



M1-04: DataFrames: Row Selection and Pandas “Cheat Sheet”

Part of the “Basics of Data Science with Python” Learning Badge

Video Walkthrough: <https://discovery.cs.illinois.edu/m1-04/>

Course Catalog Dataset

The University of Illinois offered thousands of courses across a broad range of subjects every single semester. The collection of all courses offered at Illinois is known as the “Course Catalog” and we have downloaded the full course catalog and organized it into a CSV file on github:

- The source code repository on github.com:
<https://github.com/wadefagen/datasets/tree/master/course-catalog>
- The direct link to the raw CSV file for Fall 2019:
<https://waf.cs.illinois.edu/discovery/course-catalog.csv>
(You find this URL by going to the github repository, clicking on the Fall 2019 file, and choosing “raw”.)

Here’s a sample of just a few rows from this course catalog dataset:

Year	Term	YearTerm	Subject	Number	Title	...
2019	Fall	2019-fa	AAS	100	Intro Asian American Studies	...
...						
2019	Fall	2019-fa	STAT	100	Statistics	...
2019	Fall	2019-fa	STAT	107	Data Science Discovery	...
...						

Dataset URL: <https://waf.cs.illinois.edu/discovery/course-catalog.csv>

Puzzle #1: What is the **complete** Python code that needs to be written to load and display the dataset within a DataFrame using pandas?

1	
2	
3	
4	
5	

Puzzle #2: Using the df you loaded in Puzzle #1, what code is needed to find **five random** courses offered at Illinois?

Python:	
Description:	Select five random courses offered at The University of Illinois.