

# DASC 2113

## Principles and Techniques in Data Science

Fall 2023  
Lecture: Tues/Thur 11 am – 12:15 pm. Grad Education 163

Lab: Tues/Thur 2pm – 3pm  
BELL 1108G

Instructor: Jackson Cothren  
Email: [jcothre@uark.edu](mailto:jcothre@uark.edu)  
Phone: 479-575-5421  
Office: JBHT 304  
Office Hours: Mon/Wed 10-11:30am

Teaching Assistant: Simon Chaisouang  
Email: [shchaiso@uark.edu](mailto:shchaiso@uark.edu)  
Office: JBHT 331  
Office Hours: Mon/Wed 2:30pm – 5:00pm

### Description

Principles and Techniques in Data Science is an intermediate, semester-long data science course that follows an overview of data science in today's world. This class bridges between introduction to data science and upper division data science courses as well as methods courses in other concentrations. This class equips students with essential basic elements of data science, ranging from database systems, data acquisition, storage and query, data cleansing, data wrangling, basic data summarization and visualization, and data estimation and modeling. Students will gain hands-on experience using Python and various packages in Python.

### Motivation

This course

- provides basic foundations of data, computation, and inferential thinking for functioning as a data scientist.
- provides computational and empirical practice in problem modeling and formulation to extract essential knowledge from real world data.
- prepares students for starting careers as data scientists by providing experience with tools and techniques that professional data scientists are using.
- emphasizes effective and efficient communication of data science techniques and findings to multi-disciplinary audiences and stakeholders.

### Requirements and Evaluation

Evaluations will be based on the following:

1. Data Camp (DC) exercise completion - 25%.
2. Mid-term skill/knowledge exam - 25%
3. Case Studies (CS) weekly take-home and small group (SG) activities - 25%
4. Final skill/knowledge exam - 25%

Datasets and notebooks used in weekly graded work and/or exams will be available through the class GitHub repository.

Data Camp (DC) exercises are due on the date given in Data Camp. Late submissions will lose 2% of the available points per day.

Case Studies (CS) may be individual or team assignments and are due on the Tuesday following the week they are assigned unless otherwise noted in the syllabus below. Late submissions will lose 2% of the available points per day.

### Materials

- **Python 3.8.x** with various packages including Jupyter notebook
- **VSCode** editor <https://code.visualstudio.com>
- **Data Camp** <https://learn.datacamp.com> (all interactive video/transcript lectures collected in DASC 2113 organization) | required
- **Python Data Science Handbook: Essential Tools for Working with Data**, Jake VanderPlas, 2016, <https://jakevdp.github.io/PythonDataScienceHandbook/> | optional (but highly recommended as reference)
- **Course GitHub organization** (source of Jupyter-based lectures and assignments)  
<https://github.com/DASC2113-Fall-2023>  
| required access via personal GitHub account
- **Other readings as assigned.**

## Learning Objectives

Students completing DASC2113 should...

- have the necessary foundation and context to prepare for more advanced data science topics,
- have a greater understanding of relational database management systems and their use in data acquisition, data storage and data query,
- be able to query, combine and cleanse the data to identify potential issues and resolve inconsistencies, errors and/or issues in the data,
- be able to summarize, visualize, and transform the data to understand it more deeply as well discover data patterns that may inform further analyses,
- employ various mathematical and statistical tools for modeling and estimation of the data,
- use principles and techniques in data science to communicate conclusions and patterns in the data to diverse audiences.