Student	Farley Staines	67

Part 1

	Available	Marks
Q1 The file *roller_coasters.csv* contains data on a number of roller coasters from all around the world. This shows the name, type, material, manufacturer, location and current status for each roller coaster and may also show data on the height, length, speed, and number of inversions achieved by each roller coaster. Load this data into a data frame *coasterData*, ensuring that the headers are imported and that the type, material, manufacturer, and current status are imported as factors. Your script should assume that the data file is stored in the same folder as the script file. Display the resulting data frame.	5	5
Q2: Create a new data frame *highestData* that shows the name, height, park and status of the highest operating roller coaster of each type of roller coaster. Display this data frame.	10	8
Q3: Include a function *displayComparison()* that will take the seating_type as a parameter and will produce a bar chart showing the speed of the top ten fastest operating rollercoasters of that seating type. Display this using inverted roller coasters.	15	11
Q4: Include a function *correlateLengthSpeed()* that generates a data frame *lengthSpeed* that shows the mean height, length, and coefficient of correlation of length and speed for Sit Down roller coasters These calculations should be based on records in *coasterData* that contain no NA fields.	15	10
Q5: Include a function *compareWoodenSteel()* that performs a t-test to see if there is a significant difference at the p=0.05 level between the speeds of wooden rollercoasters and steel rollercoasters. Include your hypotheses and your conclusions in your comments. These calculations should be based on records in *coasterData* that contain no NA fields.	10	7
Commenting, white space and code presentation	5	3
TOTAL		44

Comment

Your name and student ID should always be included in a comment at the top of your script. Good use of white space and comments otherwise.

Q1: Good.

Q2: Good use of filters with the loop.

Q3: Sensible idea to display the x axis labels vertically, but they still get overwritten by the axis label. Either drop the label or swap the axes so that the labels read normally down the y-axis.

Q4: The logic is good, although it would probably be better to assign the output of the function to a data frame outside of the function so it can be accessed elsewhere if needed.

Q5: Good analysis. You could also have had the function print whether or not the null hypothesis is rejected.

Part 2

	Available	Marks
Explain how your code works for each section (including flow charts) and how you have extracted the required data from the data frame where you have stored the imported data. TASK 1	30	16
A reflection on what you found challenging, what you learned, and what you would do differently next time.	10	7
TOTAL		23

Comments

Don't use a grey background for diagrams. You should aim to provide good contrast between the information pixels and the background pixels and a grey background like this doesn't provide that.
If you're documenting code it's generally useful to include sections of the code for reference as you explain them.
Your process flows are very linear. In practice it's better to diagram at a more detailed level so you can see the program path through any loops or if() structures.