

Case Study of Bike journeys

You are a Data Analyst working for the City Council who have recently introduced a bike sharing scheme and would like to understand how the bikes are being used

1. Load the data set and create a `data_dplyr` and `data_base` data frame
2. Format `trip_id`, `user_id` and `location` to be factor variables. Try to complete using the `dplyr` package and base R
3. What is the longest trip taken? Try to complete using `dplyr` and base R
4. What is the longest trip taken by location? Try to complete using `dplyr` and base R
5. Add a column with the average time for each location. Try to complete using `dplyr` and base R and time how long it takes to run each version
6. Create a data frame with only riders who have taken over 100 journeys. Do this just with `dplyr`
7. For riders who have taken over 100 journeys what is the most popular location? Do this just with `dplyr`

Here are some clues in case you are stuck with the case study:

- I. `Sys.time()` will tell you the current time. For question 6 you will need the time before you run the code and the time after the code has finished
- II. Some `dplyr` functions you will need
 - a. `group_by`
 - b. `mutate`
 - c. `summarise`
 - d. `arrange`
 - e. `filter`