Problem 3: Your title

The attached dataset, HW, refers to the bike sharing process in a U.S. city. The dataset comprises the variables listed in Table 3. The variable season includes the four seasons in the northern hemisphere: spring, summer, autumn, and winter. The variable weathersit represents four weather conditions:
'Clear skies', 'Cloudy', 'Light rain', 'Heavy rain'. The variable temp is the normalised temperature in degrees Celsius, i.e., values are divided by 41 (the maximum).

Table 4: Descrição das variáveis do dataset HW (Table 3).

TAG	DESCRIPTION
instant	Record index
dteday	Date of observation
season	Season
weathersit	Weather conditions
temp	Temperature in °C (normalised)
casual	Number of casual users
registered	Number of registered users

- 1. Load the dataset HW1_bike_sharing.csv into R. Classify the variables by type (categorical or numerical), identify the number of observations and the start and end dates of the sample.

The variables can be described as:

instant: Numerical (on the csv file it list the data in a crescent number)
dteday: Numerical (indicate the date of observation)
season: Categorical (Indicate one of the four seasons)
weathersit: Categorical (Which of the four weather conditions)
temp: Numerical (value in degrees Celsius)
casual: Numerical (value of the number of casual users)
registered: Numerical (value of registered users)

The file loaded in R indicate 731 observations, starting in 2011-01-01 and ending in 2012-12-31, which indicates 2 full years of observation.

- 3. Assign the appropriate levels to the variables season and weathersit. Create bar plots for both variables. Which season has the highest number of users? Is bike sharing usage dependent on the season? What is the most favourable weather condition?

Assigning the season levels: 1 = Winter, 2 = Spring, 3 = Summer and 4 = Autumn.
Of the weather levels: 1 = Clear Sky, 2 = Foggy, 3 = Light Rain, 4 = Heavy Rain.

Bar Plots made in R:

Visualising the graph, the numbers of users on Summer is the highest, and the Winter season the lowest, showing that bike sharing is highly dependent on seasons.

From the weather bar plot, its noticeable that bike share happens more in Clear Skies weather.

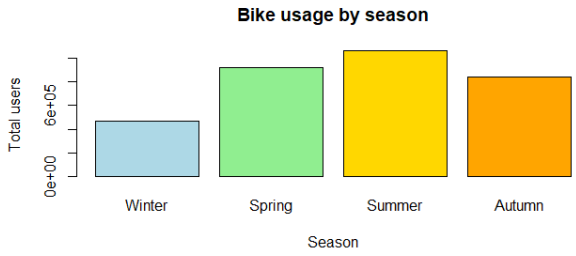

Figure 5: R's boxplot

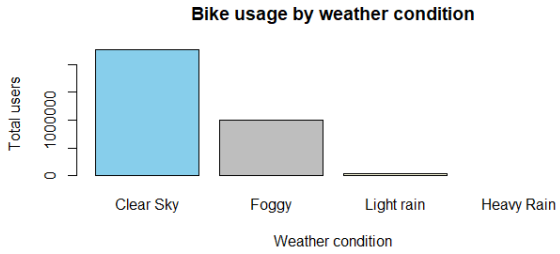

Figure 6: R's boxplot

References