

# Principles of Data Science - Course Study Sheet

Course: Principles of Data Science Cluster: Computing and Data

Integrates data acquisition, cleaning, visualization, modeling, evaluation, ethics, and communication.

## Five-module course map

| Module                         | Active evidence to produce                                                                                          |
|--------------------------------|---------------------------------------------------------------------------------------------------------------------|
| 1. data provenance             | Explain the central idea, complete an application, diagnose one plausible error, and connect it to the next module. |
| 2. cleaning and transformation | Explain the central idea, complete an application, diagnose one plausible error, and connect it to the next module. |
| 3. exploratory analysis        | Explain the central idea, complete an application, diagnose one plausible error, and connect it to the next module. |
| 4. predictive modeling         | Explain the central idea, complete an application, diagnose one plausible error, and connect it to the next module. |
| 5. evaluation and ethics       | Explain the central idea, complete an application, diagnose one plausible error, and connect it to the next module. |

## Reliable learning method

state the question, audit provenance, split or validate appropriately, build a baseline, analyze errors, and communicate limits

## Representative application

Predict an outcome from a public dataset while detecting leakage, subgroup errors, and uncertainty before making a recommendation.

## Weekly study cycle

1. Preview the objective and write two questions. 2. Attend or engage and capture decisions rather than every sentence. 3. Reconstruct the main idea within 24 hours. 4. Complete independent practice. 5. Analyze the first unsupported move on each error. 6. Use feedback. 7. Schedule a delayed mixed recheck.

## Self-check before moving on

Can you explain the idea without notes, complete a fresh task, detect a near-miss, verify the result, and state a limitation? Watch especially for: judging a model only by one accuracy score or ignoring missingness, leakage, sampling, fairness, and reproducibility.

## What to save

a reproducible notebook with data audit, baseline, model comparison, error analysis, and model card

## Use and rights

Copyright 2026 Exams.fit. This PDF is original learning support created for Exams.fit. The Exams.fit administrator may host and display it on Exams.fit. Public download is disabled by default. Do not resell, remove attribution, or present it as an official college document. Course requirements vary; the current syllabus and instructor directions control.

Open-learning verification: <https://openstax.org/subjects/computer-science> | Reviewed August 2, 2026