

## Welcome to CS 205: “Data Driven Discovery”

Course Website: <https://courses.engr.illinois.edu/cs199205/>

...or just Google Search: **cs205 uiuc**

**Description:** This course will expand on the basic programming you have learned in an introductory programming course, specifically focusing on data discovery, analysis, and visualization.

**Pilot Course:** CS 205 is a pilot course this semester, everything is subject to change. You will get to shape the experience and development of this course!

### Required Materials:

- Laptop, brought to class every day
- Colored pencils (for visualization drafts)

### Semester at a Glance:

Meta	Major Topics
#1: Structured Data	CSV/TSV Datasets Classifiers Web-based APIs
#2: Images & Natural Language	Color representation, histograms Clustering NLP (grammar trees, similarity)
#3: Geographic & Social	Heat Maps Graphs (Dijkstra’s, TSP) Social Network Trees
#4: Integrated	Solving big problems in small pieces

### Coursework and Grade

CS 205 is graded out of a total of 100 points:

- Data Driven Experience (*14 weeks x 4 pts = 56 points*)
- Projects:
  - 3 Meta Projects (*3 projects x 10 pts = 30 points*)
  - Final Project (*14 points*)
- No exam, no final exam, no scantrons!

Your final letter grade is determined by the number of points you have earned; standard breakpoints, specifics on the course website.

### Python Prerequisites

In CS 205, the primary language we will be using to process data is Python. We will be using the Miniconda Python distribution and Jupyter notebooks for small scripts.

- On the CS 205 website (Experience 0), follow the instructions in the Guidebook section on “Python Prerequisites”

### Create your CS 205 Virtual Space

Throughout CS 205, we will be working with *a lot* of files. It is strongly recommended that you work on it all in a folder on your desktop called **cs205 (no spaces)**. *All of our instructions will assume that your CS 205 space is a folder named cs205 on your desktop.*

- Create your cs205 folder before continuing!

### Navigate to your CS 205 Virtual Space

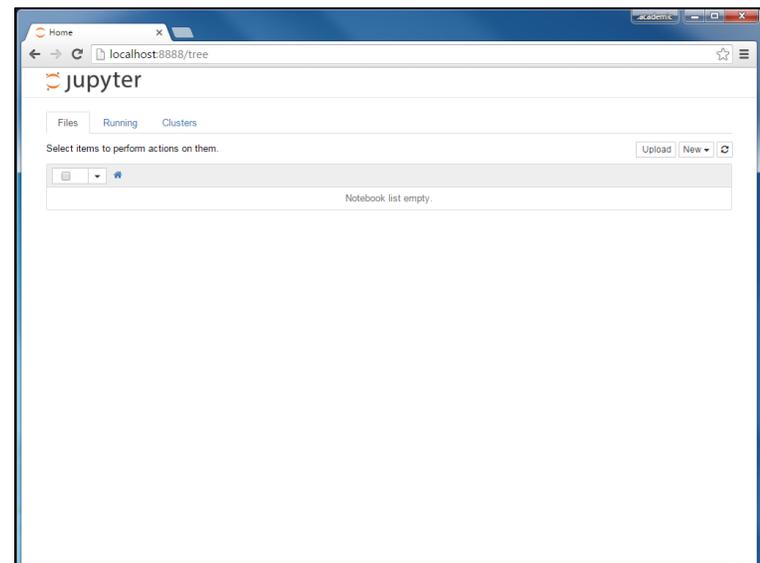
In order to interact with Python in complex ways, we need to use the Command Line.

- On the CS 205 website (Experience 0), follow the instructions on the Guidebook section “Command Line”

### Create your first Jupyter Notebook

From inside your CS 205 directory, run:

```
jupyter notebook
```



An empty Jupyter notebook

## Create your first Notebook

On the right hand side, choose:

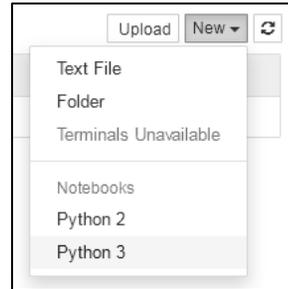
New → "Python 3"

## First Line of Python

Type your first line of Python:

```
In [1] print("Hello, world!")
```

...then press **Shift+Enter** to execute!



## By Thursday:

Complete the codecademy.com Python lesson on "Python Syntax"  
(located under the Unit 1: Python Syntax)

<https://www.codecademy.com/learn/python>

## Python: Semantics at a Glance

If you come from the world of C/Java/JavaScript, Python looks a little different. There are four major things to remember:

- Statements do not terminate with semi-colons; instead, a new line denotes the end of statement
- Functions and loops are not encapsulated with braces; instead, whitespace tabbing indicates the loop/function contained
- Variables do not need to be declared with a keyword (eg: no **var**, **int**, or otherwise), you can just start using them!
- Function calls generally remain the same; arguments are passed inside of parenthesis and

## Python: Reading a CSV File

In Python, we can import libraries to give access to specific functions.

In order to read the CSV file, we need to import the CSV library:

```
In [2] import csv
```

Next, we need to read the file:

```
In [3] f = open("cs205 - Data.csv")
```

The csv library allows us to read the entire CSV file in as a dictionary to easily access later:

```
In [4] reader = csv.DictReader(f)
```

Finally, use the following pattern to print out the names of everyone in the class:

```
In [5] for row in reader:  
    print( row["Name"] )
```