

# Written Assignment

## How-To Guide

The Written Assignment represents 100% of the overall course grade.

## Instructions

Develop a Python project to analyse real world scenarios and generate valuable insights by visualising information. The project aims to analyse data from different data sources, manipulate information and visualise to generate insights.

You can use any open-source dataset available online for analytics. Each bullet point for every learning outcome is a milestone to be achieved.

The project should be submitted on the Learn Site under the Assessments section. You will need to include two files, as described below.

1. Project ZIP File (upload to LMS)
  - o Code and Data
    - Include your entire Python project along with all the code and data files and upload as part of your submission
    - The project should cover all milestones in each learning outcome to gain full marks (see below)
  - o Project Report
    - A document containing between 1,500 and 2,000 words
    - Please use the template provided (see Assessments section to download)
    - The report describes your process, dataset, different sources, graphs and insights
    - Justify the use of each learning outcome concept, for example: Why did you use list over dictionary?
2. GitHub repository URL
  - o Create a new repository on GitHub as [UCDPA\_yourname]
  - o Keep committing to the repository
  - o Remember to include the URL of your repository at the beginning of your Project Report document

The goal of the assignment is to demonstrate how you are thinking about putting course concepts and learning into practice to demonstrate the course learning outcomes:

1. Derive insights into data sources a company can use and how to store that data
2. Outline fundamentals of Python data structures such as lists and arrays and learn powerful ways to store and manipulate data
3. Identify ways to use Python data structures, execution control statements, and DataFrames to manipulate data
4. Derive meaningful decisions using Python to compare potential projects and how to make rational, data-driven decisions

## Milestones

1. Data
  - The project should use a real-world dataset and include a reference to the source in the report
2. Importing
  - Import data from a flat file (.csv, .xls, .xlsx, .txt, etc.)
  - Retrieve data using online SQL, APIs, or web scraping
3. Preparation
  - Create pandas DataFrame
  - Sorting, indexing, grouping
  - Drop duplicates, replace missing values
  - Merge DataFrames
4. Analysis
  - Conditional statements, looping, groupby
  - Define a custom function to create reusable code
  - Use NumPy functions
  - Dictionary or Lists
5. Visualisation
  - Generate at least two charts using Matplotlib or Seaborn
6. Insights
  - Derive five valuable insights from the analysis
  - Justify your insights with reference to the charts or analysis
7. Machine Learning
  - Describe what kind of prediction you could perform in future using machine learning and/or deep learning.
  - Would you use classification or regression methods?

## Additional Guidance

Any quotes from external sources should be properly referenced. Choose a referencing style and use it consistently. Poor referencing may affect your grade, and lack of referencing makes the integrity of your entire assessment questionable. We recommend that you use the Harvard Referencing Style, which is well documented in the UCD Library pages: <https://libguides.ucd.ie/harvardstyle>.

Regarding the Project Report, a 10% tolerance above or below the stated word limit is admissible. Referencing does not count towards assessment length limits.

# How You Will Be Assessed

The following rubric describes how the essay will be assessed:

<b>1. Demonstrate an understanding of and apply key concepts and principles of various ways to import data into Python.</b>	<b>2. Show clear understanding of how to store and manipulate data in Python data structures</b>	<b>3. Show clear understanding of key concepts of Boolean logic, control flow and loops in Python.</b>	<b>4. Demonstrate ability to manipulate multiple DataFrames by combining, organising, joining and reshaping them using Pandas</b>	<b>5. Create custom visualisations of the datasets and from these generate insights of the underlying data.</b>
<b>Distinction Criteria</b>				
Demonstrates strong ability to apply key concepts and principles of various ways to import data into Python.	Demonstrates a strong understanding of how to store and manipulate data in Python data structures.	Demonstrates a strong understanding of key concepts of Boolean logic, control flow and loops in Python.	Demonstrates a strong ability to manipulate multiple DataFrames by combining, organising, joining and reshaping them using Pandas.	Demonstrates strong ability to create visualisations from datasets and derive valuable insights.
<b>Merit Criteria</b>				
Able to apply key concepts and principles of various ways to import data into Python.	Able to how to store and manipulate data in Python data structures.	Demonstrates a good understanding of key concepts of Boolean logic, control flow and loops in Python	A very good level ability to manipulate multiple DataFrames by combining, organising, joining and reshaping them using Pandas.	Demonstrates a good ability to create visualisations from datasets and derive valuable insights.
<b>Pass Criteria</b>				
Adequate use of key concepts and principles of various ways to import data into Python.	Makes adequate use Python data structures.	Makes adequate use of Boolean logic, control flow and loops in Python	Adequate manipulation of multiple DataFrames by combining, organising, joining and reshaping them using Pandas.	Adequate visualisations created and insights derived.
<b>Unsatisfactory Criteria</b>				
Response is partial or tangential. Requires greater depth, level of detail and discussion.	Inadequate application of course learning in Python data structures.	Inadequate application of course learning of Boolean logic, control flow and loops	Inadequate manipulation techniques of DataFrames.	Unsatisfactory visualisation and insights derived from data.
<b>Clear Fail Criteria</b>				
Little evidence of knowledge importing data into Python.	Little evidence of ability to apply key course concepts data structures.	Little evidence of ability to apply key course concepts in practice.	Very little evidence of self-reflection manipulation of DataFrames.	Little evidence of ability to create visualisations or derive insights.
<b>No Attempt Criteria</b>				
No submission	No submission	No submission	No submission	No submission