



---

# ASSESSMENT 2: PROJECT AND JOURNAL

Weighting: 50% (42.5% for the Jupyter Notebook report and 7.5% for the Video Journal)

Due Date: Sunday 11:59 pm end of Week 6

## Introduction: Project and Journal

### Purpose

In this project, you will identify a data set to work with (either your own or one of the recommended data sets) and build a Python program to extract and visualise information about the data set. The purpose of this assessment is for you to demonstrate your ability to apply what you have learned throughout the course in the creation of a problem to solve a problem.

### Your tasks

Your project is the major assessment in the course and demonstrates your learning across the course concepts. The concepts included here may not be clear until you reach them in the course. That is OK. They are here for reference as you reach them, and you will be working progressively on your project as you learn each new concept.

You have **two** tasks to complete.

- Jupyter Notebook report containing analysis of a data set (42.5%)
- Video Journal describing how you undertook your work (7.5%)

# Jupyter Notebook report

For this assessment project you are required to create a Jupyter notebook containing analysis of a data set. You will be selecting your own data set to work with during Week 1 and will have the opportunity to discuss possible data sets and how to find them with your tutor to make sure they meet the course requirements.

**Please note:** there are **two** separate assessment submission pages in Canvas — one for the Jupyter notebook report and one for the Video Journal.

## Requirements for your Jupyter Notebook report

The requirements for the data set used for your project are:

**Format:** We recommend you use data that is in comma separated values (CSV) format. You can use data in other formats that can be read by Pandas into data frames (check with your tutor if you are unsure).

**Size:** Minimum of 10 columns (categories) and 20,000 rows (values) of data.

**Types:** Must contain dates, numbers and strings.

**Source:** Must be a data set OTHER THAN the weather data and iris data set used in the examples in this course.

Not synthetic (made up) data — must be real data.

### Requirements for the report

**Coverage:** Must include a minimum of 5 data plots.

**Variety:** There must be at least 3 of the following different types of plots included:

- Barplot
- Scatterplot
- Pie plot
- Horizontal bar plot
- Histogram
- Line plot
- Heatmap

**Correlation:** One plot must be a Heatmap.

**Data Filtering:** Each plot must filter data on at least one attribute.

**Date Splitting:** At least one plot must include a split on a date attribute (i.e. day, month or year).

**Clarity:** Plots must be labelled with meaningful x and y axis labels and must have a title and, if appropriate, legend.

Notebook must include an English description of what each plot shows.

Notebook must include a section at the start describing the data used (format, source, size, categories, etc).

**Data Manipulation:** At least 4 plots must include application of a function to data (such as min/max/mean).

## Requirements for your Video Journal

You are required to record a short (max 10 minute) video describing one or more key plots in your notebook and how you generated the plot. For this you should choose the plot(s) that demonstrate your best work.

For each plot explain:

1. the data you used
2. the question this plot answers
3. how you filtered your data to get these plots.

When explaining your project, use correct and precise terminology:

Example of correct video journal:

*'Here I show the maximum temperature recorded in each city as a bar plot. To do this I used the Pandas groupby method to group the data in my data frame by city. I then selected the MaxTemp column from the data and found the maximum value for each city using the max method.'*

Example of insufficient detail in video journal:

*'Here I show the maximum temperature recorded in each city. I use Pandas to make the plot.'*

Example of incorrect terminology used in video journal:

*'Here I show the maximum temperature recorded in each city. I use NumPy to plot the series frame.'* (NumPy doesn't plot series and there is no such thing as a 'series frame'.)

## Directions to submit

Save and submit your Jupyter Notebook file containing your Python code and your Video Journal through the **two** Canvas submission links (one for the Jupyter Notebook report and one for the Video Journal).

Your tutor will assess your work and provide feedback.

# Rubric

Assessment Criteria	Level of performance					
<p>Data</p> <p>Must be read into Python code data to meet requirements</p> <p>Must not be weather or iris data</p>	<p><b>High Distinction (15/15)</b></p> <p>Includes data that requires parsing beyond dates or combines data from multiple data sets in analysis</p>	<p><b>Distinction (13/15)</b></p> <p>Exceeds requirements (multiple data sets used in analysis)</p>	<p><b>Credit (11/15)</b></p> <p>Meets requirements</p>	<p><b>Pass (9/15)</b></p> <p>Meets most requirements (slightly smaller than specified, missing a type)</p>	<p><b>Fail 0/15)</b></p> <p>Does not meet requirements</p>	<p><b>/15</b></p>
<p>Coverage/ Variety</p>	<p><b>High Distinction (15/15)</b></p> <p>Meets D requirements AND includes a plot not included in practice (evidence of self-study)</p>	<p><b>Distinction (13/15)</b></p> <p>Exceeds requirements</p> <p>Demonstrates appropriate use of all studied plots</p>	<p><b>Credit (11/15)</b></p> <p>Includes 5 plot totals: 1 heatmap and at least 3 other types of plots</p>	<p><b>Pass (9/15)</b></p> <p>Meets minimum requirements (5 plots including 1 heatmap and at least 2 other types of plots)</p> <p>Allow up to one minor issue in a plot</p>	<p><b>Fail 0/15)</b></p> <p>Does not meet requirements</p>	<p><b>/15</b></p>

Filtering	<b>High Distinction (15/15)</b>  Plots demonstrate ability to select meaningful filtering based on goals which include a mix of attributes (different attribute mixes for different plots)	<b>Distinction (13/15)</b>  At least one plot filters on more than one attribute where one of the attributes is a date	<b>Credit (11/15)</b>  At least one plot filters on more than one attribute	<b>Pass (9/15)</b>  Meets requirements – all plots filter on at least one attribute	<b>Fail 0/15)</b>  Does not meet requirements	<b>/15</b>
Date usage	<b>N/A</b>	<b>N/A</b>	<b>N/A</b>	<b>Pass (10/10)</b>  At least one plot is split by a date attribute (month, day, year)	<b>Fail 0/10)</b>  Does not meet requirements	<b>/10</b>

Clarity	<b>High Distinction (15/15)</b>  Distinction criteria plus: Notebook markdown cells clearly describe what is shown by the data  i.e. what conclusions can be drawn from the data) for all plots.	<b>Distinction (13/15)</b>  Credit criteria plus:  Notebook markdown cells describe what conclusions can be drawn for at least some of the plots.	<b>Credit (11/15)</b>  Pass criteria plus:  Demonstrates use of legend where appropriate  Notebook contains a clear description of each plot.	<b>Pass (9/15)</b>  All plots contain title, x & y labels  Notebook contains markdown cells with at least headers for each plot  Data description included	<b>Fail 0/15)</b>  Does not meet requirements	/15
Data Manipulation	<b>High Distinction (15/15)</b>  Video reflects the ability to use proper Python and library terminology, select appropriate visualisation for data and to use documentation to extend knowledge beyond given examples	<b>Distinction (13/15)</b>  Meets credit criteria and demonstrates use of <b>different</b> functions	<b>Credit (11/15)</b>  At least 4 plots include application of a function to the data	<b>Pass (9/15)</b>  At least one plot includes application of a function to the data	<b>Fail 0/15)</b>  Does not meet requirements	/15

Video journal	<b>High Distinction (15/15)</b>  Video reflects the ability to use proper Python and library terminology, select appropriate visualisation for data and to use document documentation to extend knowledge beyond given examples	<b>Distinction (13/15)</b>  Video reflects the ability to use proper Python and library terminology in most cases	<b>Credit (11/15)</b>  Video reflects the ability to use proper Python and library terminology, and how to use libraries to visualise data.	<b>Pass (9/15)</b>  Video reflects the ability to use proper Python and library terminology in most cases	<b>Fail 0/15)</b>  Does not meet requirements or is missing	<b>/15</b>
<b>TOTAL</b>						<b>/100</b>