

Spring 2020

## CS 301-004: Introduction to Data Science

Pan Xu

Follow this and additional works at: <https://digitalcommons.njit.edu/cs-syllabi>

---

### Recommended Citation

Xu, Pan, "CS 301-004: Introduction to Data Science" (2020). *Computer Science Syllabi*. 95.  
<https://digitalcommons.njit.edu/cs-syllabi/95>

This Syllabus is brought to you for free and open access by the NJIT Syllabi at Digital Commons @ NJIT. It has been accepted for inclusion in Computer Science Syllabi by an authorized administrator of Digital Commons @ NJIT. For more information, please contact [digitalcommons@njit.edu](mailto:digitalcommons@njit.edu).

# CS 301 (004), Introduction to Data Science, Spring 2020

**Instructor: Pan Xu**

**Location:** Central King Building 215

**Class sessions:** Wednesday/Friday, 10 AM to 11:20 AM

**Instructor Office:** GITC 4310

**Office Hours:** by appointment (**email:** [pxu@njit.edu](mailto:pxu@njit.edu))

**TA: Shaoze Fan, email:** [sf392@njit.edu](mailto:sf392@njit.edu);

**Office hours:** Wed. GITC 4324 9:00 AM to 10:00 AM;

Fri. GITC 4325 9:00 AM to 10:00 AM.

**Course Overview:** The course will focus on various practical algorithms widely used in practice related to data science including Bayes' Models, streaming algorithms, reinforcement learning and online learning algorithms. This course will focus on algorithm design and analysis.

**Prerequisite:** CS 100, Math 333

## **Course Content:**

There will be several aspects to this course:

1. Instructor lectures.
2. Individual assignments: There will be four homework assignments, involving both theory and implementation.
4. Final project: A class project will be assigned and will be completed individually.
5. Midterm and Final Exam.
6. Class participation.

Course Schedule:

<b>Dates</b>	<b>Schedule</b>
Week 1	Introduction to Data Science, Understanding different types of data
Week 2	Statistical techniques related to data science
Week 3	Basic probability theory

	<b>(Homework 1 will be offered on Wed)</b>
Week 4	Basic Python tools for data science, Major Rule
Week 5	Naïve Bayes, Linear regression
Week 6	Reinforcement learning <b>(HW2 will be offered on Wed)</b>
Week 7	Reinforcement learning
Week 8	Reinforcement learning and Mid Term <b>(Project will be posted on Friday)</b>
Week 9	Spring Break
Week 10	Streaming Algorithms
Week 11	Streaming Algorithms <b>(HW3 will be posted)</b>
Week 12	Streaming Algorithms
Week 13	Online Learning
Week 14	Online Learning <b>(HW4 will be posted)</b>
Week 15	Online Learning
Week 16	No class <b>(Since May 5 Tues is the last day of class)</b>
Week 17	<b>Final Exam, May 11 (Monday)</b>

### Class Etiquette:

It is sometimes useful to distance us from technology to obtain an environment of quality in-class discussion. Hence, kindly minimize the use of personal laptop computers in the classroom unless appropriate. Please turn off all cellphones and PDAs unless taking notes. If placed on vibrate mode, please ensure that the 'buzz' does not disturb your neighbors during the sessions. Please bring a pen (black or blue) and a writing pad to class to take notes.

Come prepared to the class in the following manner: Read the assigned reading several days prior to the session on that topic.

All assignments are to be submitted to Canvas on the date due, **unless otherwise specified in class**. If a handout is needed, make sure that your name is on each sheet that is handed in and that any hand-in with multiple pages is stapled. You should as a backup also upload your submissions to appropriate slots on Moodle to facilitate grade recording. Failure to submit /uploads to Canvas prior to deadline will result in automatic denial of grade for that work. To keep the grading process on time, extensions are not available and should not be requested.

## Grading Scheme\*:

I reserve the right to make small adjustments to grade weights, or to add small assignments as the need arises.

Item	Grade (% of final grade)	Due date
Homework Assignments (total 4)	10*4=30%	Homework would be posted on Weds. It would be due on Fridays after 1-week.
Final Project	15%	TBD
Midterm Exam	20%	TBD
Final Exam	30%	TBD
Class participation and attendance	5%	

**All exams are open book and notes.**

**\*20% late submission penalty if submitted within one day of deadline. Beyond that, late submission is not entertained. Unless otherwise stated, all submissions are due by 11:59 pm eastern on the due date.**

## Grading Scale

A: Top 20%; B+: Top 21-40% B: Top 41-60%; C+/C/D/F/W: TBD

## Honor and Policy

- Students found cheating or plagiarizing will be immediately referred to the Dean of Students and the NJIT Committee on Professional Conduct and subject to Disciplinary Probation, a permanent marking on the record, possible dismissal, and an "F" grade in the course. All submitted assignments will be checked for similarities, and plagiarism and guilty students identified.

- In the exam, each student is required to sign the Honor Code Agreement "On my honor, I pledge that I have not violated the provision of the NJIT Student Honor Code."