Friedrich-Alexander-Universität
Technische Fakultät

Department of Computer Science
Computer Science 6
(Data Management)

Knowledge Discovery in Databases with Exercises
Summer Semester 2025

Exercise Sheet 3: Frequent Patterns

About this Exercise Sheet

This exercise sheet focuses on the content of lecture *6. Mining Frequent Patterns, Associations and Correlations*.

It includes both a practical data science exercise (Exercise 1) and theoretical exercises on Apriori (Exercise 2) and FP-growth (Exercise 3).

The exercise sheet is designed for a two-week period, during which the tasks can be completed flexibly (Exercise 1 is planned for the first exercise session, and Exercises 2 and 3 for the second session).

The sample solution will be published after the two weeks have elapsed.

Preparation

Before participating in the exercise, you must prepare the following:

1. **Install Python and pip on your computer**
   - Detailed instructions can be found in `1-Introduction-Python-Pandas.pdf`.
2. **Download provided additional files**
   - Download `Additional-Files-Student.zip` from StudOn
   - Extract it to a folder of your choice.
3. **Install required Python packages**
   - Open a terminal and navigate to the folder where you extracted the files.
   - Run the command `pip install -r requirements.txt` within the extracted additional files folder to install the required Python packages.

Exercise 1: Mining Frequent Patterns

This exercise comprises practical data science tasks and thus utilizes a Jupyter Notebook:

1. Open `Mining-Frequent-Patterns.ipynb`.
2. Take a look at the tasks (blue boxes) in the notebook and try to solve them.

If you are unfamiliar with how to open a Jupyter Notebook, please refer to Exercise 1 of `1-Introduction-Python-Pandas.pdf`.

Exercise 2: Apriori

Given is a **transactional dataset**:

ID	Transaction
1	Apple, Banana, Cherry
2	Banana, Cherry
3	Cherry, Apple
4	Dragonfruit, Apple, Banana
5	Apple, Dragonfruit

Use **Apriori** to find all frequent itemsets for a **minimum support count** of **2**.

Write down **all** intermediate steps.

Exercise 3: FP-growth

Given is a **transactional dataset**:

ID	Transaction
1	Apple, Banana
2	Banana, Cherry
3	Cherry, Apple
4	Apple, Banana
5	Apple, Dragonfruit

Use **FP-growth** to find all frequent itemsets for a **minimum support count** of **2**.

Write down **all** intermediate steps. This **includes** the header table for each FP-tree.