DATA AND
ARTIFICIAL INTELLIGENCE

Profile the Data

Learning Objectives

By the end of this lesson, you will be able to:

- List the parameters of data anomalies
- Analyze data structures
- Modify column properties
- Examine data statistics

Data Anomalies

Anomaly detection

It helps to enhance line charts by automatically detecting anomalies in time series data.

Anomalies can be viewed in Power BI Desktop and Power BI service.

Format Anomalies

Five parameters that can be configured are:

1. Shape
2. Size
3. Colour
4. Parameter of the algorithm
5. Style and transparency of expected range

Limitations of Anomaly Detection

- It supports line chart visual containing time series data.
- It does not support Forecast, Min, Max, Average, Median, and Percentile lines.
- It does not support legends, multiple values, or secondary values in line chart visual.
- It requires a minimum of four data points.

Limitations of Anomaly Detection

- It does not support DirectQuery over SAP data source, Power BI Report Server, live connection to Azure Analysis Services, and SQL Server Analysis Services.
- Anomaly explanations do not work with **Show Value As** options.
- It does not support drilling down to go to the next level in the hierarchy.

Simplify Data Structure

Simplify Data Structure

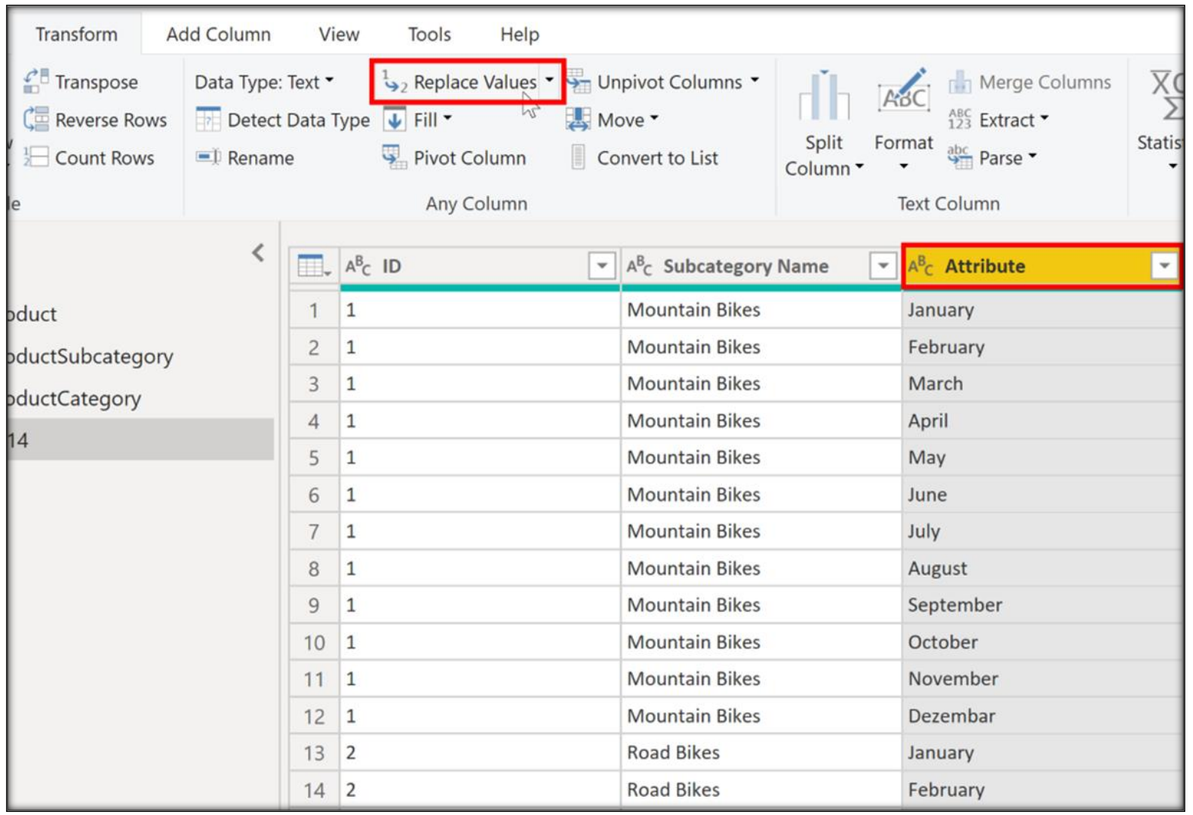
- Data imported from multiple sources into Power BI Desktop retains its predefined table and column names.
- Some column names can be changed to ensure sources are in a consistent format.
- Power Query Editor in Power BI Desktop can be used to change column names and simplify data structure.

Apply User-Friendly Value Replacements

Rename query	Changes uncommon or unhelpful query names for more understandability
Replace values	Changes any value with another value in a selected column
Replace null values	Replaces null values in data sources to make it consistent
Remove duplicates	Eliminates duplicates to keep unique values in a selected column

Apply User-Friendly Value Replacements

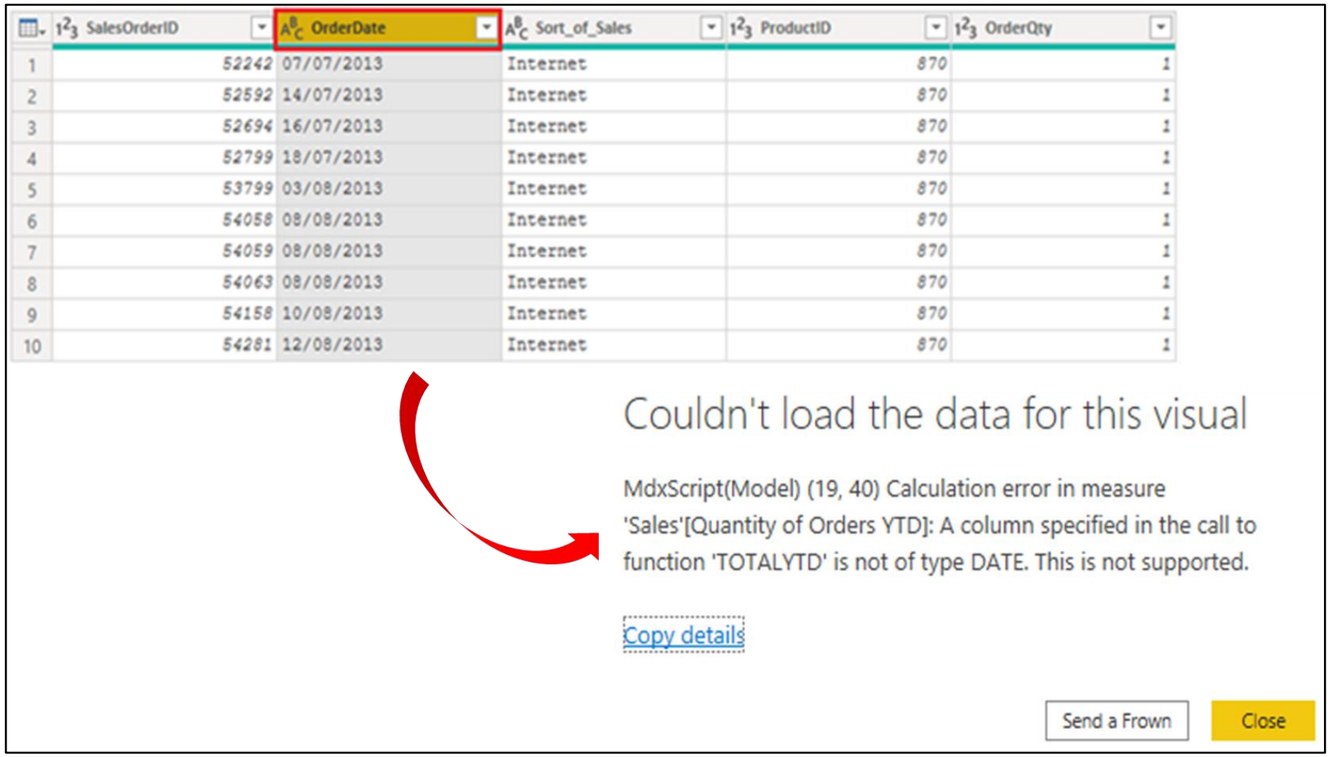
With the help of Power BI Desktop, you will be able to use rename query, replace values, replace null values, and remove duplicates.



Interrogate Column Properties

Interrogate Column Properties

Power BI Desktop automatically scans the first 1,000 rows and detects the type of data in the columns whenever a table is imported.



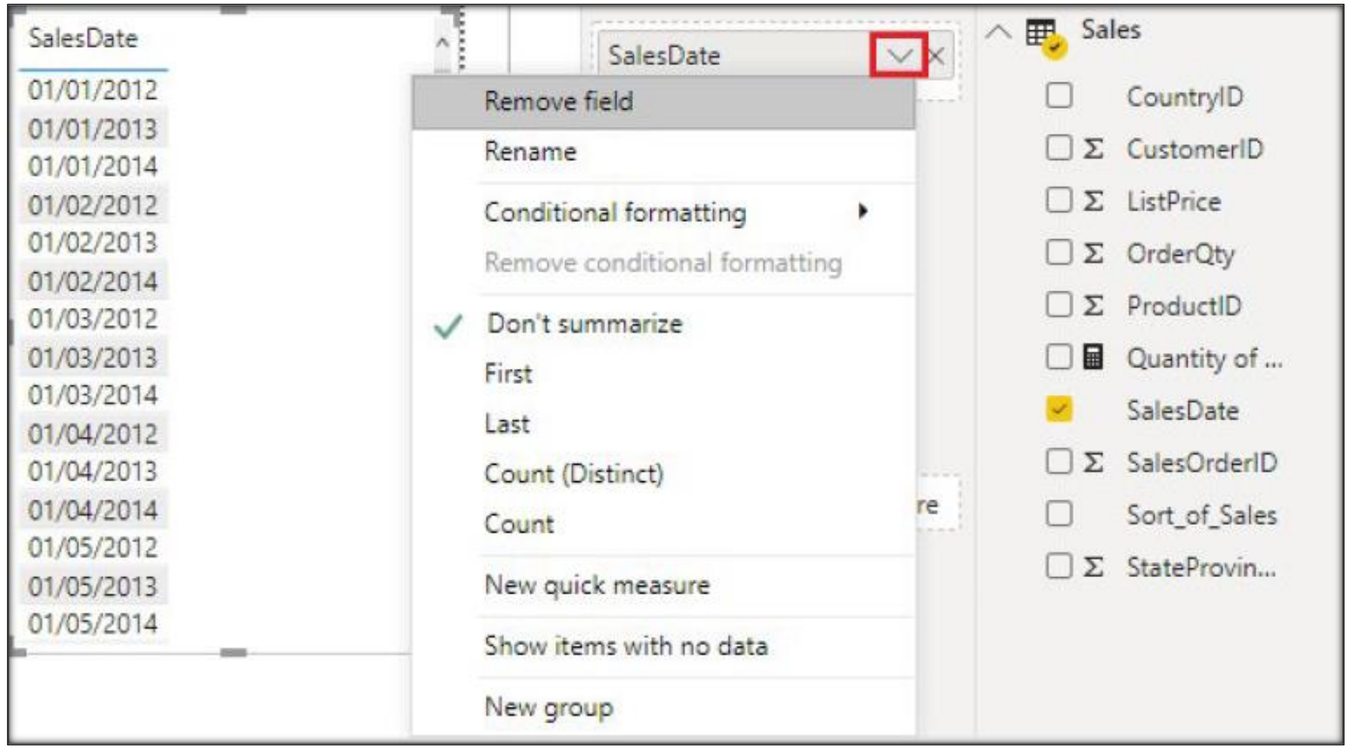
Evaluate Column Data Types

Before loading the data into a Power BI data model, evaluate the column data types in Power Query Editor.

	SalesOrderID	OrderDate	Sort_of_Sales	ProductID	OrderQty
1	52242	07/07/2013	Internet	870	1
2	52592	14/07/2013	Internet	870	1
3	52694	16/07/2013	Internet	870	1
4	52799	18/07/2013	Internet	870	1
5	53799	03/08/2013	Internet	870	1
6	54058	08/08/2013	Internet	870	1
7	54059	08/08/2013	Internet	870	1
8	54063	08/08/2013	Internet	870	1
9	54158	10/08/2013	Internet	870	1
10	54261	12/08/2013	Internet	870	1

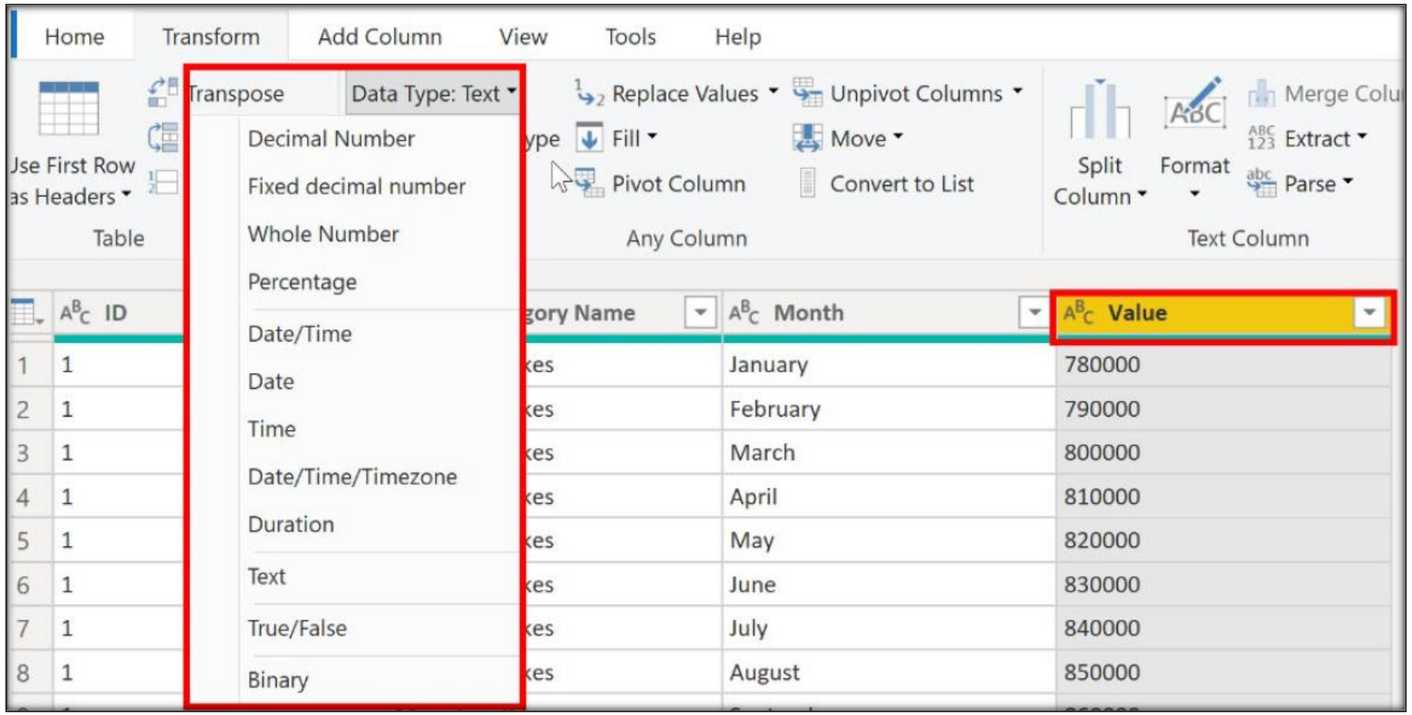
Implications of Incorrect Data Types

Incorrect data types help to prevent certain calculations, derived hierarchies, or proper relationships with the other tables.



Change the Column Data Type

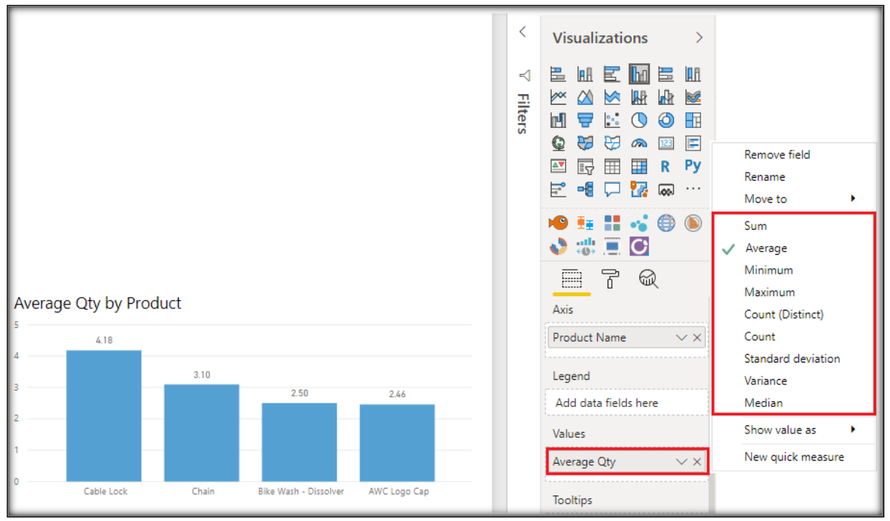
Column data type can be modified using Power Query Editor and Power BI Desktop Report view.



DATA AND
ARTIFICIAL INTELLIGENCE

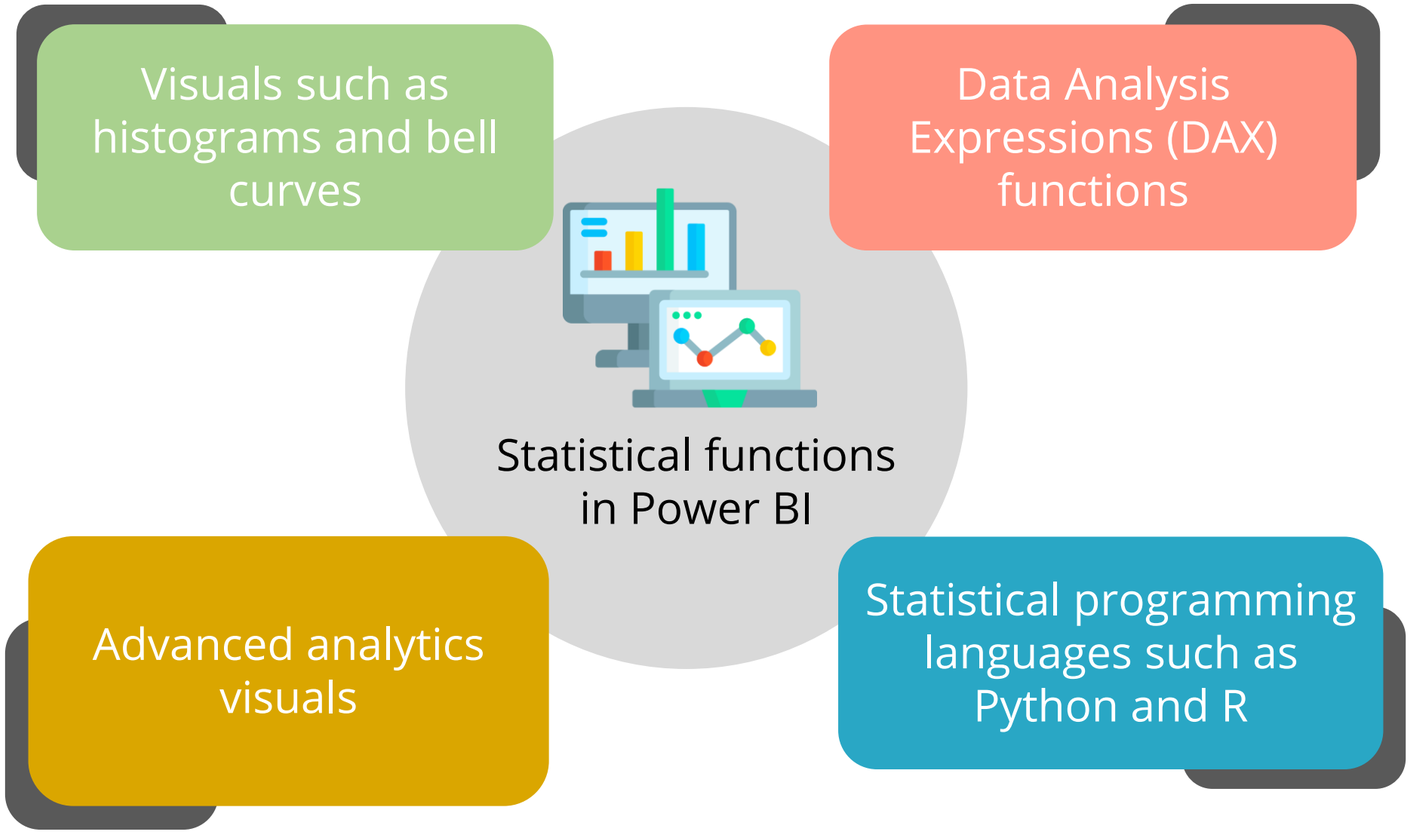
Interrogate Data Statistics

Interrogate Data Statistics



- Statistics plays a vital role in exploring data.
- It helps to visualize the distribution of data.
- It helps to identify key takeaways, trends, and outliers.

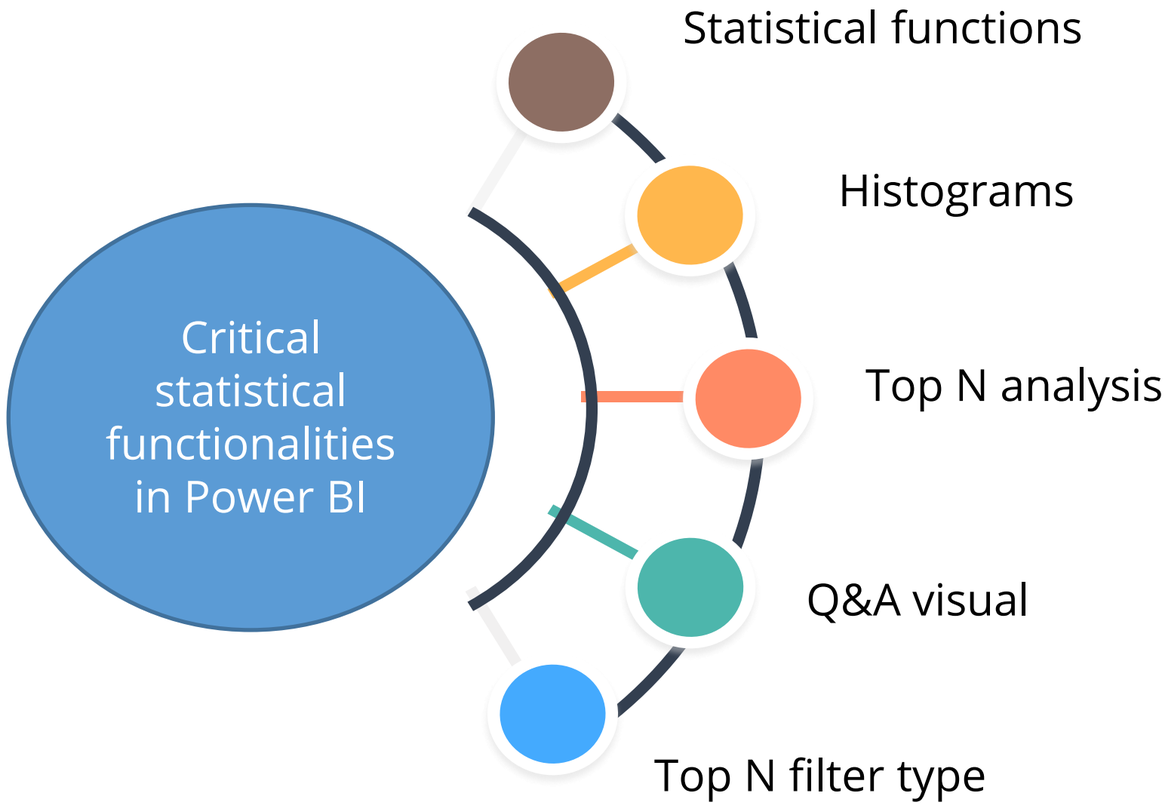
Explore Statistical Summary



Benefits of Exploring Statistical Summary

- It provides end-users a high-level view of datasets.
- It helps to visualize clusters, patterns on behavioral data, and data averages.
- It provides insights about data that will drive business decisions.

Types of Statistical Functionality



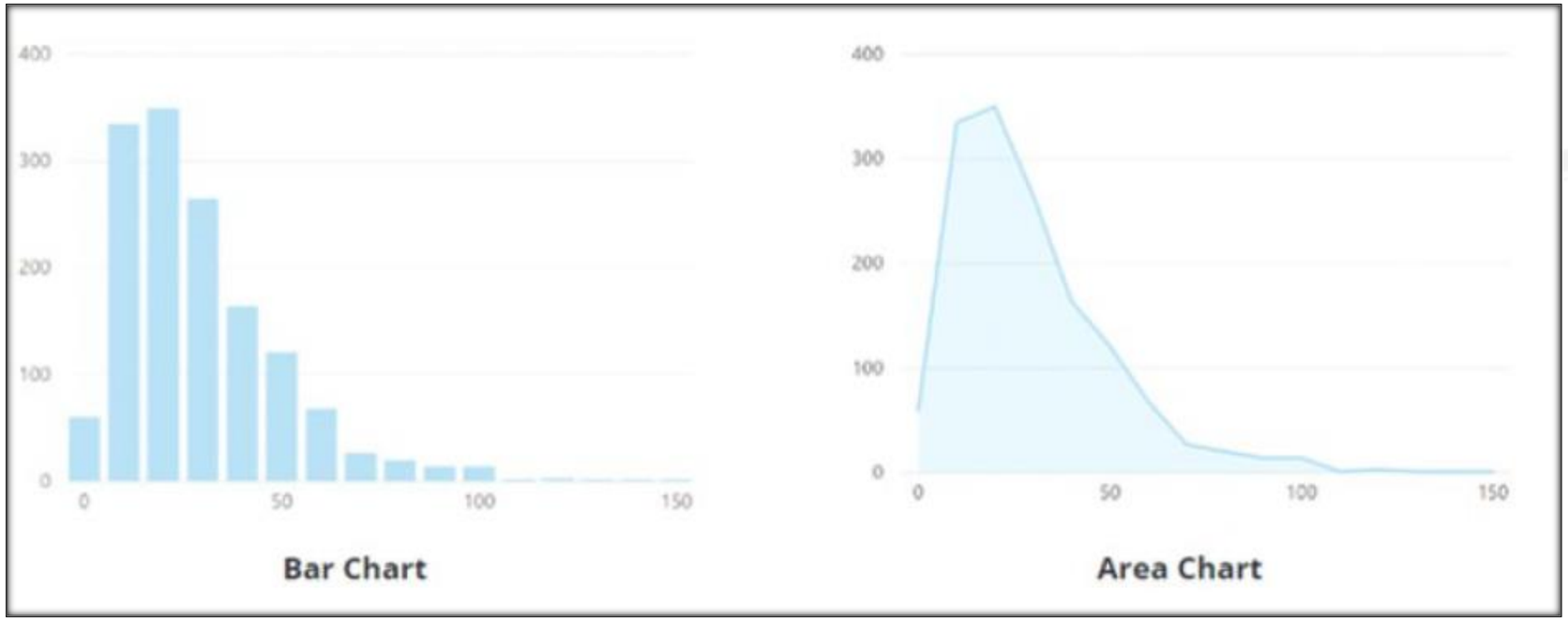
Statistical Functions

- DAX functions help to provide quick statistics based on data.
- Quick functions can be accessed by right-clicking on the value field in the visualization pane.

Histograms

It is a representation of data points into ranges.

Datasets can be represented using common charts such as histograms and bell curves.

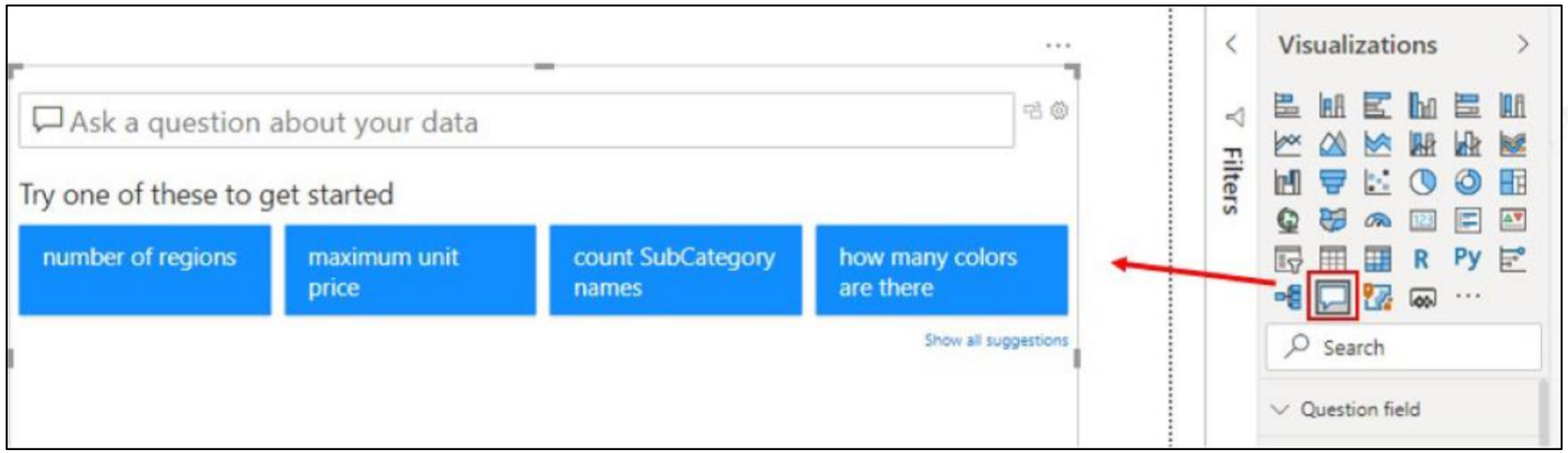


Top N Analysis

- TOPN DAX function returns the top N rows of a specified table.
- Top N analysis is a great way to represent list of ten important data.

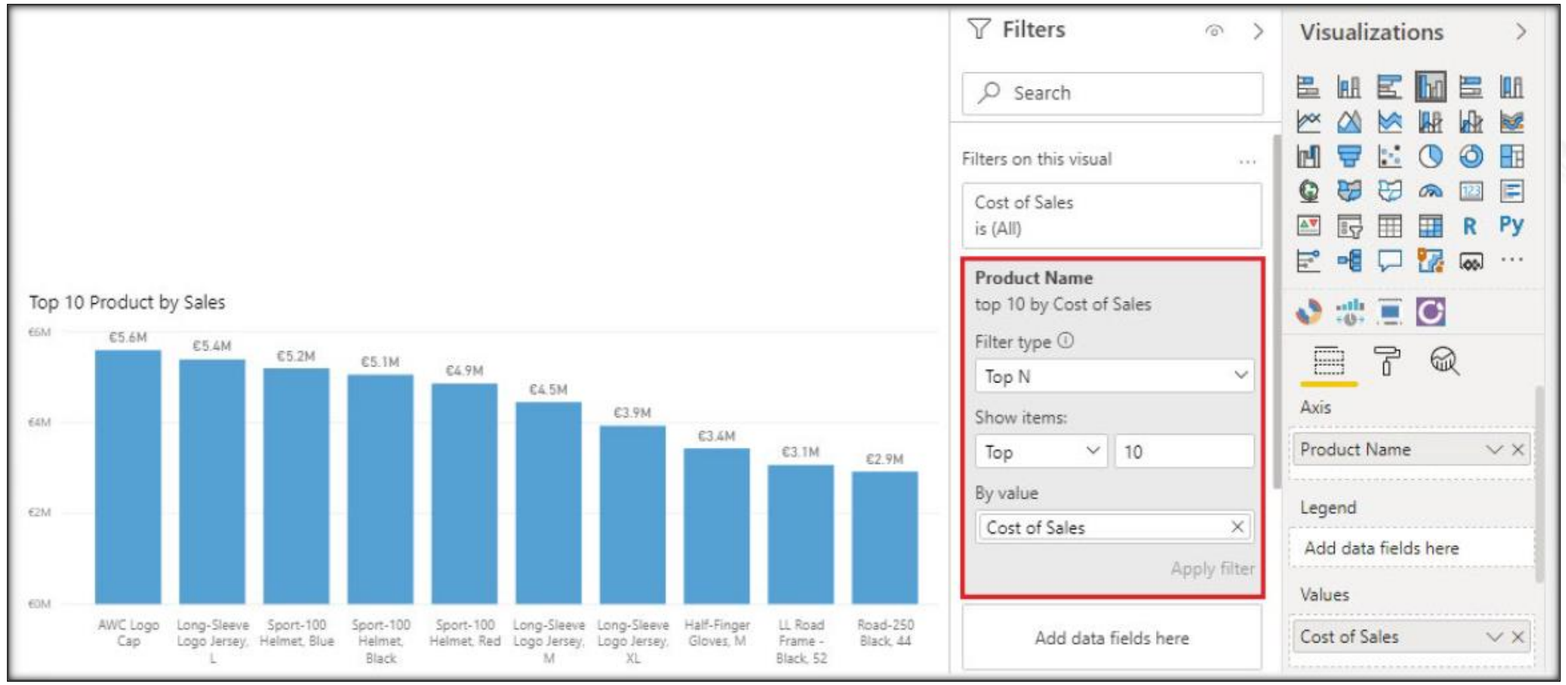
Q&A Visual

Power BI has a built-in Q&A visual that allows users to ask questions and get answers.



Top N Filter Type

Top N is a filtering option available on the filters pane to display the top or bottom N in a list.



Key Takeaways

- Anomalies help to enhance line charts by automatically detecting anomalies in time series data.
- Power Query Editor in Power BI Desktop is used to change column names and simplify data structure.
- Column data type can be modified using Power Query Editor and Power BI Desktop Report view.
- Critical statistical functionalities in Power BI includes statistical functions, histograms, Top N analysis, Q&A visual, and Top N filter type.

Knowledge Check

Knowledge Check

1

Which of the following feature in Power BI helps to ask questions and get answers?

A. Q&A Visual

B. DAX function

C. Top N analysis

D. Top N filter type

Knowledge Check

1

Which of the following function in Power BI helps to ask questions and get answers?

A. Q&A Visual

B. DAX function

C. Top N analysis

D. Top N filter type

The correct answer is **A**

Q&A Visual function in Power BI helps to ask questions and get answers.

Knowledge Check

2

Which of the following function in Power BI extracts list of ten important data?

A. Q&A Visual

B. DAX function

C. Top N analysis

D. Top N filter type

Knowledge Check

2

Which of the following function in Power BI extracts list of ten important data?

A. Q&A Visual

B. DAX function

C. Top N analysis

D. Top N filter type

The correct answer is **C**

Top N analysis function in Power BI extracts list of ten important data.

Knowledge Check

3

Which of the following helps to enhance line charts by automatically detecting anomalies in time series data?

A. Q&A Visual

B. Anomaly detection

C. Top N analysis

D. Top N filter type

Knowledge Check

3

Which of the following helps to enhance line charts by automatically detecting anomalies in time series data?

A. Q&A Visual

B. Anomaly detection

C. Top N analysis

D. Top N filter type

The correct answer is **B**

Anomaly detection helps to enhance line charts by automatically detecting anomalies in time series data.