# Designing and Scaling a Large-Scale Data Science Course

Andy Gunawardena

Computer Science

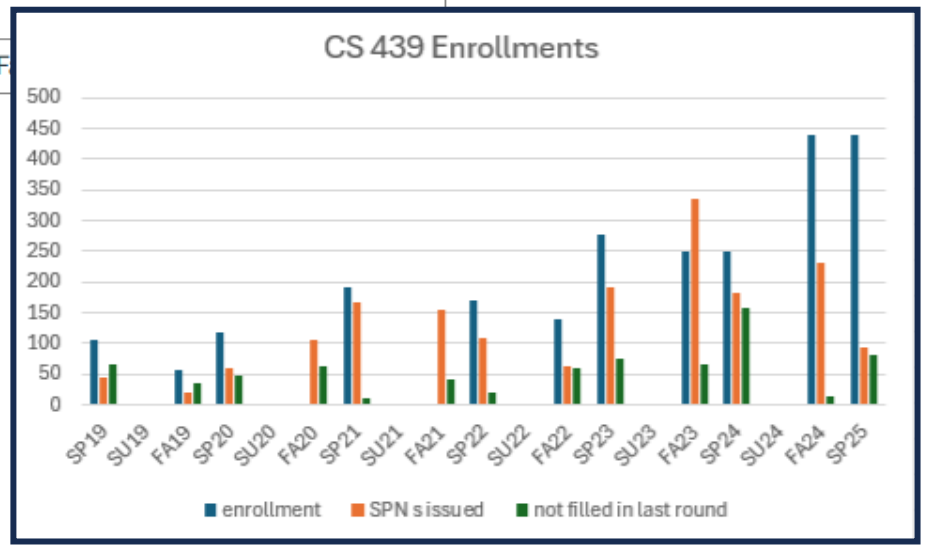
andy.guna@cs.rutgers.edu

# Designing and Scaling a Large-Scale Data Science Course

RUTGERS–NEW BRUNSWICK
Artificial Intelligence
and Data Science Collaboratory

| Module | Topics Covered |
|---|---|
| [1] Tools for Data Science | Python and Pandas, Data Manipulation with Pandas |
| [2] Data Collection, Management and Visualization | Data Collection, EDA, Data Visualization, Kernel Density Estimators, Text Data/Regex |
| [3] Mathematical and Statistical Foundations | Linear Algebra Foundations, Matrix Decompositions, SVD, PCA, Dimension Reduction, probability and statistics |
| [4] Statistical Modeling and Machine Learning | Naive Bayes, Linear Regression, Gradient Descent, Feature Engineering, Regularization, Logistic Regression, Multiple Logistic Regression. Clustering |
| [5] Advanced Modeling Techniques | Deep Learning, Neural Networks, Recommender Systems, |
| [6] Additional Topics (time permitting) | Map Reduce, Tensor Flow, Decision Trees*, Big data platforms and technology*, Human F |

**https://dstf.acm.org/Example-Courses/Introduction_to_Data_Science_Rutgers.pdf**




Current Enrollment – **~500** per semester (8 sections, 2 faculty)
SPNs (not enrolled) – **50-100** per semester

- Number of Graded Labs/Projects – **8 per semester**
- **4000+ weekly** assignment submissions and grading management – 12+ PTL's and Graders

**Handle Scale?** Design a new platform

# Designing and Scaling a Large-Scale Data Science Course

RUTGERS–NEW BRUNSWICK
Artificial Intelligence
and Data Science Collaboratory

RUTGERS–NEW BRUNSWICK
School of Arts and Sciences

**https://codebench.cs.rutgers.edu/**

## Codebench

Codebench is an integrated platform for learning data science through Jupyter notebooks. It assists in desigining, developing,releasing, and overseeing project submissions. During the assignment completion phase, users can leverage generative AI for responsible AI assistance.

See About Codebench for more information about Codebench.

### Links to codebench

There is a separate instance of Codebench for each course that uses it.

For Spring of 2025:

- CS 210, sections 1-4
- CS 210, sections 5-8
- CS 439, section 1-4 (Naina Chaturvedi)
- CS 439, section 5-8 (Ruixiang Tang)

**Over 1000+ students**


Integrated System with **Jupyter** + **nbgrader** + **AI-assistant** + **Grading Management**

Faculty can **design**, **release**, **manage** assignments
Students can **fetch**, **complete**, **seek responsible AI help**, **submit** assignments
Graders can **run**, **annotate**, **release grades** to students

**More Information?**
andy.guna@cs.rutgers.edu

**ALL IN ONE PLATFORM**

**Winner of 2024 Grossman Innovation Award**