DATA SCIENCE AND ANALYTICS

DATA WRANGLING

COURSE
SYLLABUS

TABLE OF CONTENTS

COURSE DETAILS

Duration

40 hours • 10 Modules

Course Description

Learners develop core data-wrangling skills by expanding their Python programming skills. Learners explore a series of data analysis processes from sourcing, curating, and importing data, exploratory data analysis, data-cleansing techniques, and data visualization techniques. Next, learners expand their toolbox by using industry-standard software to automate data-wrangling processes.

Learning Goals

By the end of this course, learners will know how to clean messy data as part of the extract, transform, load (ETL) process.

Learning Objectives

- Data collection
- Data cleaning and manipulation
- Data ethics
- An introduction to time series data analysis

REQUIREMENTS

Course Completion Requirements

- Receive a pass recommendation on the course project.
- Receive a 50% or above on each quiz.
- Show passion, aptitude, and potential.
- Instructor recommendation
- Attendance of 80% or higher

Software

To complete the course, learners need the following:

- Computer with internet access
- Web browser
- Google Drive
- Google Colaboratory

GRADING

Assessment Scale

Assessment	Points	% of Grade	# of Assessments	Cumulative Pts
Projects	55	61%	1	55
Quizzes	5	39%	7	35
			Points Possible	**90**

Project Rubric

Criteria	Novice (1)	Trained (6)	Experienced (8)	Score
Effort	Failed to meet all requirements	Met all requirements	Went above and beyond assignment ask	
Craftsmanship	Errors are noticeable, are distracting, or detract from the work.	Work has some errors, but they don't detract from the work.	Work has no errors and exceeds expectations.	
Contrast	Elements clash or are difficult to read.	Text is legible, and elements are mostly not clashing.	All text is completely readable and accessible; no elements clash.	
Repetition	No elements repeat; every element is different, creating confusion.	Some elements repeat, but there may be some confusion for the viewer.	Elements are intentionally repeated, creating visual hierarchy and patterns for the viewer.	
Alignment	Elements are misaligned, either vertically or horizontally.	Elements are mostly aligned, but some may be vertically or horizontally unaligned.	All elements perfectly align on all axes.	
Proximity	Elements are unintentionally placed, causing confusion.	Elements are intentionally placed, but their purpose may be confusing.	All elements are very clear in their purpose and intentionally placed.	

Presentation	Did not meet time limit or show persuasive storytelling techniques	Met time limit and showed persuasive storytelling techniques	Met time limit, showed persuasive storytelling techniques, and went above and beyond on presentation creativity	
			TOTAL	**/55**

COURSE OUTLINE

Class Details

#	Module Name	Lessons
01	Introduction to Data Wrangling	**Lesson 1:** Introduction to Data Wrangling **Lesson 2**: Pandas **Lesson 3**: Series and DataFrame Case Study **Async:** Series and DataFrames
02	Data Collection	**Lesson 1**: Data Access **Lesson 2**: Converting File Formats **Lesson 3**: Merging Files in DataFrames **Async:** Data Collection
03	Data Issues	**Lesson 1**: Data Exploration **Lesson 2**: Data Issues **Lesson 3:** Missing Data **Async:** Data Issues
04	Data Cleaning	**Lesson 1**: Data Cleaning **Lesson 2**: Conditional Statements **Lesson 3**: The Apply Function **Async:** Unit Conversions with Functions
05	Data Manipulation	**Lesson 1**: Column Manipulation **Lesson 2:** String Operations **Lesson 3:** Concat and Merge **Async:** Data Manipulation
06	Time Series Data I	**Lesson 1**: Time Series Review

		Lesson 2: Python Packages and Libraries **Lesson 3**: Calculating Differences in Dates and Times **Async:** Time Series Data I
07	Time Series Data II	**Lesson 1**: Imputing Missing Time Series Data **Lesson 2:** Wildfire Exploration **Lesson 3:** Stationary Processes **Async:** Series and DataFrames
08	Ethics I	**Lesson 1**: Ethics in Data Wrangling **Lesson 2**: PII and Confidentiality **Lesson 3**: Cognitive and Statistical Bias **Async:** Consent and Ethical Human Subjects Research
09	Ethics II	**Lesson 1**: Implicit Bias **Lesson 2**: Data Governance **Lesson 3**: Data Representation and Application **Async:** Same Data, Two Interpretations
10	Final Project Workshop	Howie's Hotel Project

Class Assignments

#	Module Name	To Do Before Next Class	Canvas Assignments
DS-01	Introduction to Data Wrangling	Lesson 1 Activity: Introduction to Data Wrangling	Submit your completed questions / answers.
DS-01	Introduction to Data Wrangling	Lesson 2 Activity: Pandas	Submit your completed questions / answers.
DS-01	Introduction to Data Wrangling	Lesson 3 Activity: Series and DataFrame Case Study	Submit your completed questions / answers.

DS-02	Data Collection	Lesson 1 Activity: Data Access	Submit your completed questions / answers.
DS-02	Data Collection	Lesson 2 Activity: Converting File Formats	Submit your completed questions / answers.
DS-02	Data Collection	Lesson 3 Activity: Merging Files in DataFrames	Submit your completed questions / answers.
DS-03	Data Issues	Lesson 1 Activity: Data Exploration	Submit your completed questions / answers.
DS-03	Data Issues	Lesson 2 Activity: Data Issues	Submit your completed questions / answers.
DS-03	Data Issues	Lesson 3 Activity: Missing Data	Submit your completed questions / answers.
DS-04	Data Cleaning	Lesson 1 Activity: Data Cleaning	Submit your completed questions / answers.
DS-04	Data Cleaning	Lesson 2 Activity: Conditional Statements	Submit your completed questions / answers.
DS-04	Data Cleaning	Lesson 3 Activity: The Apply Function	Submit your completed questions / answers.
DS-05	Data Manipulation	Lesson 1 Activity: Column Manipulation	Submit your completed questions / answers.
DS-05	Data Manipulation	Lesson 2 Activity: String Operations	Submit your completed questions / answers.
DS-05	Data Manipulation	Lesson 3 Activity: Concatenate and Merge	Submit your completed questions / answers.
DS-06	Time Series Data	Lesson 1 Activity: Time Series Review and Case Study	Submit your completed questions / answers.
DS-06	Time Series Data	Lesson 2 Activity: Python Packages and Libraries	Submit your completed questions / answers.
DS-06	Time Series Data	Lesson 3 Activity: Calculating Differences in Dates and Times	Submit your completed questions / answers.
DS-07	Time Series Data II	Lesson 1 Activity: Imputing Missing Time Series Data	Submit your completed questions / answers.
DS-07	Time Series Data II	Lesson 2 Activity: Wildfire Exploration	Submit your completed questions / answers.
DS-07	Time Series Data II	Lesson 3 Activity: Stationary Processes	Submit your completed questions / answers.

DS-08	Ethics	Lesson 1 Activity: Ethics in Data Wrangling	Submit your completed questions / answers.
DS-08	Ethics	Lesson 2 Activity: PII and Confidentiality	Submit your completed questions / answers.
DS-08	Ethics	Lesson 3 Activity: Cognitive and Statistical Bias	Submit your completed questions / answers.
DS-09	Ethics II	Lesson 1 Activity: Implicit Bias	Submit your completed questions / answers.
DS-09	Ethics II	Lesson 2 Activity: Data Governance	Submit your completed questions / answers.
DS-09	Ethics II	Lesson 3 Activity: Data Representation and Applications	Submit your completed questions / answers.
DS-10	Final Project Workshop	Howie's Hotel Project	Submit your completed questions / answers.