Download data from website given in the Case Study file.

Start Data cleaning steps

1. Loaded each month’s data into excel
2. Formatted Columns “started\_at” and “ended\_at” to **dd-mm-yyyy hh:mm** format.
3. In a new column checked whether ended-at is greater than started\_at or not. Because we cannot go back in time. Found values as **True** and **False.**
4. Deleted rows containing False values
5. Next, created one column called “ride\_length”.
6. Calculated the length of rides in minutes.

Formula Used: =(D2-C2)\*1440

1. Next, created One more column called “day\_of\_the\_week”
2. Extracted the Day name using the “started\_at” column.

Formula Used: =TEXT(C2, “dddd”)

1. Save the clean data sheet.
2. Next, process and further clean the data using R Studio.