

Student	Wael Baghdadi	71
---------	---------------	----

Part 1

	Available	Marks
Generate WineSamples dataframe	5	1
Add proline categories	10	7
Add ratio of hue to colour intensity	5	5
Generate WineSummary	5	2
Add descriptive data for alcohol and phenols	10	7
Identify high/v. high proline levels	10	9
Commenting and use of white space	5	4
Comments Why change the filename? And why not simply use StringsAsFactors=TRUE? <pre>> wineData <- read.csv("wineSamplesCSV.csv", header=TRUE, stringsAsFactors = FALSE)</pre> Error in file(file, "rt") : cannot open the connection In addition: warning message: <pre>In file(file, "rt") : cannot open file 'wineSamplesCSV.csv': No such file or directory</pre> Remember to make both your code and comments narrow enough to fit on a normal screen – 80-90 characters wide is good. Don't remove NA data – work around it later. If you do plan to remove it then you must justify that decision in the comments. Re-ordering the columns to put everything related to the ratio together is a good idea. You have 4 rows in wineSummary rather than 3. The aggregate() function is more efficient for calculating the descriptive statistics but your solution works. Good use of which(). Good use of white space and comments other than the comment length. You could have moved quite a lot of this into functions.		

Part 2

	Available	Marks
Create the countryData dataframe	5	1
Create the CovidSymmary dataframe	15	11
Check the data for normality and process accordingly	10	8
Perform the appropriate t-tests	5	4
Perform the appropriate ANOVA tests	10	7
Commenting and use of white space	5	5
Comments Good use of external packages, but the comments must also explain if any extra packages are used, rather than just trying to load them. Why change the filename? <pre>> library("dplyr") # Load dplyr package Error in library("dplyr") : there is no package called 'dplyr' > > ## Loading the data by converting the xlsx file to csv file then using read.csv function > countryData <- read.csv("CovidDataCSV.csv", header=TRUE, stringsAsFactors = FALSE) Error in file(file, "rt") : cannot open the connection In addition: Warning message: In file(file, "rt") : cannot open file 'CovidDataCSV.csv': No such file or directory</pre> CovidSummary works, although it's a bit of a long way to do it. The outlier identification and elimination process is good, and the double boxplot display is effective. The t-test is good, but remember to state the hypotheses before carrying out the test. ANOVA works, but it would have been better to just run the ANOVA and the Tukey post-hoc test rather than add the extra stuff with descriptive statistics. Good commenting and use of white space. You could have moved quite a lot of this into functions.		