

Student	Christopher Hopkins	39
---------	---------------------	----

Part 1

	Available	Marks
Import data and generate data frame	5	3
Create a new data frame <i>cumulativeData</i> that shows the cumulative rainfall and cumulative number of hours of sunshine for each city. Display this data frame.	10	7
Include a function <i>displayGraph()</i> that will take the city name weather parameter and will produce a line graph of that weather parameter over the year for that city. Display this using humidity in Wuhan and rainfall in Swansea.	15	10
Include a function <i>compareCities()</i> that will take the start month and end month and two weather parameters, and will plot the data for those parameters for both cities on the same graph. Mean temperature and humidity, if displayed, should be shown as line graphs, while all other weather parameters should be shown as bar plots. Appropriate labelling and colours should be used. Display this using mean temperature and hours of sunshine.	15	0
Commenting, white space and code presentation	5	5
TOTAL		25
Comment Q1: If you make any changes to the source datafiles you should include them in your uploaded submission. The code worked when appropriate files were created. You could also have added the cities before binding together, using a recycling vector. Q2: Good generation of cumulative summary, although a City/Rain/Sun format would be better. Q3: Displaying the humidity works well. The function manages the choice of weather parameter well, but not the choice of city. Q4: Nothing submitted. Excellent commenting and use of white space.		

Part 2

	Available	Marks
Load the data into a data frame <i>covidData</i> , ensuring that the headers are imported. Your script should assume that the data file is stored in the same folder as the script file.	5	3
Plot the number of cases and number of deaths by month for both countries on the same appropriately-labelled graph.	10	6
Create a data frame <i>CovidSummary</i> that shows descriptive statistics for the total number of new cases in China and the UK for April 2020, October 2020, April 2021 and October 2021. This should contain the following measures: mean, standard deviation, median, max, min, and range.	20	0
Perform a t-test to identify whether there is a significant difference in the numbers of new cases in the UK between October 2020 and October 2021. You should state your experimental hypotheses, the result of the test, and whether you accept or reject the null hypothesis in your comments.	10	0
Commenting and use of white space	5	5
TOTAL		14
Comments Q1: If you make any changes to the source datafiles you should include them in your uploaded submission. The code worked when appropriate files were created. Q2: The graph is good, although it would be better to use four different colours for the four datasets. Similarly, you should put cases on one vertical axis and deaths on the other. Q3: Nothing submitted Q4: Nothing submitted Excellent commenting and use of white space.		