Q2A (2 POINTS)

1 point for each of the following:

- Two histograms appear in the same plot device.
- A legend differentiates the two histograms

Q2B (5 POINTS)

1 point awarded for commenting on each of the following:

- The casual riders' distribution has a sharp peak at 1000. The distribution may be bimodal.
- The casual rider distribution is skewed right.
- There is a long right tail with a small set of unusually large daily counts around 2500.
- The distribution of registered riders has a more symmetric distribution centered around almost 4000 daily riders.
- The distribution for registered riders spread much wider than the casual riders.

Q2C (2 POINTS)

1 point awarded for commenting on each of the following:

- The points should be small to avoid over plotting.
- There should be a legend to differentiate the two sets of points

Q2D (2 POINTS)

1 point for each of the following statements:

- There appears to be a linear relationship between the counts for registered and casual riders, but this relationship depends on whether the day is a work day or a weekend day.
- Given the overplotting, it is difficult to discern the relationships because for example, we can't see if there might be blue points that have been plotted over by green points which do not follow the linear relationship noted above.

Q3A (1 POINT)

- 1 point if the two contours have shapes like the ones given in the instructions

Q3B (3 POINTS)

1 point for each of the following statements:

- On the weekends, the casual and registered riders have a roughly linear relationship.
- During the work week, the relationship also appears roughly linear, but the slope is much higher.
- Here we can see the shape of the relationship because high density regions that we were not able to see due to overplotting in the scatter plot are revealed. That is, we can verify that the linear relationship is maintained in the regions where the working day and weekend overlap.

Q3C (1 POINT)

1 point for recreating the plot

Q4A (2 POINTS)

1 point for each of the following

- The two lines seem to match the ones in the instructions
- A legend differentiates the two histograms

Q4B (2 POINTS)

1 point for each of the following (max 2 points):

- We see daily patterns in rental with what appear to be regular peaks each day. It's hard to see exactly when these occur.
- The weekends have far fewer registered users and more casual riders.
- The number of casual riders is always less than the number of registered riders, but on the week end, the counts are much closer together.

Q5A (1 POINT)

1 point for recreating the plot

Q5B (2 POINTS)

1 point for each of the following observations:

- The casual riders appear to ride throughout the day (rather than night) with roughly the same number of riders from 11 a.m. to 5 p.m.
- Alternatively, while the registered riders also ride more during the day than at night, there are very strong spikes during the morning and evening commute hours and a small bump during lunch.

Q6B (3 POINTS)

1 point for each of the following:

- Seven lines of proportion casual against temperature. (It’s ok if they have multiplicity from not sorting x---the lines will look “thicker”)
- A legend differentiates the seven lines
- The x-axis has the correct scale (values are in Celsius with a range roughly 0-42)

Q6C (2 POINTS)

1 point for each of the following

- As temperature increase, the proportion of casual riders increase as well.
- Weekends (Saturday, Sunday) have higher proportion of casual riders for all temperatures except the coldest.