

Student	Ross Meek	68.5
---------	-----------	------

Part 1

	Available	Marks
Load <i>iris.xlsx</i> data into data frame	3	3
Headers are correct	1	1
Species names are a factor	1	1
Loop for <i>iris Setosa</i>	4	4
Loop for <i>iris Versicolour</i>	4	4
Loop for <i>iris Virginica</i>	4	4
Binding results to <i>irisSummary</i>	3	0
Pearson correlation	5	5
Labelled scatterplot	5	4
Correct use of function	3	1
Calculation for <i>iris Setosa</i>	3	0
Calculation for <i>iris Versicolour</i>	3	0
Calculation for <i>iris Virginica</i>	3	0
Binding results to <i>irisSummary</i>	3	0
Commenting, white space, and code presentation	5	3
Comments The strings should be a factor, using <i>stringsAsFactors=TRUE</i> No <i>irisSummary</i> data frame has been created. Correlation is good (and well commented). You should use <i>plot()</i> rather than <i>scatter.smooth()</i> as the question asks for a plain scatterplot rather than one with a smoothed best-fit curve. The labelling etc. is fine though. The structure of the function is correct but the logic within it isn't. It should be called using <i>apply()</i> and should check to see what the species is, then calculate the area accordingly. Comments overall should be more detailed. There should be commenting at the top to state the script title and author, as a start.		

Part 2

	Available	Marks
Load <i>Kicks.xlsx</i> into <i>kickData</i> data frame	5	5
Generate statistics for air-filled balls	7	7
Generate statistics for helium-filled balls	7	7
Binding results to <i>kickSummary</i>	1	1
Removing the outliers	10	6
Performing the t-test	5	2
Extract data for Neil	4	4
Extract data for Jonny	4	4
Test for significant difference	2	0.5
Commenting, white space, and code presentation	5	2
Comments While the Shapiro test is performed correctly, there is no discussion in the comments of what these answers mean and how they have impacted your coding decisions. The air data is already normal (as confirmed by the Shapiro test) and therefore does not need outliers removed from it. What are the hypotheses for the t-test? You should not perform a t-test without defining the hypotheses first. The interpretation is incorrect – as $p > 0.05$ there is no significant difference between them Filters are good. Again, the t-test needs the hypotheses defined and has been misinterpreted. Although the code call is right, $p > 0.05$ so the null hypothesis (there is no difference between the means) must be accepted.		