR_2016
FROM 'C:\Users\pkavi\Documents\MCB_Assignment\Data Files\HR_2016.csv'
WITH
(

       FIRSTROW=2, /* Import of data starts as from row 2, else header will be imported as well. */
       FORMAT='CSV'

)
```



```
SELECT *
FROM Raw_WHR_2016
```

100 %

Results | Messages

	Country	Happiness_Score	Lower_Confidence_Interval	Upper_Confidence_Interval	GDP_per_Capita	Family	Health	Freedom	Trust	Generosity	Dystopia
1	Denmark	7.526	7.46000	7.59200	1.44178	1.16374	0.79504	0.57941	0.44453	0.36171	2.73939
2	Switzerland	7.509	7.42800	7.59000	1.52733	1.14524	0.86303	0.58557	0.41203	0.28083	2.69463
3	Iceland	7.501	7.33300	7.66900	1.42666	1.18326	0.86733	0.56624	0.14975	0.47678	2.83137
4	Norway	7.498	7.42100	7.57500	1.57744	1.12690	0.79579	0.59609	0.35776	0.37895	2.66465
5	Finland	7.413	7.35100	7.47500	1.40598	1.13464	0.81091	0.57104	0.41004	0.25492	2.82596
6	Canada	7.404	7.33500	7.47300	1.44015	1.09610	0.82760	0.57370	0.31329	0.44834	2.70485
7	Netherlands	7.339	7.28400	7.39400	1.46468	1.02912	0.81231	0.55211	0.29927	0.47416	2.70749
8	New Zealand	7.334	7.26400	7.40400	1.36066	1.17278	0.83096	0.58147	0.41904	0.49401	2.47553
9	Australia	7.313	7.24100	7.38500	1.44443	1.10476	0.85120	0.56837	0.32331	0.47407	2.54650
10	Sweden	7.291	7.22700	7.35500	1.45181	1.08764	0.83121	0.58218	0.40867	0.38254	2.54734
11	Israel	7.267	7.19900	7.33500	1.33766	0.99537	0.84917	0.36432	0.08728	0.32288	3.31029
12	Austria	7.119	7.04500	7.19300	1.45038	1.08383	0.80565	0.54355	0.21348	0.32865	2.69343

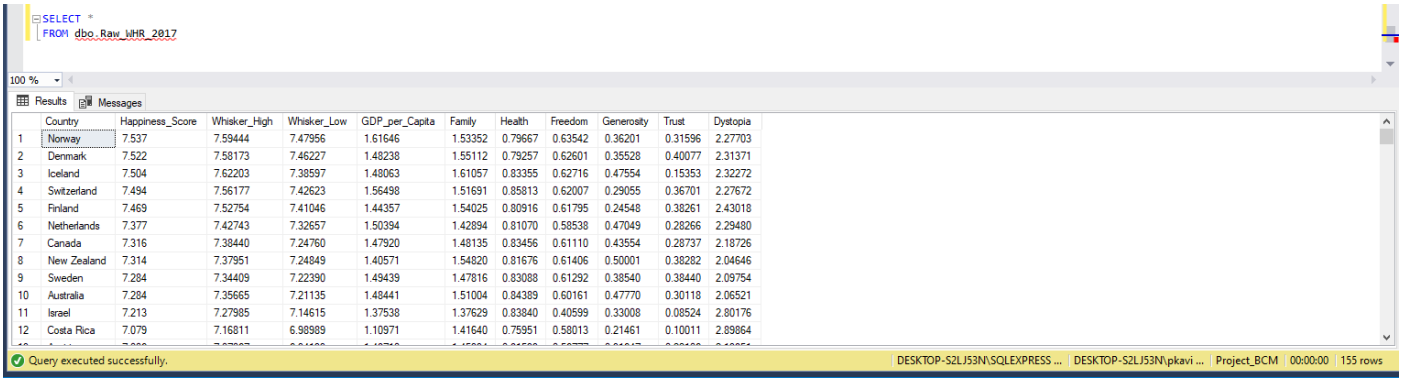
Query executed successfully. | DESKTOP-S2LJ53N\SQLEXPRESS ... | DESKTOP-S2LJ53N\pkavi ... | Project_BCM | 00:00:00 | 157 rows

```
/* (b) - Inserting data into "dbo.Raw_WHR_2017" from "happiNess_report_2017.csv". */

BULK INSERT dbo.Raw_WHR_2017
FROM 'C:\Users\pkavi\Documents\MCB_Assignment\Data Files\happiNess_report_2017.csv'
WITH
(

      FIRSTROW=2, /* Import of data starts as from row 2, else header will be imported as well. */
      FORMAT='CSV'

)
```



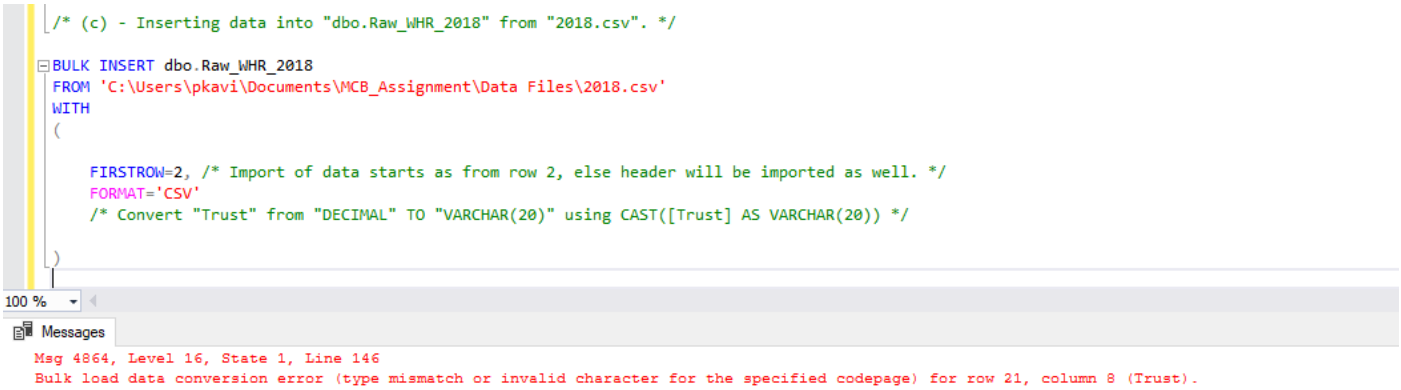
```
SELECT *
FROM dbo.Raw_WHR_2017
```

100 %

Results | Messages

	Country	Happiness_Score	Whisker_High	Whisker_Low	GDP_per_Capita	Family	Health	Freedom	Generosity	Trust	Dystopia
1	Norway	7.537	7.59444	7.47956	1.61646	1.53352	0.79667	0.63542	0.36201	0.31596	2.27703
2	Denmark	7.522	7.58173	7.46227	1.48238	1.55112	0.79257	0.62601	0.35528	0.40077	2.31371
3	Iceland	7.504	7.62203	7.38597	1.48063	1.61057	0.83355	0.62716	0.47554	0.15353	2.32272
4	Switzerland	7.494	7.56177	7.42623	1.56498	1.51691	0.85813	0.62007	0.29055	0.36701	2.27672
5	Finland	7.469	7.52754	7.41046	1.44357	1.54025	0.80916	0.61795	0.24548	0.38261	2.43018
6	Netherlands	7.377	7.42743	7.32657	1.50394	1.42894	0.81070	0.58538	0.47049	0.28266	2.29480
7	Canada	7.316	7.38440	7.24760	1.47920	1.48135	0.83456	0.61110	0.43554	0.28737	2.18726
8	New Zealand	7.314	7.37951	7.24849	1.40571	1.54820	0.81676	0.61406	0.50001	0.38282	2.04646
9	Sweden	7.284	7.34409	7.22390	1.49439	1.47816	0.83088	0.61292	0.38540	0.38440	2.09754
10	Australia	7.284	7.35665	7.21135	1.48441	1.51004	0.84389	0.60161	0.47770	0.30118	2.06521
11	Israel	7.213	7.27985	7.14615	1.37538	1.37629	0.83840	0.40599	0.33008	0.08524	2.80176
12	Costa Rica	7.079	7.16811	6.98989	1.10971	1.41640	0.75951	0.58013	0.21461	0.10011	2.89864

Query executed successfully. | DESKTOP-S2LJ53N\SQLEXPRESS ... | DESKTOP-S2LJ53N\pkavi ... | Project_BCM | 00:00:00 | 155 rows



```
/* (c) - Inserting data into "dbo.Raw_WHR_2018" from "2018.csv". */

BULK INSERT dbo.Raw_WHR_2018
FROM 'C:\Users\pkavi\Documents\MCB_Assignment\Data Files\2018.csv'
WITH
(

    FIRSTROW=2, /* Import of data starts as from row 2, else header will be imported as well. */
    FORMAT='CSV'
    /* Convert "Trust" from "DECIMAL" TO "VARCHAR(20)" using CAST([Trust] AS VARCHAR(20)) */

)
```

100 %

Messages

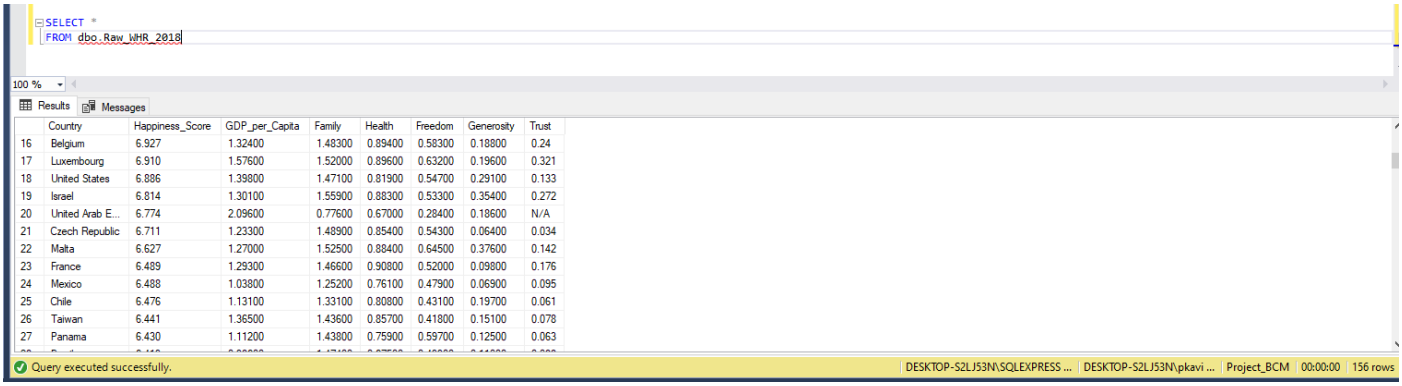
```
Msg 4864, Level 16, State 1, Line 146
Bulk load data conversion error (type mismatch or invalid character for the specified codepage) for row 21, column 8 (Trust).
```

Error:

"Msg 4864, Level 16, State 1, Line 146
Bulk load data conversion error (type mismatch or invalid character for the specified codepage) for row 21, column 8 (Trust).

Completion time: 2022-03-26T07:51:15.5477724+04:00"

This is because of "N/A" value in table "Trust" for "Country – United Arab Emirates".
We have to use "CAST" to convert "Trust" as VARCHAR(20) to populate the dataset.



```
SELECT *
FROM dbo.Raw_WHR_2018
```

100 %

Results | Messages

	Country	Happiness_Score	GDP_per_Capita	Family	Health	Freedom	Generosity	Trust
16	Belgium	6.927	1.32400	1.48300	0.89400	0.58300	0.18800	0.24
17	Luxembourg	6.910	1.57600	1.52000	0.89600	0.63200	0.19600	0.321
18	United States	6.886	1.39800	1.47100	0.81900	0.54700	0.29100	0.133
19	Israel	6.814	1.30100	1.55900	0.88300	0.53300	0.35400	0.272
20	United Arab E...	6.774	2.09600	0.77600	0.67000	0.28400	0.18600	N/A
21	Czech Republic	6.711	1.23300	1.48900	0.85400	0.54300	0.06400	0.034
22	Malta	6.627	1.27000	1.52500	0.88400	0.64500	0.37600	0.142
23	France	6.489	1.29300	1.46600	0.90800	0.52000	0.09800	0.176
24	Mexico	6.488	1.03800	1.25200	0.76100	0.47900	0.06900	0.095
25	Chile	6.476	1.13100	1.33100	0.80800	0.43100	0.19700	0.061
26	Taiwan	6.441	1.36500	1.43600	0.85700	0.41800	0.15100	0.078
27	Panama	6.430	1.11200	1.43800	0.75900	0.59700	0.12500	0.063

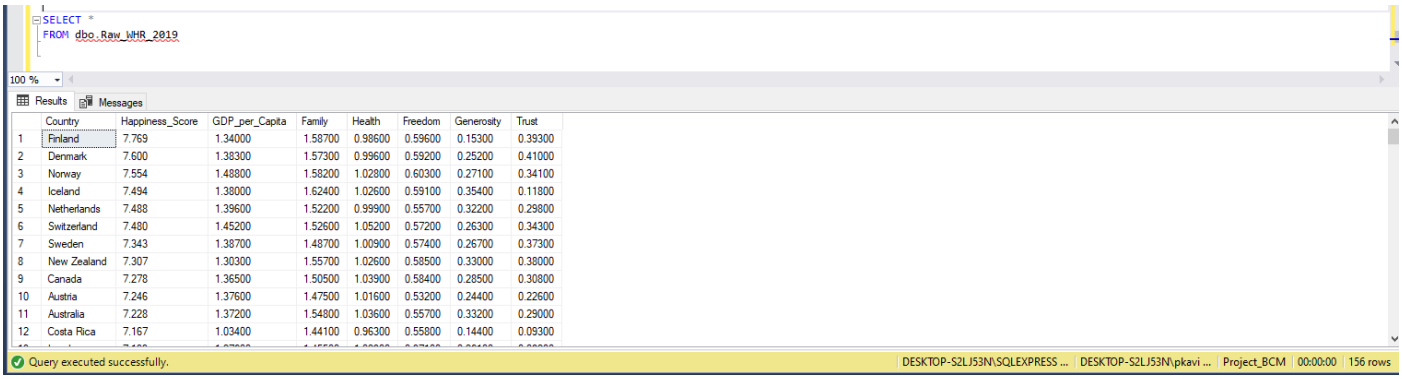
Query executed successfully. | DESKTOP-S2LJ53N\SQLEXPRESS ... | DESKTOP-S2LJ53N\pkavi ... | Project_BCM | 00:00:00 | 156 rows

```
/* (d) - Inserting data into "dbo.Raw_WHR_2019" from "report_2019.csv". */

BULK INSERT dbo.Raw_WHR_2019
FROM 'C:\Users\pkavi\Documents\MCB_Assignment\Data Files\report_2019.csv'
WITH
(

	FIRSTROW=2, /* Import of data starts as from row 2, else header will be imported as well. */
	FORMAT='CSV'

)
```



```
SELECT *
FROM dbo.Raw_WHR_2019
```

100 %

Results | Messages

	Country	Happiness_Score	GDP_per_Capita	Family	Health	Freedom	Generosity	Trust
1	Finland	7.769	1.34000	1.58700	0.98600	0.59600	0.15300	0.39300
2	Denmark	7.600	1.38300	1.57300	0.99600	0.59200	0.25200	0.41000
3	Norway	7.554	1.48800	1.58200	1.02800	0.60300	0.27100	0.34100
4	Iceland	7.494	1.38000	1.62400	1.02600	0.59100	0.35400	0.11800
5	Netherlands	7.488	1.39600	1.52200	0.99900	0.55700	0.32200	0.29800
6	Switzerland	7.480	1.45200	1.52600	1.05200	0.57200	0.26300	0.34300
7	Sweden	7.343	1.38700	1.48700	1.00900	0.57400	0.26700	0.37300
8	New Zealand	7.307	1.30300	1.55700	1.02600	0.58500	0.33000	0.38000
9	Canada	7.278	1.36500	1.50500	1.03900	0.58400	0.28500	0.30800
10	Austria	7.246	1.37600	1.47500	1.01600	0.53200	0.24400	0.22600
11	Australia	7.228	1.37200	1.54800	1.03600	0.55700	0.33200	0.29000
12	Costa Rica	7.167	1.03400	1.44100	0.96300	0.55800	0.14400	0.09300

Query executed successfully. | DESKTOP-S2LJ53N\SQLEXPRESS ... | DESKTOP-S2LJ53N\pkavi ... | Project_BCM | 00:00:00 | 156 rows

Next, we will create tables “dbo.WHR_2016”, “dbo.WHR_2017”, “dbo.WHR_2018” and “dbo.WHR_2019” to perform manipulations and analysis.

```
/* (a) - Inserting data into "dbo.WHR_2016" from "Raw_WHR_2016". */


INSERT INTO dbo.WHR_2016

SELECT Country, Happiness_Score,

	RANK() OVER(

		ORDER BY Happiness_Score DESC
		) AS Happiness_Rank,

		GDP_per_Capita, Family, Health,
		Freedom, Trust, Generosity, Dystopia

FROM dbo.Raw_WHR_2016


/* List the data which has been inserted in table "dbo.WHR_2016". */

SELECT *
FROM dbo.WHR_2016
```

```sql
/* (b) - Inserting data into "dbo.WHR_2017" from "Raw_WHR_2017". */


INSERT INTO dbo.WHR_2017

SELECT Country, Happiness_Score,

       RANK() OVER(

              ORDER BY Happiness_Score DESC
              ) AS Happiness_Rank,

              GDP_per_Capita, Family, Health,
              Freedom, Trust, Generosity, Dystopia

FROM dbo.Raw_WHR_2017


/* List the data which has been inserted in table "dbo.WHR_2017". */

SELECT *
FROM dbo.WHR_2017


/* (c) - Inserting data into "dbo.WHR_2018" from "Raw_WHR_2018". */


INSERT INTO dbo.WHR_2018

SELECT Country, Happiness_Score,

       RANK() OVER(

              ORDER BY Happiness_Score DESC
              ) AS Happiness_Rank,

              GDP_per_Capita, Family, Health,
              Freedom, Trust, Generosity

FROM dbo.Raw_WHR_2018

/* List the data which has been inserted in table "dbo.WHR_2018". */

SELECT *
FROM dbo.WHR_2018


/* (d) - Inserting data into "dbo.WHR_2019" from "Raw_WHR_2019". */


INSERT INTO dbo.WHR_2019

SELECT Country, Happiness_Score,

       RANK() OVER(

              ORDER BY Happiness_Score DESC
              ) AS Happiness_Rank,

              GDP_per_Capita, Family, Health,
              Freedom, Trust, Generosity

FROM dbo.Raw_WHR_2019


/* List the data which has been inserted in table "dbo.WHR_2019". */

SELECT *
FROM dbo.WHR_2019
```

Data from “dbo.WHR_2016” with “Happiness_Rank” for each year.

```
SELECT *
FROM WHR_2016
```

100 %

Results | Messages

	Country	Happiness_Score	Happiness_Rank	GDP_per_Capita	Family	Health	Freedom	Trust	Generosity	Dystopia
1	Denmark	7.526	1	1.44178	1.16374	0.79504	0.57941	0.44453	0.36171	2.73939
2	Switzerland	7.509	2	1.52733	1.14524	0.86303	0.58557	0.41203	0.28083	2.69463
3	Iceland	7.501	3	1.42666	1.18326	0.86733	0.56624	0.14975	0.47678	2.83137
4	Norway	7.498	4	1.57744	1.12690	0.79579	0.59609	0.35776	0.37895	2.66465
5	Finland	7.413	5	1.40598	1.13464	0.81091	0.57104	0.41004	0.25492	2.82596
6	Canada	7.404	6	1.44015	1.09610	0.82760	0.57370	0.31329	0.44834	2.70485
7	Netherlands	7.339	7	1.46468	1.02912	0.81231	0.55211	0.29927	0.47416	2.70749
8	New Zealand	7.334	8	1.36066	1.17278	0.83096	0.58147	0.41904	0.49401	2.47553
9	Australia	7.313	9	1.44443	1.10476	0.85120	0.56837	0.32331	0.47407	2.54650
10	Sweden	7.291	10	1.45181	1.08764	0.83121	0.58218	0.40867	0.38254	2.54734
11	Israel	7.267	11	1.33766	0.99537	0.84917	0.36432	0.08728	0.32288	3.31029
12	Austria	7.119	12	1.45038	1.08383	0.80565	0.54355	0.21348	0.32865	2.69343

Query executed successfully. DESKTOP-S2LJ53N\SQLEXPRESS ... DESKTOP-S2LJ53N\pkavi ... Project_BCM 00:00:00 157 rows

Data from “dbo.WHR_2017” with “Happiness_Rank” for each year.

```
SELECT *
FROM WHR_2017
```

100 %

Results | Messages

	Country	Happiness_Score	Happiness_Rank	GDP_per_Capita	Family	Health	Freedom	Trust	Generosity	Dystopia
1	Norway	7.537	1	1.61646	1.53352	0.79667	0.63542	0.31596	0.36201	2.27703
2	Denmark	7.522	2	1.48238	1.55112	0.79257	0.62601	0.40077	0.35528	2.31371
3	Iceland	7.504	3	1.48063	1.61057	0.83355	0.62716	0.15353	0.47554	2.32272
4	Switzerland	7.494	4	1.56498	1.51691	0.85813	0.62007	0.36701	0.29055	2.27672
5	Finland	7.469	5	1.44357	1.54025	0.80916	0.61795	0.38261	0.24548	2.43018
6	Netherlands	7.377	6	1.50394	1.42894	0.81070	0.58538	0.28266	0.47049	2.29480
7	Canada	7.316	7	1.47920	1.48135	0.83456	0.61110	0.28737	0.43554	2.18726
8	New Zealand	7.314	8	1.40571	1.54820	0.81676	0.61406	0.38282	0.50001	2.04646
9	Sweden	7.284	9	1.49439	1.47816	0.83088	0.61292	0.38440	0.38540	2.09754
10	Australia	7.284	9	1.48441	1.51004	0.84389	0.60161	0.30118	0.47770	2.06521
11	Israel	7.213	11	1.37538	1.37629	0.83840	0.40599	0.08524	0.33008	2.80176
12	Costa Rica	7.079	12	1.10971	1.41640	0.75951	0.58013	0.10011	0.21461	2.89864

Query executed successfully. DESKTOP-S2LJ53N\SQLEXPRESS ... DESKTOP-S2LJ53N\pkavi ... Project_BCM 00:00:00 155 rows

Data from “dbo.WHR_2018” with “Happiness_Rank” for each year.

```
SELECT *
FROM WHR_2018
```

100 %

Results | Messages

	Country	Happiness_Score	Happiness_Rank	GDP_per_Capita	Family	Health	Freedom	Trust	Generosity
1	Finland	7.632	1	1.30500	1.59200	0.87400	0.68100	0.393	0.20200
2	Norway	7.594	2	1.45600	1.58200	0.86100	0.68600	0.34	0.28600
3	Denmark	7.555	3	1.35100	1.59000	0.86800	0.68300	0.408	0.28400
4	Iceland	7.495	4	1.34300	1.64400	0.91400	0.67700	0.138	0.35300
5	Switzerland	7.487	5	1.42000	1.54900	0.92700	0.66000	0.357	0.25600
6	Netherlands	7.441	6	1.36100	1.48800	0.87800	0.63800	0.295	0.33300
7	Canada	7.328	7	1.33000	1.53200	0.89600	0.65300	0.291	0.32100
8	New Zealand	7.324	8	1.26800	1.60100	0.87600	0.66900	0.389	0.36500
9	Sweden	7.314	9	1.35500	1.50100	0.91300	0.65900	0.383	0.28500
10	Australia	7.272	10	1.34000	1.57300	0.91000	0.64700	0.302	0.36100
11	United Kingdom	7.190	11	1.24400	1.43300	0.88800	0.46400	0.082	0.26200
12	Austria	7.139	12	1.34100	1.50400	0.89100	0.61700	0.224	0.24200

Query executed successfully. DESKTOP-S2LJ53N\SQLEXPRESS ... DESKTOP-S2LJ53N\pkavi ... Project_BCM 00:00:00 156 rows

Data from “dbo.WHR_2019” with “Happiness_Rank” for each year.

```
SELECT *
FROM WHR_2019
```

100 %

Results | Messages

	Country	Happiness_Score	Happiness_Rank	GDP_per_Capita	Family	Health	Freedom	Trust	Generosity
1	Finland	7.769	1	1.34000	1.58700	0.98600	0.59600	0.39300	0.15300
2	Denmark	7.600	2	1.38300	1.57300	0.99600	0.59200	0.41000	0.25200
3	Norway	7.554	3	1.48800	1.58200	1.02800	0.60300	0.34100	0.27100
4	Iceland	7.494	4	1.38000	1.62400	1.02600	0.59100	0.11800	0.35400
5	Netherlands	7.488	5	1.39600	1.52200	0.99900	0.55700	0.29800	0.32200
6	Switzerland	7.480	6	1.45200	1.52600	1.05200	0.57200	0.34300	0.26300
7	Sweden	7.343	7	1.38700	1.48700	1.00900	0.57400	0.37300	0.26700
8	New Zealand	7.307	8	1.30300	1.55700	1.02600	0.58500	0.38000	0.33000
9	Canada	7.278	9	1.36500	1.50500	1.03900	0.58400	0.30800	0.28500
10	Austria	7.246	10	1.37600	1.47500	1.01600	0.53200	0.22600	0.24400
11	Australia	7.228	11	1.37200	1.54800	1.03600	0.55700	0.29000	0.33200
12	Costa Rica	7.167	12	1.03400	1.44100	0.96300	0.55800	0.09300	0.14400

Query executed successfully. DESKTOP-S2LJ53N\SQLEXPRESS ... DESKTOP-S2LJ53N\pkavi ... Project_BCM 00:00:00 156 rows

```
/* Creating table "dbo.CountriesVerification" to verify consistency of data in the four datasets. */

CREATE TABLE dbo.CountriesVerification(Country VARCHAR(100) NULL)


/* Inserting data in "dbo.CountriesVerification" table from WHR_2016, WHR_2017, WHR_2018, WHR_2019".
*/

INSERT INTO dbo.CountriesVerification

       SELECT DISTINCT Country
       FROM dbo.WHR_2016
UNION
       SELECT DISTINCT Country
       FROM dbo.WHR_2017
UNION
       SELECT DISTINCT Country
       FROM dbo.WHR_2018
UNION
       SELECT DISTINCT Country
       FROM dbo.WHR_2019


/* List the data which has been inserted in table "dbo.CountriesVerification". */

SELECT *
FROM dbo.CountriesVerification
```

The total count for combined list is 170 which is higher than WHR_2016 (157), WHR_2017 (155), WHR_2018 (156) and WHR_2019 (156).

This difference implies that there are few countries which are not a part of all four datasets, hence including them in the final combined dataset will result in null values.

-As such we will have to combine the dataset so as to get the countries which are not part of the four datasets:

```
/* Creating table "dbo.Combined1" to assign year to the data from the four datasets. */

CREATE TABLE dbo.Combined1(Year INT, Country VARCHAR(100) NULL,
                                                        Happiness_Score DECIMAL(6,3),
                                                        Happiness_Rank INT,
                                                        GDP_per_Capita DECIMAL(7,5),
                                                        Family DECIMAL(7,5),
                                                        Health DECIMAL(7,5),
                                                        Freedom DECIMAL(7,5),
                                                        Trust DECIMAL(7,5),
                                                        Generosity DECIMAL(7,5)
)


/* Inserting data in "dbo.Combined1" for comparison. */

INSERT INTO dbo.Combined1

       SELECT 2016 as Year, Country,
              Happiness_Score,

              RANK() OVER(

              ORDER BY Happiness_Score DESC
              ) AS Happiness_Rank,

              GDP_per_Capita, Family, Health,
              Freedom, Trust, Generosity

       FROM dbo.WHR_2016
```

```
UNION

	SELECT 2017 as Year, Country,
		Happiness_Score,

		RANK() OVER(

		ORDER BY Happiness_Score DESC
		) AS Happiness_Rank,

		GDP_per_Capita, Family, Health,
		Freedom, Trust, Generosity

	FROM dbo.WHR_2017

UNION

	SELECT 2018 as Year, Country,
		Happiness_Score,

		RANK() OVER(

		ORDER BY Happiness_Score DESC
		) AS Happiness_Rank,

		GDP_per_Capita, Family, Health,
		Freedom, Trust, Generosity

	FROM dbo.WHR_2018

UNION

	SELECT 2019 as Year, Country,
		Happiness_Score,

		RANK() OVER(

		ORDER BY Happiness_Score DESC
		) AS Happiness_Rank,

		GDP_per_Capita, Family, Health,
		Freedom, Trust, Generosity

	FROM dbo.WHR_2019


/* List the data which has been inserted in table "dbo.Combined1". */

SELECT *
FROM dbo.Combined1



/* Creating table "dbo.Combined2" to verify missing data from the four datasets. */

CREATE TABLE dbo.Combined2(Year INT DEFAULT NULL, Country VARCHAR(100) NULL,
                                                        Happiness_Score DECIMAL(6,3),
                                                        Happiness_Rank INT DEFAULT NULL,
                                                        GDP_per_Capita DECIMAL(7,5),
                                                        Family DECIMAL(7,5),
                                                        Health DECIMAL(7,5),
                                                        Freedom DECIMAL(7,5),
                                                        Trust DECIMAL(7,5),
                                                        Generosity DECIMAL(7,5)
)
```

```sql
/* Inserting data in "dbo.Combined2" for comparison to obtain the list of missing countries and year
from the four datasets. */

INSERT INTO dbo.Combined2

SELECT a.Year, a.Country,
       b.Happiness_Score, b.Happiness_Rank,
       b.GDP_per_Capita, b.Family, b.Health,
       b.Freedom, b.Trust, b.Generosity

FROM

(
SELECT DISTINCT c.Country, d.Year

FROM

       (SELECT e.Country
        FROM dbo.WHR_2016 e

        INNER JOIN

        dbo.WHR_2017 f
        ON e.Country = f.Country

        INNER JOIN

        dbo.WHR_2018 g
        ON e.Country = g.Country

        INNER JOIN

        dbo.WHR_2019 h
        ON e.Country = h.Country) c

       FULL JOIN

       (SELECT DISTINCT Year, Country
        FROM dbo.Combined1) d
        ON c.Country = d.Country
        WHERE c.Country is not NULL AND d.country is not NULL) a

LEFT JOIN

        dbo.Combined1 b
ON a.Country = b.Country AND a.Year = b.Year



/* List the data from table "dbo.Combined2" */

SELECT *
FROM dbo.Combined2
```

Task 3:

```sql
SELECT Country, Happiness_Score,

       (CASE WHEN Happiness_Score < 2.6 THEN 'RED'
                  WHEN Happiness_Score BETWEEN 2.6 AND 5.6 THEN 'AMBER'
                  WHEN Happiness_Score > 5.6 THEN 'GREEN'
             END) AS Happiness_Status,

             GDP_per_Capita, Family, Health,
       Freedom, Trust, Generosity

FROM dbo.Combined2
```

```sql
SELECT Country, Happiness_Score,

    (CASE WHEN Happiness_Score < 2.6 THEN 'RED'
          WHEN Happiness_Score BETWEEN 2.6 AND 5.6 THEN 'AMBER'
          WHEN Happiness_Score > 5.6 THEN 'GREEN'
      END) AS Happiness_Status,

      GDP_per_Capita, Family, Health,
  Freedom, Trust, Generosity

FROM dbo.Combined2
```

100 %

Results | Messages

	Country	Happiness_Score	Happiness_Status	GDP_per_Capita	Family	Health	Freedom	Trust	Generosity
1	Afghanistan	3.360	AMBER	0.38227	0.11037	0.17344	0.16430	0.07112	0.31268
2	Albania	4.655	AMBER	0.95530	0.50163	0.73007	0.31866	0.05301	0.16840
3	Algeria	6.355	GREEN	1.05266	0.83309	0.61804	0.21006	0.16157	0.07044
4	Argentina	6.650	GREEN	1.15137	1.06612	0.69711	0.42284	0.07296	0.10989
5	Armenia	4.360	AMBER	0.86086	0.62477	0.64083	0.14037	0.03616	0.07793
6	Australia	7.313	GREEN	1.44443	1.10476	0.85120	0.56837	0.32331	0.47407
7	Austria	7.119	GREEN	1.45038	1.08383	0.80565	0.54355	0.21348	0.32865
8	Azerbaijan	5.291	AMBER	1.12373	0.76042	0.54504	0.35327	0.17914	0.05640
9	Bahrain	6.218	GREEN	1.44024	0.94397	0.65696	0.47375	0.25772	0.17147
10	Bangladesh	4.643	AMBER	0.54177	0.24749	0.52989	0.39778	0.12583	0.19132
11	Belarus	5.802	GREEN	1.13062	1.04993	0.63104	0.29091	0.17457	0.13942
12	Belgium	6.929	GREEN	1.42539	1.05249	0.81959	0.51354	0.26248	0.24240

Query executed successfully. | DESKTOP-S2LJ53N\SQLEXPRESS ... | DESKTOP-S2LJ53N\pkavi ... | Project_MCB | 00:00:00 | 576 rows