unionbank

TURBOCHARGED
FOR SUCCESS

Introduction to Data Visualization using Power Bi

September 2021

Table Of Content

Introduction to Power Bi

What is Power BI?

• Power BI is a cloud-based business analysis and intelligence service by Microsoft. It is a collection of business intelligence and data visualization tools such as software services, apps and data connectors.

Microsoft offers three types of Power BI platforms:

- Power BI Desktop (A desktop application)
- Power BI Service (SaaS i.e., Software as a Service)
- Power BI Mobile (For iOS and Android devices)

Introduction to Power Bi

History of Power BI

- Power BI is a Microsoft's product initially released on 11th July 2011. It was originally designed and created by Ron George in 2010, who released it with the name "Project Crescent". Later in September of 2013, Microsoft changed the name to Power BI and launched it for the public.
- Power BI's first general public release was on July 24th, 2015. As of 2019, Power BI has been officially declared as one of the leading BI tools by 2019 *Gartner Magic Quadrant for Analytics and Business Intelligence Platform.*

Connecting to a Data Source

Connecting to Data Sources

Data Sources

Power BI can already connect to over 110 different data sources and connection types, with more being added. As well as connecting to Text/CSV files.

Commonly used Power BI data sources include:

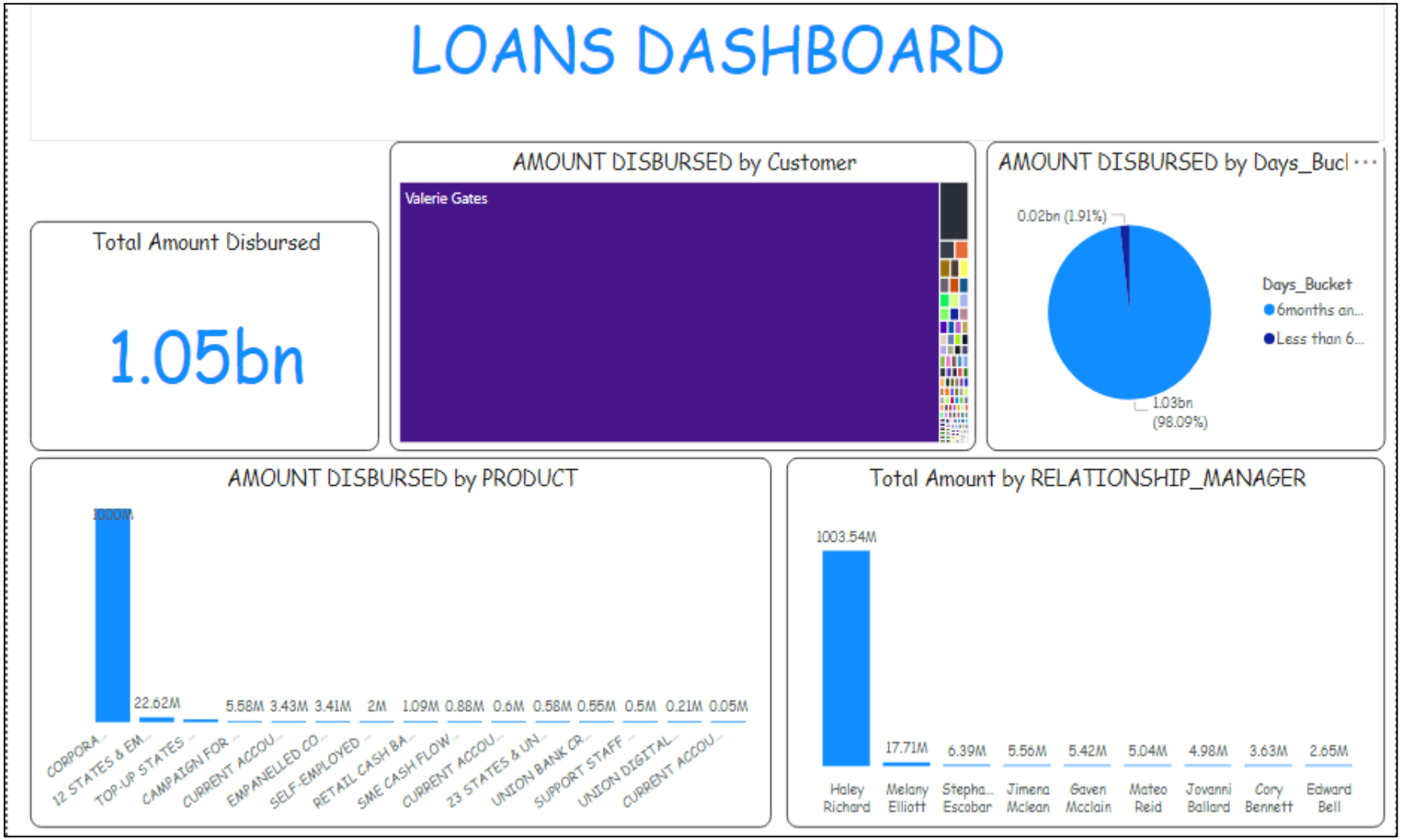
- File (Excel, Text/CSV, XML, JSON, PDF).
- Database (SQL Server, Oracle, IBM DB2, MySQL, PostgreSQL, Snowflake, etc).
- Power Platform (Power BI datasets, Power BI Dataflows, Common Data Services)
- Azure (SQL Database, Synapse Server, Analysis Services, Blob Storage, Data Lake, Cosmos DB, etc).

Connecting Modes

Some data sources allow you to choose the Data Connectivity mode (i.e. connecting directly to data). There are four options available:

- Import
- DirectQuery
- Mixed Mode
- Connect Live

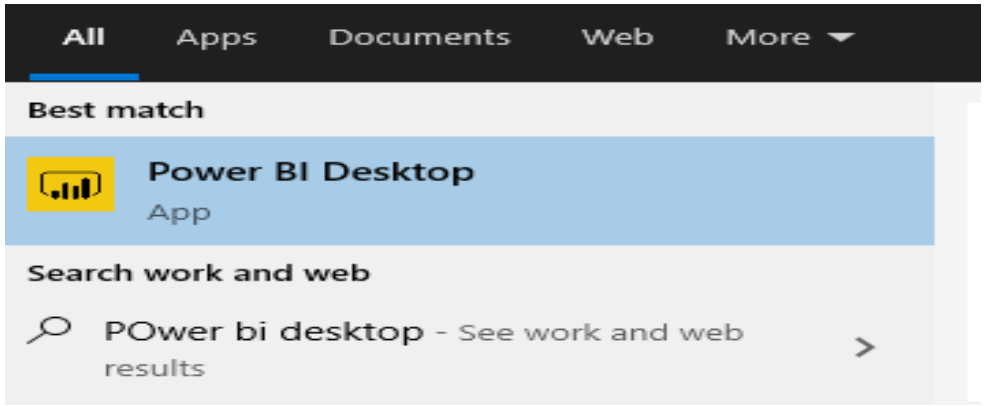
Sample Dashboard



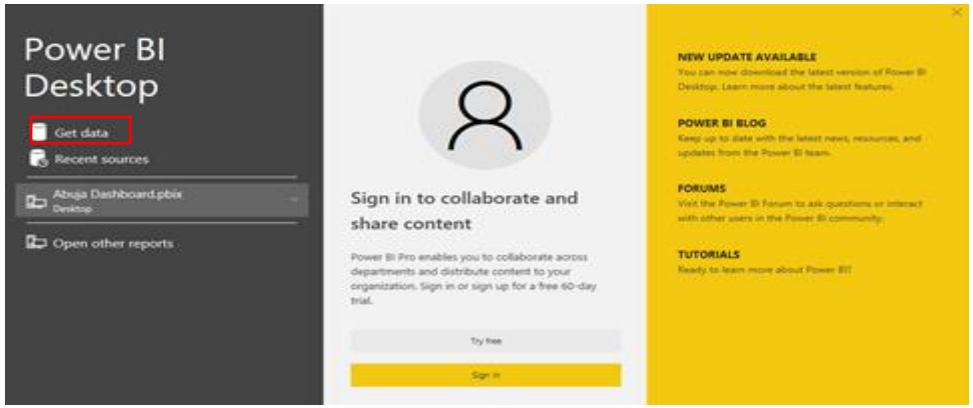
Importing Data

Importing Data

- Open **Power BI Desktop** as shown in the Figure below.



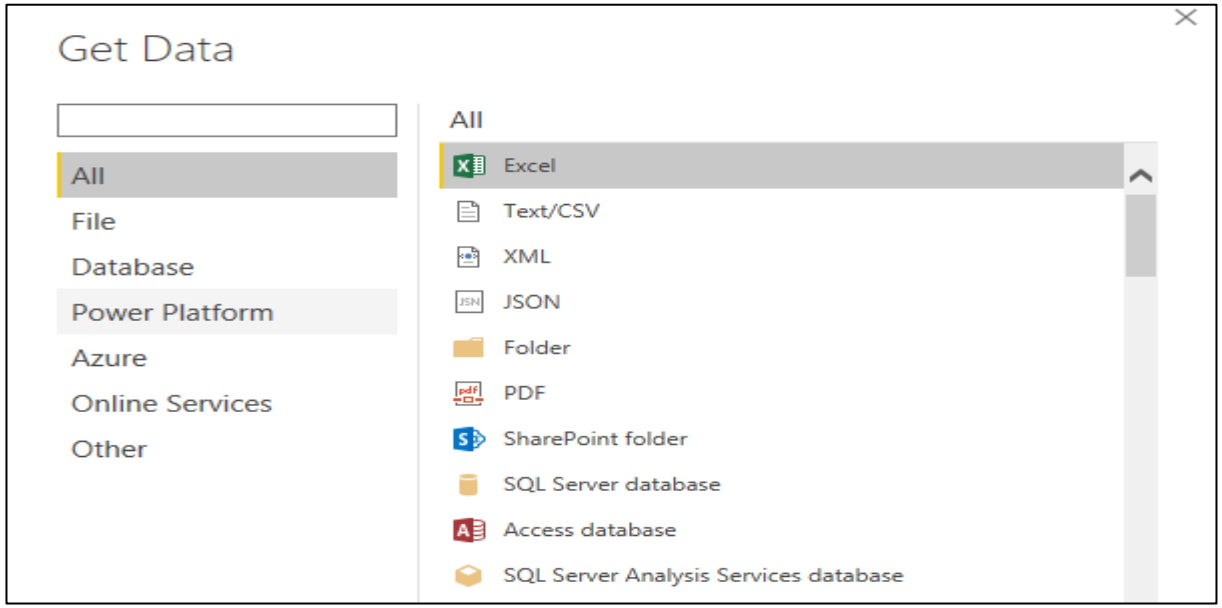
- When you launch Power BI Desktop, a welcome splash screen is displayed
- To connect to the sample data for this exercise, select **Get data**



Importing Data

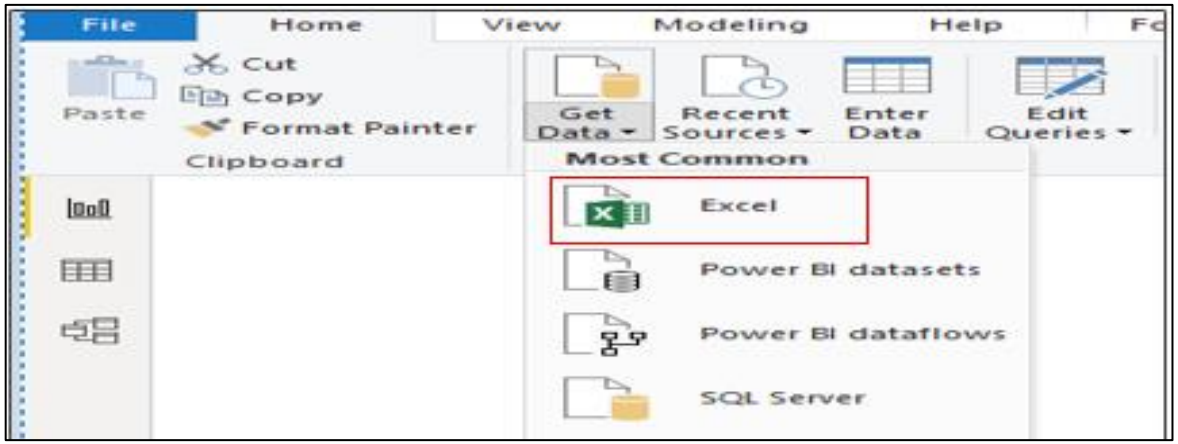
- Alternatively, click the Get data button from the Home tab on the ribbon bar

- Selecting the down arrow on the Get data button shows the most common data sources menu. Click on get more to see the list..

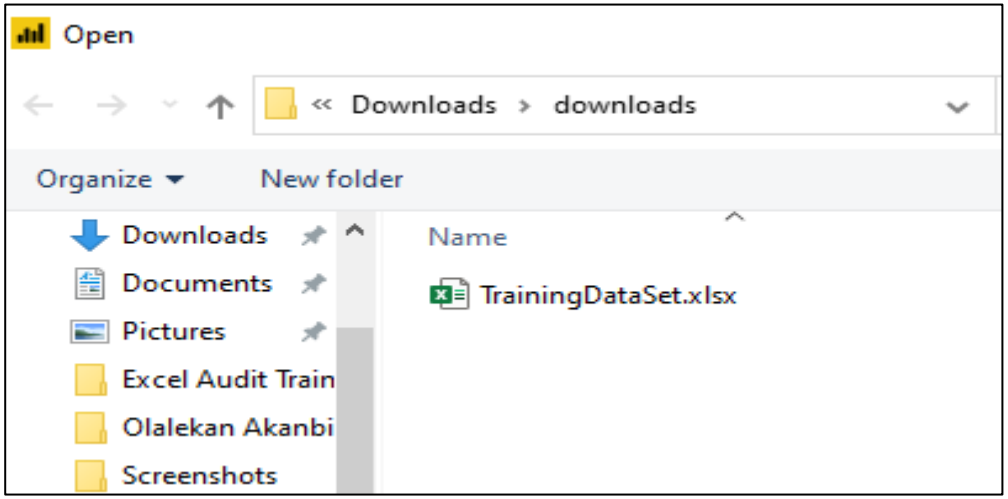


Importing Data

Select **Excel** from the list



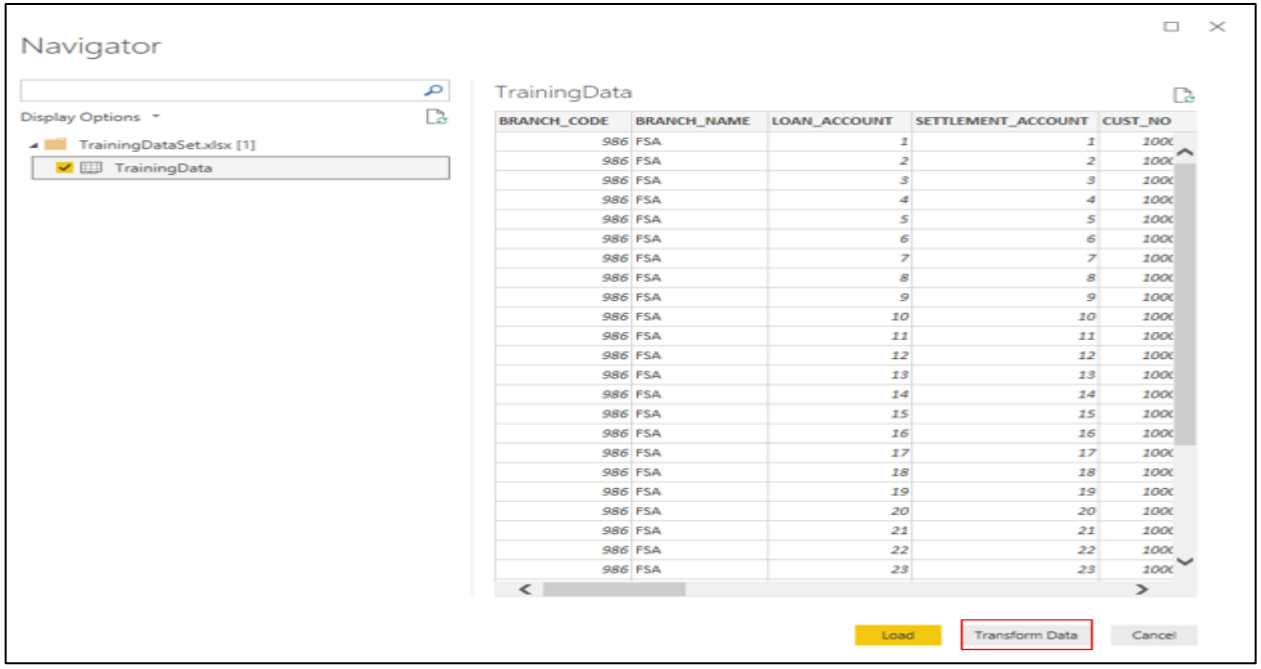
- Browse to the **download** files, select the file **TrainingDataSet.xlsx** and click **Open.**



Transforming Data

Transforming Data

- We have two options, **Load** or **Transform Data**. We want to transform our data, so we click the **Transform Data** button.



- Clicking Load will import the tables exactly as is. It is still possible to transform the data later by clicking the Transform Data button from the Home tab in Power BI Desktop.

Transforming Data

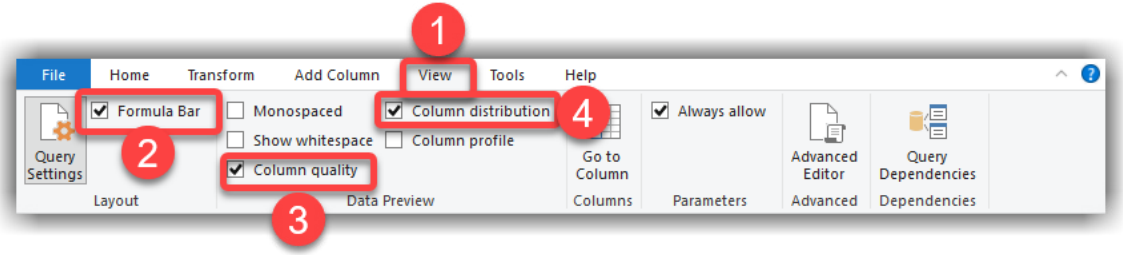
Clicking Transform Data opens a separate window - the Power Query Editor: a powerful data profiling and data preparation tool.

Let's take a moment to explore the Power Query interface.

The Application Ribbson contains all options and settings. Complete the steps:

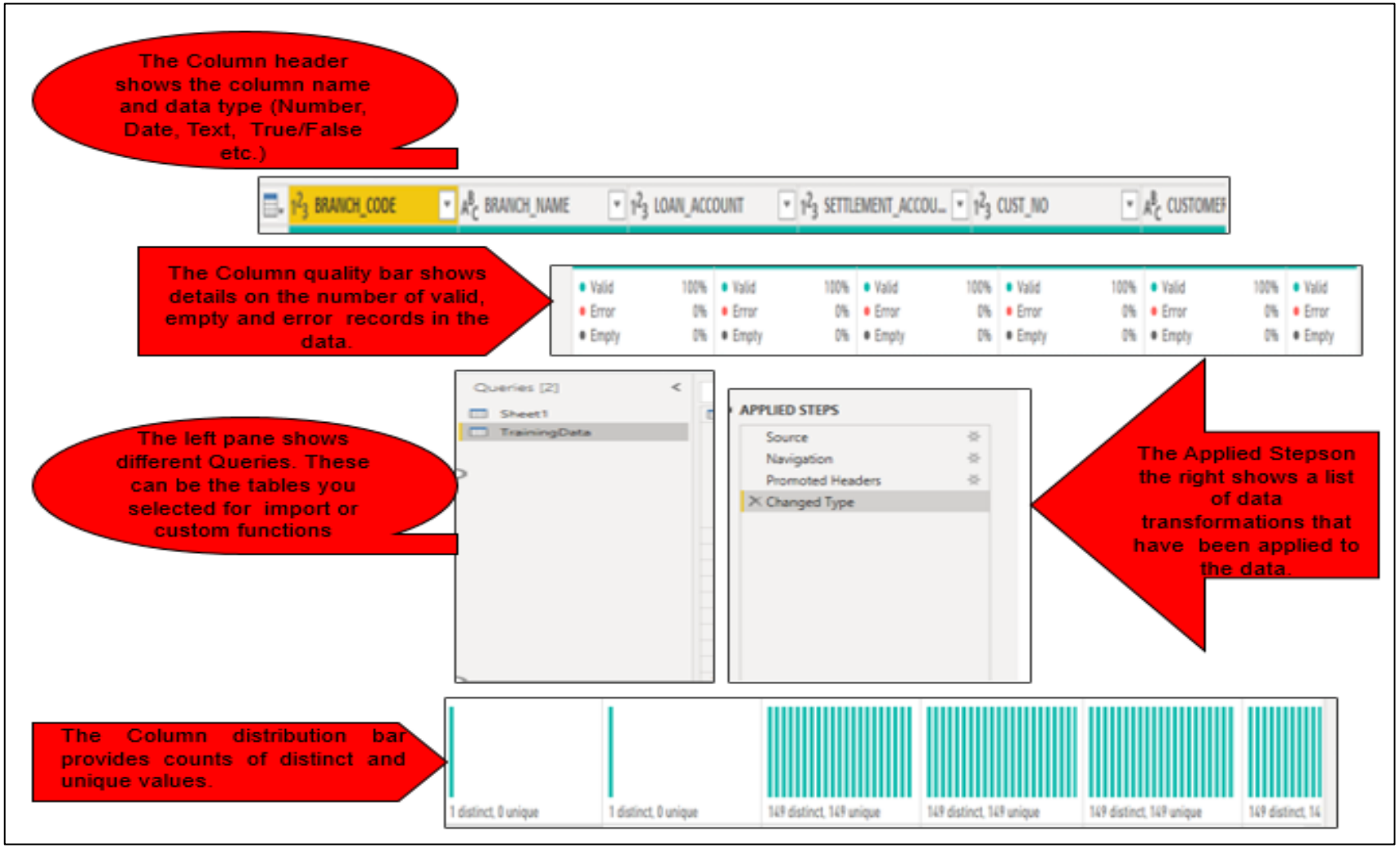
1. Click the View tab from the ribbon. Make sure the following are ticked:
2. Formula Bar
3. Column quality
4. Column distribution

Note: If show white space is ticked, you can leave as is.



Transforming Data

The application ribbons contain all options and settings, transformations, and other settings configurations.

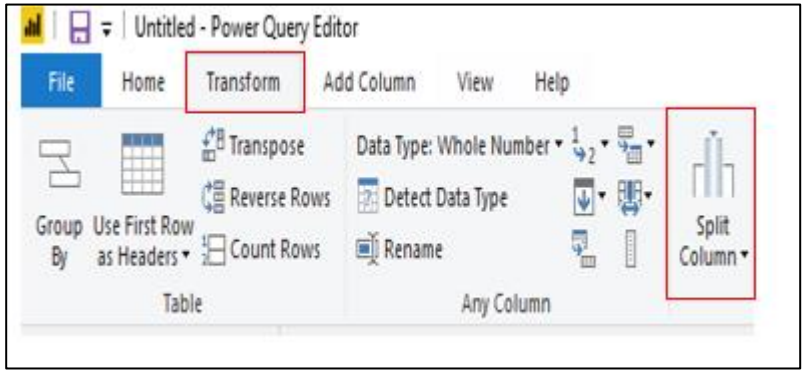


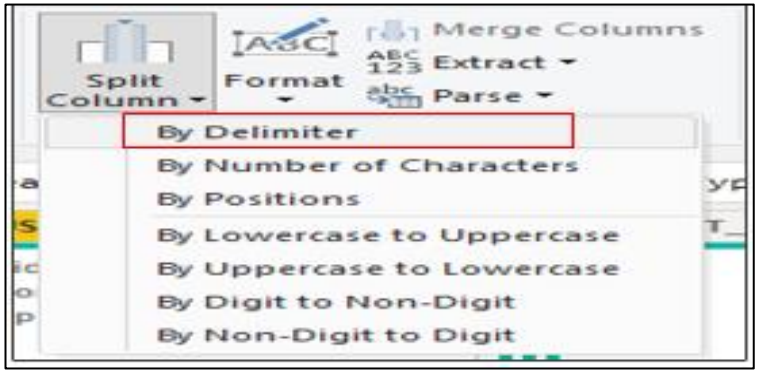
Transforming Data

Power Query allows us to split a column into one or more other columns. There are various options available under the Split Column dropdown.

We will be using By Delimiter to split the Name into First and Last Name.

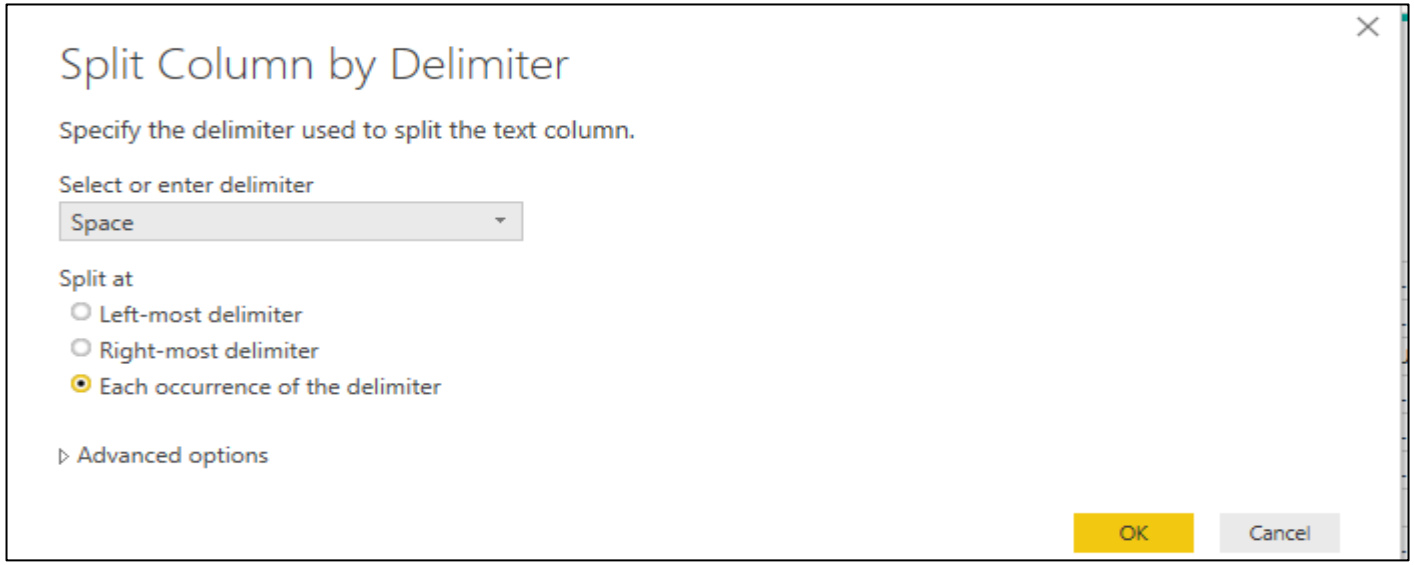
- Go to the Queries pane on the left and select TrainingData.
- Select the Customer Name column.
- From the ribbon bar, go to the Transform tab, click Split Column and select By **Delimiter**.





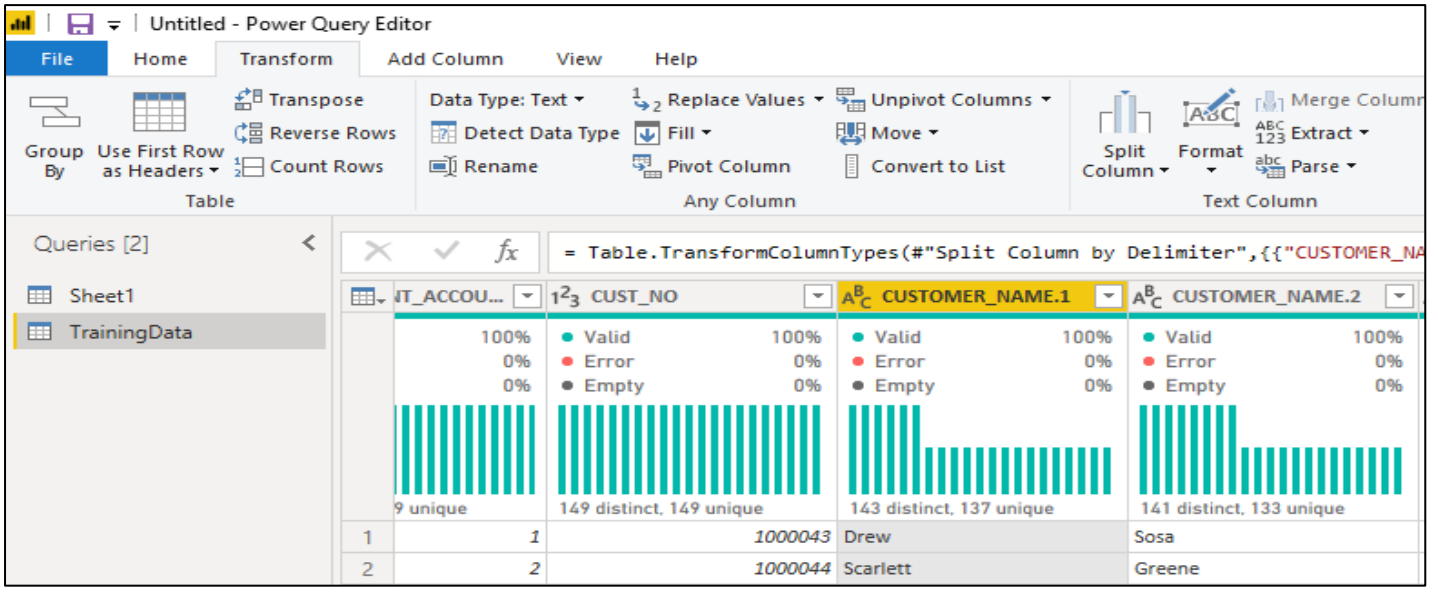
Transforming Data

- In the dialog box, ensure **Select or enter delimiter** is set to **Space**.
- Select **Each occurrence of the delimiter** and click **OK**.



- The Customer_Name column is split into Customer_Name.1 that has the First Name and Customer_Name.2 that has the Last Name. The splitting column by delimiter action above creates a new applied step as shown in the next figure:

Transforming Data

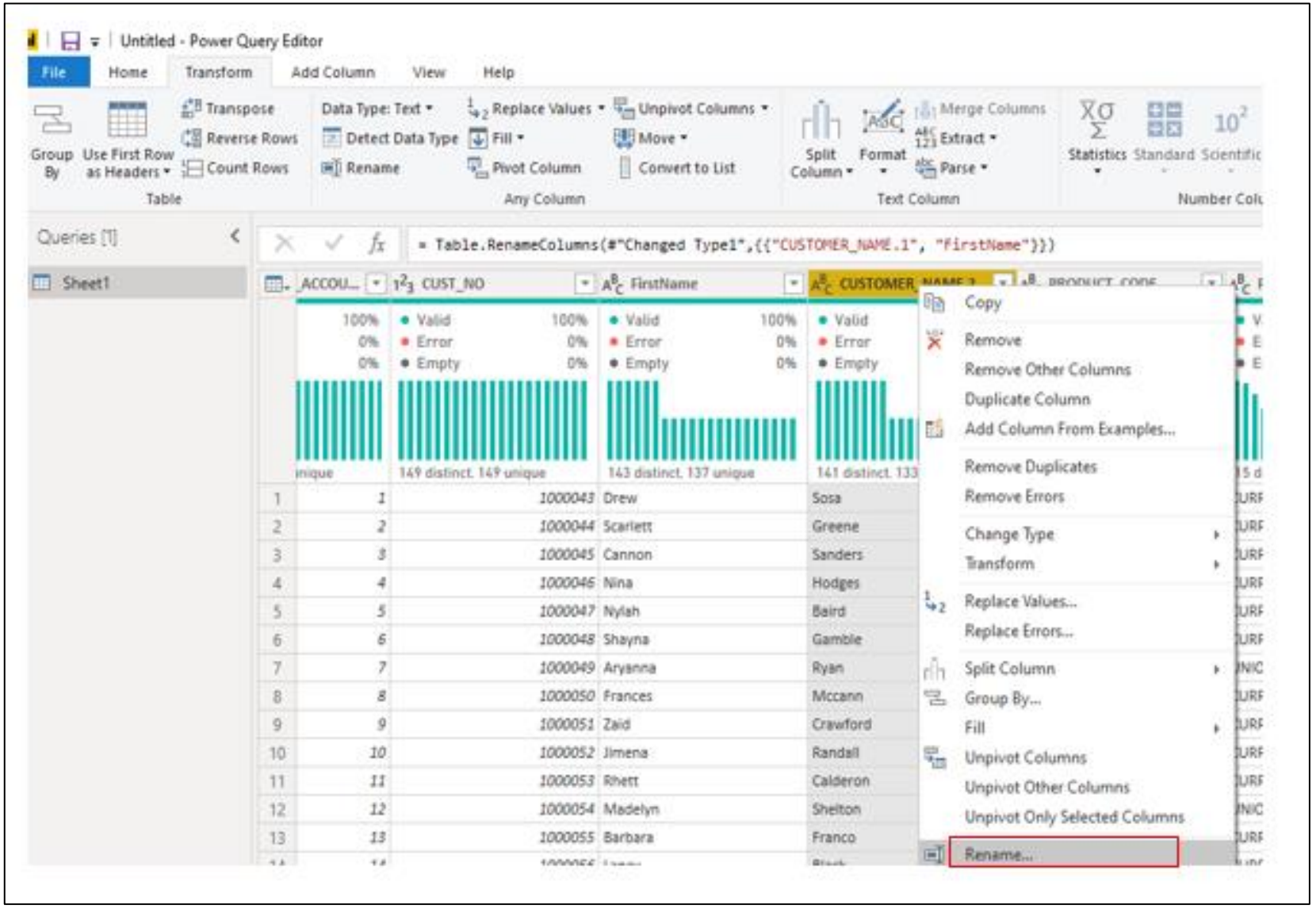


Renaming Columns

We will rename the two columns created by the splitting of the Name column. There are a couple of ways to do this: either double click the Header name or right-click the header and select rename.

- Double click the Customer_Name.1 Header and rename it to FirstName.
- Right click the Customer _Name.2 Header and rename it to LastName.

Transforming Data



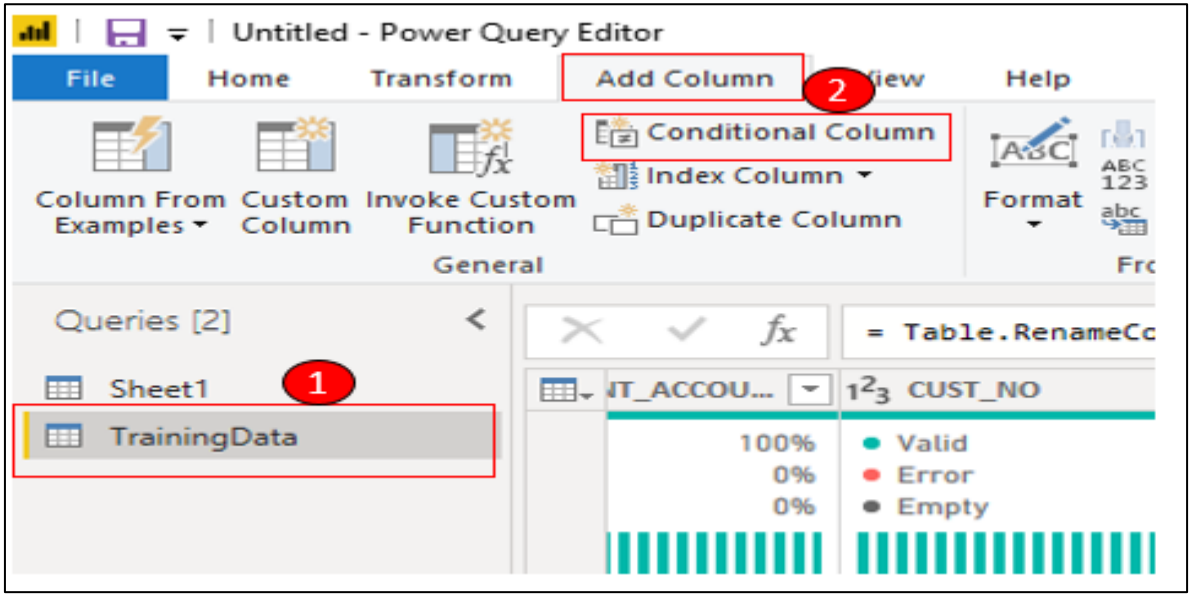
Transforming Data

Adding Columns

There is a dedicated tab in the Ribbon bar for adding columns. In the next few steps, we use a Conditional Column and a Column from Example to add new columns.

In the steps below, we will add a new conditional column DaysBucket.

1. Go to the Queries pane, select the TrainingData query.
2. Go to the Ribbon bar, select Add Column, select Conditional Column.

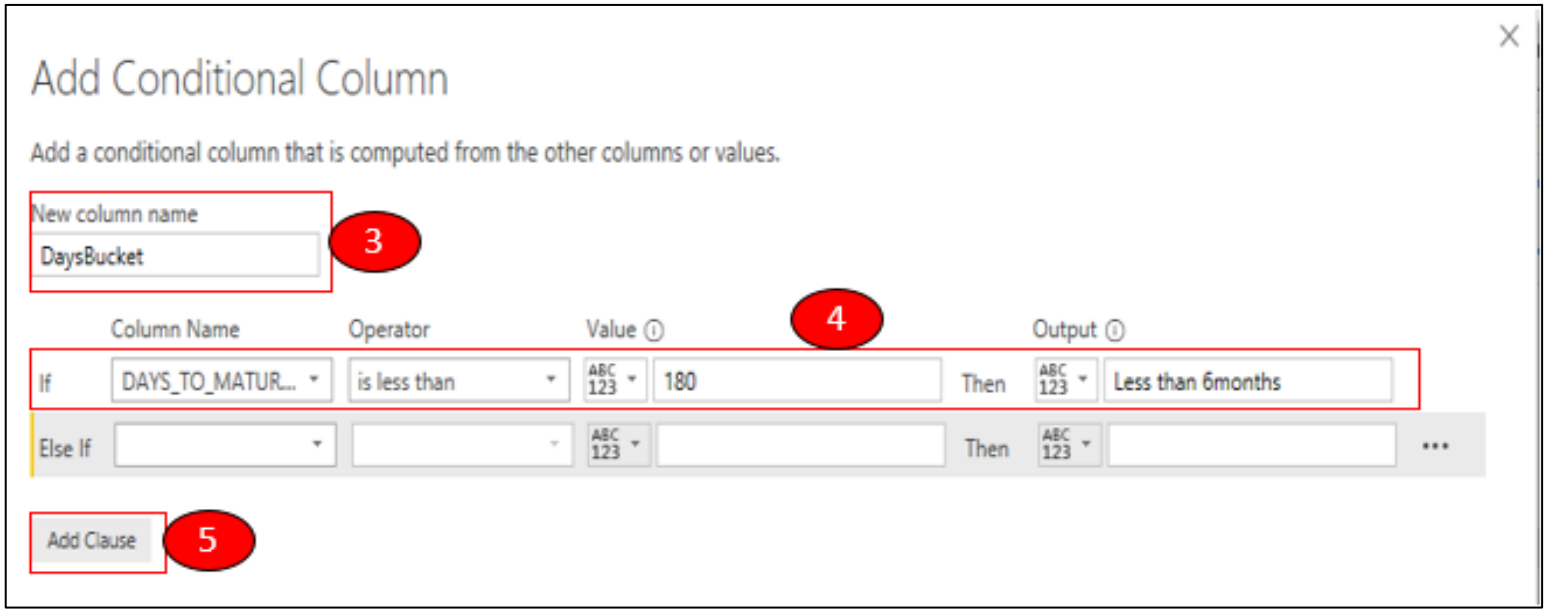


Transforming Data

3. In the dialog box, change column name from Custom to DaysBucket.

4. Modify the If statement to read: If Days_to_Maturity less than 180 then less than 6months.

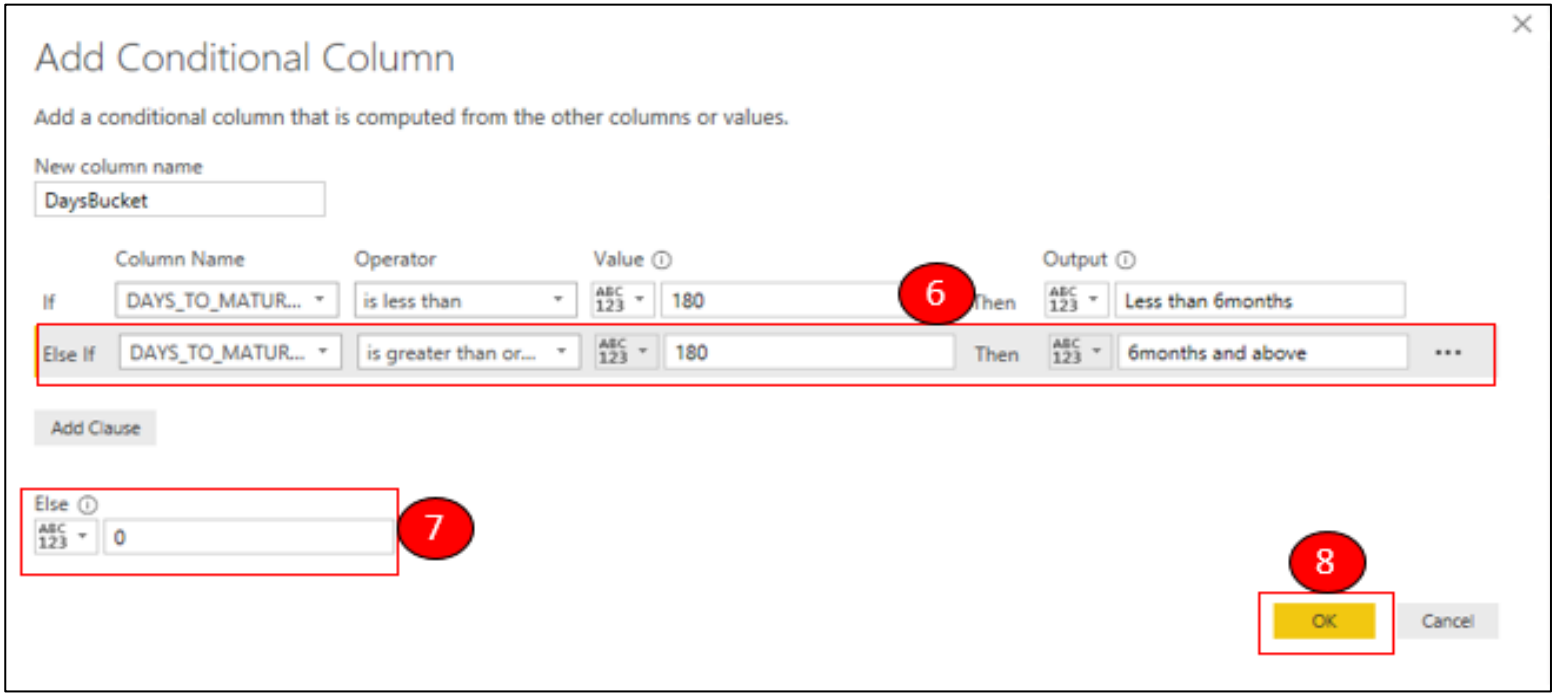
5. Click Add Clause.



Transforming Data

6. Add the additional ‘if’ clauses for the remaining values.
7. Enter 0 for Else, this will assign 0 to any remaining records including those with the value ‘None’.
8. Click OK and our new column will be added.

*Important: Note that Power Query is CAsE sEnSiTive.



Transforming Data

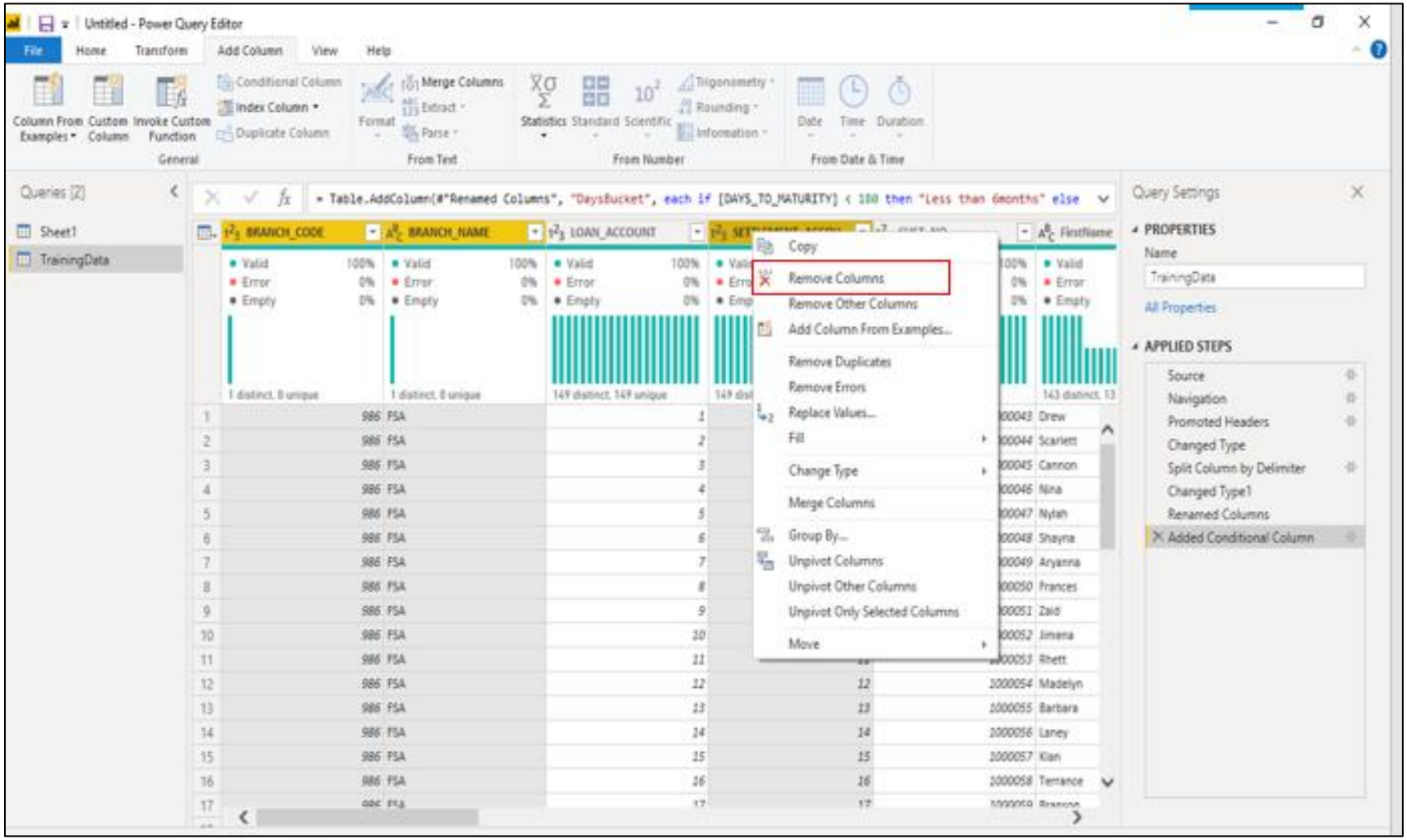
Removing Columns

We can remove columns that we no longer need e.g. duplicate columns. You can remove a column that has previously been used to create a new column, It does not impact on the data in the new column.

- Select the TrainingData query from the Queries panel.
- Hold Ctrl and click the column headers for Rm_code, start_date, Product_code, settlement_account, cust_no, branch_code and Branch_name.
- Right click in any of the selected column headers and click Remove Columns.

Transforming Data

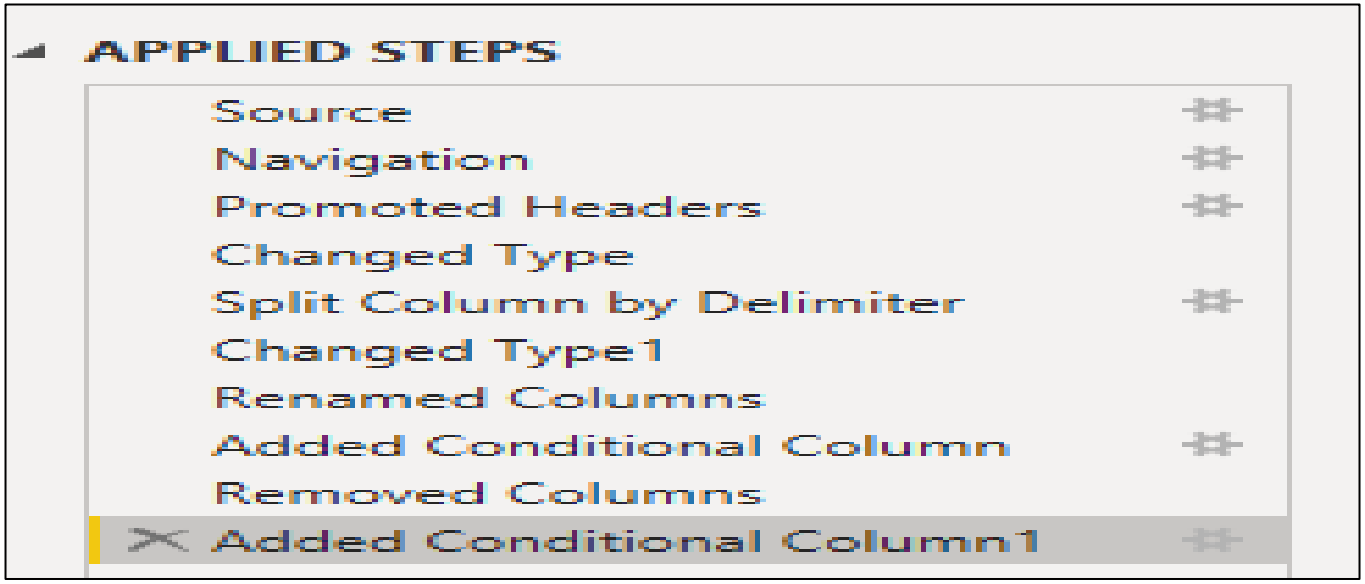
Removing Columns



Transforming Data

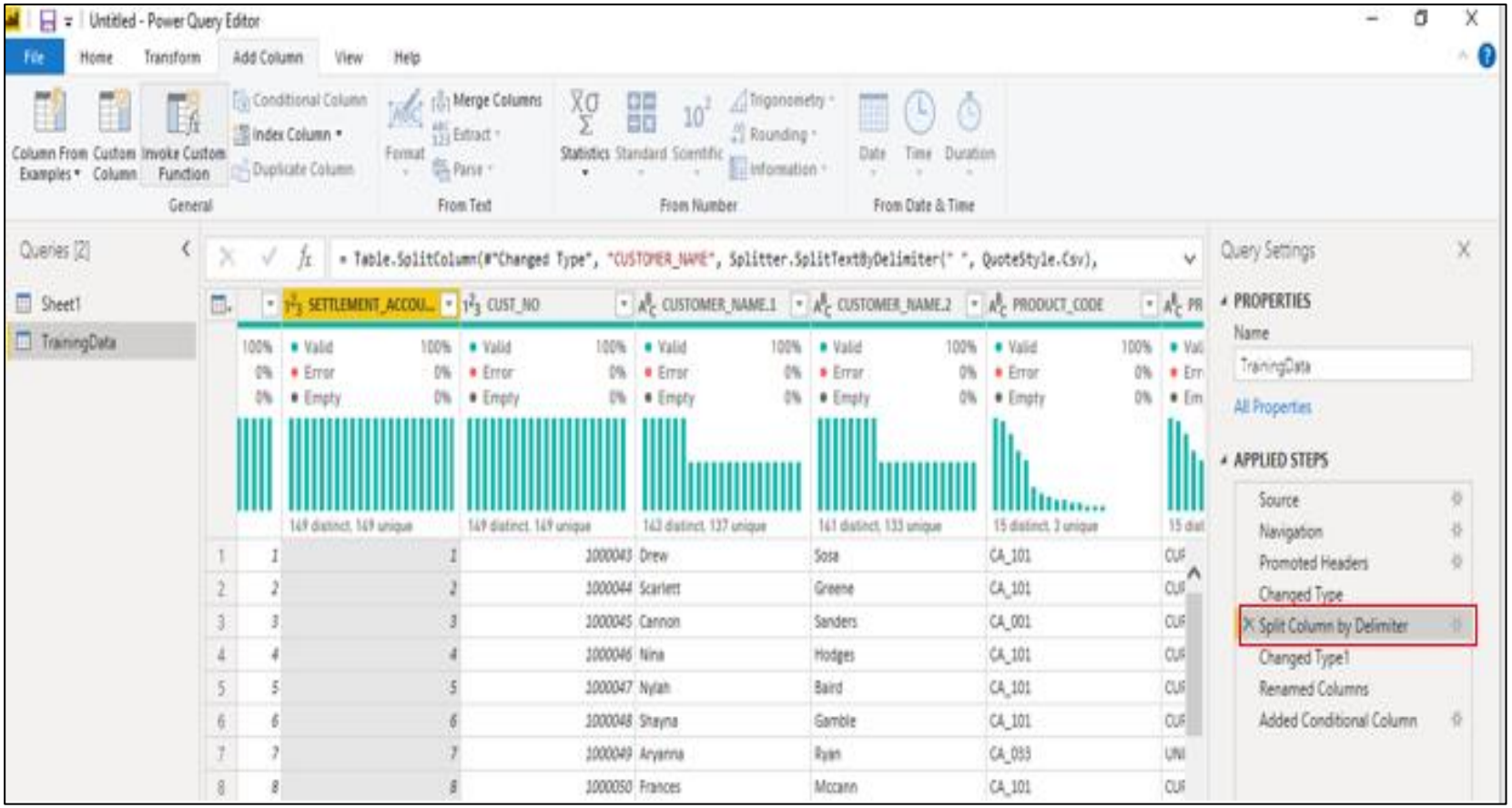
Transformation Steps

All transformations are applied as a series of steps shown in the **Query Settings** pane as shown in this Figure.



Transforming Data

It's possible to see the resulting data of each step by clicking on a desired step from the **Query Settings** pane as shown in Figure below. This makes it easy to visually inspect the data after applying each transformation step.



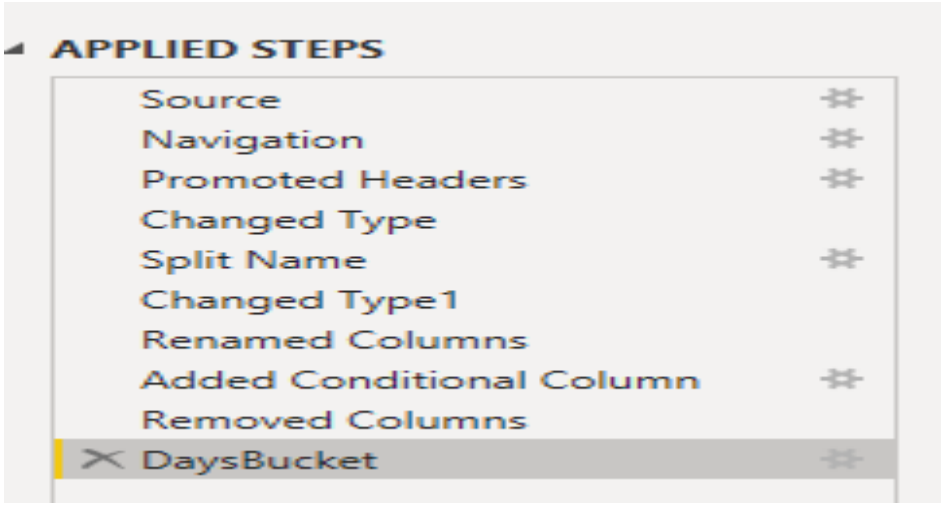
Transforming Data

Renaming an applied step

The existing steps can be renamed within the Power Query Editor. This shows what the step did and makes future modification easier.

- Select the TrainingData query in the Queries panel.
- Right click the Split Column by Delimiter step.
- Click Rename from the context menu and change to Split Name.
- Right click the Added Conditional Column step.
- Click Rename and change to Add DaysBucket.

The next Figure shows the renamed steps.



Transforming Data

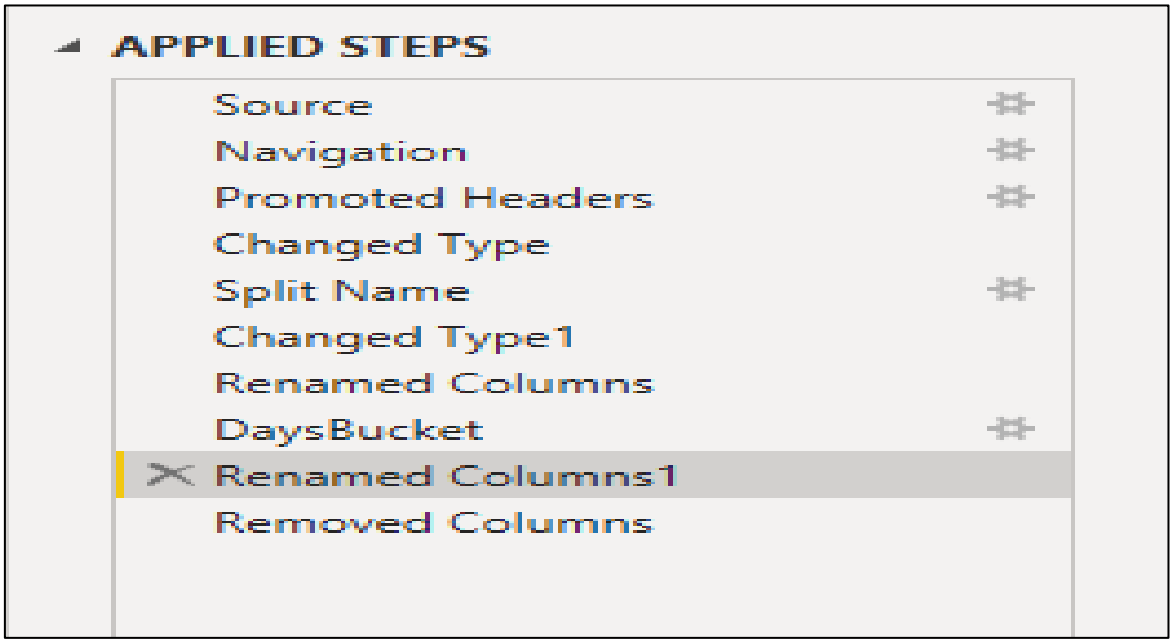
Adding a step between an applied step

The results of the last step contain the data that will be imported into the data model and available for visualisation. Let's add a step that renames DaysBucket to Days_Bucket, but we'll add it after the DaysBucket column was added.

1. Select the TrainingData query in the Queries pane.
2. Go to the Applied Steps pane and click Add DaysBucket.
3. Right click DaysBucket column
4. Select Rename
5. Type Days_Bucket and press enter
6. A popup will ask if you are sure you want to insert the step. Click Insert.

You should see a step called Renamed Columns1 is added between Removed Columns and Add DaysBucket.

Transforming Data

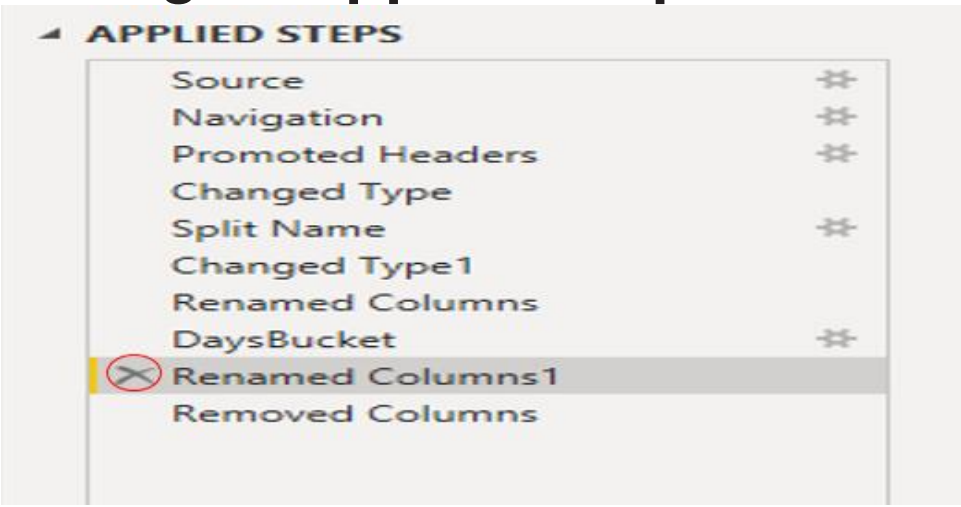


Deleting an applied step

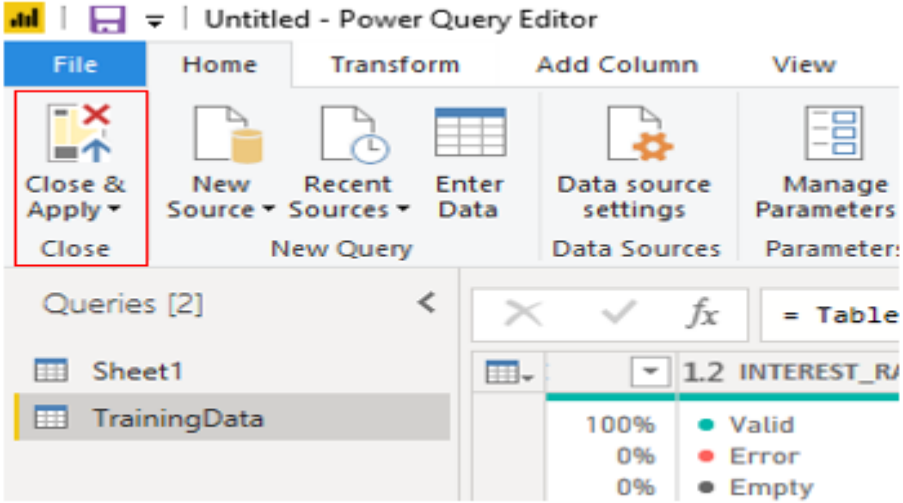
All steps can be deleted except the source steps. You can either delete just the selected step or all steps from that one until the end.

Transforming Data

Deleting an applied step



- Click on close & apply.



Modelling the Data

Modelling the Data

Data modelling is one of the most important aspects of data analysis, regardless of the tools we use, and Power BI is not an exception. After we import data into the data model, we need to create analytical calculations and implement the business logics available for data visualization.

Creating Calculated Columns and Measures with DAX

A data model consists of Tables and their Relationships. There are also other elements included in the data modelling:

1. Calculated Tables
2. Calculated Columns
3. Measures

All these can be created programmatically using DAX.

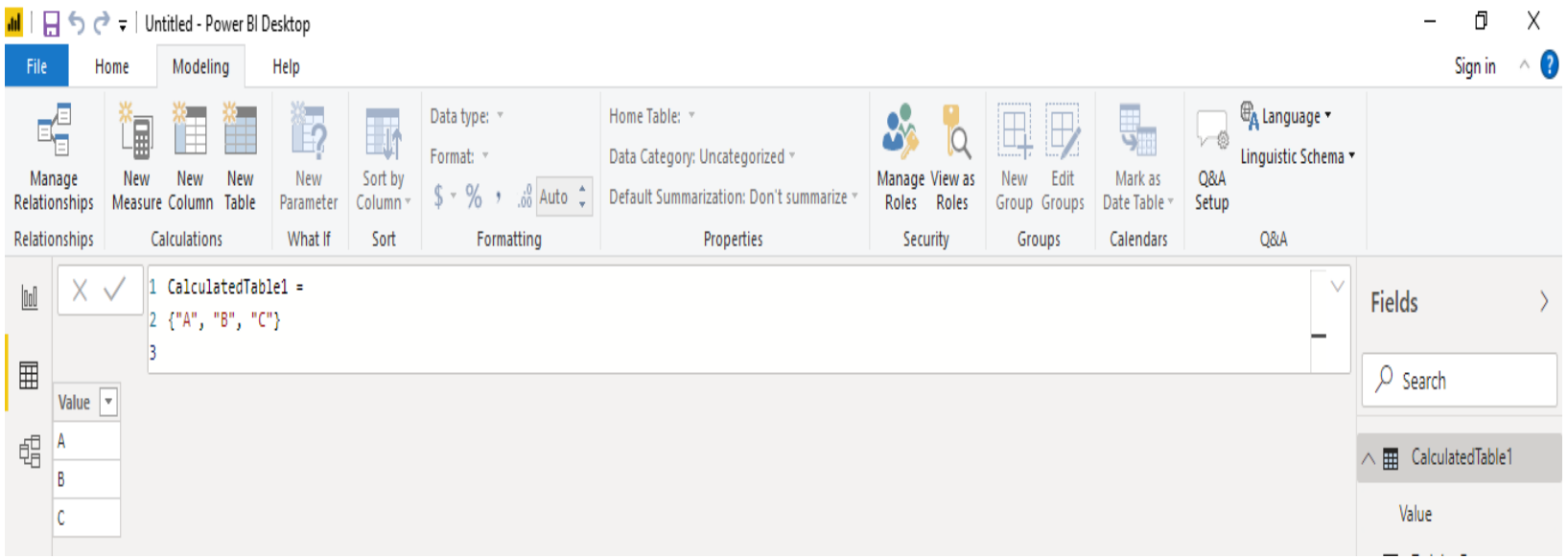
Modelling the Data

Calculated Tables

On some occasions, you need to add new tables based on data you've already loaded into the model. These tables can be created using DAX.

You can also use table constructor in DAX to create a calculated table. Table constructor isn't a function; it's a set of characters that allow you to create a table in DAX. For example, the following DAX expression creates a table with one column.

```
CalculatedTable1 =
  {"A", "B", "C"}
```



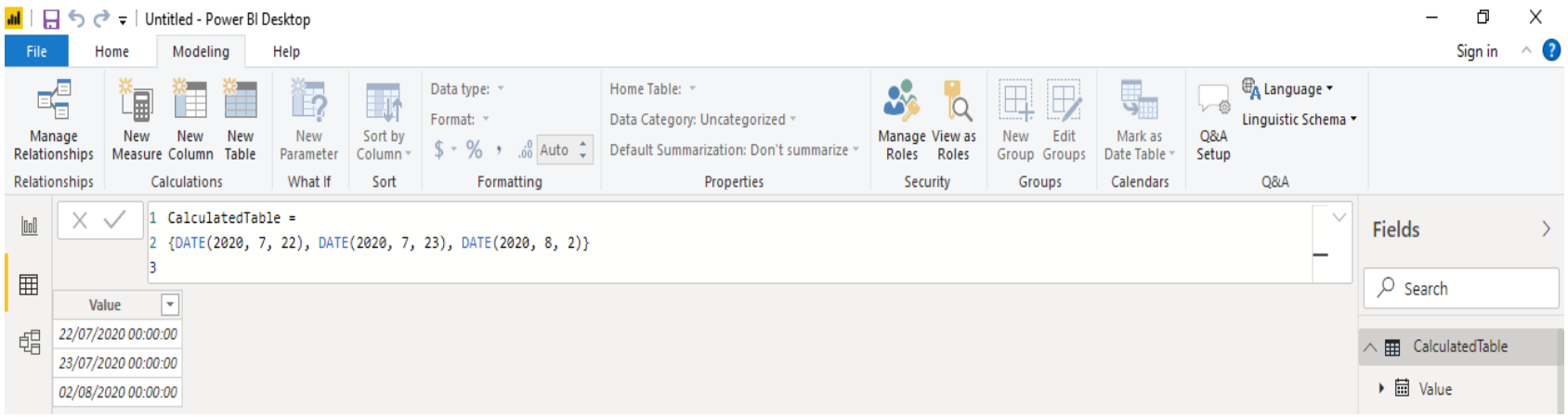
Modelling the Data

You can use curly brackets {} to construct a table. This is a simple form of using table constructor in DAX.

- Start of table construct - open curly bracket {
- “A”, “B” and “C” are the values of a single column
- End of table construct - close curly bracket }

You can use any scalar DAX expressions in the values. The example below creates another calculated table using the table constructor and the DATE() function:

```
CalculatedTable=
  {DATE(2020, 7, 22), DATE(2020, 7, 23), DATE(2020, 8, 2)}
```



Modelling the Data

Calculated columns

Calculated columns are the new columns created in the data model using DAX. There are many scenarios when you want to create a calculated column; the general rule of thumb is that you only create a new calculated column if:

- There's a complex scenario and you want to create calculated columns to use them in other calculations like measures.
- You need to create a new calculated column to be used in a slicer or as a filter on a report element.

Other than that, you should avoid creating new calculated columns as there's risk of performance degradation.

How To Create Calculated Columns:

A. Right click on a table from the Fields pane and select New Column.

B. Select a table from the Fields pane: select New Column from the Table Tools tab.

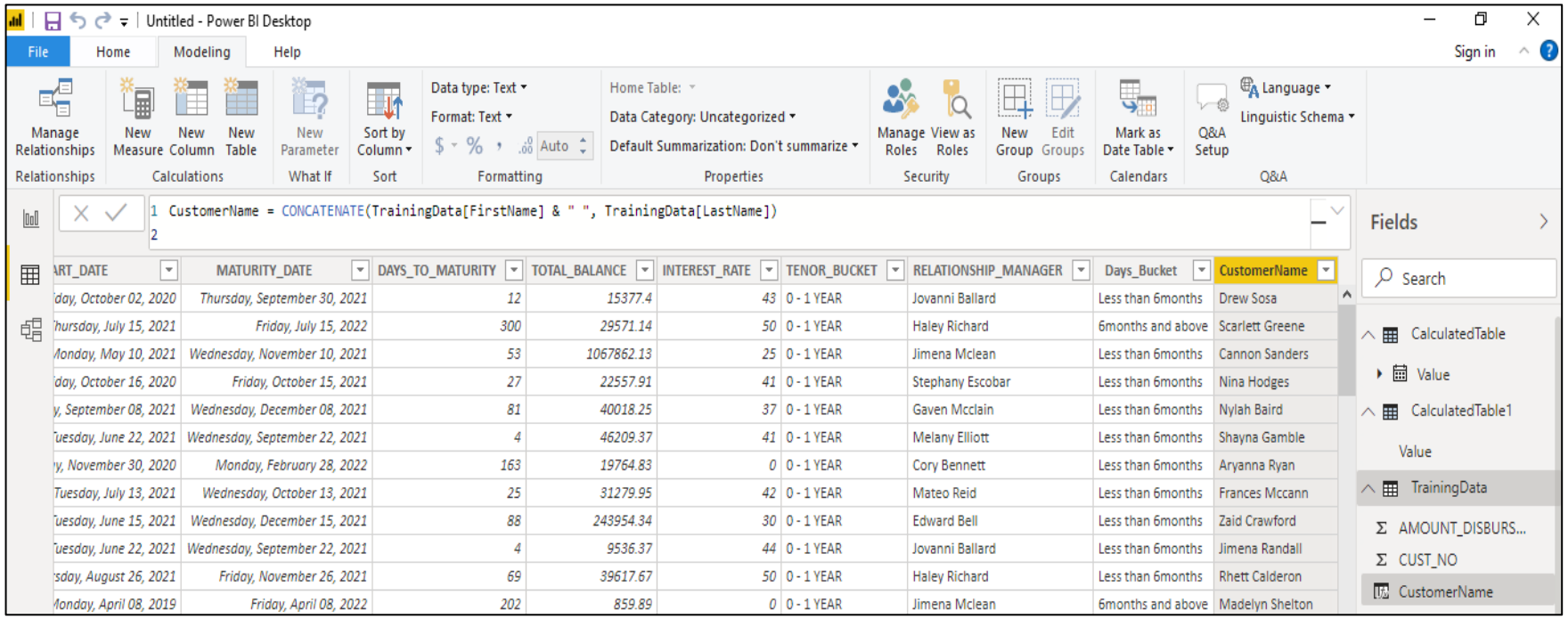
C. Select a table from the Fields pane: select New Column from the Modelling tab.

Modelling the Data

Let's have a look at Calculated Columns in action. Follow the steps below:

1. Click the **Data view** tab from the left pane.
2. Go to the **Fields** column on the right, select **TrainingData.**
3. Right click, go to **New column**. Type in the expression below.
4. Press enter.

```
CustomerName = CONCATENATE(TrainingData[FirstName] & " ", TrainingData[LastName])
```



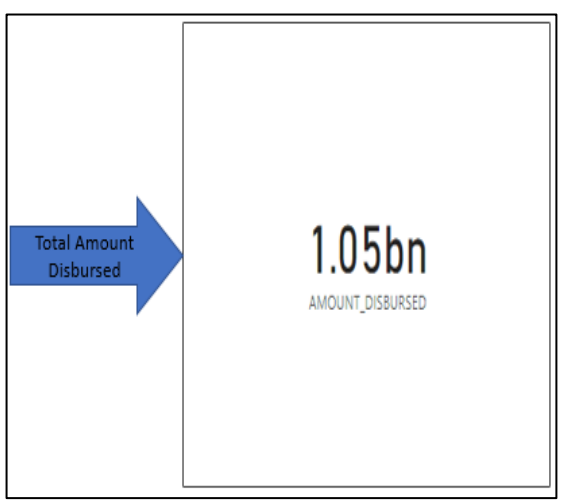
ART_DATE	MATURITY_DATE	DAYS_TO_MATURITY	TOTAL_BALANCE	INTEREST_RATE	TENOR_BUCKET	RELATIONSHIP_MANAGER	Days_Bucket	CustomerName
day, October 02, 2020	Thursday, September 30, 2021	12	15377.4	43	0 - 1 YEAR	Jovanni Ballard	Less than 6months	Drew Sosa
hursday, July 15, 2021	Friday, July 15, 2022	300	29571.14	50	0 - 1 YEAR	Haley Richard	6months and above	Scarlett Greene
Monday, May 10, 2021	Wednesday, November 10, 2021	53	1067862.13	25	0 - 1 YEAR	Jimena Mclean	Less than 6months	Cannon Sanders
day, October 16, 2020	Friday, October 15, 2021	27	22557.91	41	0 - 1 YEAR	Stephany Escobar	Less than 6months	Nina Hodges
y, September 08, 2021	Wednesday, December 08, 2021	81	40018.25	37	0 - 1 YEAR	Gaven Mcclain	Less than 6months	Nylah Baird
uesday, June 22, 2021	Wednesday, September 22, 2021	4	46209.37	41	0 - 1 YEAR	Melany Elliott	Less than 6months	Shayna Gamble
y, November 30, 2020	Monday, February 28, 2022	163	19764.83	0	0 - 1 YEAR	Cory Bennett	Less than 6months	Aryanna Ryan
Tuesday, July 13, 2021	Wednesday, October 13, 2021	25	31279.95	42	0 - 1 YEAR	Mateo Reid	Less than 6months	Frances Mccann
uesday, June 15, 2021	Wednesday, December 15, 2021	88	243954.34	30	0 - 1 YEAR	Edward Bell	Less than 6months	Zaid Crawford
uesday, June 22, 2021	Wednesday, September 22, 2021	4	9536.37	44	0 - 1 YEAR	Jovanni Ballard	Less than 6months	Jimena Randall
sday, August 26, 2021	Friday, November 26, 2021	69	39617.67	50	0 - 1 YEAR	Haley Richard	Less than 6months	Rhett Calderon
Monday, April 08, 2019	Friday, April 08, 2022	202	859.89	0	0 - 1 YEAR	Jimena Mclean	6months and above	Madelyn Shelton

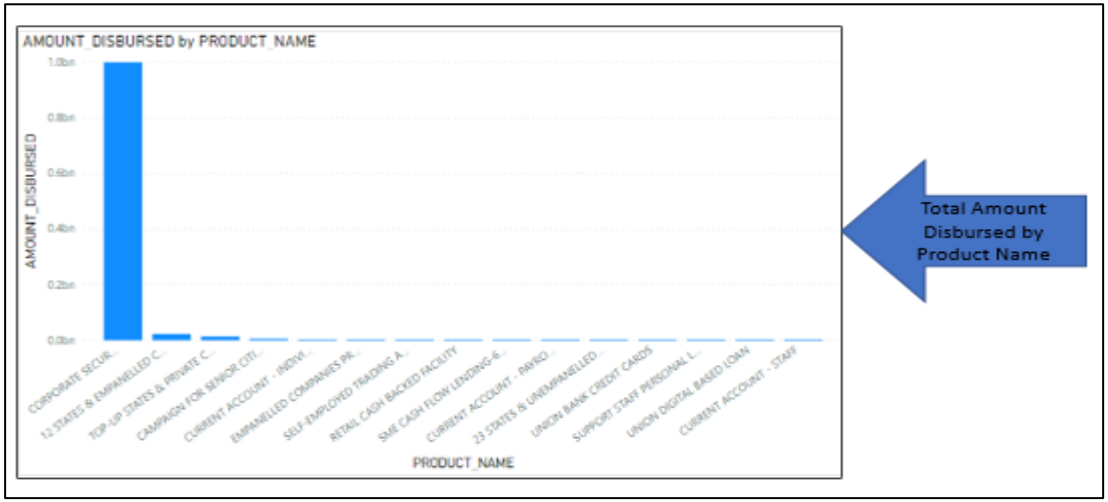
Modelling the Data

Measures

Measures are normally analytical calculations: summations, calculating averages, minimum, maximum, counts and so on. You can use the measures for visuals in Power BI.

The result of the measures change depending on how we interact with them in different visuals. For instance, you create a measure to calculate Amount_disbursed. If you put the Amount_disbursed on a Card visual, it shows total Amount disbursed over the whole data. If you use the same measure in a Column Chart with Product_Name on the Axis, the measure always calculates the correct results for each category.





Modelling the Data

You can create measures in various ways, just the way you create calculated columns - either from the Report view or from the Data view.

A. Right click on a table from the **Fields** pane and select **New measure.**

B. Select a table from the **Fields** pane and select **New measure** from the **Table Tools** tab.

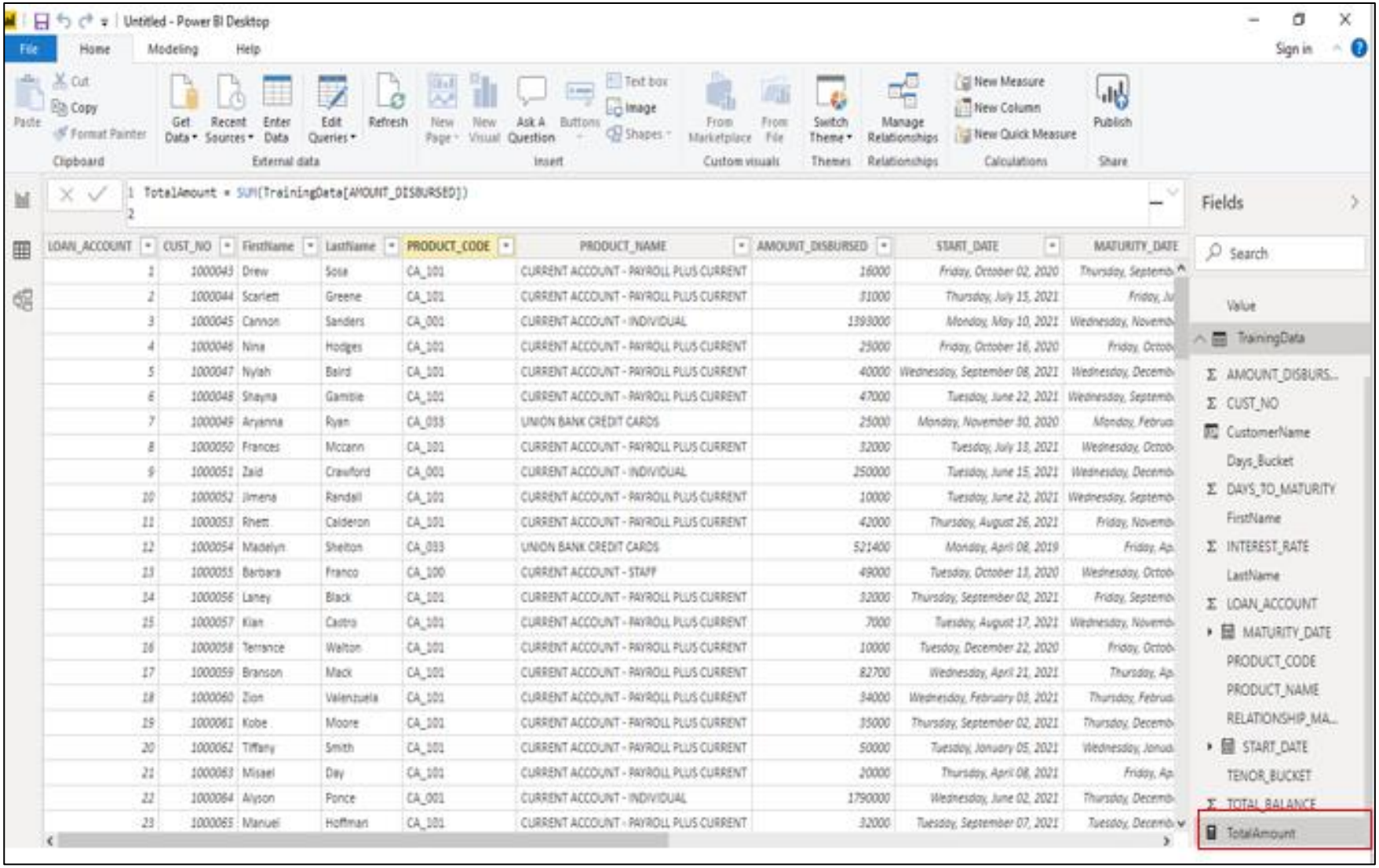
C. Select a table from the **Fields** pane and select **New measure** from the **Modelling** tab.

Follow these steps below:

1. Click the **Report** view on the left pane.
2. Right click the **TrainingData** table in the **Fields** pane.
3. Click **New Measure.**
4. Use the DAX expression below. This creates a **TotalAmount** measure in the **TrainingData** table.

```
TotalAmount = SUM(TrainingData[AMOUNT_DISBURSED])
```

Modelling the Data



LOAN_ACCOUNT	CUST_NO	FirstName	LastName	PRODUCT_CODE	PRODUCT_NAME	AMOUNT_DISBURSED	START_DATE	MATURITY_DATE
1	1000043	Drew	Sosa	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	16000	Friday, October 02, 2020	Thursday, Septemb
2	1000044	Scarlett	Greene	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	31000	Thursday, July 15, 2021	Friday, Ju
3	1000045	Cannon	Sanders	CA_001	CURRENT ACCOUNT - INDIVIDUAL	1393000	Monday, May 10, 2021	Wednesday, Novemb
4	1000046	Nina	Hodges	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	25000	Friday, October 16, 2020	Friday, Octob
5	1000047	Nylah	Baird	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	40000	Wednesday, September 08, 2021	Wednesday, Decemb
6	1000048	Shayna	Gamble	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	47000	Tuesday, June 22, 2021	Wednesday, Septemb
7	1000049	Aryanna	Ryan	CA_033	UNION BANK CREDIT CARDS	25000	Monday, November 30, 2020	Monday, Februa
8	1000050	Frances	Mccann	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	32000	Tuesday, July 13, 2021	Wednesday, Octob
9	1000051	Zaid	Crawford	CA_001	CURRENT ACCOUNT - INDIVIDUAL	250000	Tuesday, June 15, 2021	Wednesday, Decemb
10	1000052	Jimena	Randall	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	10000	Tuesday, June 22, 2021	Wednesday, Septemb
11	1000053	Rhett	Calderon	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	42000	Thursday, August 26, 2021	Friday, Novemb
12	1000054	Madelyn	Shelton	CA_033	UNION BANK CREDIT CARDS	521400	Monday, April 08, 2019	Friday, Ap
13	1000055	Barbara	Franco	CA_100	CURRENT ACCOUNT - STAFF	49000	Tuesday, October 13, 2020	Wednesday, Octob
14	1000056	Laney	Black	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	32000	Thursday, September 02, 2021	Friday, Septemb
15	1000057	Kian	Castro	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	7000	Tuesday, August 17, 2021	Wednesday, Novemb
16	1000058	Terrance	Walton	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	10000	Tuesday, December 22, 2020	Friday, Octob
17	1000059	Branson	Mack	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	82700	Wednesday, April 21, 2021	Thursday, Ap
18	1000060	Zion	Valenzuela	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	34000	Wednesday, February 03, 2021	Thursday, Februa
19	1000061	Kobe	Moore	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	35000	Thursday, September 02, 2021	Thursday, Decemb
20	1000062	Tiffany	Smith	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	50000	Tuesday, January 05, 2021	Wednesday, Janua
21	1000063	Misael	Day	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	20000	Thursday, April 08, 2021	Friday, Ap
22	1000064	Alyson	Ponce	CA_001	CURRENT ACCOUNT - INDIVIDUAL	1790000	Wednesday, June 02, 2021	Thursday, Decemb
23	1000065	Manuel	Hoffman	CA_101	CURRENT ACCOUNT - PAYROLL PLUS CURRENT	32000	Tuesday, September 07, 2021	Tuesday, Decemb

Reporting on the Data - Visualization

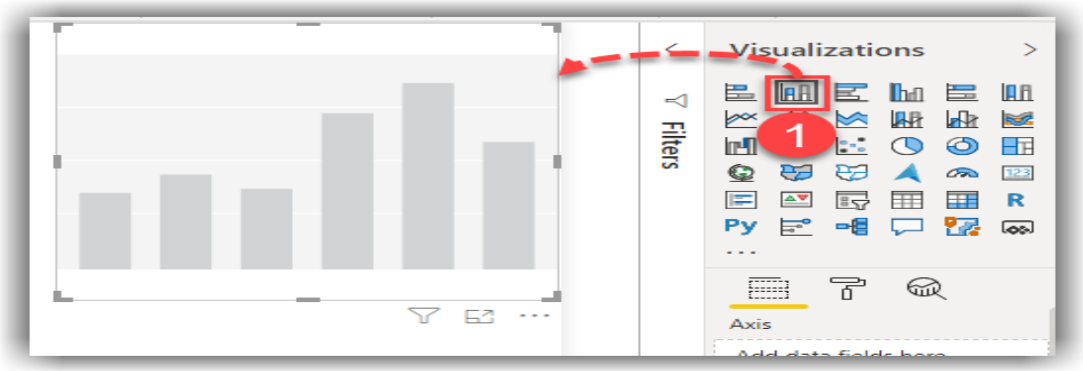
Reporting on the Data – Visualizations

Building Basic Visualisations

Let's start with a graph to support a simple business scenario.

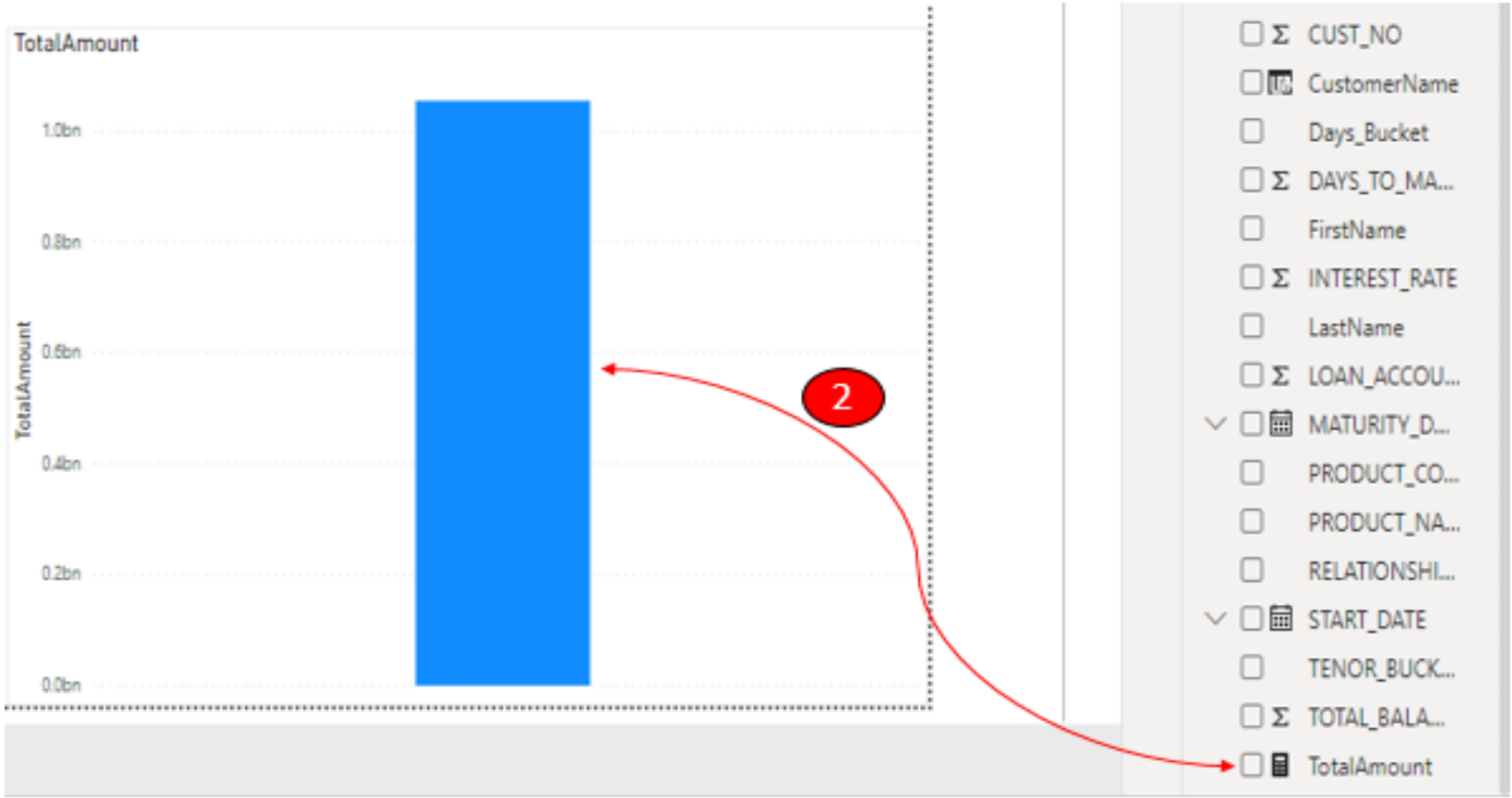
SCENARIO: Follow these steps:
Our BSM wants to track the value of loans disbursed by each Relationship Manager.

1. From the **Report** view, select the **Stacked Column Chart** visual from the **Visualizations** pane (This will add a grey "place holder" graphic to the Report Canvas).



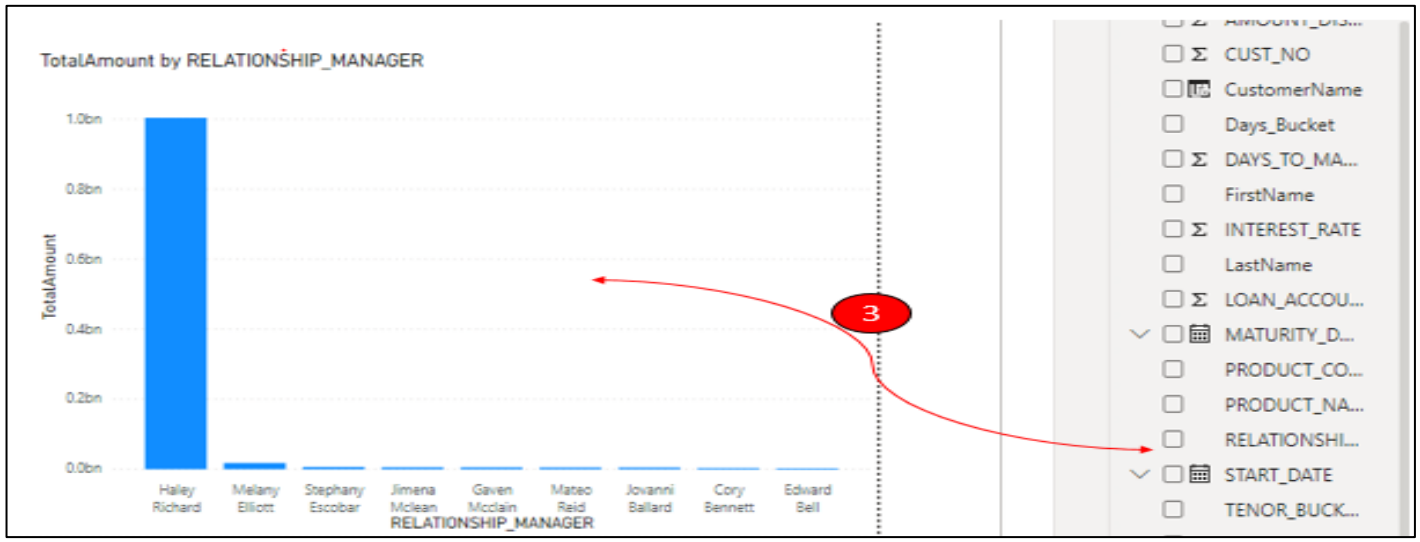
Reporting on the Data – Visualizations

2. Drag the **TotalAmount** measure, drop it on the visual place holder (this creates one column).



Reporting on the Data – Visualizations

3. Drag the **Relationship_Manager** column from the **TrainingData** table and drop it on the visual (this adds the RM to the visual's Axis).

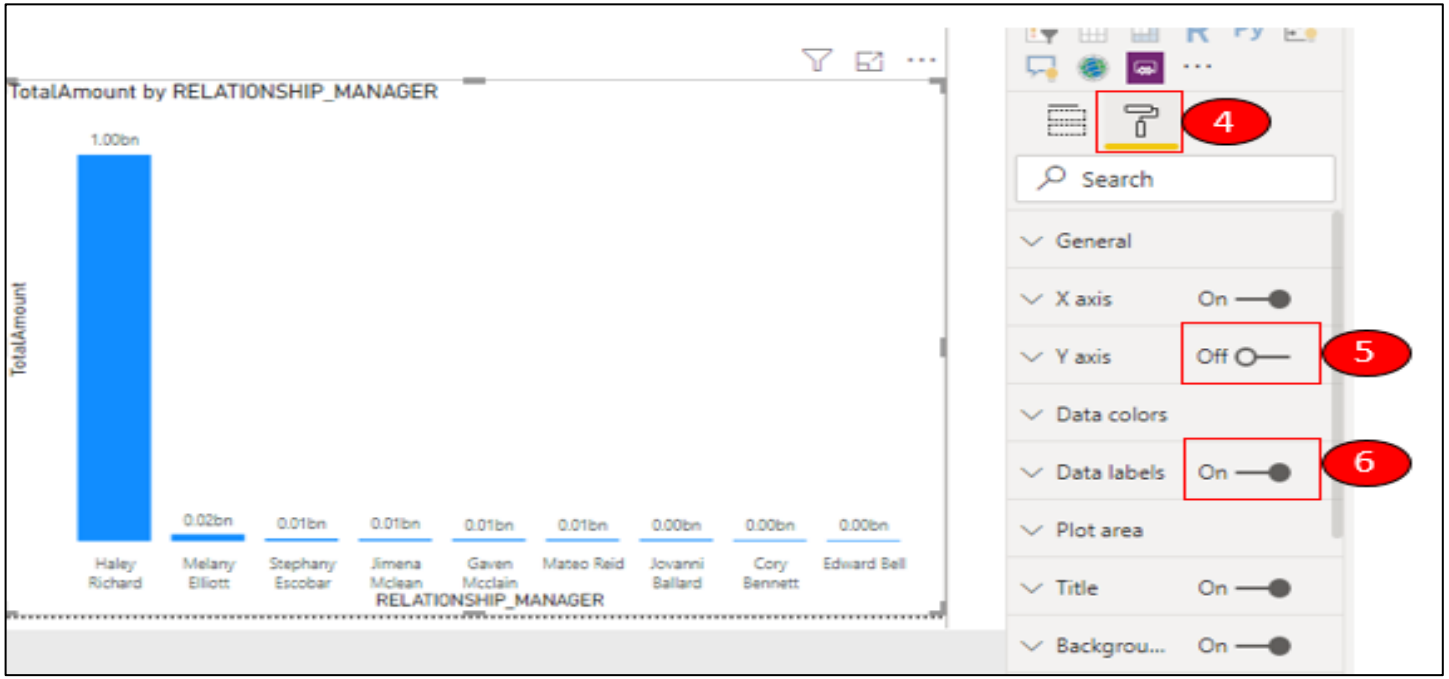


4. Click the chart, go to the the Visualizations pane, click the **format** icon from the visualizations pane.

5. Disable Y Axis.

6. Enable **Data Labels** to show values on each bar. This makes the visual more readable.

Reporting on the Data – Visualizations

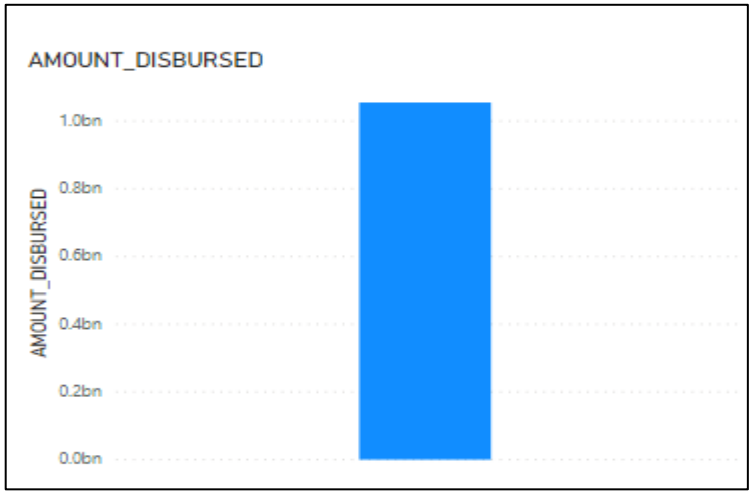


Next, our BSM would like to know how much was disbursed for the two different days_buckets.

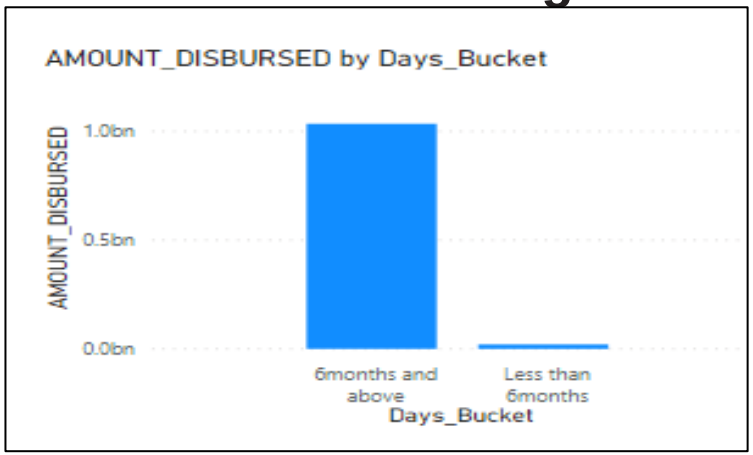
We'll follow a slightly different approach to build this visualisation:

Reporting on the Data – Visualizations

7. Drag the **Amount_disbursed** column from the **TrainingData** and drop it onto a blank space of the report canvas. This creates a new Visualization using the default Stacked Column Chart.

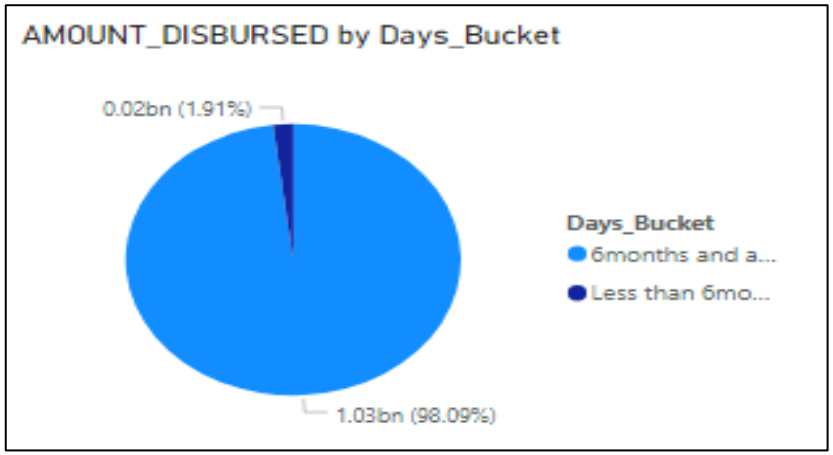


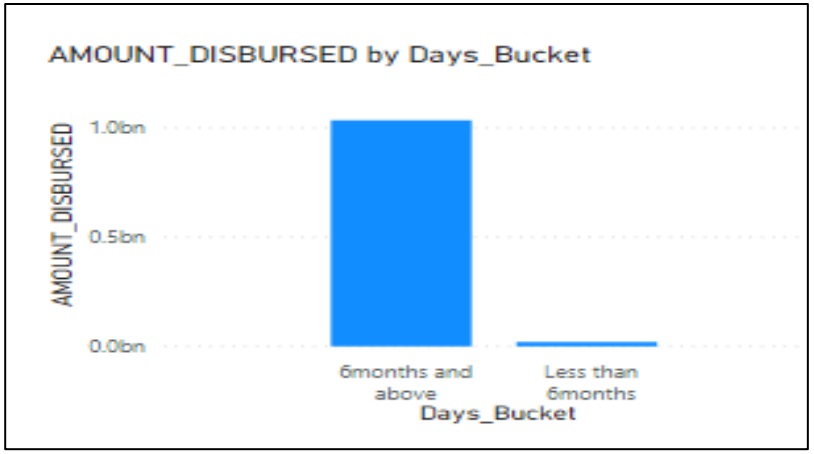
8. Drag the **Days_bucket** column from the **TrainingData** and drop it on the newly created chart.



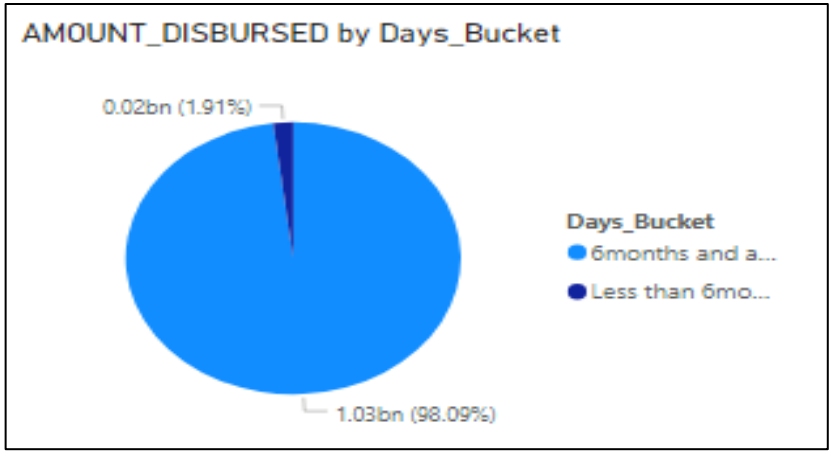
Reporting on the Data – Visualizations

9. Make sure the column chart is still on focus, then click on **Pie Chart** visual from the Visualizations pane to turn the column chart to a Pie chart.





From this



To this

Publish to Power BI Service

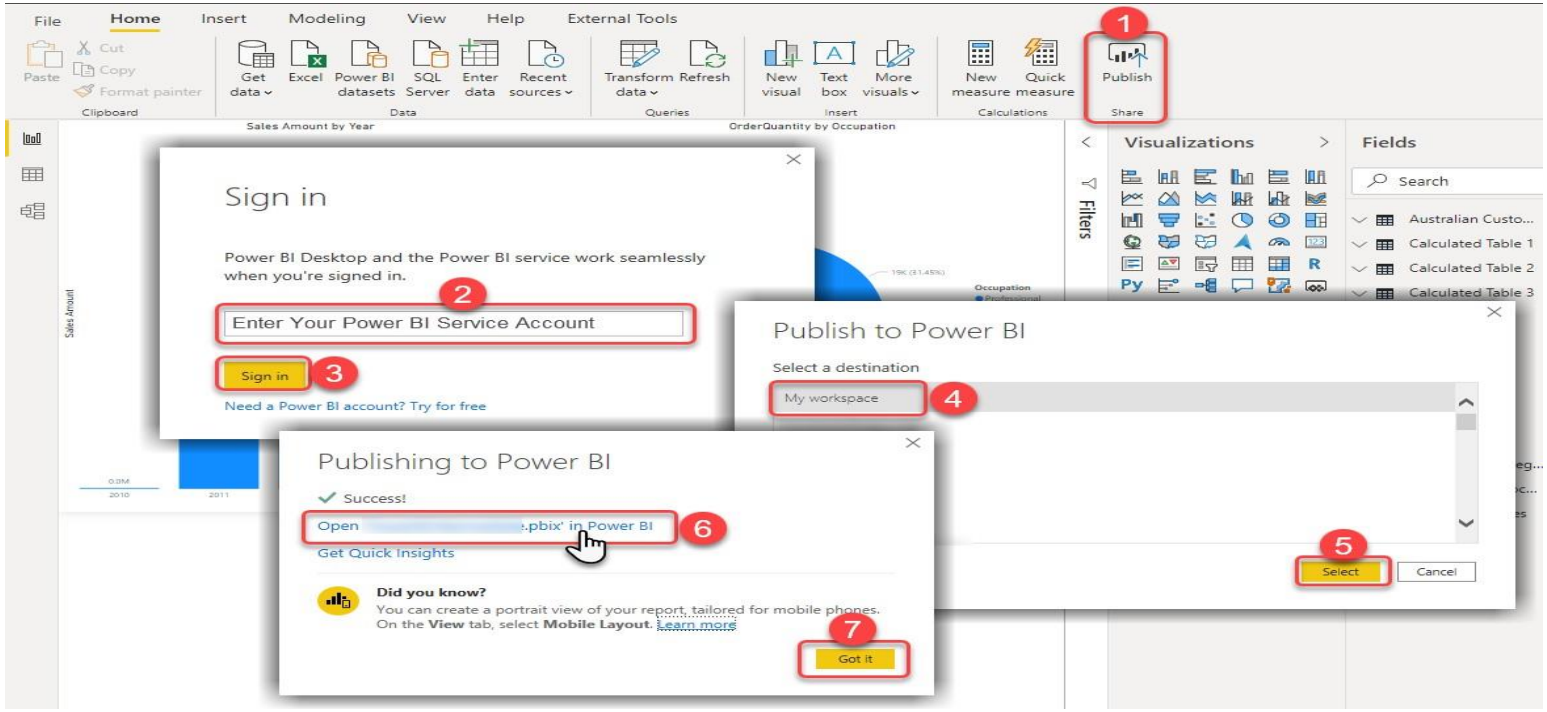
Publishing to Power BI Service

After you've finished your data visualizations in Power BI Desktop, it's time to Publish the report to Power BI Service.

If you have permission to publish reports to the service, it's easy. If you don't have access/permission, you won't be able to publish and share your reports in the cloud. You can still use your reports in Power BI Desktop or share it with others by sending them the .pbix file.
Follow the steps below to publish your report to Power BI Service

1. Click the **Publish** button from the ribbon bar.
2. Type in your Power BI Service credentials.
3. Click **Sign in.**
4. Select a workspace you want to publish your report to.
5. Click **Select.**
6. After your report is successfully published, you can click the report link on the Publishing to Power BI window.
7. Click **Got it.**

Publishing to Power BI Service



Questions

THANK YOU