

# TOP: Urban Data Science 002 Spring 2026

## Syllabus

USP 410/510 Urban Data Science · Spring 2026

---

**Time:** Spring 2026, Thursdays 3:30-5:10pm

**Office Hours:** Mondays 11:00am - 1:00pm, URBN 350D

**Credits:** 3

**Modality:** In Person, URBN 220

**Instructor:** Dr. Liming Wang (lmwang@pdx.edu)

**Course Website:** <https://usp510.github.io/> 

This course introduces urban data science, an interdisciplinary approach to understanding, managing, and designing the city using data-driven theories and methods. Urban data science builds on the science and technologies of information processing, information systems, computer science, and statistics to develop applications to cities.

In this project-based class, students have an opportunity to develop applications that combine technical skills and domain knowledge and use information processing, analysis, and presentation to support problem solving in cities. It will introduce students to basic coding, data processing and analysis, visualization and mapping. Students will also learn to work effectively with large language models (LLMs) and AI agents as tools to accelerate data science workflows — from writing and debugging code to exploring data and generating visualizations.

There are no prerequisites, but it requires some tolerance for experimentation, self-directed trial and error, and an interest in learning to write computer code and work with AI tools.

## Synopsis and Objectives

This course is designed to provide students with a toolkit of technical skills for quantitative problem solving. Through project-based hands-on learning, the course aims to achieve these objectives for students:

- An introduction to the fundamentals of computer code to automate tasks;
- Learning to use LLMs and AI agents as tools for coding, data analysis, and problem solving;
- Familiarity with workflow and project management best practices working with data;
- Developing skills of accessing, cleaning, visualizing, and analyzing urban data;
- Learning to combine quantitative technical skills and domain knowledge to support problem solving in cities

## Textbook and Readings

There is no specific textbook for the class. The course will draw on materials from a wide range of sources and will provide students with book excerpts, technical reports, and journal papers as

appropriate to supplement lecture notes. The following textbooks are recommended as general references:

- Yu and Barter, 2024, Veridical Data Science: The Practice of Responsible Data Analysis and Decision Making, MIT Press. (Available online at <https://vdsbook.com/> ↗)
- Downey, 2024, Think Python, 3rd Edition. (Available online at <https://alldowney.github.io/ThinkPython/> ↗)
- Turrell, Python for Data Science. (Available online at <https://aeturrell.github.io/python4DS/> ↗)
- McKinney, 2022, Python for Data Analysis, 3rd Edition, O'Reilly. (Available online at <https://wesmckinney.com/book/> ↗)
- Rey, Arribas-Bel, and Wolf, 2023, Geographic Data Science with Python, CRC Press. (Available online at <https://geographicdata.science/book/> ↗)

## Grade

Component	USP 410	USP 510
DataCamp exercises (4 x 5pts)	20%	20%
Data science show & tell (2 x 5pts)	10%	10%
Assignments	30%	20%
Project presentation	10%	10%
Project report	30%	40%
<b>Total</b>	<b>100%</b>	<b>100%</b>

**DataCamp exercises:** Each DataCamp course is approximately 2-4 hours and you will have two weeks to complete each one.

**Data science show & tell:** Students will take turns sharing examples of good and bad data science projects/products at the beginning of each class (~5 minutes per presentation). Each student will present twice during the quarter. Sign up for your preferred weeks and submit your entry in the [shared Google Doc](#) ↗ following the template provided. Submissions are due by the end of Monday before your scheduled class.

**Class project:** The final product can be in the form of a project report, an infographic, or a dashboard, generated using Python and/or Quarto. Follow the best practices in creating infographics/dashboard & report. Submit your final product in appropriate (html/pdf/png) format and the accompanying Quarto document (& Python script if any). Your project presentation will be no more than 20 minutes in length with 5 minutes for Q&A.

- **USP 410 (undergraduate):** Projects should demonstrate competency in the core skills covered in class — data cleaning, visualization, and presentation of findings using at least one dataset relevant to an urban issue.
- **USP 510 (graduate):** Projects are expected to go beyond the core skills and demonstrate a higher level of analytical rigor. This includes integrating multiple data sources, applying spatial analysis or API-sourced data, and providing deeper interpretation of results that connects findings to relevant urban policy or planning contexts.

**Assignments:** Both assignments use Oregon Department of Transportation (ODOT) crash data. Use of AI tools such as coding agents (e.g., Claude Code, GitHub Copilot, Cursor) is encouraged and recommended — these assignments are an opportunity to practice the AI-assisted data science workflow introduced in class.

- **Assignment 1 — Exploring ODOT Crash Data:** Use ODOT's crash data to investigate questions of your choosing. Example questions include: Does the spring Daylight Saving Time change increase crashes? Do pedestrian fatalities exhibit a nighttime pattern similar to what was [reported by the New York Times in 2023](#) ↗? You are free to explore other questions that interest you. Your

submission should include a Quarto or Jupyter notebook with clear visualizations and a written narrative explaining your findings.

- **Assignment 2 — Interactive Crash Map:** ODOT provides a basic [fatal crash map viewer](#) ↗, but it is a general-purpose data browser. Your task is to create an interactive map (using `folium`, `plotly`, `streamlit`, or a similar library) that goes beyond data browsing — it should have a specific point of view, question, or audience. Example directions include: Which Portland corridors (e.g., 82nd Ave, Powell Blvd) are the most dangerous for pedestrians, and has that changed over time? Are crash hotspots concentrated in lower-income communities or communities of color? Where are the most dangerous intersections near schools, and what does a “safe routes” map look like for parents? How do crash patterns shift across the hours of the day — can you animate a 24-hour cycle to reveal when and where risk peaks? You are free to pursue other questions. Deploy or export your map as a standalone HTML file.

## AI Policy

This course teaches you to work with AI tools as part of the data science workflow. However, different assignments have different goals, and the AI policy reflects that:

**DataCamp exercises — AI tools are not permitted.** These exercises build foundational skills. Using ChatGPT, Copilot, or other AI assistants to complete them undermines the learning process. You need to develop the mental models that make you an effective user of AI tools later.

**Show & tell — AI may be used for research only.** You may use AI tools to help discover examples of data science projects, but your written explanation and in-class discussion should reflect your own understanding and judgment.

**Assignments — AI use is encouraged and recommended.** You may (and are encouraged to) use AI coding agents and assistants to help with data processing, analysis, visualization, and writing code. Include a brief note describing which AI tools you used and for what tasks. You must understand and be able to explain any code or analysis in your submission.

**Final project — AI use is encouraged, with disclosure.** Using AI tools for your project mirrors real-world data science practice. The following requirements apply:

- Include an **AI Use Statement** as an appendix to your final product. Describe which AI tools you used (e.g., ChatGPT, Claude, Copilot, Cursor), what tasks you used them for (e.g., writing code, debugging, generating visualizations, drafting text), and how you verified the outputs.
- You must be able to **explain every part of your project** during your presentation and Q&A. If you cannot explain code or analysis that appears in your submission, it will not receive credit.
- AI-generated analyses must be **validated for correctness** — consistent with the Veridical Data Science framework’s emphasis on truthful, reproducible results.

## Topics and Schedule (Tentative)

Week	Date	Topic	Readings
W1	04/02	<a href="#">Slides</a> · Overview, Computer Setup, <a href="#">Introduction to Python</a>	Yu & Barter, <a href="#">Chapter 2: The Data Science Life Cycle</a> ↗; Downey, <a href="#">Chapter 1</a> ↗
W2	04/09	Learning and working with LLMs and AI agents	Karpathy, <a href="#">How I Use LLMs</a> ↗



Week	Date	Topic	Readings
			; Evkaya & de Carvalho, <a href="#">Using ChatGPT for Data Science Analyses</a> (HDSR, 2026)
W3	04/16	All about data: Data import/export, cleaning & processing	Yu & Barter, <a href="#">Chapter 4: Data Preparation</a> ; McKinney, <a href="#">Chapters 6-8</a>
W4	04/23	Workflow & project management	Yu & Barter, <a href="#">Chapter 3: Setting Up Your Data Science Project</a> ; Turrell, <a href="#">Workflow chapters</a>
W5	04/30	Exploring and visualizing data	Yu & Barter, <a href="#">Chapter 5: Exploratory Data Analysis</a> ; Turrell, <a href="#">Visualize chapter</a>
W6	05/07	Reproducible research/work; Quarto & Jupyter Notebooks	Turrell, <a href="#">Quarto for Python</a> ; <a href="#">jupyter</a>
W7	05/14	Working with spatial data and maps	Rey et al., <a href="#">Chapters 1-4</a>
W8	05/21	Accessing public data from the web and via APIs	<a href="#">censusdis documentation</a> ; <a href="#">Web Scraping with BeautifulSoup</a>
W9	05/28	Developing infographics and dashboard	<a href="#">Ultimate Infographic Design Guide</a> ; <a href="#">Streamlit documentation</a>
W10	06/04	Project workshop	
W11	06/11	Project presentation	

## DataCamp Schedule

#	Course	Hours Assigned	Due
DC1	<a href="#">Introduction to Python</a>	4h	W1 (04/02) W3 (04/16)
DC2	<a href="#">Data Manipulation with pandas</a>	4h	W3 (04/16) W5 (04/30)
DC3	<a href="#">Intro to Data Visualization with Seaborn</a>	4h	W5 (04/30) W7 (05/14)
DC4	<a href="#">Working with Geospatial Data in Python</a>	4h	W7 (05/14) W9 (05/28)

## Assignment Schedule

#	Assignment	Assigned	Due
A1	Exploring ODOT Crash Data	W2 (04/09)	W4 (04/23)
A2	Interactive Crash Map	W6 (05/07)	W8 (05/21)

## Project Milestones

Milestone	Due
Project idea	W3 (04/16)
Project proposal (1 page)	W6 (05/07)
Progress update	W8 (05/21)
Project presentation	W11 (06/11)
Final project submission	W11 06/12

## Resources

DataCamp: Students will be able to take DataCamp courses free of charge courtesy of the DataCamp Classroom program.

## Key Python Libraries


Purpose	Library	Notes
Data manipulation	<code>pandas</code> , <code>numpy</code>	Core data wrangling (replaces dplyr/tidyr)
Visualization	<code>matplotlib</code> , <code>seaborn</code> , <code>plotly</code>	seaborn for statistical plots; plotly for interactive
Reproducible documents	Quarto, Jupyter Notebooks	Quarto renders .qmd and .ipynb to HTML/PDF/Word
Web scraping	<code>requests</code> , <code>beautifulsoup4</code>	HTTP requests and HTML parsing
Census data	<code>censusdis</code> , <code>pygris</code>	Census API access and TIGER/Line shapefiles
Spatial analysis	<code>geopandas</code> , <code>folium</code> , <code>contextily</code>	Vector data, interactive maps, basemaps
Dashboards	<code>streamlit</code>	Pure Python dashboards with free cloud hosting

## PSU Policies and Resources

### Academic Integrity & Grading Policies

- [PSU Academic Calendar](#)
- [PSU Grading System](#)
- [Student Code of Conduct](#)
- [Incomplete Grades Policy](#)

### Student Support Resources

- [How to Find Help at PSU](#) 
  - [Access and Inclusion for Students with Disabilities](#)
  - [Understanding Sexual Misconduct](#)
  - [Title IX Reporting](#)
  - [Religious Accommodations](#)
- 
-