

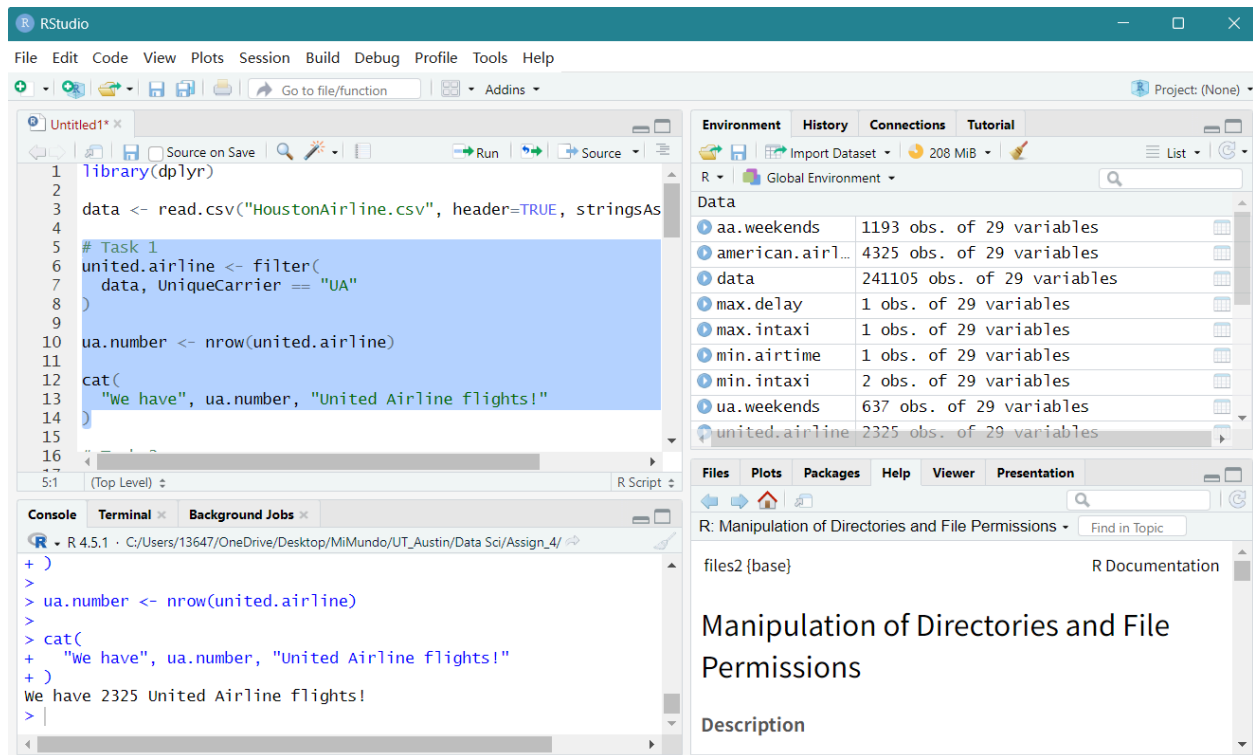
# Assignment 4

Rafael Ignacio Gonzalez Chong & Zhaowen Fan

UT Austin GSI

## Task 1: Number of United Airline Flights in the Dataset

Determine the number of flights operated by United Airlines in the given dataset.



According to the console output, the number of United Airline Flights in the dataset is 2325.

## Task 2: Comparison of United and American Airline Weekend Flights

Compare the number of United Airlines flights on weekends to the number of American Airlines flights on weekends.

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for Task 2.
 

```

15 # Task 2
16 american.airline <- filter(
17   data, UniqueCarrier == "AA"
18 )
19
20 aa.number <- nrow(american.airline)
21
22 aa.weekends <- filter(
23   american.airline, DayOfWeek == 6 | DayOfWeek == 7
24 )
25
26 ua.weekends <- filter(
27   united.airline, DayOfWeek == 6 | DayOfWeek == 7
28 )
29
30 cat(
31   "UA's weekends flights", nrow(ua.weekends)
32 )
33
34 cat(
35   "AA's weekends flights", nrow(aa.weekends)
36 )
37
38 if (nrow(ua.weekends) > nrow(aa.weekends)) {
39   print("UA flights on weekends are larger!")
40 } else if (nrow(ua.weekends) == nrow(aa.weekends)) {
41   print("Both flights on weekends are the same!")
42 } else {
43   print("AA flights on weekends are larger!")
44 }
45
46 # Task 3
47
161 (Top Level) z
      
```
- Environment:** Lists objects in the Global Environment.
 

Object	Class	Attributes
aa.weekends	data.frame	1193 obs. of 29 variables
american.airline	data.frame	4325 obs. of 29 variables
data	data.frame	241105 obs. of 29 variables
max.delay	numeric	1 obs. of 29 variables
max.intaxi	numeric	1 obs. of 29 variables
min.airtime	numeric	1 obs. of 29 variables
min.intaxi	numeric	2 obs. of 29 variables
ua.weekends	data.frame	637 obs. of 29 variables
united.airline	data.frame	2325 obs. of 29 variables
- Console:** Shows the output of the R code.
 

```

R - R 4.5.1 - C:/Users/13647/OneDrive/Desktop/MIMundo/UT_Austin/Data Sci/Assign_A/
+ "AA's weekends flights", nrow(aa.weekends)
+ )
AA's weekends flights 1193>
> if (nrow(ua.weekends) > nrow(aa.weekends)) {
+ print("UA flights on weekends are larger!")
+ } else if (nrow(ua.weekends) == nrow(aa.weekends)) {
+ print("Both flights on weekends are the same!")
+ } else {
+ print("AA flights on weekends are larger!")
+ }
[1] "AA flights on weekends are larger!"
> |
      
```
- Documentation:** Shows the documentation for the 'files2' package, specifically the 'Manipulation of Directories and File Permissions' section.
 

**Manipulation of Directories and File Permissions**

**Description**

These functions provide a low-level interface to the computer's file system.

**Usage**

```

dir.exists(paths)
dir.create(path, showWarnings = TRUE, recursive = FALSE, mode = "0777")
Sys.chmod(paths, mode = "0777", use_umask = TRUE)
Sys.umask(mode = NA)
      
```

As we can see, the number of United Airline flights on weekends are less than the number of American Airline flights, and I printed that in the output console.

### Task 3: Largest Departure Delay

Identify the flight with the longest departure delay in the dataset.

### Task 4: Shortest Flight

Determine the flight with the shortest duration in the dataset.

```
> # Task 3
> max.delay <- data %>% slice_max(order_by = DepDelay, n = 1)
>
> print("The flight with the longest departure delay flight is:")
[1] "The flight with the longest departure delay flight is:"
> print(max.delay)
  Year Month DayOfMonth DayOfWeek DepTime CRSDepTime ArrTime CRSArrTime UniqueCarrier
1 2008     7          31           4      8      555      102      700             MQ
  FlightNum TailNum ActualElapsedTime CRSElapsedTime AirTime ArrDelay DepDelay Origin Dest
1    3574   N811AE           54           65      40    1082    1093    HOU   DFW
  Distance TaxiIn TaxiOut Cancelled CancellationCode Diverted CarrierDelay WeatherDelay
1    247     6     8         0              0         0      1082              0
  NASDelay SecurityDelay LateAircraftDelay
1         0             0                 0
>
> # Task 4
> min.airtime <- data %>% slice_min(order_by = AirTime, n = 1)
>
> print("The flight with the shortest duration:")
[1] "The flight with the shortest duration:"
> print(min.airtime)
  Year Month DayOfMonth DayOfWeek DepTime CRSDepTime ArrTime CRSArrTime UniqueCarrier
1 2008     7          23           3    1159      1155      1505      1505             WN
  FlightNum TailNum ActualElapsedTime CRSElapsedTime AirTime ArrDelay DepDelay Origin Dest
1     17   N636WN           126           130      1         0         4    HOU   JAX
  Distance TaxiIn TaxiOut Cancelled CancellationCode Diverted CarrierDelay WeatherDelay
1     816     4    121         0              0         0      NA              NA
  NASDelay SecurityDelay LateAircraftDelay
1      NA             NA                 NA
>
```

As we can see from the output console, the flight with the longest departure delay in the dataset is the flight number 3574, tail number N811AE; and the shortest duration in the dataset is the flight number 17, tail number N636WN.

### Task 5: Longest and Shortest Taxi In Times

Find the flights with the longest and shortest Taxi In times in the dataset.

```
> # Task 5
> max.intaxi <- data %>% slice_max(order_by = TaxiIn, n = 1)
> min.intaxi <- data %>% slice_min(order_by = TaxiIn, n = 1)
>
> print("The flights with the longest Taxi In time:")
[1] "The flights with the longest Taxi In time:"
> print(max.intaxi)
  Year Month DayOfMonth DayOfWeek DepTime CRSDepTime ArrTime CRSArrTime UniqueCarrier
1 2008     3         25          2   1230      1233    1908      1616             NW
  FlightNum TailNum ActualElapsedTime CRSElapsedTime AirTime ArrDelay DepDelay Origin Dest
1    1630   N347NB          338          163      141      172      -3    IAH   DTW
  Distance TaxiIn TaxiOut Cancelled CancellationCode Diverted CarrierDelay WeatherDelay
1    1076    184     13         0          0          0          0          0          0
  NASDelay SecurityDelay LateAircraftDelay
1     172          0          0
> print("The flights with the shortest Taxi In time:")
[1] "The flights with the shortest Taxi In time:"
> print(min.intaxi)
  Year Month DayOfMonth DayOfWeek DepTime CRSDepTime ArrTime CRSArrTime UniqueCarrier
1 2008     8         27          3   1507      1514    1841      1849             OH
2 2008     9          3          3   1745      1750    2050      2056             9E
  FlightNum TailNum ActualElapsedTime CRSElapsedTime AirTime ArrDelay DepDelay Origin Dest
1    5100   N815CA          154          155      136      -8       -7    IAH   CVG
2    2023   91879E          125          126      113      -6       -5    HOU   ATL
  Distance TaxiIn TaxiOut Cancelled CancellationCode Diverted CarrierDelay WeatherDelay
1     871         0      18         0          0          0          NA          NA
2     696         0      12         0          0          0          NA          NA
  NASDelay SecurityDelay LateAircraftDelay
1      NA          NA          NA
2      NA          NA          NA
> |
```

As we can see, the flight with the longest Taxi In time in the dataset is the flight number 1630, tail number N347NB; the shortest Taxi In time in the dataset are the flight number 5100 and 2023, tail number N815CA and 91879E.

