**Assignment 3**

1. **Introduction**

In this assignment, I analyse two tables: Expenditure and Data using Power BI. The goal is to apply advanced Power BI features such as calculated columns, measures, and visualizations to gain insights into the dataset. The dataset provided is [Expenditure and Data](https://drive.google.com/file/d/1iX5Sw1h4GDgOKQEcj8Zqw0SH0UvZUU36/view), which contains information about different country, including institute\_type, direct\_expenditure\_type and other relevant details.

1. **Methodology**

Importing the Dataset

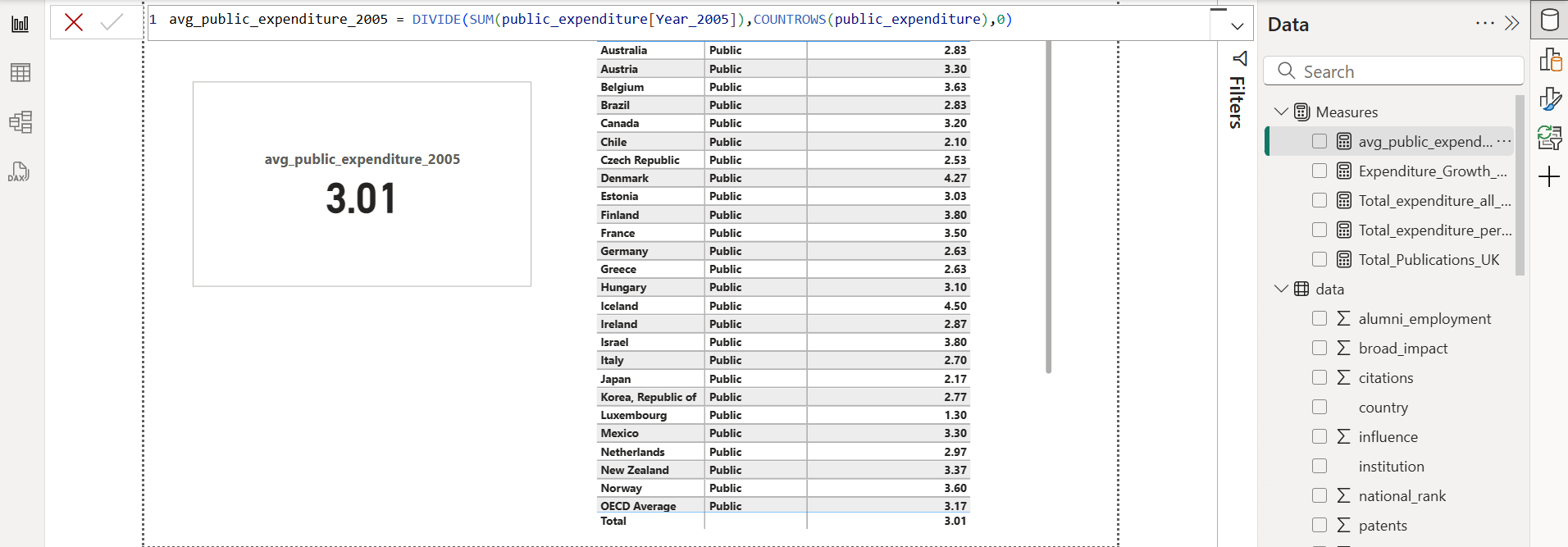
* Open Power BI Desktop.
* Click on Get Data → CSV → Select Data.csv and expenditure.csv
* Load the dataset into Power BI.

1. **Tasks (1):**

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

**Explication: (1)**

1. Click on table view 🡪 new table
2. Used this DAX function **public\_expenditure = SELECTCOLUMNS(FILTER(expenditure,expenditure[direct\_expenditure\_type]="Public"),"Country",expenditure[country],"Expenditure\_Type",expenditure[direct\_expenditure\_type],"Year\_2005",expenditure[2005]).** Now we have new table “public\_expenditure” table.
3. FILTER(...) ensures only rows with "Public" in direct\_expenditure\_type are selected.
4. SELECTCOLUMNS(...) picks only Country, 2005 values and expenditure\_type = public
5. Then in report view in measure column 🡪 click on new measure
6. Used the following DAX formula in formula bar : **avg\_public\_expenditure\_2005 = DIVIDE(SUM(public\_expenditure[Year\_2005]),COUNTROWS(public\_expenditure),0)**
7. SUM(public\_expenditure[Year\_2005]) This sums up all the values in the Year\_2005 column from the public\_expenditure table.
8. COUNTROWS(public\_expenditure) This counts the total number of rows in public\_expenditure.
9. DIVIDE This divides the sum by the count of rows. If the denominator is zero (to avoid division errors), it returns 0 instead of an error.
10. for view Inserted a Card Visual (to show a single value) and a Table Visual (to compare by country).
11. For card visual drag Avg\_Public\_Expenditure\_2005 to data and for table visual drag country and Avg\_Public\_Expenditure\_2005 column to add data.

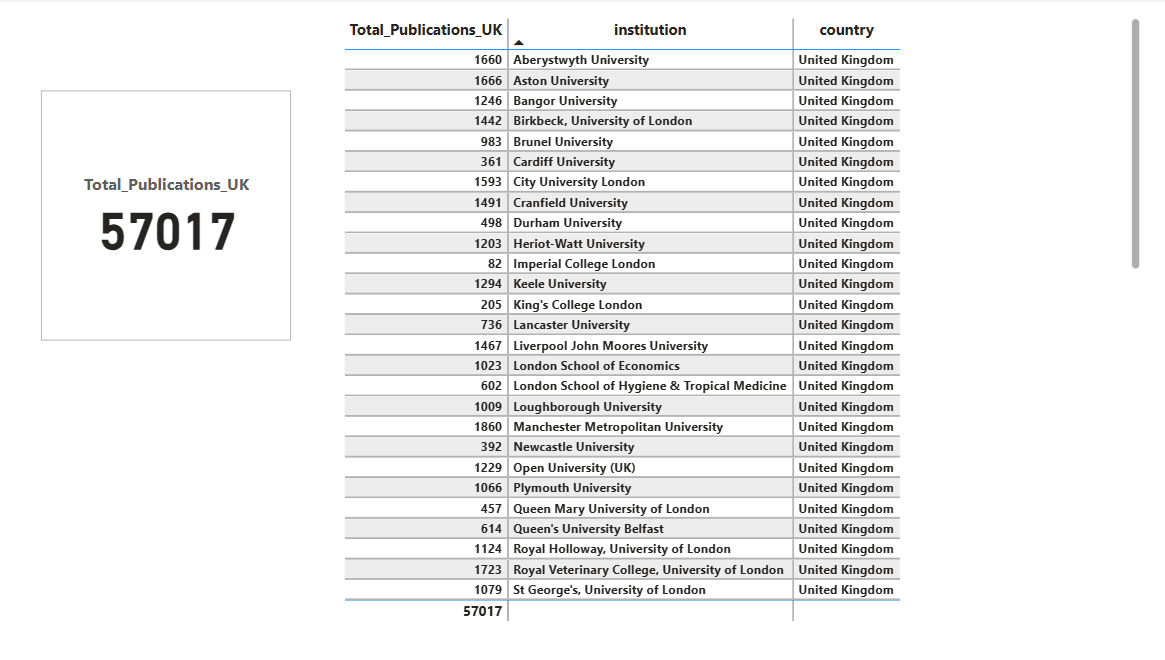


**Tasks (2):**

Calculate the total publications for institutions in the UK.

**Explication: (2)**

1. Click "New Measure" and enter this DAX formula Total\_Publications\_UK = **CALCULATE(SUM(data[publications]),data[country]="United Kingdom")**
2. SUM(data[publications]): Adds up all publication values.
3. CALCULATE(...): Applies a filter.
4. data[country] = "United Kingdom": Filters only rows where country = "United Kingdom".
5. Card Visual → To show the total publications.
6. Table Visual → To break down by institution, publications and United kingdom.

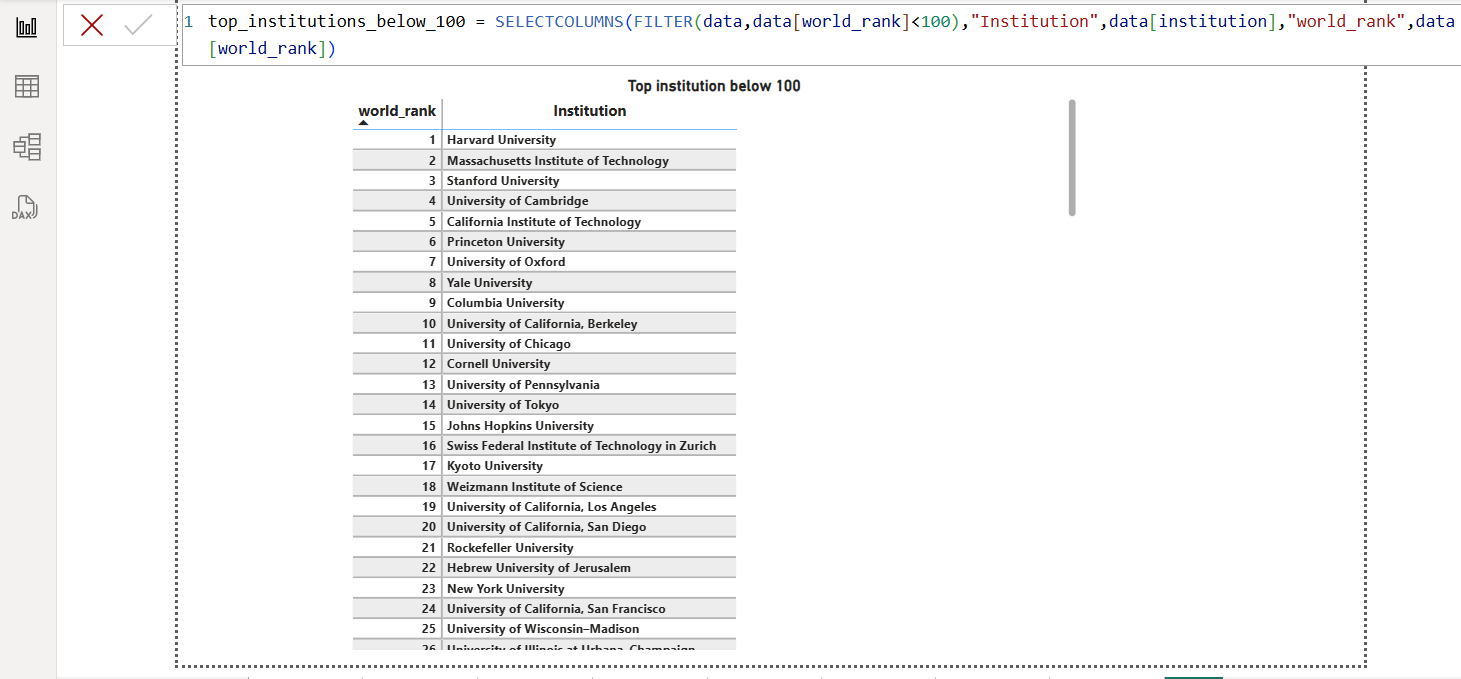


**Tasks (3):**

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

**Explication: (3)**

1. Remove the duplicates from institution column.
2. Go to table view 🡪 new table for Top\_institutions\_below\_100 (Table Name)
3. In the DAX formula I used this formula top\_Institutions\_below\_100 = **SELECTCOLUMNS(FILTER(data,data[world\_rank]<100),"Institution",data[institution],"world\_rank",data[world\_rank])**
4. FILTER(data, data[world\_rank] < 100): Filters rows where world\_rank is less than 100.
5. SELECTCOLUMNS(...): Extracts only the institution and world\_rank columns.
6. This created a new table named Top\_Institutions\_below\_100 with only two columns: Institution and World Rank (filtered to only values below 100).

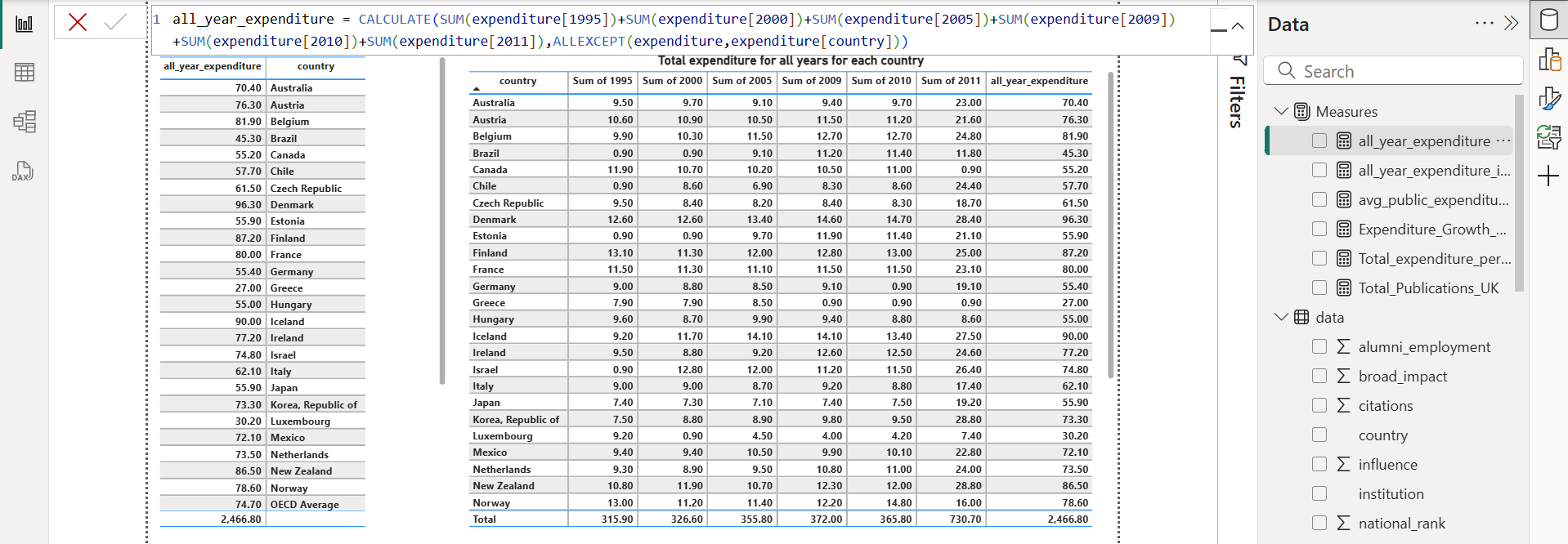


**Tasks (4):**

Calculate the total expenditure for all years for each country.

**Explication: (4)**

1. Right click in measure table 🡪 select new measure
2. In DAX use this formula **all\_year\_expenditure = CALCULATE(SUM(expenditure[1995])+SUM(expenditure[2000])+SUM(expenditure[2005])+SUM(expenditure[2009])+SUM(expenditure[2010])+SUM(expenditure[2011]),ALLEXCEPT(expenditure,expenditure[country]))**
3. CALCULATE() → Modifies the filter context to ensure totals are calculated correctly.
4. SUM() → Summing each year column individually.
5. ALLEXCEPT() → Ensures grouping happens by country only (ignores other filters).
6. For visualization selected card and table from visualization pane and drag all\_year\_expenditure measure to data column and country column from expenditure table.



**Tasks (5):**

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

**Explication: (5)**

1. DAX Formula: **all\_year\_expenditure\_ignore\_all\_filters=CALCULATE([all\_year\_expenditure],ALL(expenditure))**
2. CALCULATE() is that it allows us to add, remove, or modify filters.
3. ALL() is a filter modifier that **removes all filters** from the specified table or columns.



**Tasks (6):**

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

**Explication: (6)**

1. Click on data 🡪 new measure.
2. For calculating the growth rate formula is: 2000 – 1995 / 1995 \* 100 (Austria)
3. In the formula bar used this formula: Expenditure\_Growth\_Austria = **DIVIDE(CALCULATE(SUM(expenditure[2000]),expenditure[country]="Austria") -**

**CALCULATE(SUM(expenditure[1995]),expenditure[country]="Austria"),**

**CALCULATE(SUM(expenditure[1995]),expenditure[country]="Austria"),0) \* 100**

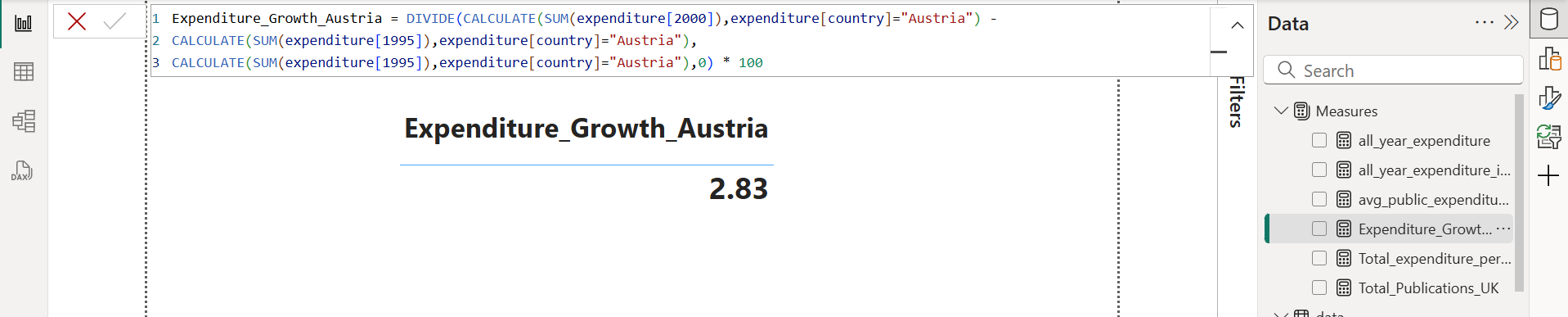
1. DIVIDE() ensures stability by avoiding division errors.
2. CALCULATE(SUM(expenditure[2000]), expenditure[country] = "Austria")

→ Filters the data to include only Austria and calculates the sum of expenditure in 2000.

1. CALCULATE(SUM(expenditure[1995]), expenditure[country] = "Austria")

→ Filters the data to include only Austria and calculates the sum of expenditure in 1995**.**

1. 0 in DIVIDE(): Acts as a fallback value in case the denominator is zero (prevents errors).
2. Multiplying by 100. This converts the result from a decimal to a percentage.

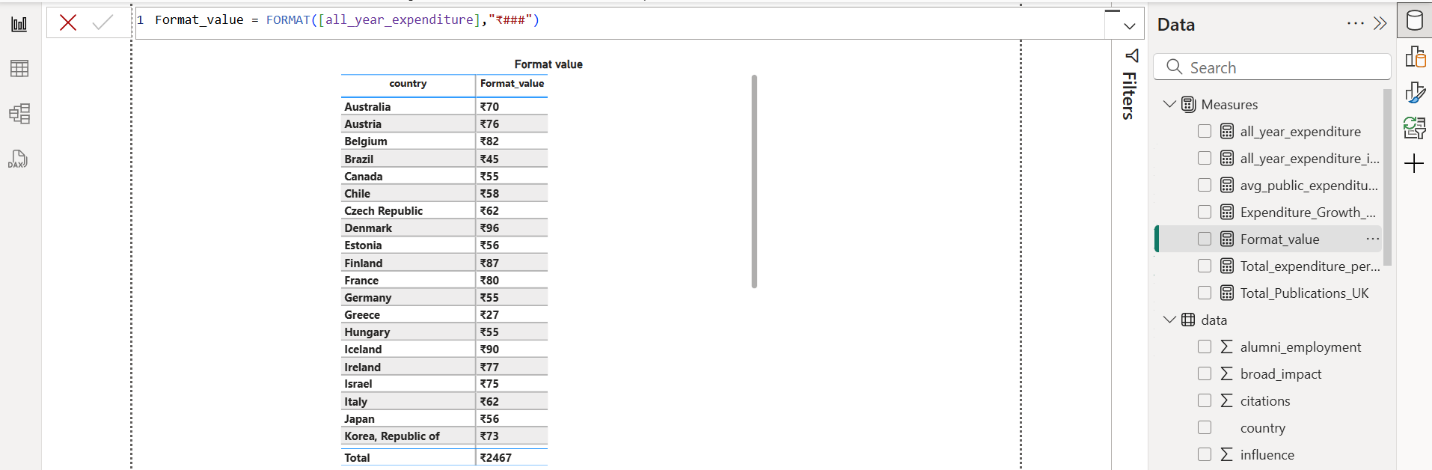


**Tasks (7):**

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

**Explication: (7)**

1. Right click on the measure data 🡪 new measure.
2. In the DAX I used this formula: **Format\_value = FORMAT([all\_year\_expenditure],"₹###").**
3. For visualization selected table from visualization pane then drag country column from expenditure data and format\_value from measure data.

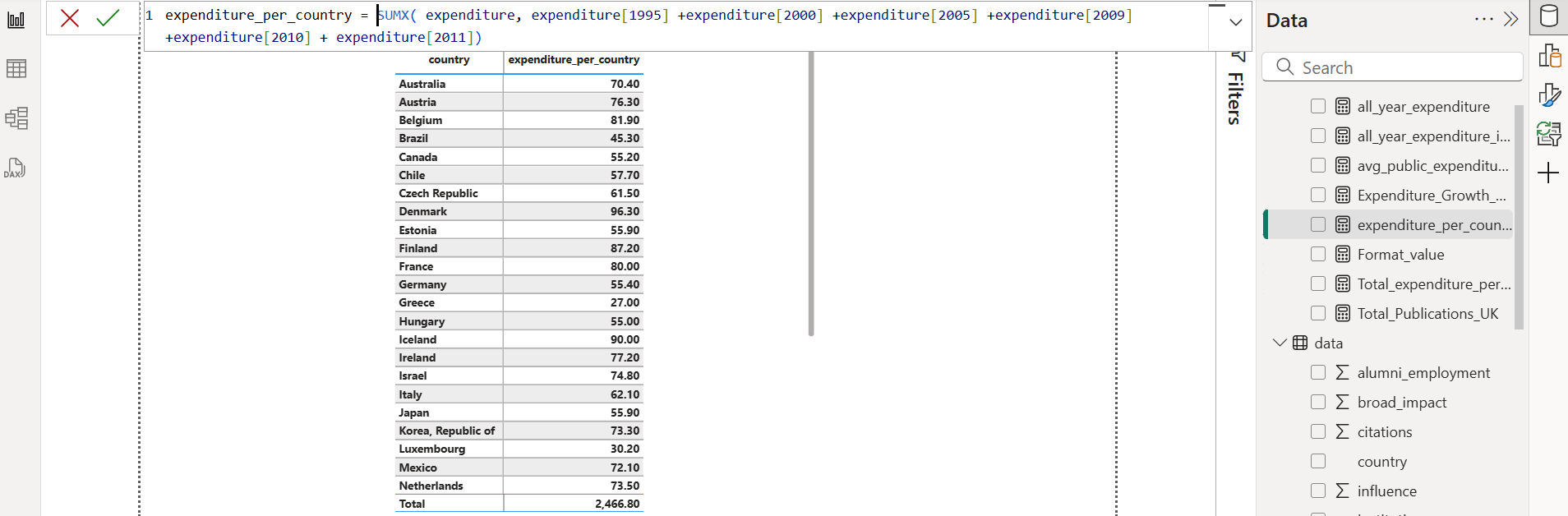


**Tasks (8):**

Calculate the total expenditure for each country.

**Explication: (8)**

1. Right click on the measure data 🡪 new measure.
2. In DAX used this formula: expenditure\_per\_country **= SUMX( expenditure, expenditure[1995] +expenditure[2000] +expenditure[2005] +expenditure[2009] +expenditure[2010] + expenditure[2011])**
3. For visualization selected table from visualization pane and drag expenditure\_per\_country column to data and country columns from expenditure table



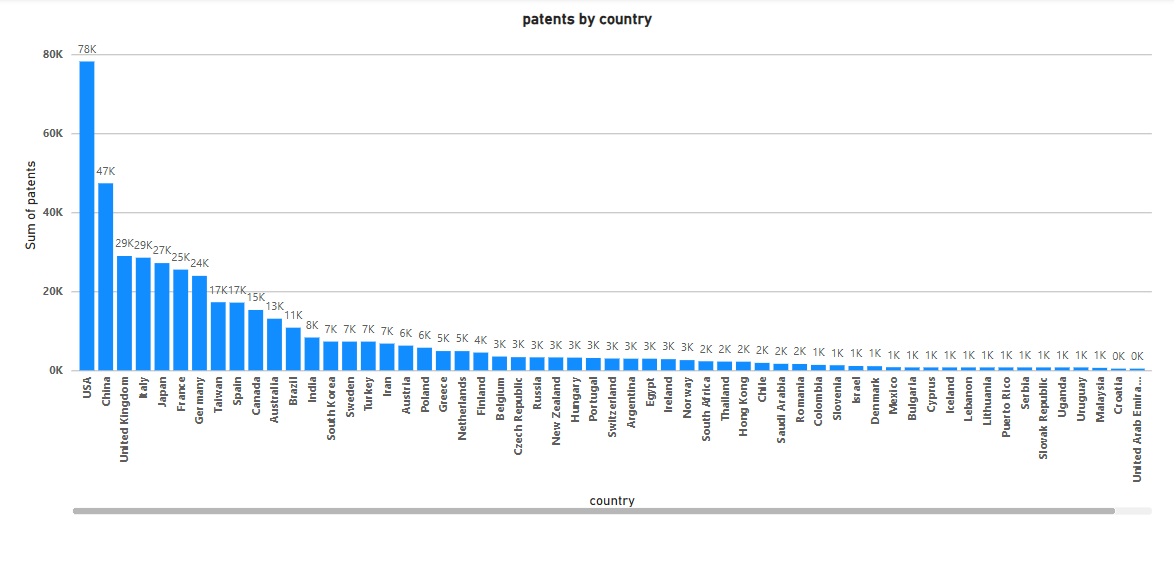
**Tasks (9):**

1. 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.

**Explication: (9)(a)**

* Visual 1 - Column Chart (Total Patents by Country)

1. Selected the column Chart visual from the Visualizations pane.
2. Drag the Country column to the x-Axis field.
3. Drag the Number of Patents column to the y-axis field.
4. USA country have 78K patents



* Visual 2 - Stacked Bar Chart (Country + Faculty Quality Breakdown)

1. Selected tree map from visualization pane.
2. Drag country 🡪 categories, quality of faculty 🡪 details and patents 🡪 values



* **Analyse**

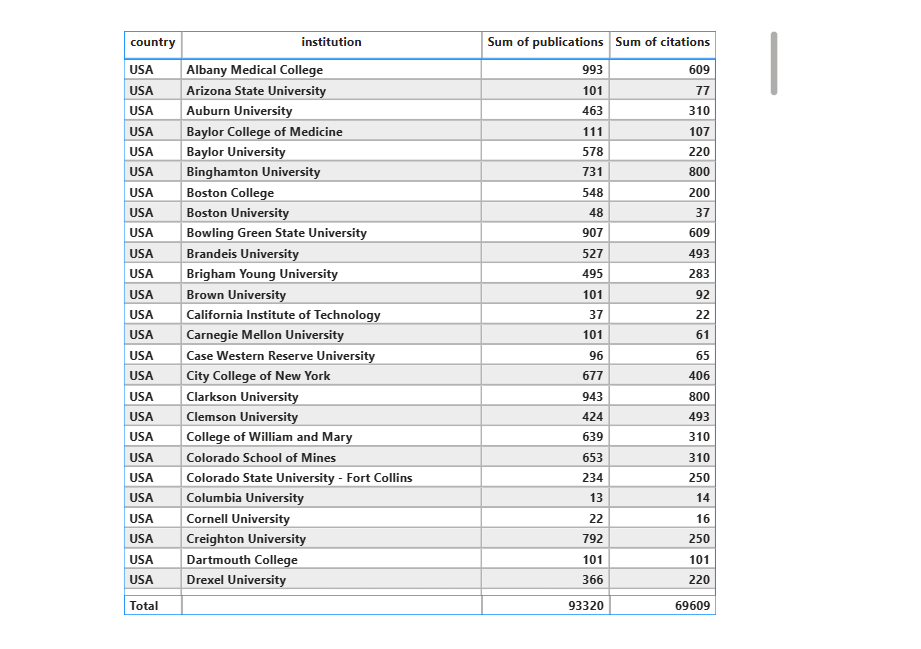
1. USA country have 78K patents.
2. **USA**, **China**, and the **United Kingdom** lead with the highest number of patents, indicating these countries have strong research ecosystems.
3. Countries like the **USA, United Kingdom**, and **Japan** show a strong presence in institutions with high-quality faculty, suggesting a clear link between faculty strength and patent output.
4. Countries with smaller patent counts often have lower-quality faculty segments in the visual, highlighting a potential gap**.**

**Tasks (9):**

1. 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.

**Explication: (9)(b)**

1. In Power BI Desktop, go to the Visualizations pane.
2. Select the Q&A visual, Click inside the visual's text box.
3. Type 1: "Show total publications and total citations for institutions in the USA in a table."
4. Type 2: "Bar chart of total publications and total citations by institution in the USA."
5. The Q&A result will automatically convert into our chosen visual format.



**A screenshot of a graph

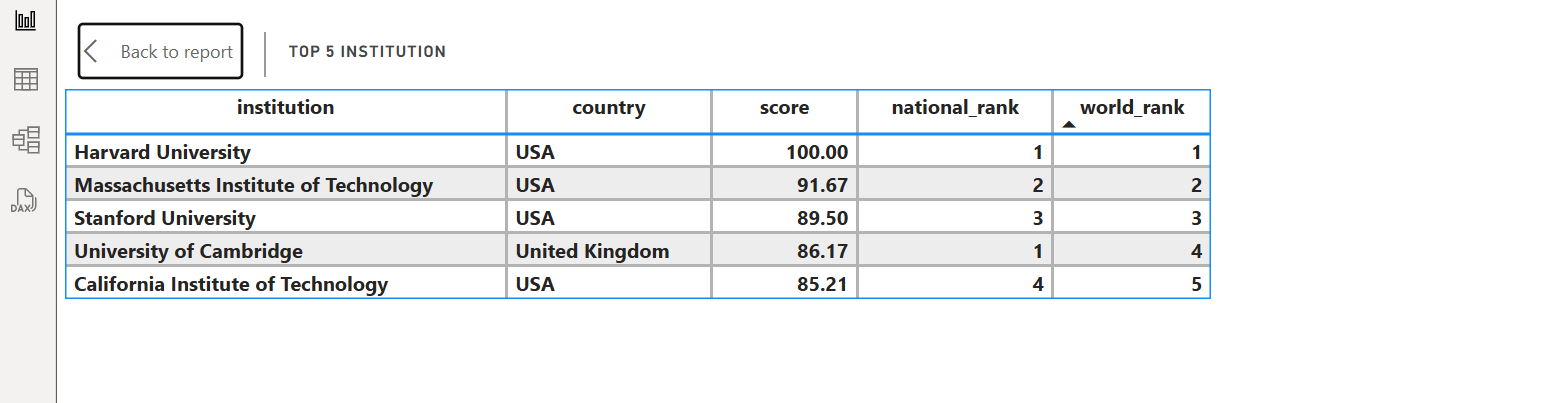
AI-generated content may be incorrect.**

**Tasks (9):**

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

**Explication: (9)(c)**

1. Go to the visualisation pane 🡪 selected table
2. Drag country, institution, score, national score and world rank column to add data column.
3. Score, national sore and world rank set to don’t summarize.
4. In the filter section go to the score column🡪 Top N 🡪 put 5 and drag score column to value section and select maximum aggregation.



**Tasks (9):**

1. 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.

**Explication: (9)(d)**

1. Click to the transform data
2. Select the columns: 1995, 2000, 2005, 2009, 2010, 2011. (expanditure\_copy) table.
3. Click on Transform → Unpivot Columns.
4. This will create two new columns: attribute 🡪 rename to (Year) and expenditure 🡪 rename to 🡪 Value.
5. Click Close & Apply.
6. Clicked on data 🡪 expanditure\_copy 🡪 right click new measure
7. Used this DAX formula = all\_direct\_expenditure\_type\_2011 = **CALCULATE(SUM('expediture\_copy'[Value]),'expediture\_copy'[Year]= 2011,FILTER('expediture\_copy','expediture\_copy'[direct\_expenditure\_type] IN {"Public","Private"}))**
8. CALCULATE() modifies the filter context of the calculation. Filter for Year = 2011 Filter for direct\_expenditure\_type as either "Public" or "Private".
9. SUM calculates the **total expenditure** in the Value column.
10. FILTER() function Filters only rows where direct\_expenditure\_type is either "Public" or "Private".
11. The IN function checks if the specified value exists in a list of values.
12. For OECD average created new measure in expenditure\_copy.
13. Formula used: OECD\_Average\_2011 = **CALCULATE(AVERAGE('expediture\_copy'[Value]),'expediture\_copy'[Year]=2011,'expediture\_copy'[country]="OECD Average",'expediture\_copy'[direct\_expenditure\_type] IN {"Public","Private"})**
14. For visualisation clicked stacked bar chart.
15. Drag country 🡪 y-axis, all\_direct\_expediture\_type\_2011 to x-axis and legend to direct\_expediture columns.
16. For OECD\_Average\_2011 put to reference line.

A blue and white graph

AI-generated content may be incorrect.

**Tasks (10):**

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

**Explication: (10)**

1. Publish the report to Power BI Service.
2. Go to Datasets 🡪 Select the dataset.
3. Go to the setting 🡪 Power BI setting.
4. Select 🡪 Semantic Model 🡪 Refresh.
5. Select the time zone 🡪 Refresh frequency select daily.
6. Set the time zone 5:00 PM 🡪 Apply.

