



08/13/2020

Jans Johnson

has successfully completed

**Capstone: Retrieving, Processing, and Visualizing
Data with Python**

an online non-credit course authorized by University of Michigan and offered through
Coursera

A handwritten signature in black ink, appearing to read 'Charles', followed by a horizontal line.

Charles Severance
Clinical Professor, School of Information
University of Michigan

**COURSE
CERTIFICATE**



Verify at coursera.org/verify/6PA43V4KL2AK

Coursera has confirmed the identity of this individual and
their participation in the course.

WEEK

1



1 hour to complete

Welcome to the Capstone

Congratulations to everyone for making it this far. Before you begin, please view the Introduction video and read the Capstone Overview. The Course Resources section contains additional course-wide material that you may want to refer to in future weeks.



4 videos (Total 36 min), 5 readings, 1 quiz [SEE ALL](#)

WEEK

2



2 hours to complete

Building a Search Engine

This week we will download and run a simple version of the Google PageRank Algorithm and practice spidering some content. The assignment is peer-graded, and the first of three optional Honors assignments in the course. This is a continuation of the material covered in Course 4 of the specialization, and is based on Chapter 16 of the textbook.



6 videos (Total 61 min), 2 readings, 1 quiz [SEE ALL](#)

WEEK

3



1 hour to complete

Exploring Data Sources (Project)

The optional Capstone project is your opportunity to select, process, and visualize the data of your choice, and receive feedback from your peers. The project is not graded, and can be as simple or complex as you like. This week's assignment is to identify a data source and make a short discussion forum post describing the data source and outlining some possible analysis that could be done with it. You will not be required to use the data source presented here for your actual analysis.



2 videos (Total 9 min), 2 readings [SEE ALL](#)

WEEK

4



2 hours to complete

Spidering and Modeling Email Data

In our second optional Honors assignment, we will retrieve and process email data from the Sakai open source project. Video lectures will walk you through the process of retrieving, cleaning up, and modeling the data.



5 videos (Total 49 min), 1 reading, 1 quiz [SEE ALL](#)

WEEK

5



25 minutes to complete

Accessing New Data Sources (Project)

The task for this week is to make a discussion thread post that reflects the progress you have made to date in retrieving and cleaning up your data source so can perform your analysis. Feedback from other students is encouraged to help you refine the process.



1 video (Total 5 min), 1 reading [SEE ALL](#)

WEEK

6



2 hours to complete

Visualizing Email Data

In the final optional Honors assignment, we will do two visualizations of the email data you have retrieved and processed: a word cloud to visualize the frequency distribution and a timeline to show how the data is changing over time.



3 videos (Total 29 min), 1 reading, 1 quiz [SEE ALL](#)

WEEK

7



1 hour to complete

Visualizing new Data Sources (Project)

This week you will discuss the analysis of your data to the class. While many of the projects will result in a visualization of the data, any other results of analyzing the data are equally valued, so use whatever form of analysis and display is most appropriate to the data set you have selected.



2 videos (Total 14 min), 2 readings

[SEE ALL](#)