Data Pre-Processing in Power BI

Cleaning and Transformation Techniques

Introduction to Data Pre-Processing

Definition

Importance in the context of Power BI

DATA QUALITY

EFFICIENCY

USABILITY

Importance of Data Pre-Processing

Steps in Data Pre-Processing

- Data Cleaning:
- Handling Missing Values
- Managing Outliers
- Removing Duplicates

Data Transformation

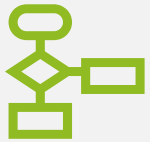
- Normalization
- Standardization
- Encoding
- Feature Engineering

Data Reduction

Why and when?

Power Query Editor

Overview and importance



Interface and features

Query settings and applied steps

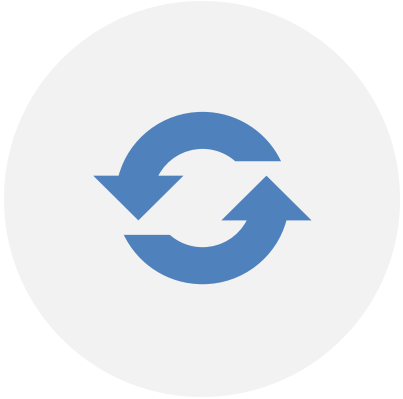
Common Transformations in Power Query Editor

COLUMN OPERATIONS: ADD, REMOVE, SPLIT, RENAME

ROW OPERATIONS: FILTER, SORT, REMOVE DUPLICATES

TEXT OPERATIONS: CLEAN, EXTRACT, REPLACE

Advanced Transformations in Power Query Editor



PIVOTING AND UNPIVOTING

MERGING QUERIES

GROUPING DATA

Error Handling in Power Query Editor

Identifying	Removing	Replacing
Identifying errors	Removing error rows	Replacing errors

Data Transformation

Why transform data?

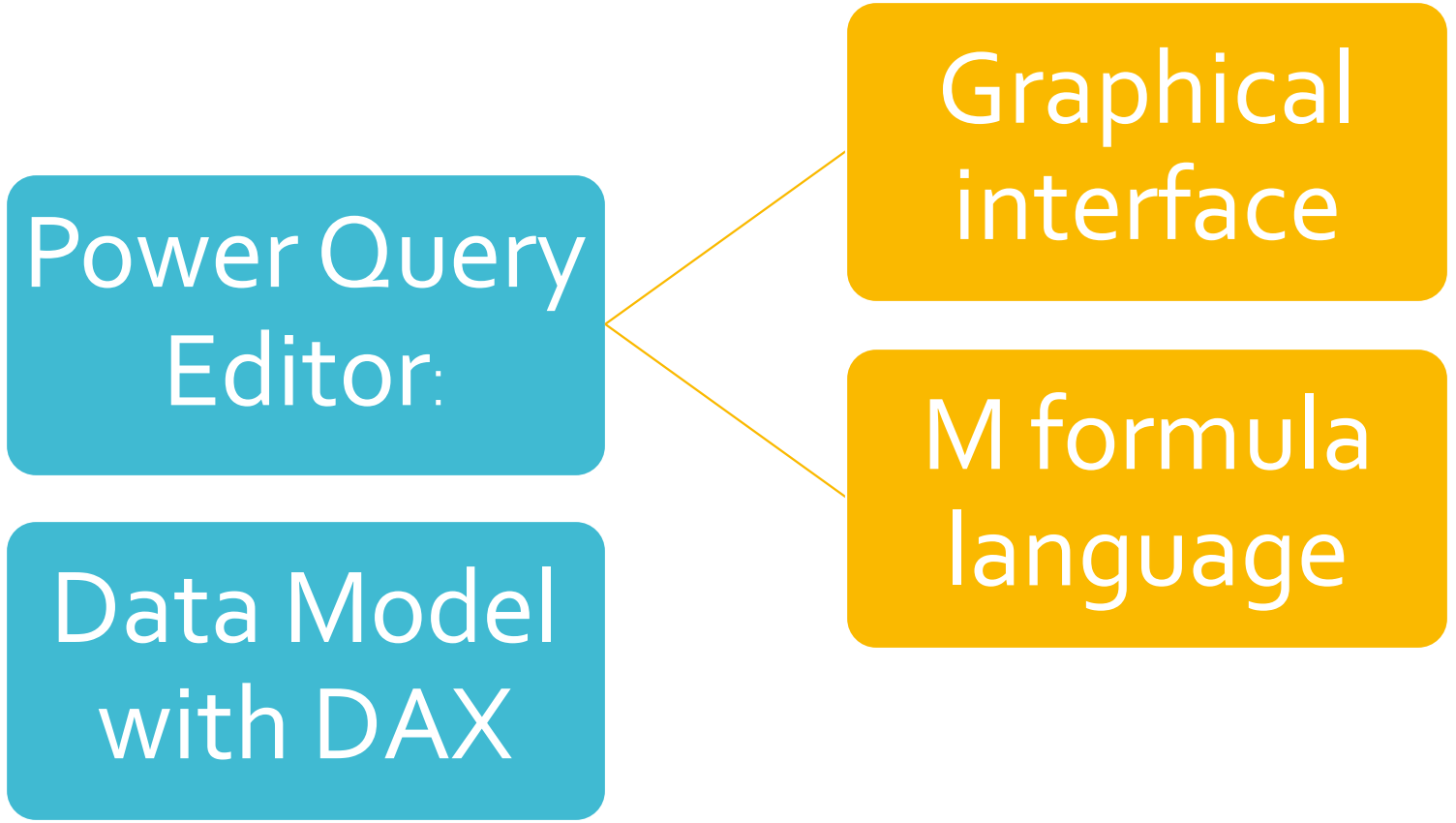
- Types of transformations:
  - Binning
  - Aggregation
  - Pivoting
  - Encoding
  - Scaling and normalization

Data Normalization

Purpose of normalization ?

- Methods:
  - Min-Max scaling
  - Z-score normalization
  - Decimal scaling

Tools in Power BI



Lab Introduction

OVERVIEW OF THE SAMPLE SALES DATASET

HIGHLIGHTING THE CHALLENGES (MISSING VALUES, DUPLICATES, ETC.)

Conclusion

Summary of key points

Importance of consistent data pre-processing for accurate insights

Q&A

Open floor for questions and discussions