

Table of contents

1	Syllabus	1
1.1	Modification Log	1
1.2	Course Information	2
1.2.1	Course Name	2
1.2.2	Location	2
1.2.3	Professor	2
1.2.4	Office Hours	2
1.3	Course Details	2
1.3.1	Catalog Description	2
1.3.2	Course Description	3
1.3.3	Learning Outcomes	3
1.3.4	Learning Objectives	3
1.3.5	Textbook and Materials	3
1.4	Assessments	3
1.4.1	Grade Distribution	3
1.4.2	Assessment Breakdown	4
1.5	Course Schedule	4
1.6	Course Policies	5
1.6.1	Attendance	5
1.6.2	Late Work	5
1.6.3	Academic Honesty	5
1.7	Willamette Policies	5



1 Syllabus

1.1 Modification Log

Date	Modification
1/12/2025	Initial publishing

1.2 Course Information

1.2.1 Course Name

DATA 503 Fundamentals of Data Engineering

1.2.2 Location

Campus	Room	Day
Portland	GPC 210	Monday
Salem	Ford 102	Wednesday

1.2.3 Professor

Lucas Cordova, Ph.D. Email: LPCordova@willamette.edu Office: Ford 210 (Salem)

1.2.4 Office Hours

Office hours are available by 15-minute appointments via the [Office Hours](#) page or on a drop-in basis when available. Multiple modalities are offered (in-person, phone, Google Meet). If the scheduled times do not align with your availability, please don't hesitate to contact me.

Monday	Tuesday	Wednesday	Thursday	Friday
10:20–11:20 (Salem)		10:20–11:20 (Salem)		
4:30–5:30 (Portland)		4:30–5:30 (Salem)		

1.3 Course Details

1.3.1 Catalog Description

Data management is core to applied computer science and data science. This course introduces relational databases, file-based databases, cloud storage, and data streaming, with an emphasis on selecting appropriate architectures for different data problems.

1.3.2 Course Description

The course focuses on data acquisition, storage, and organization. Topics include the data engineering pipeline and relational database design and querying using PostgreSQL.

1.3.3 Learning Outcomes

Upon completion, students will be able to:

1. Understand the role of a data engineer.
2. Query relational databases using SQL.
3. Analyze database design decisions.
4. Execute advanced SQL queries.
5. Evaluate alternatives to relational databases.

1.3.4 Learning Objectives

Students will be able to:

1. Explain relational database principles.
2. Design and implement relational schemas.
3. Write advanced SQL queries.
4. Use SQL for text and spatial data.
5. Implement a full data pipeline.

1.3.5 Textbook and Materials

- *Practical SQL* (2nd ed.), Anthony DeBarros. **Required**
- *CodeGrade Subscription* (Bookstore or self-purchase) **Required**
- *Fundamentals of Data Engineering*, Reis & Housley. Optional
- *The Data Warehouse Toolkit* (3rd ed.), Kimball & Ross. Optional

1.4 Assessments

1.4.1 Grade Distribution

Percentage	Grade
92	A
90–91.99	A-

Percentage	Grade
88–89.99	B+
82–87.99	B
80–81.99	B-
78–79.99	C+
72–77.99	C
70–71.99	C-
68–69.99	D+
62–67.99	D
60–61.99	D-
59.99	F

1.4.2 Assessment Breakdown

Deliverable	Individual Weight	Total	Description
In-Class Exercises	~1% each	10%	Attendance and participation based.
Homework	~5% each	30%	Design, coding, and analysis assignments.
Midterm	15%	15%	Culmination of in-class exercises and homework assignments.
Project	—	30%	Semester-long team project with milestones.
Final Exam	15%	15%	Comprehensive exam covering course material.
Total		100%	

1.5 Course Schedule

The schedule is tentative.

Week	Date	Topics	Homework	Project
1	Jan 12	Course overview, pipelines	—	—
2	Jan 19	Modeling, normalization	Assignment 1	—
3	Jan 26	SQL basics	Assignment 2	Matchmaking Workshop

Week	Date	Topics	Homework	Project
4	Feb 2	SQL joins	Assignment 3	—
5	Feb 9	SQL constraints	Assignment 4	Proposal
6	Feb 16	Shell, grouping	—	—
7	Feb 23	Midterm	—	—
8	Mar 2	Data generation	Assignment 6	Milestone Checkpoint 1
9	Mar 9	Docker, JSON	Assignment 7	—
10	Mar 16	Subqueries, windows	Assignment 8	—
11	Mar 23	Spring break	—	Milestone Checkpoint 2
12	Mar 30	Regex, text	—	—
13	Apr 6	APIs, views	Assignment 10	—
14	Apr 13	PostGIS, MongoDB	—	Presentation
15	Apr 20	Project presentations	—	Final Write-Up

1.6 Course Policies

1.6.1 Attendance

Consistent attendance and participation are expected. In-class activities cannot be made up.

1.6.2 Late Work

Students receive three late tokens for assignments. Projects incur a 5% per-day penalty up to five days. Late tokens cannot be used for projects. Late token requests are made through Canvas.

1.6.3 Academic Honesty

All work must be your own. Limited collaboration and AI use for idea generation is permitted, but submitting AI-generated work is not allowed.

1.7 Willamette Policies

Includes inclusive classroom practices, accessibility, Title IX, religious accommodations, land acknowledgement, SOAR Center resources, and intellectual property guidelines.