Assignment 3

Data: Data

We have two tables: Expenditure and Data

Dataset Overview:

1. **Expenditure Dataset:**
    - **Columns:**
        - country: The name of the country.
        - institute_type: Type of institution (e.g., all institutions).
        - direct_expenditure_type: The type of expenditure (e.g., public).
        - 1995, 2000, 2005, 2009, 2010, 2011: Direct expenditure values for respective years.
2. **Data Dataset:**
    - **Columns**
        - world_rank: Global ranking of the institution.
        - institution: Name of the institution.
        - country: Country where the institution is located.
        - national_rank: National ranking of the institution.
        - quality_of_education: Quality score of education.
        - alumni_employment: Alumni employment rank.
        - quality_of_faculty: Quality rank of faculty.
        - publications, influence, citations, broad_impact, patents: Various impact and research-related metrics.
        - score: Overall score.
        - year: The year of the ranking.

Task 1.

Calculate the average public expenditure for the year 2005 across all countries.

Key:

1) Create Measures

- **average_public_expenditure_2005 = CALCULATE(AVERAGE(expenditure[Value]),expenditure[Year]=2005, expenditure[direct_expenditure_type]="Public")**

2) Add a card visual

- In Report View, click on card from the Visualization pane.
- Drag average_public_expenditure_2005 into the Fields section.

Task 2.

Calculate the total publications for institutions in the UK.

Key:

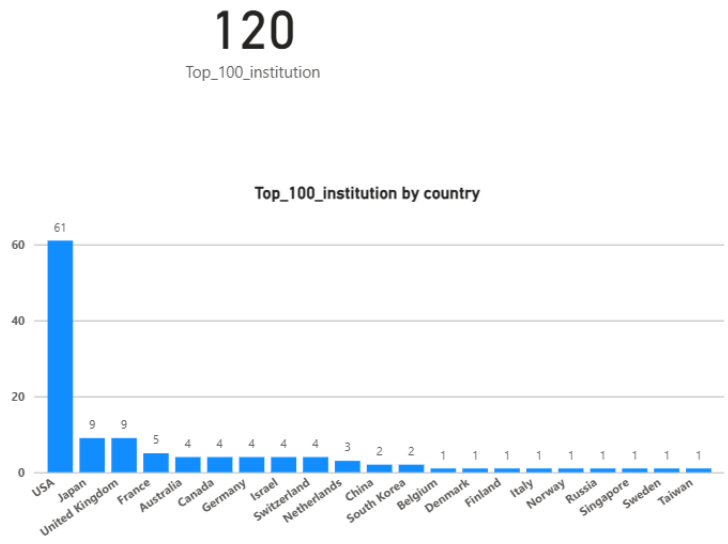
1) Create Measures
    - **total_publications_institutions_UK = CALCULATE(SUM(data[publications]),data[country]="United Kingdom")**
2) Add a Card Visual
    - In Report View, click on card from the Visualization pane.
    - Drag total_publications_institutions_UK into the Fields section.

Task 3.

Filter the dataset to show only institutions with a world rank below 100.

Key:

1) Create Measures
    - **Top_100_institution = CALCULATE(DISTINCTCOUNT(data[institution]),data[world_rank]<100)**
2) Add a Card Visual
    - In Report View, click on card from the Visualization pane.
    - Drag Top_100_institution into the Fields section.
3) Insert a Bar Chart
    - In Report View, click on Bar Chart from the Visualizations pane.
4) Add Fields
    - Drag Country into the X-axis field.
    - Drag Top_100_institution into the Y-axis field.
5) Insert Table Visual
    - Click on Table from the Visualizations pane.
    - Drag Institution and Country into the column section.
    - In the Filters pane, select institution → change the filter type to Top N → set Show items to 100 → drag the Top_100_institution measure into the By value field → click Apply filter.



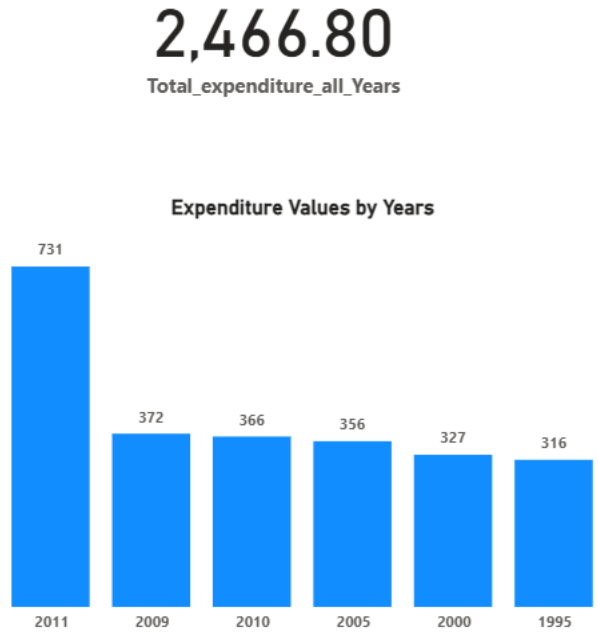
institution	country
Arizona State University	USA
Australian National University	Australia
Boston University	USA
Brown University	USA
California Institute of Technology	USA
Carnegie Mellon University	USA
Case Western Reserve University	USA
Columbia University	USA
Cornell University	USA
Dartmouth College	USA
Duke University	USA
École normale supérieure - Paris	France
École Polytechnique	France
Emory University	USA
Erasmus University Rotterdam	Netherlands
Georgia Institute of Technology	USA
Harvard University	USA
Hebrew University of Jerusalem	Israel
Imperial College London	United Kingdom
Indiana University - Bloomington	USA
Johns Hopkins University	USA
Karolinska Institute	Sweden
Katholieke Universiteit Leuven	Belgium
Keio University	Japan
King's College London	United Kingdom
Kyoto University	Japan
Kyushu University	Japan
Leiden University	Netherlands
Lomonosov Moscow State University	Russia
Ludwig Maximilian University of Munich	Germany

Task 4.

Calculate the total expenditure for all years for each country.

Key:

1) Create Measures
   - **Total_expenditure_all_Years = SUM(expenditure[Value])**
2) Added a Card Visual to display total expenditure.
3) Inserted a Stacked Column Chart to compare expenditure by year.
4) Inserted a Table Visual to show detailed expenditure values.



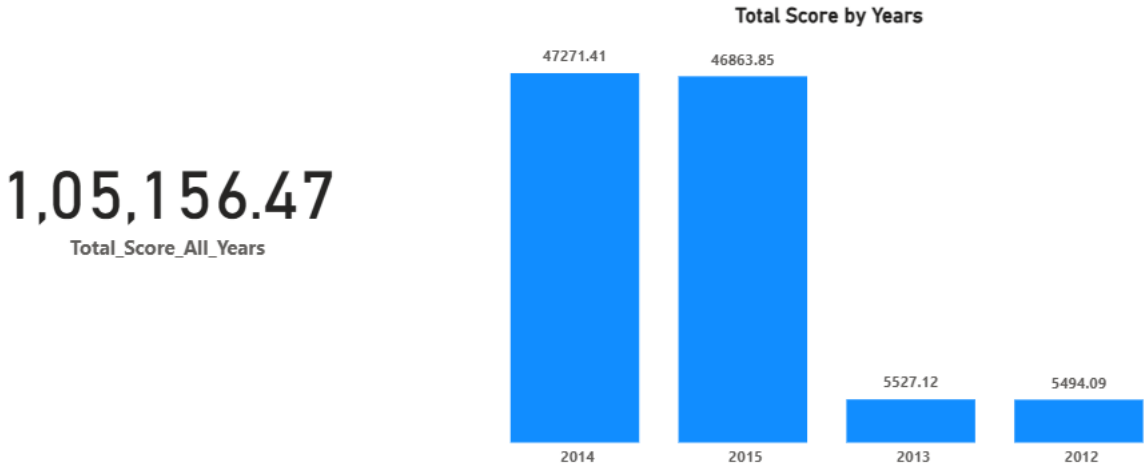
country	Year	Sum of Value
Australia	1995	9.50
Australia	2000	9.70
Australia	2005	9.10
Australia	2009	9.40
Australia	2010	9.70
Australia	2011	23.00
Austria	1995	10.60
Austria	2000	10.90
Austria	2005	10.50
Austria	2009	11.50
Austria	2010	11.20
Austria	2011	21.60
Belgium	1995	9.90
Belgium	2000	10.30
Belgium	2005	11.50
Belgium	2009	12.70
Belgium	2010	12.70
Belgium	2011	24.80
Brazil	1995	0.90
Brazil	2000	0.90
Brazil	2005	9.10
Brazil	2009	11.20
Brazil	2010	11.40
Total		**2,466.80**

Task 5.

Write a DAX formula to ignore any filters on the year column and calculate the total score across all years.

Key:

1) Create Measure
   - **Total_Score_All_Years = CALCULATE(SUM(data[score]),ALL(data[year]))**
2) Added a Card Visual to display Total_Score_All_Years.
3) Inserted a Stacked Column Chart to compare score by year.
4) For Column chart created new measure
   - Total_data_score = SUM(data[score])



Task 6.

Calculate the growth in expenditure for Austria from 1995 to 2000.

Key:

1) Create Measures
   - **Growth_Austria_1995_2000 =**

```
VAR Expend2000 =
   CALCULATE ( SUM ( expenditure[Value] ),
      expenditure[country] = "Austria",
      expenditure[Year] = 2000
   )
VAR Expend1995 =
   CALCULATE ( SUM ( expenditure[Value] ),
      expenditure[country] = "Austria",
      expenditure[Year] = 1995
   )
RETURN
DIVIDE ( Expend2000 - Expend1995, Expend1995 ) * 100
```

2) Added a Card Visual to display Growth_Austria_1995_2000.

Task 7.

Format the expenditure values to include a currency symbol and zero decimal places.

Key:

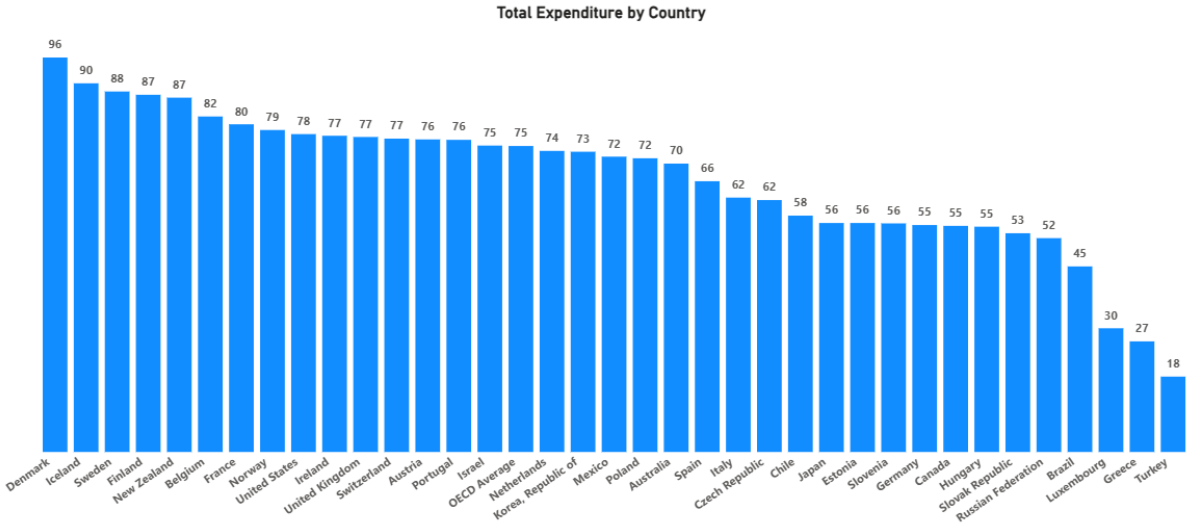
1) Go to Table View in Power BI.
2) Click on the table → expenditure.
3) Select the column Value.
4) On the Ribbon (Column tools) → under Formatting:
    - Change Data type (keep as Decimal/Whole Number).
    - Change Format → choose Currency ($).
    - Set Decimal places = 0.

Task 8.

Calculate the total expenditure for each country.

Key:

1) Add a Stacked Column chart
    - In Report View, click on stacked column chart from the Visualization pane.
    - Drag country column into the X-axis and Total_expenditure_all_Years into the Y-axis Fields section.



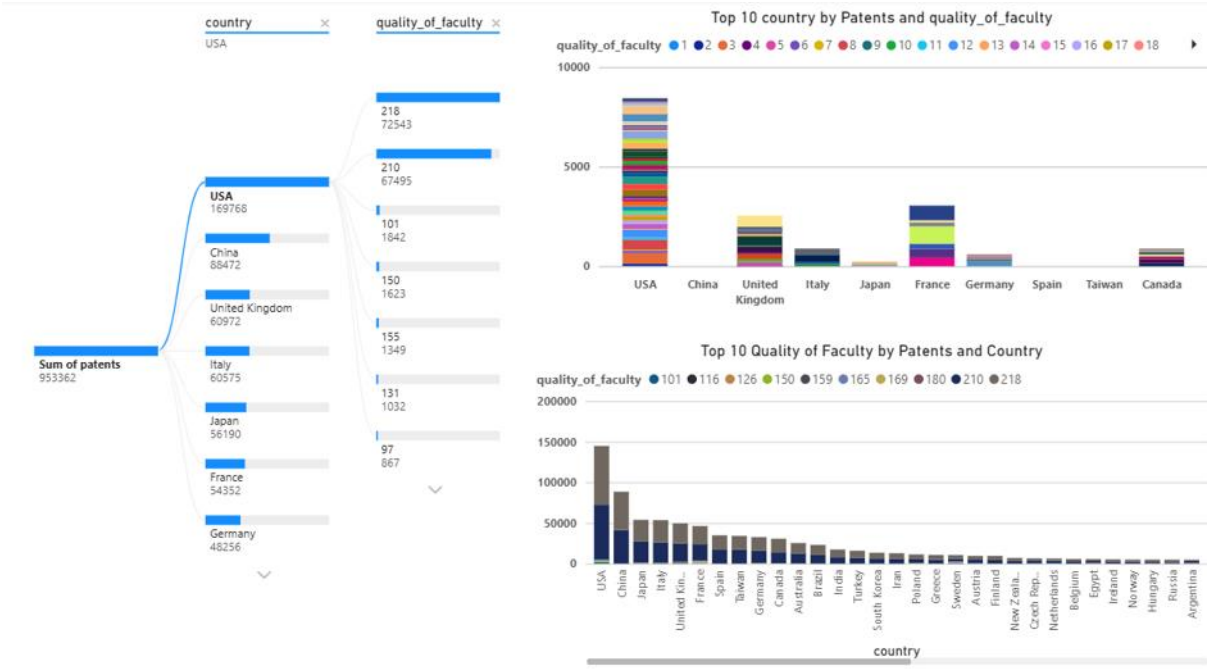
Task 9(a).

Break down the total patents of institutions by country and then by quality of

faculty. Analyze which factors contribute most to the number of patents across different countries.

Key:

1) Add a Decomposition Tree and Stacked Column chart
    - In Decomposition Tree
        - Drag patents column to Analyze, country and quality_of_faculty to Explain by Fields Section.
    - In Stacked Column chart
        - Drag country column into the X-axis, patents into the Y-axis and quality_of_faculty into the Legend Fields section.
        - In the Filters pane, select country → change the filter type to Top N → set Show items to 10 → drag the patents into the by value field → click Apply filter.
    - In another Stacked Column chart
        - Same Process for X-axis, Y-axis and Legend.
        - In the Filters pane, select quality_of_faculty → change the filter type to Top N → set Show items to 10 → drag the patents into the by value field → click Apply filter.
2) **Factors contributing most across countries**
    - USA leads by a large margin, followed by China and Japan.
    - Together, USA and China account for the biggest share of patents.
    - Quality of Faculty 218 and 210 contribute the highest patent counts across countries.
    - In the USA and China, these two faculty categories alone make up most of the patents.
    - Other countries (UK, Italy, France, Germany, etc.) contribute fewer patents overall, and the effect of faculty quality is weaker compared to USA and China.



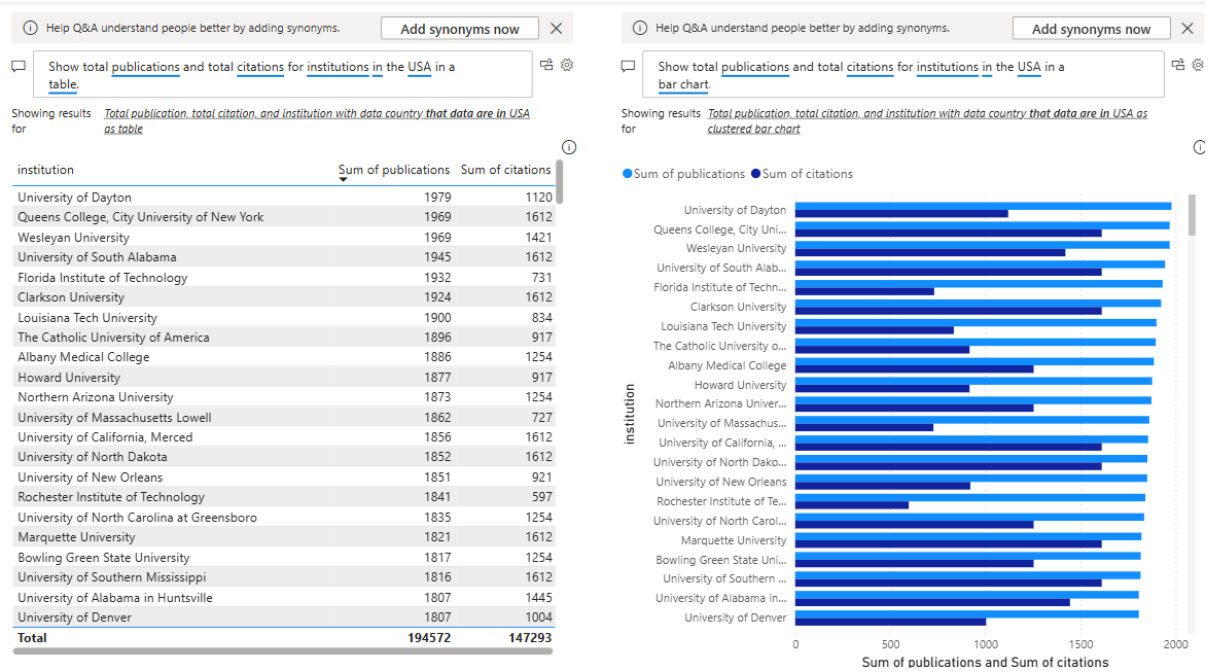
Task 9(b).

Use the Q&A feature in Power BI to answer the question: "What is the total publications and citations for institutions in the USA?" and display the results in a table and bar chart format.

Key:

1) Access the Q&A Feature

- Go to the Visualizations pane.
- Click on the Q&A icon.
- In the Q&A text box, type: “Show total publications and total citations for institutions in the USA in a table”.
- In the other Q&A text box, type: “Show total publications and total citations for institutions in the USA in a bar chart”.
- Power BI will automatically create a visualization based on our data.



Task 9(c).

Display key metrics for the top 5 institutions by world rank, including fields such as institution, country, score, and national rank.

Key:

1) Create a Table Visual
    - Insert Table Visual
    - In Report View, click on Table from the Visualizations pane
2) Add Fields to the Table
    - Drag the following fields into the Columns area
        - World_rank
        - Institution
        - Country
        - Score
        - National rank
3) In the Filters pane, select world_rank → change the filter type to Top N → set Show items bottom to → drag the world_rank into the By value field → Set Aggregation = Max of world_rank → click Apply filter.

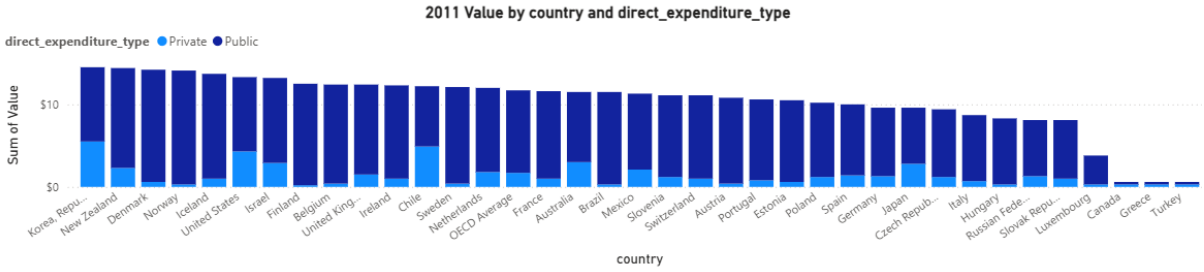
world_rank	institution	country	Sum of score	national_rank
1	Harvard University	USA	400.00	1
2	Massachusetts Institute of Technology	USA	91.67	2
2	Stanford University	USA	291.69	2
3	Massachusetts Institute of Technology	USA	196.23	3
3	Stanford University	USA	89.50	3
3	University of Oxford	United Kingdom	92.54	1
4	Massachusetts Institute of Technology	USA	91.45	3
4	University of Cambridge	United Kingdom	280.62	1
5	California Institute of Technology	USA	85.21	4
5	University of Cambridge	United Kingdom	90.24	2
5	University of Oxford	United Kingdom	193.97	2
Total			**1,903.12**	

Task 9(d).

Represent the distribution of direct_expenditure_type (e.g., public vs. private) for the year 2011 across all countries. Highlight the OECD Average as a separate segment.

Key:

1) Insert chart
   - In Report View, click on Stacked Column Chart from the Visualizations pane.
   - Card and Donut chart
2) In the Filter pane
   - Filters on this page → Drag direct_expenditure_type and year columns.
   - From direct_expenditure_type → Select Basic Filtering → Select Public and Private.
   - From Year → Basic Filtering → Select only 2011.
3) Add Fields for Stacked Column chart
   - Drag country into the X-axis field.
   - Drag value into the Y-axis field.
   - Drag direct_expenditure_type into the Legend.
4) Add Fields for Donut chart
   - Drag direct_expenditure_type into the Legend.
   - Drag value into the Values.
5) Add fields for card
   - Drag value into the Fields.



Task 10.

Create a workspace "Institution Analysis" and set up a schedule to refresh the datasets every day at 6 AM.

Key:

1. **Power BI Service**
   - Publish the data and expenditure report to Power BI Service.
   - Go to Datasets → Select the dataset.
   - Go to the setting → Power BI setting.
   - Select → Semantic Model → Refresh.

- Select the time zone → Refresh frequency select daily.
- Set the time zone 6:00 AM → Apply.