1. **Name of Course:** Data Management

2. **Instructor: Petra Kralj Novak**

3. **No. of Credits (US/ECTS): 4 ECTS**

4. **Academic term:** *Winter*

5. **Course level:** *BA/BSc*

6. **Relationship with other courses**: A prerequisite for this course are basic programming skills (e.g. Introduction to Programming in Python) and algorithms (e.g. Algorithms and Data Structures).

7. **Course type:** *Mandatory*

8. **Name of the module in which the course is offered:** Data Science and Society

9. **Short description and the overall aim of the course:**

   This course aims to introduce students to the key principles of data management, with an emphasis on database management, as databases are the foundation of the modern information society. They allow multiple users to access large amounts of integrated data simultaneously. This course introduces the design, use, and application of database management systems. We will focus on the SQL query language, but also become familiar with NoSQL databases. We will conclude the course by discussing data managements plans.

10. **The learning outcomes of the course:**

    Upon completion of this course, students will understand the key role of a database in supporting concurrent data access by multiple users to large volumes of integrated data. Students will be able to connect to databases and perform queries to quickly and efficiently access the data they need. Students will know the basics of SQL and NoSQL databases and have hands-on experience with contemporary database software. They will be able to design and develop a simple relational database as well as connect to existing databases and perform complex queries. They will be able to use both interfaces and programming access to query the databases.

11. **Learning activities and teaching methods:** Lectures, hands-on exercises, project work.

12. **Assessment:**
    - Weekly homework: 10%
    - Assignment 1/2: 20%
    - Assignment 2/2: 20%
    - Written exam: 50%

13. **Course contents:**

    Week 1: Introduction
    - Why databases
    - Multi-user, client-server, safety, redundancy, read/write access, transactions
    - Client-server architecture

- Postgres SQL
- Intro to SQL select statement

Week 2: SQL Select statement
- Plain
- Aggregation
- Joins

Week 3: SQL join
- Entity-relation diagram
- Different types of JOINS

Week 4: Programming access
- Connecting to a database programmatically
- Hands-on

Week 5: Database design and normalization
- 1^{st} – 4^{th} normal forms
- Create clause
- Insert clause
- Primary key

Week 6: Indexing
- The role of indexes, foreign key
- Update and delete clauses

Weeks 7 & 8: Introduction to NoSQL databases
- Graph database: Neo4J

Weeks 9 & 10: Text database:
- Elasticsearch with Kibana

Week 11: Data management plan

Week 12: Final exam

14. Further items:

The course is mostly based on selected chapters of the book: Lemahieu, W., vanden Broucke, S., & Baesens, B. (2018). Principles of database management: the practical guide to storing, managing and Analyzing big and small Data. Cambridge University Press. https://www.pdbmbook.com/

Material available on Moodle.

Technical requirement: A computer with Python