

Data Science II (11.08200) (2021)

Adopted 2021

Demonstrate how Data Science can be used to impact school, work, and leisure time DS2-1

- 1. Demonstrates how students currently use data science in their lives, and how data science impacts their lives and possible career choices** DS2-1.1
- 2. Identify and differentiate between different data governance standards and argue why governance is important.** DS2-1.2
- 3. Identify ethical issues in data science.** DS2-1.3
- 4. Identify and compare potential bias issues in data science** DS2-1.4

Formulate Questions to Clarify the problem at hand and formulate 1 or more questions that can be answered with data DS2-2

- 1. Identify the objectives of data and information management** DS2-2.1
- 2. Determine whether a problem involves categorical or quantitative data** DS2-2.2
- 3. Frame a statistical question of interest in terms of measurable data** DS2-2.3

Design and implement a plan to collect appropriate data to answer the research question DS2-3

- 1. Describe the factors that must be considered in distributing data effectively and how a simple model can be used to obtain at least a first-cut distribution** DS2-3.1

Analyze data by selecting appropriate graphical and numerical methods DS2-4

- 1. Implement advanced spreadsheet functions, automation, and dynamic reporting.** DS2-4.1
- 2. Utilize various tools (such as the ARIMA model) to analyze time series data.** DS2-4.2
- 3. Demonstrate the ability to take data and create a dashboard that provides insight to solve real world problems.** DS2-4.3
- 4. Use graphical and numerical displays to foster further investigation into question of interest** DS2-4.4

Identify the general concepts of databases/data tools and how to utilize design thinking to produce solutions that are clean and thoughtful. [DS2-5](#)

1. Identify and distinguish between variations of techniques (Artificial Intelligence, Machine Learning, Deep Learning, etc.) [DS2-5.1](#)
2. Provide definitions of key terms and concepts that describe the database environment [DS2-5.2](#)
3. Describe and build the major components of the database environment and explain how these components interact with each other [DS2-5.3](#)
4. Provide a review of systems development methodologies, particularly the waterfall method and agile programming development and show how database development fits with these methodologies. [DS2-5.4](#)
5. Generate Entity Relationship logical models to represent organization data and plan for database development and infrastructure [DS2-5.5](#)
6. Assess end user data and information requirements and develop a logical model to fit those requirements [DS2-5.6](#)
7. Describe the concept of supertype/subtype relationships and recognize when to use these relationships in data modeling. [DS2-5.7](#)
8. Describe the use of specialization (top-down perspective) and generalization (bottom-upper perspective) as complementary techniques for defining supertype/subtype relationships and understand relationship constraints when modelling. [DS2-5.8](#)
9. Describe the position of logical database design within the overall database development process [DS2-5.9](#)
10. Describe the relational model including the properties of relations, integrity constraints, and well-structured relations. [DS2-5.10](#)
11. Describe the principles and detailed steps involved in mapping Enhanced Entity Relationship diagrams to relations. [DS2-5.11](#)
12. Understand data normalization, functional dependency, and develop a fully normalized Entity Relationship Diagram. Evaluate the normality of a logical data model and correct any anomalies [DS2-5.12](#)

Build a database based on designed model, identify implementation policies and procedures, and establish plans for testing/debugging a data science solution. [DS2-6](#)

1. Describe a Database Management System Language (DMBS) like SQL and summarize its basic operators. [DS2-6.1](#)
2. Illustrate data definition language (DDL) commands for creating tables and views as well as for modifying and dropping tables. [DS2-6.2](#)
3. Formulate single table DMBS (SQL) queries. [DS2-6.3](#)
4. Formulate DMBS (SQL) queries that use functions. [DS2-6.4](#)

-
5. Show how to establish referential integrity using DMBS (SQL). [DS2-6.5](#)
 6. Use of the "group by" and "order by" clauses in DMBS (SQL) queries. [DS2-6.6](#)
 7. Demonstrate (DBMS) SQL capabilities such as multiple-table data retrieval (join and other operators such as difference, union, and intersection), explicit and implicit joining, and built-in functions. [DS2-6.7](#)
 8. Illustrate the differences between the joining and subquery approaches to manipulating multiple tables in DMBS (SQL) [DS2-6.8](#)
 9. Describe triggers and stored procedures and provide examples of how these might be used [DS2-6.9](#)
 10. Generate and implement a testing plan for a data management solution implementation [DS2-6.10](#)
-

Deploy a data science solution in a production environment, follow implementation procedures, and develop a plan for long term maintenance. [DS2-7](#)

1. Describe the differences between the processes of deployment and implementation of solutions. [DS2-7.1](#)
 2. Understand the components and key steps to a successful deployment. [DS2-7.2](#)
 3. Build and deploy a data management system solution implementation. [DS2-7.3](#)
 4. Generate and implement a maintenance plan for a data management solution implementation [DS2-7.4](#)
-

Analyze results by interpreting the information provided by the data and how its interpretation supports possible answers to question or problem being investigated. [DS2-8](#)

1. Utilize visual reporting and statistic tools to perform, understand, and interpret statistic such as regression analysis, ANOVA, hypothesis testing, and sampling distributions [DS2-8.1](#)
2. Identify and express areas for further study or investigation based on results [DS2-8.2](#)
3. Create a dashboard with appropriate high-level charts, such as heat plots, box and whiskers, etc. to express the data that is being analyzed [DS2-8.3](#)