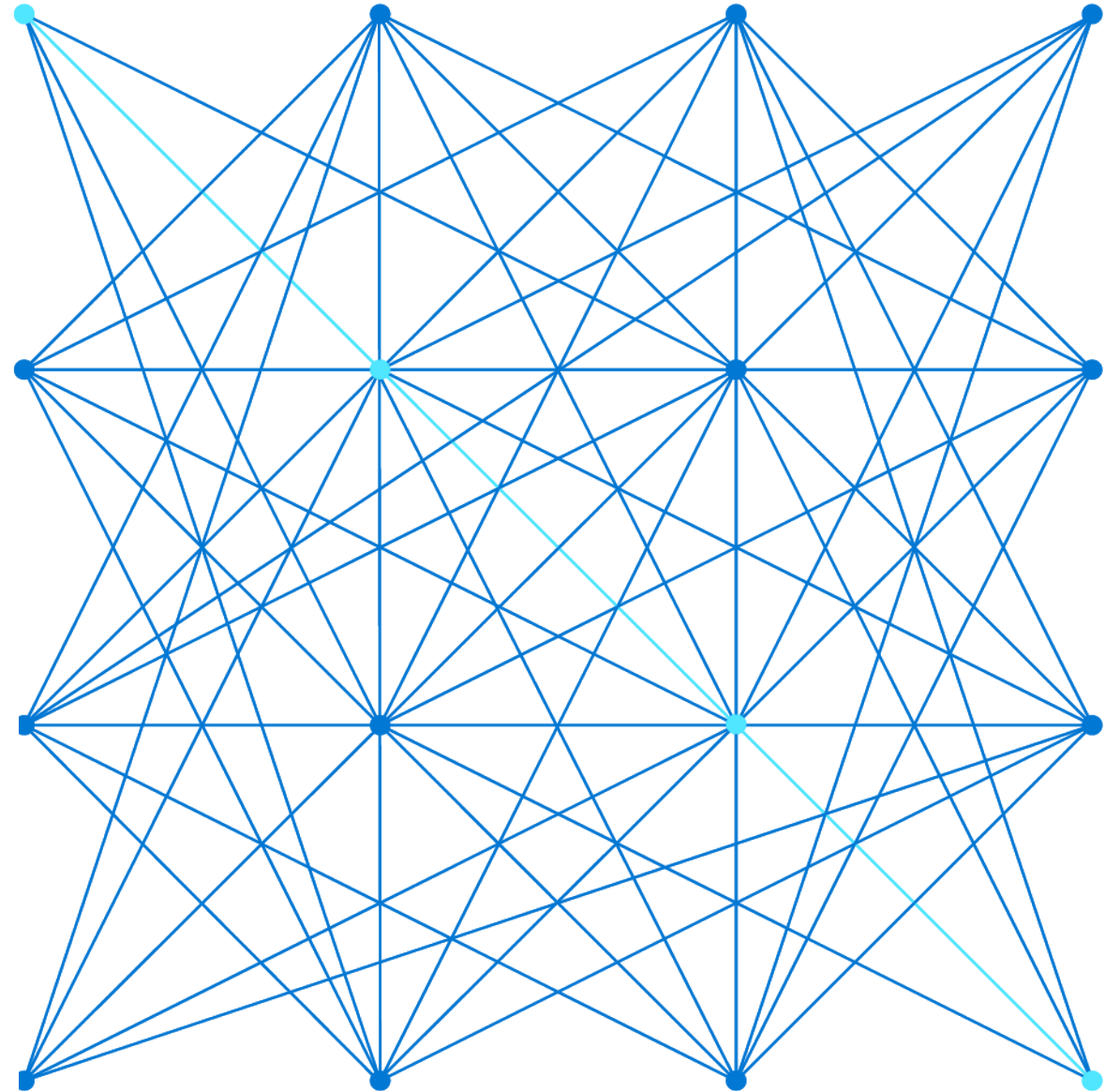# DA-100 Analyzing Data with Power BI 3

Prof. Ernesto Lee

BROWARD COLLEGE

Module 3: Clean, Transform, and Load Data In Power BI

Learning Objectives

You will learn the following concepts:

- Profiling the Data
- Shaping the Data
- Enhancing the structure of the data

Lesson 1:
Profiling the Data

Profiling Data and Examining Structures

Data profiling is understanding the state and structure of the data you are working with.

Review Questions

- Q01 – How many rows does Power Query scan to detect the type of data in the columns?
- A01 – 1000

- Q02 – Data profiling is defined as what?
- A02 – Studying the nuances of the data

- Q03 – What is the risk of having null values in a numeric column?
- A03 – DAX expressions that AVERAGE data will be incorrect.
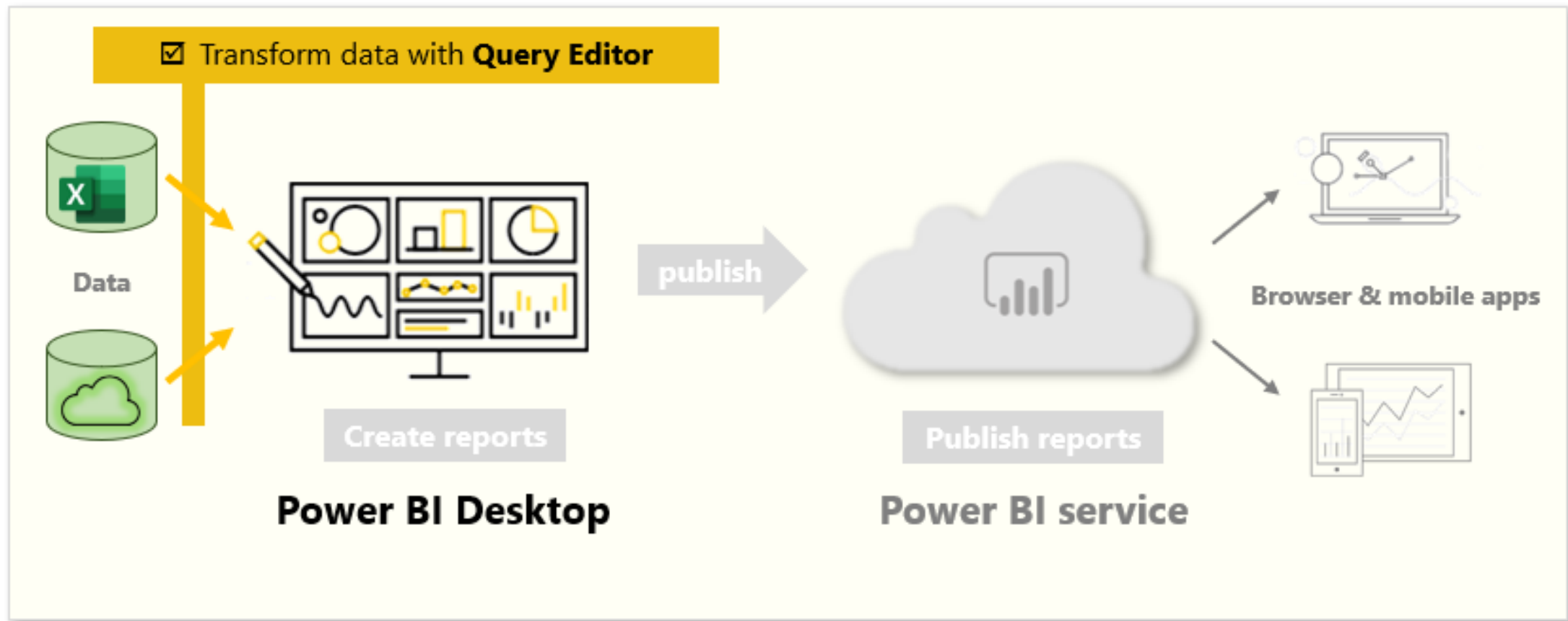
Lesson 2:
Shaping the Data

Introduction

- Benefits of clean data:
  - More accurate results
  - Better organized tables
  - Simpler data navigation
  - Human-readable values

Identify column headers and names

Use Power Query Editor to clean up and shape data.

Shaping Table Structure

Shape the data to meet reporting needs.

Unpivot and Pivot columns

Transfer data from rows to columns, and columns to rows.

Review Questions

- Q01 – The primary data preparation tool in Power BI is called what?
- A01 – Power Query Editor

- Q02 – The process of shaping data by converting your flat data into a table that contains an aggregation value for each unique value in a column is called what?
- A02 – Pivot (pivoting a column)

- Q03 – What can be achieved by removing unnecessary rows and columns?
- A03 – Deleting unnecessary rows and columns will reduce a dataset size and its good practice to load only necessary data into your data model.
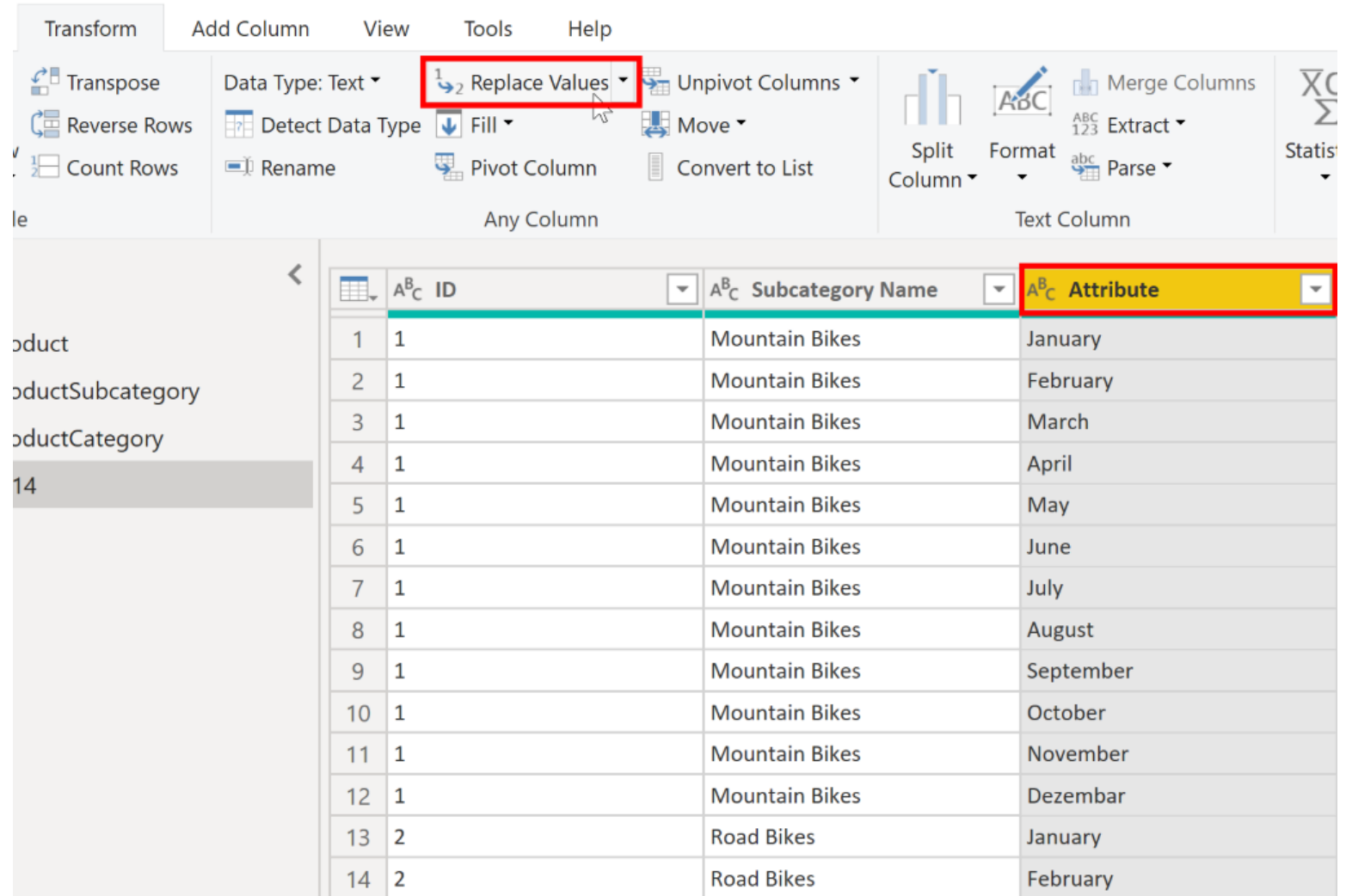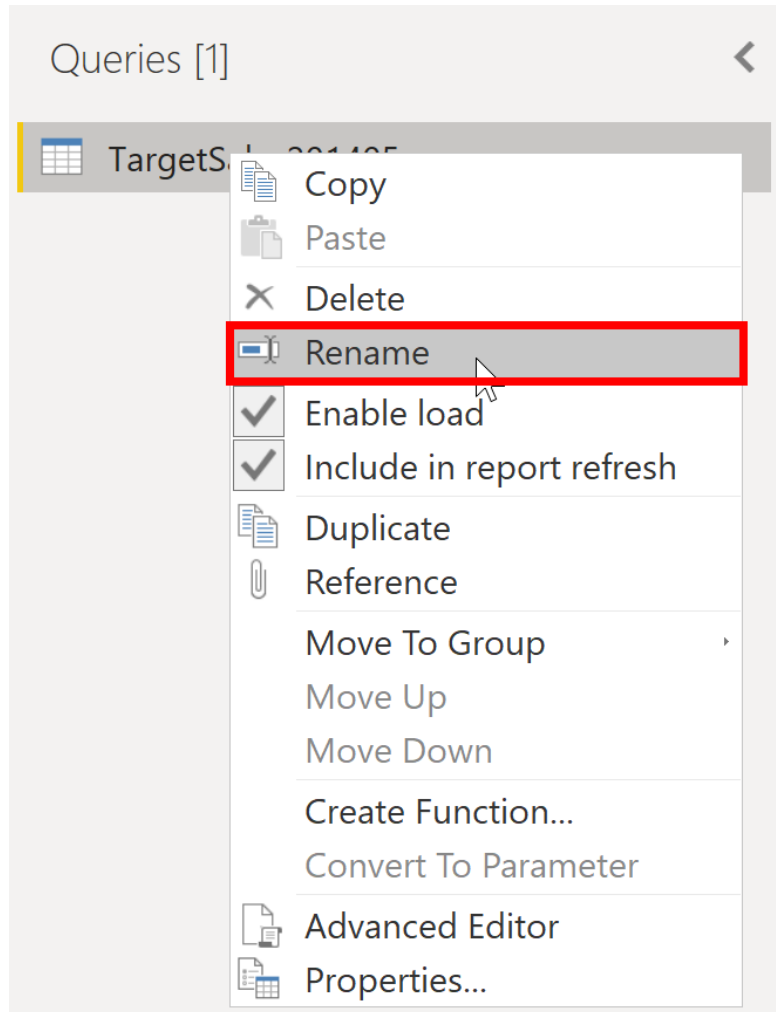
Lesson 3:
Enhance the Data Structure

Apply user-friendly value replacements

Make information user-friendly.

Evaluate and Change Column Data Types

Combine Multiple Tables into a Single Table

Append

Concatenate rows from three or more tables into a single table.

○ Two tables ● Three or more tables

Available tables

| Production Suppliers |
| Sales Customers |
| HR Employees |

Add >>

Tables to append

| Production Suppliers |
| Sales Customers |
| HR Employees |

OK Cancel

Merge

Select a table and matching columns to create a merged table.

Sales Orders

orderid	custid	empid	orderdate	requireddate	shippeddate	shipperid	freight	shipname
10248	85	5	7/4/2014	8/1/2014	7/16/2014	3	32.38	Ship to 85-B
10249	79	6	7/5/2014	8/16/2014	7/10/2014	1	11.61	Ship to 79-C
10250	34	4	7/8/2014	8/5/2014	7/12/2014	2	65.83	Destination SCQ
10251	84	3	7/8/2014	8/5/2014	7/15/2014	1	41.34	Ship to 84-A

Sales OrderDetails ▾

orderid	productid	unitprice	qty	discount
10248	11	14.00	12	0
10248	42	9.80	10	0
10248	72	34.80	5	0
10249	14	18.60	9	0
10249	51	42.40	40	0

Join Kind

Left Outer (all from first, matching from second) ▾

☐ Use fuzzy matching to perform the merge

▷ Fuzzy matching options

✓ The selection matches 830 of 830 rows from the first table.

OK Cancel

Use Advanced Editor to Modify M Code

See the code that Power Query Editor is creating with each step.

Sales Orders

Display Options ▼ ❓

```
let
    Source = Sql.Database("localhost", "tsqlv4"),
    Sales_Orders = Source{[Schema="Sales",Item="Orders"]}[Data],
    #"Split Column by Delimiter" = Table.SplitColumn(Sales_Orders, "shipaddress", Splitter.SplitTextByDelimiter(",", QuoteStyle.Csv), {"shipa
    #"Changed Type" = Table.TransformColumnTypes(#"Split Column by Delimiter",{{"shipaddress.1", type text}, {"shipaddress.2", type text}})
in
    #"Changed Type"
```

✔ No syntax errors have been detected.

Done   Cancel

Review Questions

- Q01 – What is not a best practice for naming conventions in Power BI?
- A01 – Abbreviated column names.

- Q02 – What functionality lets you see the code that is generated as part of each transformation step?
- A02 – Advanced Editor

- Q03 – If you have two queries that contain different data with the same structure, and you want to combine them into one query, which operation should you perform?
- A03 - Append

Lab: Shape and Combine Data

Module Overview

We covered the following concepts:

- Profiling the Data
- Shaping the Data
- Enhancing the structure of the data

References

- DA-100 Clean, transform, and load data in Power BI
https://docs.microsoft.com/en-us/learn/modules/clean-data-power-bi/