

Carleton University
Department of Economics
ECON 3880, Summer 2024

Assignment 4

Due date: 17 June 2024, 11:59 pm

Professor: Firouz Fallahi

Download and read the data set **flights2022.rds** from Brightspace and save it in the same folder as your R Notebook file. Then read it into RStudio and call it **df22** for simplicity. Notice that this dataset is different from what we have used in Assignment 3!

This data set contains flights data for 2022, which is similar to the New York flights dataset and has been extracted using a package called **anyflights** in R.

1 Using df22:

- calculate the number and percentage of cancelled flights in each month (note that for cancelled flights the air_time is missing) and show them in a table in the R notebook
- create a bar chart to show the percentages of canceled flights in different months, fill them based on month
- remove the legend from the plot
- add “Proportion of Cancelled Flights Each Month” and “proportion cancelled” to the plot in appropriate places.

2 Write code that shows the **number** of flights from different airports **to** MYR, OAK, SNA, and PDX airports in df22.

3 How many missing observations do we have in each column of the df22 (note that we need the number of missing in every single column!)? Write code to count the missing observations and show them in the notebook.

4 Use the df22 data:

- drop all the flights to EGE,
- keep only origin, dest, and distance columns,
- exclude all the rows with any missing observation,
- remove all the duplicated rows (i.e., rows with exactly the same values),
- create a wide dataset that shows the destinations in rows and the origins in the columns,
- show the top 10 rows of the wide dataset.

5 Read a data set called **Money** from the package **Ecdat**, which shows Money, GDP and Interest Rate in Canada and assign it to **can_money**:

- use `as.data.frame()` function to convert `can_money` into a data frame
- create a new variable called **my** as `m * y` (i.e., `m` multiplied by `y`) and add that to the dataset
- show the first 2 rows of the data.

Good luck!