

Course Competency

CAP 3321C Data Wrangling

Course Description

This course offers a broad introduction to data wrangling, data retrieval, and the use of Python in data analytics. Students will use Python and other business intelligence tools to retrieve data from various sources, clean the dataset, and prepare it for data analysis tasks. Prerequisites: CAP1788 and CAP2704C. (3 hr. lecture; 2 hr. lab).

Course Competency	Learning Outcomes
<p>Competency 1:The student will demonstrate an understanding of the importance of data wrangling by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Computer / Technology Usage
<p>a) Explaining the need for high-quality and accurate data in analytics. b) Describing the typical data wrangling steps. c) Identifying various tools and languages that can perform data wrangling. d) Explaining why Python and various associated packages are ideal for data wrangling.</p>	
<p>Competency 2:The student will demonstrate an understanding of Python data structures and file management by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Computer / Technology Usage
<p>a) Identifying the four built-in data types in Python that are used for data collection (lists, sets, tuples, dictionaries) and writing code using these data types. b) Describing the function of advanced data types such as stacks and queues. c) Opening, reading, writing, and closing a file.</p>	
<p>Competency 3:The student will explore various Python packages for data analytics by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data

	<ol style="list-style-type: none"> 2. Critical thinking 3. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Creating an array and performing mathematical operations on it using NumPy. b) Creating series and dataframes using Pandas. c) Performing basic column and row operations on dataframes. d) Creating various charts using Matplotlib. e) Applying various statistical measures such as min, max, and median. 	
<p>Competency 4:The student will demonstrate an understanding of data cleaning and preparation using Python by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Describing the importance of extracting only the subset of a dataset required for a specific business question. b) Sorting, filtering, indexing, and grouping a dataset. c) Finding missing values, duplicate values, and outliers, and performing various operations to handle them. d) Applying functions to entire rows and columns. e) Identifying the data types in each column and converting one to another when necessary f) Concatenating and joining various datasets. g) Using Pandas methods to create pivot tables and perform additional data manipulation/ transformation tasks. 	
<p>Competency 5:The student will demonstrate an understanding of data retrieval by:</p>	<ol style="list-style-type: none"> 1. Numbers / Data 2. Critical thinking 3. Computer / Technology Usage
<ol style="list-style-type: none"> 1. a) Describing the various sources of data such as databases, data warehouses, and data lakes. b) Exploring various websites with open data. c) Downloading a dataset from the web. d) Importing various file 	

formats such as CSV, JSON, XML, XLSX, PDF, and HTML. e) Scraping data from a website using a Python library. f) Connecting to an API and retrieving data using Python.

Updated: FALL TERM 2022