

Internshala Trainings

Data Science

PowerBI Assignment 3

**Prepared By,
Utkarsh Anand**

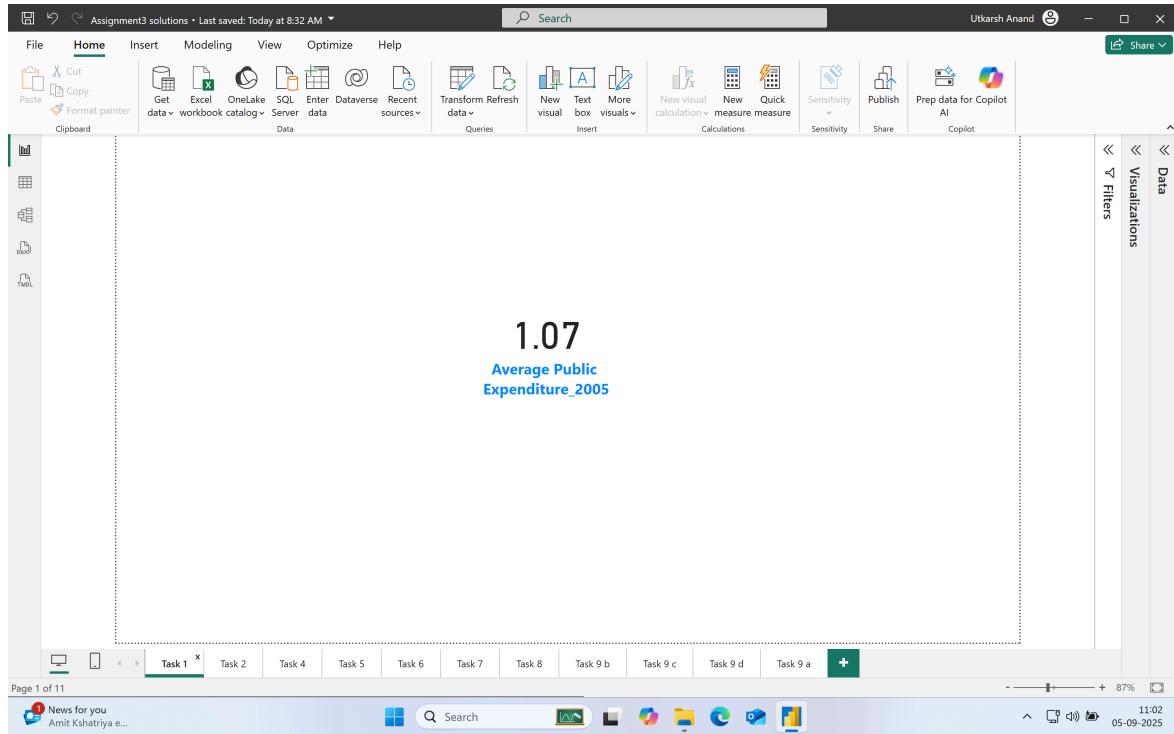
Note on data shape: In Expenditure, the year values are separate columns (1995, 2000, 2005, 2009, 2010, 2011). For easier analysis, I first unpivoted the year columns so that we get a tidy structure: [country] [institute_type] [direct_expenditure_type] [Year] [Expenditure]. (Steps are included below and should be done once before the tasks.)

Pre-Step (done once): Unpivot year columns in Expenditure

Steps taken to reshape the Expenditure table for analysis:

1. In Power BI Desktop, click Home → Transform data to open Power Query Editor.
2. In the Queries pane (left), select the Expenditure table.
3. In the preview grid, Ctrl + Click to select the non-year columns:
 - country
 - institute_type
 - direct_expenditure_type
4. On the ribbon, click Transform → Unpivot Columns → Unpivot Other Columns.
5. Rename the two new columns:
 - Attribute → Year
 - Value → Expenditure
6. Change data types:
 - Set Year to Whole Number
 - Set Expenditure to Decimal Number
7. Click Home → Close & Apply to save the transformation.

Task 1 - Average public expenditure(year 2005 across all countries):



Steps taken to calculate the average public expenditure for the year 2005 across all countries:

1. Open Power BI Desktop → Home → Transform data (Power Query) and ensure Expenditure is loaded and Year is a column.
2. Go to Report view.
3. Create a New measure in Expenditure table.

DAX formula is below:

```
Avg Public Expenditure 2005 :=  
AVERAGEX(  
    FILTER(  
        Expenditure,  
        Expenditure[Year] = 2005  
        && Expenditure[direct_expenditure_type] = "public"  
    ),  
    Expenditure[Expenditure]  
)
```

4. Insert a Card visual → add [Avg Public Expenditure 2005] to Values.
5. Title the card “Average Public Expenditure (2005)”.

Task 2 – Total Publications for UK Institutions

The screenshot shows a Microsoft Power BI report titled "Assignment3 solutions". The report contains two card visualizations:

- Total Publications_UK**: Value **57017**
- Total Publications_UK**: Value **1011799**

The report has a navigation bar at the bottom with tabs: Task 1, Task 2 (selected), Task 4, Task 5, Task 6, Task 7, Task 8, Task 9 b, Task 9 c, Task 9 d, Task 9 a, and a plus sign. The status bar at the bottom right shows "Page 2 of 11", "News for you Amit Kshatriya e...", "11:03", and "05-09-2025".

Steps taken to calculate the total publications for institutions in the UK:

1. Go to Report view.
2. Create a New measure in Data table.

DAX formula is below:

```
Total Publications (UK) :=  
CALCULATE(  
    SUM(Data[publications]),  
    Data[country] = "United Kingdom"  
)
```

3. Insert a Card visual → add [Total Publications (UK)] to Values.
4. Title the card “Total Publications – UK”.

Task 3 – Show Only Institutions with World Rank < 100

The screenshot shows the Power BI desktop interface. In the center, there's a table titled "Table.AddColumn(#"Removed Columns", "YearR.1", each #date([year], 1, 1))". The table contains columns: institution, country, national_rank, quality_of_education, and al. The data includes rows for Harvard University, Massachusetts Institute of Technology, Stanford University, and many others. A "Filter Rows" dialog is open over the table, with the condition "Keep rows where 'world_rank' is less than 100". On the right side, the "Query Settings" pane is visible, showing properties like "Name: data" and applied steps such as "Added Custom", "Changed Type1", "Removed Columns1", and "Renamed Columns". The bottom status bar shows "15 COLUMNS, 999+ ROWS" and "Column profiling based on top 1000 rows". The system tray indicates it's 05-09-2025 at 08:25, with a weather icon showing 28°C and mostly cloudy.

Steps taken to show only institutions with world rank below 100:

1. Insert a Table visual → add fields: institution, country, world_rank, score.
2. In Filters (for this visual): world_rank → Advanced filtering → is less than → 100 → Apply filter.
(Optional count via DAX for a KPI card)

DAX formula is below:

```
Institutions <100 (Count) :=  
CALCULATE(  
    DISTINCTCOUNT(Data[institution]),  
    Data[world_rank] < 100  
)
```

Assignment3 solutions • Last saved: Today at 11:06 AM

File Home Help

Paste Cut Copy Format Comment Uncomment Find Replace Command palette Copilot (preview)

Clipboard Editing Copilot

Run Update model with changes (0) Share feedback Data

1 // Welcome to DAX query view! Learn more about DAX queries at <https://aka.ms/dax-queries>.
2 // Right-click on tables, columns, or measures in the data pane to access quick queries, or ask Copilot for help writing DAX.
3
4 // Select "Run" to try this sample DAX query.
5 EVALUATE
6 | TOPN(100, 'data')
7

Tables Model

Search

> Measure (2)
> data
> expenditure
> Filtered institutions world rank

Results

Your query results will show up here

Query 1 +

NIFTY -0.14%

Search

11:10 05-09-2025

This screenshot shows the Power BI DAX Query view interface. The main area displays a sample DAX query:

```
1 // Welcome to DAX query view! Learn more about DAX queries at https://aka.ms/dax-queries.
2 // Right-click on tables, columns, or measures in the data pane to access quick queries, or ask Copilot for help writing DAX.
3
4 // Select "Run" to try this sample DAX query.
5 EVALUATE
6 | TOPN(100, 'data')
7
```

The interface includes a 'Clipboard' ribbon tab, a 'Copilot (preview)' button, and a 'Data' pane on the right showing tables and models. Below the main area, it says 'Your query results will show up here'. At the bottom, there's a navigation bar with 'Query 1' and a '+' button, along with system status icons.

Task 4 – Total Expenditure for All Years (per Country)

The screenshot shows a Microsoft Power BI interface. At the top, the ribbon is visible with tabs like File, Home, Insert, Modeling, View, Optimize, and Help. The Home tab is selected. Below the ribbon, there are various icons for data sources (Clipboard, Get data from Excel, OneLake, SQL Server, Enter data, Dataverse, Recent sources), queries (Transform data, Refresh data, New visual, Text box, More visuals), calculations (New visual calculation, New measure, Quick measure), sensitivity (Sensitivity), publishing (Publish, Prep data for Copilot AI), and sharing (Copilot). On the left, there's a navigation pane with icons for Home, Data, Visualizations, and Filters. The main area displays two tables and a clustered bar chart.

Table 1:

country	Total Expenditure
Australia	70.40
Austria	76.30
Belgium	81.90
Brazil	45.30
Canada	55.20
Chile	57.70
Czech Republic	61.50
Denmark	96.30
Estonia	55.90
Finland	87.20
France	80.00
Germany	55.40
Greece	27.00
Hungary	55.00
Iceland	90.00
Ireland	77.20
Israel	74.80
Italy	62.10
Japan	55.90
Montenegro	73.70
Total	2,466.80

Table 2:

Total_Expenditure	country
70.40	Australia
76.30	Austria
81.90	Belgium
45.30	Brazil
55.20	Canada
57.70	Chile
61.50	Czech Republic
96.30	Denmark
55.90	Estonia
87.20	Finland
80.00	France
55.40	Germany
27.00	Greece
55.00	Hungary
90.00	Iceland
77.20	Ireland
74.80	Israel
62.10	Italy
2,466.80	

Bar Chart: A clustered bar chart showing the total expenditure for each country. The bars are light blue with dark blue outlines. The x-axis labels are the country names, and the y-axis represents the total expenditure.

Steps taken to calculate total expenditure across all years for each country:

1. Create a New measure in Expenditure table.

DAX formula is below:

Total Expenditure (All Years) :=
SUM(Expenditure[Expenditure])

2. Insert a Clustered bar chart.
3. Axis = Expenditure[country]; Values = [Total Expenditure (All Years)].
4. Turn Data labels ON; title “Total Expenditure by Country (All Years)”.

Task 5 – Total Score Ignoring Any Year Filters

Year	Sum of score	Total_Score_All_Years
2012	5,494.09	1,05,156.47
2013	5,527.12	1,05,156.47
2014	47,271.41	1,05,156.47
2015	46,863.85	1,05,156.47
Total	1,05,156.47	1,05,156.47

Steps taken to compute total score across all years while ignoring filters on Year:

1. Create a New measure in Data table.

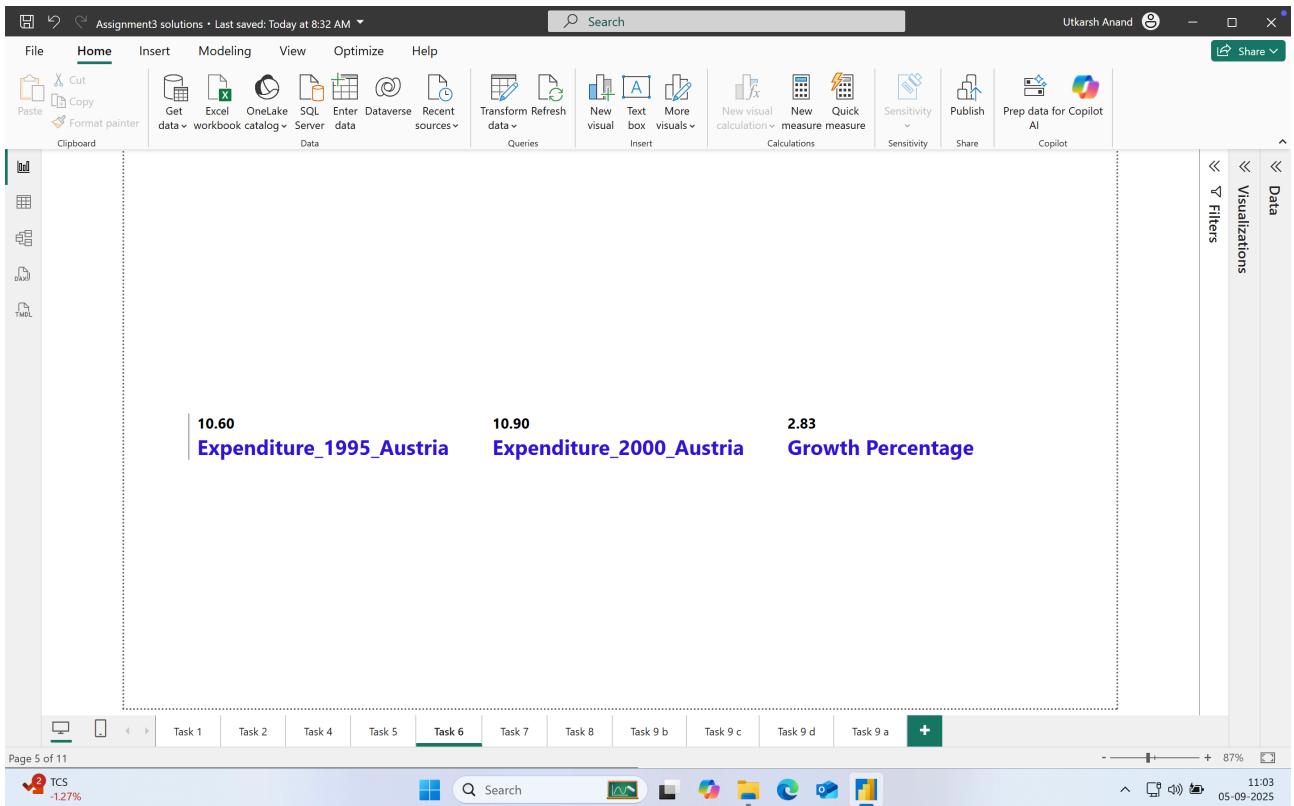
DAX formula is below:

```
Total Score (All Years, Ignore Year Filters) :=  
CALCULATE(  
    SUM(Data[score]),  
    ALL(Data[year])  
)
```

2. Insert a Card visual → add [Total Score (All Years, Ignore Year Filters)].

3. Verify slicers on Year do not change the card.

Task 6 – Growth in Expenditure for Austria (1995 → 2000)



Steps taken to calculate absolute and percentage growth for Austria from 1995 to 2000:

1. Create measures:

DAX formula is below:

```
Expenditure Austria 1995 :=  
CALCULATE(  
    SUM(Expenditure[Expenditure]),  
    Expenditure[country] = "Austria",  
    Expenditure[Year] = 1995  
)
```

Expenditure Austria 2000 :=

```
CALCULATE(  
    SUM(Expenditure[Expenditure]),  
    Expenditure[country] = "Austria",  
    Expenditure[Year] = 2000  
)
```

Growth Austria (2000 - 1995) :=

```
[Expenditure Austria 2000] - [Expenditure Austria 1995]
```

Growth % Austria (1995→2000) :=

```
DIVIDE(  
    [Growth Austria (2000 - 1995)],  
    [Expenditure Austria 1995]  
)
```

2. Insert a Table or two Cards → show [Growth Austria (2000 - 1995)] and [Growth % Austria (1995→2000)].

3. Format the % measure as Percentage.

Task 7 – Format Expenditure Values with Currency (0 decimals)

The screenshot shows the Microsoft Power BI desktop application. The ribbon menu is visible at the top, with the 'Home' tab selected. The main workspace displays a single visual titled 'Total_Expenditure'. Below the title, there are two numerical values: '2,466.80' and '€2,467'. The interface includes a sidebar on the right with sections for 'Data', 'Visualizations', and 'Filters'. At the bottom, there is a navigation bar with tabs labeled 'Task 1' through 'Task 9 a', and a '+' button. The status bar at the bottom right shows the date '05-09-2025' and time '11:04'.

Steps taken to format the Expenditure field:

1. Data view → select Expenditure[Expenditure].
2. Modeling tab → Data type = Decimal Number.
3. Format = Currency; Decimal places = 0; choose symbol (₹/\$ etc.).
4. All visuals now display currency with 0 decimals.

Task 8 – Total Expenditure for Each Country (Visual)

The screenshot shows the Microsoft Power BI desktop interface. The ribbon menu is visible at the top, with the 'Home' tab selected. The main workspace displays a clustered bar chart titled 'Total Expenditure country'. The chart has two columns: 'country' and 'Total Expenditure country'. The data is sorted in descending order of expenditure. The total expenditure for all countries shown is 2,466.80. The chart is set against a light blue background with white bars. The right side of the screen features a 'Filters' pane and a 'Visualizations' pane. The bottom of the screen shows the taskbar with various icons and the date/time.

country	Total Expenditure country
Turkey	18.40
Greece	27.00
Brazil	45.30
Slovak Republic	53.40
Hungary	55.00
Canada	55.20
Germany	55.40
Slovenia	55.80
Japan	55.90
Estonia	55.90
Chile	57.70
Czech Republic	61.50
Italy	62.10
Spain	66.10
Australia	70.40
Poland	71.70
Total	2,466.80

Steps taken to create a visual for total expenditure per country:

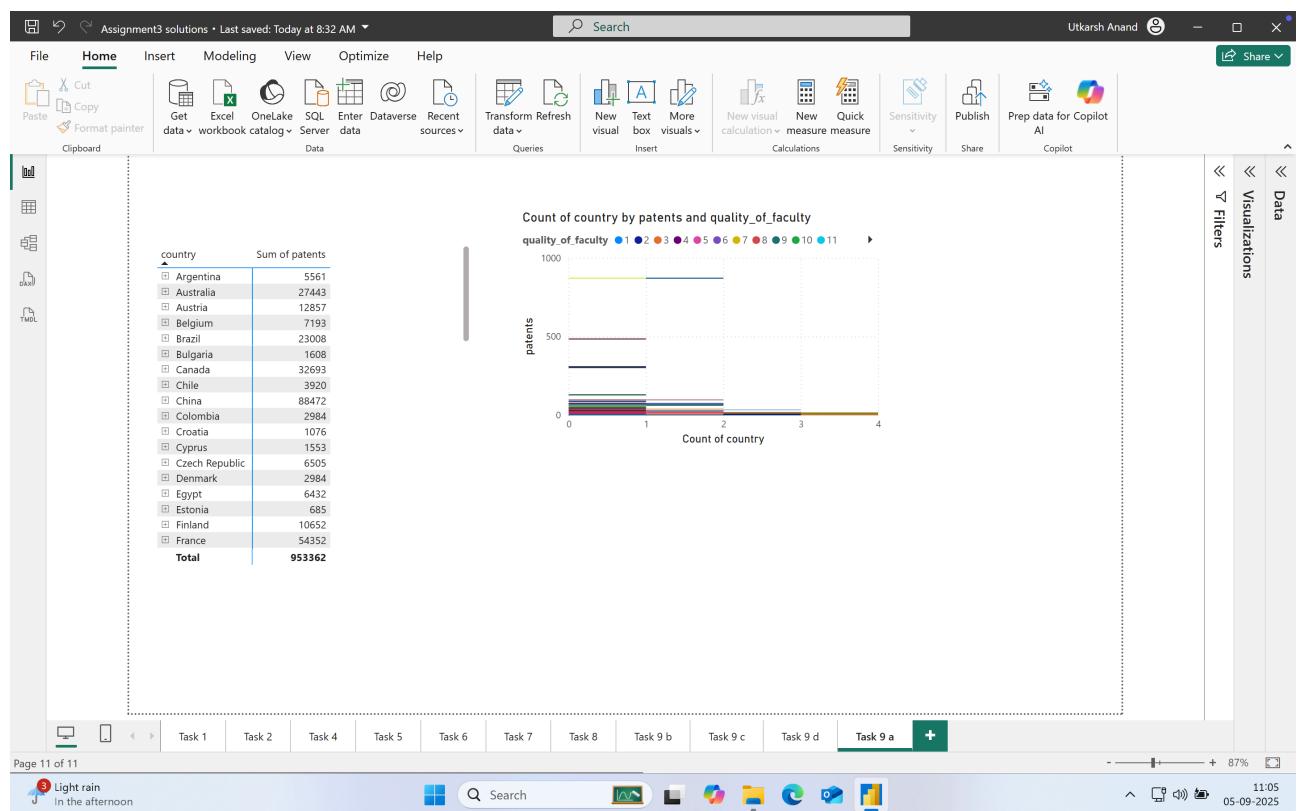
1. Insert a Clustered column/bar chart.
2. Axis = Expenditure[country]; Values = [Total Expenditure (All Years)].
3. Sort Descending; turn Data labels ON.
4. Title “Total Expenditure by Country (All Years)”.

Task 9 – Report Pages (a-d)

Steps taken to create visuals on separate report pages:

(a) Patents by Country then by Quality of Faculty

1. New page → insert Decomposition Tree.
2. Analyze = $\text{SUM}(\text{Data}[\text{patents}])$; Explain by = $\text{Data}[\text{country}], \text{Data}[\text{quality_of_faculty}]$.
3. Expand branches to identify which factors drive higher patents.



(b) Q&A: "What is the total publications and citations for institutions in the USA?"

1. New page → insert Q&A visual.
2. Type the question exactly: total publications and citations for institutions in the USA.
3. Convert the result to a Table and also a Bar chart via visual switcher.

The screenshot shows a Microsoft Power BI interface with the following components:

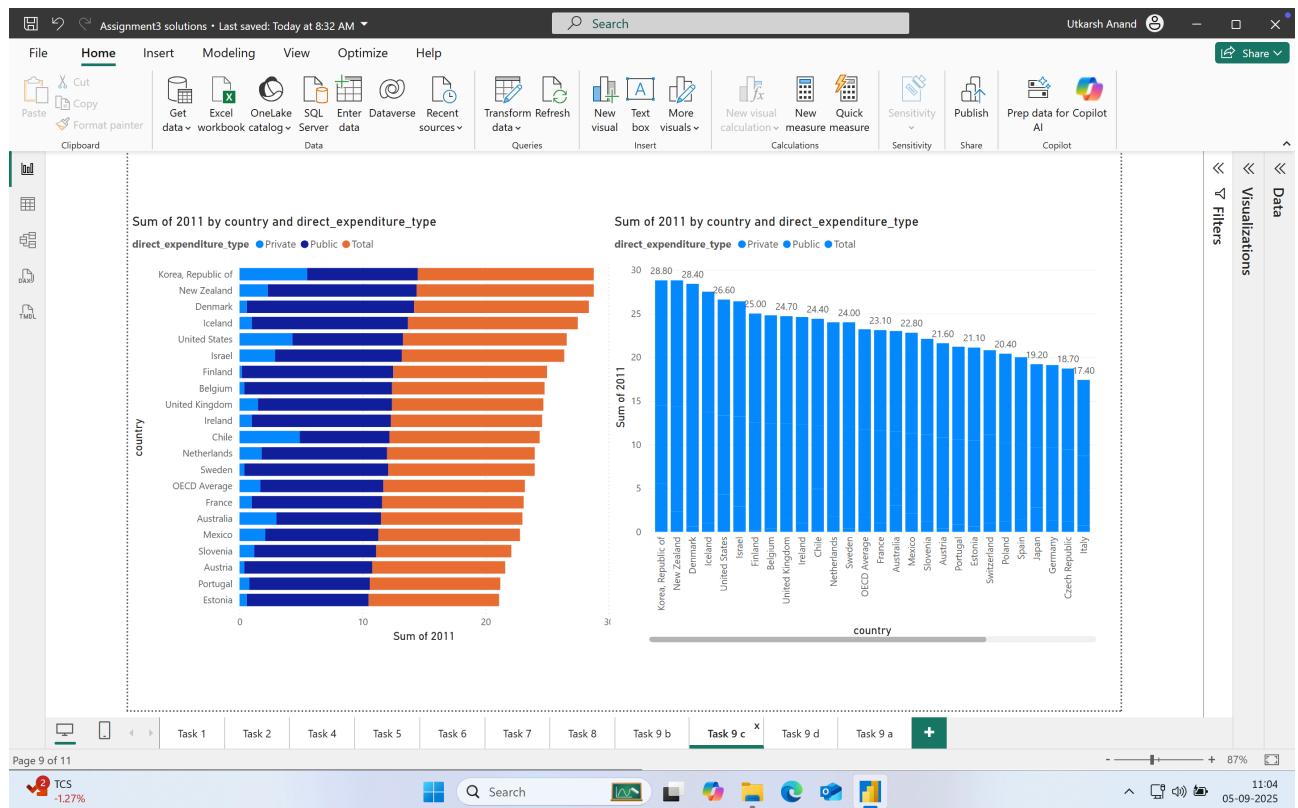
- Q&A Visual:** A card at the top left contains the question: "What is the total publications and citations for institutions in the USA show results in a table and bar chart". Below it, a note says: "Showing Total filtered institutions world rank publication, filtered institutions world rank results for citation, and filtered institutions world rank institution with filtered institutions for world rank country that filtered institutions world ranks are in USA as table".
- Table:** A table titled "citations institution Sum of publications" lists the following data:

institution	Sum of publications
115 Rockefeller University	372
107 Rockefeller University	347
125 Dartmouth College	246
136 Dartmouth College	246
182 University of Notre Dame	242
159 University of Notre Dame	241
78 Carnegie Mellon University	211
52 Carnegie Mellon University	207
101 Dartmouth College	202
93 Tufts University	166
79 Tufts University	164
76 Arizona State University	143
78 University of Texas Southwestern Medical Center	131
81 University of Texas Southwestern Medical Center	127
55 Brown University	123
76 Brown University	121
68 University of Rochester	114
Total	12542

- Bar Chart:** A bar chart titled "Sum of publications by institution and citations" displays the same data as the table, with bars representing the sum of publications for each institution. The x-axis is labeled "Sum of publications" and ranges from 0 to 200.
- Visualizations:** A sidebar on the right shows "Data", "Visualizations", and "Filters".
- Page Navigation:** At the bottom, there are tabs for "Task 1" through "Task 9 b" (the active tab), and "Task 9 c" through "Task 9 d".
- System Status:** On the far left, there is a TCS status icon showing "-1.27%".
- System Information:** At the bottom right, it shows the date "05-09-2025", time "11:04", and a zoom level of "87%".

(c) Top 5 Institutions by World Rank (show institution, country, score, national_rank):

1. New page → insert a Table visual.
2. Filters (on this visual) → institution → Top N → Top 5 by MIN(Data[world_rank]).
3. Add columns: institution, country, score, national_rank, world_rank. Sort by world_rank ascending.



(d) Distribution of direct_expenditure_type in 2011 (highlight OECD Average):

1. New page → insert 100% Stacked column chart.
2. Filters → Year = 2011.
3. Axis = direct_expenditure_type; Legend = country; Values = SUM(Expenditure[Expenditure]).
4. In Data colors, set a distinct standout color for “OECD Average”.

The screenshot shows the Microsoft Power BI desktop interface. The ribbon menu at the top includes File, Home, Insert, Modeling, View, Optimize, and Help. The Home tab is selected. The Data ribbon group contains options like Get data, OneLake, SQL Server, Enter data, Dataverse, Recent sources, Transform data, Refresh data, New visual, Text box, More visuals, New visual calculation, New measure, Quick measure, Sensitivity, Publish, Prep data for Copilot AI, and Copilot. The main workspace displays a data grid with the following columns: country, First institution, Sum of national_rank, Sum of score, and Count of world_rank. The data rows include Brazil, Federal University of Ceará, 18, 44.03, 1; Egypt, Alexandria University, 4, 44.03, 1; Portugal, University of the Algarve, 7, 44.03, 1; Spain, University of A Coruña, 40, 44.02, 1; Taiwan, National Dong Hwa University, 49, 88.48, 2; China, China Pharmaceutical University, 332, 176.64, 3; and a total row for Total, Alexandria University, 450, 441.23, 5. To the right of the grid, there are sections for Data, Visualizations, and Filters. The bottom of the screen shows a navigation bar with tabs for Task 1 through Task 9 d, a plus sign for creating new items, and a search bar. The status bar at the bottom indicates "Page 10 of 11", "Light rain in the afternoon", "11:05", and "05-09-2025".

country	First institution	Sum of national_rank	Sum of score	Count of world_rank
Brazil	Federal University of Ceará	18	44.03	1
Egypt	Alexandria University	4	44.03	1
Portugal	University of the Algarve	7	44.03	1
Spain	University of A Coruña	40	44.02	1
Taiwan	National Dong Hwa University	49	88.48	2
China	China Pharmaceutical University	332	176.64	3
Total	Alexandria University	450	441.23	5

Task 10 – Create Workspace “Institution Analysis” & Schedule Daily Refresh (6 AM)

The screenshot shows the Power BI Home interface. On the left, there's a sidebar with various navigation options like Home, Create, Browse, OneLake catalog, Apps, Metrics, Monitor, Learn, Real-Time, Workspaces, My workspace, and Power BI. The main area displays 'Recommended' sections for 'My workspace' and 'Pro Project'. Below these are tabs for 'Recent', 'Favorites', 'My apps', and 'From external orgs'. A table lists recent items: 'PowerBlandPowerBIServiceFinalProject1' (Report, 2 months ago), 'PowerBlandPowerBIServiceFinalProject1' (Report, 3 months ago), 'PowerBlandPowerBIServiceFinalProject1' (Semantic model, 3 months ago), 'Pro Project' (Workspace, 3 months ago), and 'My workspace' (Workspace, 3 months ago). To the right, a modal window titled 'Create a workspace' is open, asking for a 'Name' (Institution Analysis) which is noted as available, and a 'Description' (PowerBI Assignment 3). It also includes fields for 'Domain' (optional) and 'Workspace image' (with upload and reset buttons). At the bottom of the modal are 'Apply' and 'Cancel' buttons.

The screenshot shows the 'Institution Analysis' workspace. The sidebar on the left is identical to the home page. The main area has a title bar with 'Power BI Institution Analysis' and 'PowerBI Assignment 3'. It features a toolbar with 'New item', 'New folder', 'Import', 'Migrate', 'Create app', 'Manage access', and 'Workspace settings'. Below this is a search bar and filter controls. A central area prompts the user to 'Choose from predesigned task flows or add a task to build one' with buttons for 'Select a predesigned task flow' and 'Add a task'. A large circular placeholder image with a paperclip icon and the text 'There's nothing here yet' is centered. At the bottom, it says 'Add something new, or upload something to see them here.'

Steps taken to publish and schedule refresh:

1. Publish your PBIX to app.powerbi.com.
2. In the Service → Workspaces → Create workspace → Name: Institution Analysis → Save.
3. Upload/Publish the report to this workspace.
4. Workspace → Datasets + dataflows → ... (dataset) → Settings.
5. If source is local Excel: configure On-premises Data Gateway and map sources.
6. Scheduled refresh:
 - Keep data up to date = On
 - Frequency = Daily
 - Time = 6:00 AM
 - Select correct Time zone
7. Click Apply and verify Next refresh time shows 6:00 AM daily.

The screenshot shows the Power BI service interface for the 'Institution Analysis' workspace. The left sidebar includes icons for Home, Create, Browse, OneLake catalog, Apps, Metrics, Monitor, Learn, Real-Time, Workspaces, and the current workspace, 'Institution Analysis'. The main content area displays the 'Gateway and cloud connections' section. It shows a 'Personal Gateway' listed under 'Gateway connections' with the status 'Running on UTKARSHANAN0605'. A tooltip indicates that updates to the 'C:\Users\utkarshanand\OneDrive - kii.ac.in\Desktop\pbiasignmentdata\expenditure.xlsx' data source have been applied. Below this, the 'Cloud connections' section shows 'No cloud connections'. At the bottom of the page are 'Apply' and 'Discard' buttons. Other sections visible include 'Data source credentials' (with links to edit credentials for 'data.xlsx' and 'expenditure.xlsx'), 'Parameters', 'Query Caching', 'Refresh', 'Server settings', 'Data connections', and 'Q&A'.

The screenshot shows the Power BI app settings for the 'Institutions Analysis' workspace. The left sidebar includes icons for Home, Create, Browse, OneLake catalog, Apps, Metrics, Monitor, Learn, Real-Time, Workspaces, Institutions Analysis, and a Power BI logo. The main content area has a search bar at the top. A trial message indicates 59 days left. A notification box says 'Assignment3 solutions refresh schedule updated' with a checkmark, stating 'Your updates to the Assignment3 solutions refresh schedule changes have been applied'. The settings section includes:

- Data source credentials: 'data.xlsx' and 'expenditure.xlsx' with 'Edit credentials' and 'Show in lineage view' options.
- Parameters: A collapsed section.
- Query Caching: A collapsed section.
- Refresh:
 - Time zone: '(UTC+05:30) Chennai, Kolkata, Mumbai, India'
 - Configure a refresh schedule: A note states 'Time zone configuration is applied not only to determine the schedule refresh time but also to establish the current date and time for incremental refresh models during on-demand and API refreshes.' It includes a link to 'Learn more'.
 - Toggle switch: 'On'.
 - Refresh frequency: 'Daily'.
 - Time: '6 00 AM'.
 - Add another time: A link.
- Send refresh failure notifications to:
 - Semantic model owner
 - These contacts: 'Enter email addresses' input field.

At the bottom are 'Apply' and 'Discard' buttons.

**Prepared By,
Utkarsh Anand**