

Specialization: **Data Mining** Course: **Data Mining Pipeline** 

with Dr. Qin Lv

## Module 3, Lesson 1: Data Cleaning, Data Integration

## Attributions

- Slide 8: Data Cleaning
  - Data cleaning illustration: https://thenounproject.com/search/? q=data+cleaning&i=2942558
- Slide 10: Noisy Data
  - Linear regression example: https://commons.wikimedia.org/wiki/ File:LineareFunktionRegression01B.png
  - Clustering example: https://commons.wikimedia.org/wiki/File:EM-Gaussian-data.svg
- Slide 15: Correlation Analysis (3)
  - https://en.wikipedia.org/wiki/Correlation\_does\_not\_imply\_causation