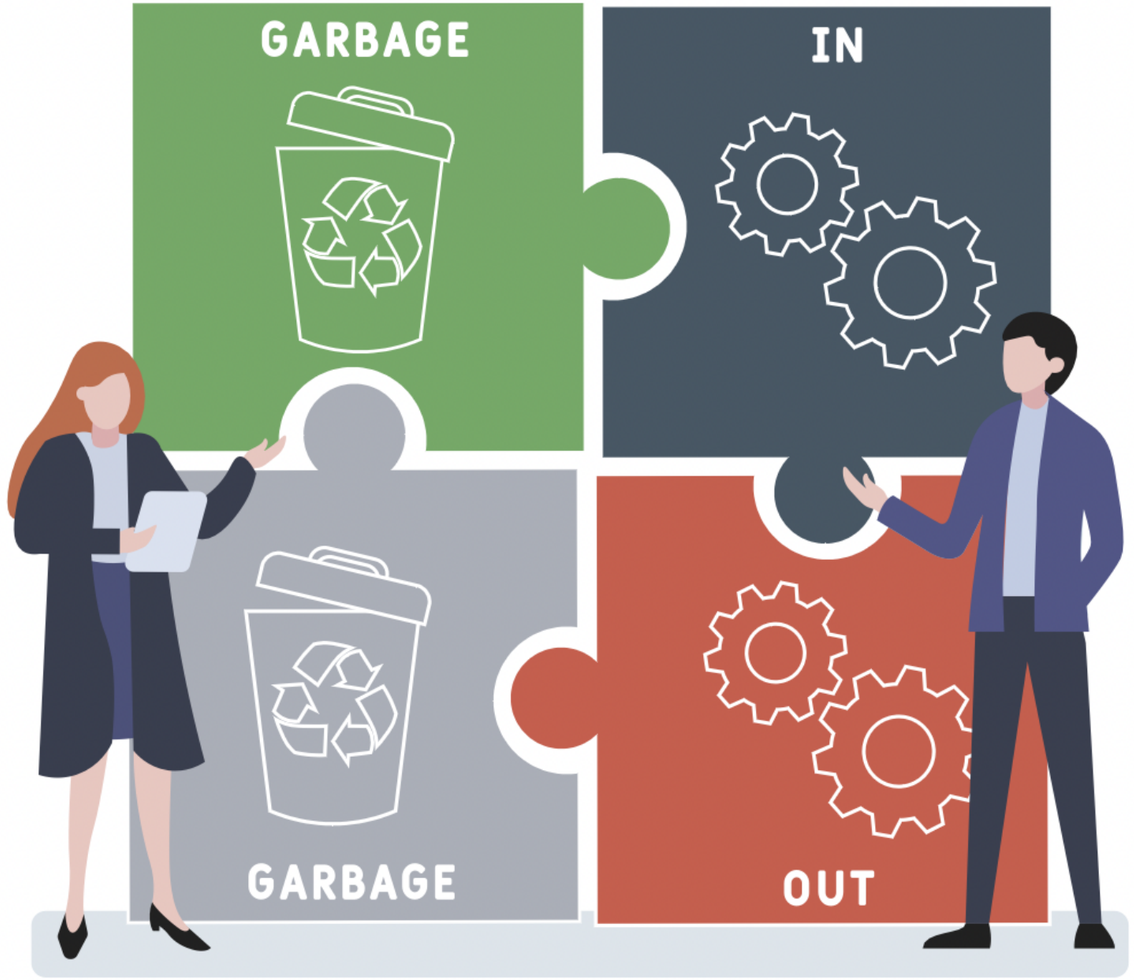
# Introduction to Power Query

DATA PREPARATION IN POWER BI

**Khaled Choucri**
Strategic Analyst - Trilogy

datacamp

# Why prepare data?



# What is clean data?

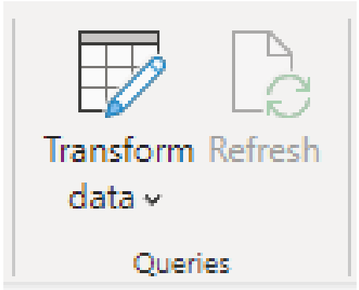
- Free from missing values
- Typos/data entry errors are corrected
- No duplicated data
- Irrelevant data is excluded
- Outliers are dealt with
- Correct data types for all columns
- Column & table names are short and descriptive

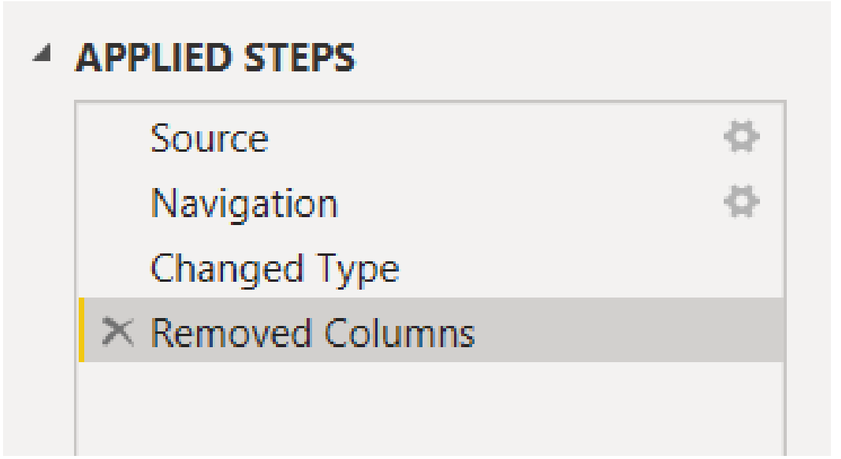
| Student ID | Student Name | Age | GPA | Classification |
|---|---|---|---|---|
| 100122014 | Joseph | 21 | 3.5 | Junior |
| 100232015 | Patrick | 200 | 3.2 | Sophomore |
| 100122012 | Seller | 24 | 3.0 | Senior |
| 100342013 | Roger | 23 | 234 | Senior |
| 100942012 | Davis | 2.8 | 3.7 | Sophomore |
| | Travis | 23 | 3.4 | Sr |
| 100982015 | Alex | 27 | | Sophomore |
| 100982013 | Trevor | -22 | 4.0 | Senior |
| AUC2016XC | Aman | 30 | 3.5 | Jr |

Missing Data  Inconsistent Data  Noisy Data

# How does data preparation work in Power Query?

- To load & transform data, we use the *Power Query* interface in Power BI
- Transformation steps are added sequentially
- Steps are saved in a "recipe" called "Applied Steps" that executes every time your data is loaded/refreshed
- To apply the transformations, we choose "Close & Apply" in the *Power Query* editor



# Data types in Power Query

- Numerical Data (Whole Number / Decimal / Percentage)
  - `123`, `45.23`, `23%`
- Date / Time Data
  - `01/01/2022`, `09:30 AM`
- Text Data
  - `ABC`
- Logical Data
  - `TRUE`
- Binary Data
  - `RGF0YWNhbXA=`

Power Query will auto detect data types by default

# Basic structural transformation of data

- Promote headers
- Reorder columns
- Sort columns
- Remove columns/rows
- Rename columns

# Let's practice!

DATA PREPARATION IN POWER BI

# Introduction to Power Query demo

DATA PREPARATION IN POWER BI

**Khaled Choucri**
Strategic Analyst - Trilogy

datacamp

# Let's practice!

DATA PREPARATION IN POWER BI