

# Assignment: data.table

Owen Auston-Babcock and Bradley Collins

2025-03-02

## Owen's section

### Question 1

Shows a table of year, month, day and hour from imported data (data.table)

```
nycdata[, .(year, month, day, hour)]
```

```
##           year month   day  hour
##      <int> <int> <int> <int>
##    1:  2014     1     1     9
##    2:  2014     1     1    11
##    3:  2014     1     1    19
##    4:  2014     1     1     7
##    5:  2014     1     1    13
##      ---
## 253312: 2014    10    31    14
## 253313: 2014    10    31     8
## 253314: 2014    10    31    11
## 253315: 2014    10    31    11
## 253316: 2014    10    31     8
```

### Question 2

Shows a table of DL flights from JFK to SEA (data.table)

```
nycdata[carrier=="DL" & origin=="JFK" & dest=="SEA"]
```

```
##           year month   day dep_delay arr_delay carrier origin  dest air_time
##      <int> <int> <int>    <int>    <int>    <char> <char> <char>    <int>
##    1:  2014     1     1        86        79      DL   JFK   SEA        347
##    2:  2014     1     1        -2        -4      DL   JFK   SEA        347
##    3:  2014     1     2         0        11      DL   JFK   SEA        339
##    4:  2014     1     2        -3         9      DL   JFK   SEA        337
##    5:  2014     1     2        21        19      DL   JFK   SEA        337
##      ---
## 1074: 2014    10    30         -3       -15      DL   JFK   SEA        339
## 1075: 2014    10    31         -6       -26      DL   JFK   SEA        317
## 1076: 2014    10    31         -1        -8      DL   JFK   SEA        338
## 1077: 2014    10    31         -1       -23      DL   JFK   SEA        326
## 1078: 2014    10    31          4       -27      DL   JFK   SEA        318
##           distance  hour
##      <int> <int>
##    1:    2422     9
```

```
##      2:      2422      18
##      3:      2422      15
##      4:      2422       7
##      5:      2422      18
##      ---
## 1074:      2422      18
## 1075:      2422       9
## 1076:      2422       6
## 1077:      2422      15
## 1078:      2422      18
```

### Question 3

Shows a table of DL flights in March that were under 330 minutes flight time (data.table)

```
nycdata[carrier=="UA" & month==3 & air_time<330]
```

```
##      year month   day dep_delay arr_delay carrier origin  dest air_time
##      <int> <int> <int>    <int>    <int>  <char> <char> <char>    <int>
##      1:  2014     3     1        11        43     UA   EWR   STT      209
##      2:  2014     3     1        47        13     UA   EWR   PBI      133
##      3:  2014     3     1        39        10     UA   EWR   MIA      139
##      4:  2014     3     1        -2       -12     UA   EWR   IAH      197
##      5:  2014     3     1        34        36     UA   EWR   DEN      256
##      ---
## 3785:  2014     3    31         6        -8     UA   EWR   FLL      155
## 3786:  2014     3    31         7        -9     UA   EWR   PBI      135
## 3787:  2014     3    31         1       -21     UA   EWR   RSW      145
## 3788:  2014     3    31         0       -19     UA   EWR   IAH      196
## 3789:  2014     3    31        18        -7     UA   EWR   ORD      108
##      distance  hour
##      <int> <int>
##      1:    1634     9
##      2:    1023    19
##      3:    1085    17
##      4:    1400     5
##      5:    1605    16
##      ---
## 3785:    1065    16
## 3786:    1023    10
## 3787:    1068    14
## 3788:    1400    16
## 3789:     719     6
```

### Question 4

Shows a table of DL flights in March that were under 330 minutes flight time (tidyverse)

```
nycdata %>% filter(carrier=="UA" & month==3 & air_time<330)
```

```
## Index: <dest__origin__carrier>
##      year month   day dep_delay arr_delay carrier origin  dest air_time
##      <int> <int> <int>    <int>    <int>  <char> <char> <char>    <int>
##      1:  2014     3     1        11        43     UA   EWR   STT      209
##      2:  2014     3     1        47        13     UA   EWR   PBI      133
##      3:  2014     3     1        39        10     UA   EWR   MIA      139
```

```
##      4: 2014      3      1      -2      -12      UA      EWR      IAH      197
##      5: 2014      3      1      34      36      UA      EWR      DEN      256
##      ---
## 3785: 2014      3     31       6      -8      UA      EWR      FLL      155
## 3786: 2014      3     31       7      -9      UA      EWR      PBI      135
## 3787: 2014      3     31       1     -21      UA      EWR      RSW      145
## 3788: 2014      3     31       0     -19      UA      EWR      IAH      196
## 3789: 2014      3     31      18      -7      UA      EWR      ORD      108
##           distance  hour
##           <int> <int>
##      1:      1634      9
##      2:      1023     19
##      3:      1085     17
##      4:      1400      5
##      5:      1605     16
##      ---
## 3785:      1065     16
## 3786:      1023     10
## 3787:      1068     14
## 3788:      1400     16
## 3789:       719      6
```

## Bradley's section

### Question 5

Calculates airspeed in miles per hour (data.table)

```
nycdata[, speed := distance / (air_time/60)]
```

### Question 6

Calculates airspeed in miles per hour (tidyverse)

```
nycdata <- nycdata %>% mutate(air_speed = distance/(air_time/60))
```

### Question 7

#### 7a

Changes the abbreviation to UnitedAir using data.table

```
nycdata[carrier == 'UA', carrier := "UnitedAir"]
```

#### 7b

Changes the abbreviation to UnitedAir using tidyverse

```
nycdata <- nycdata %>% mutate(carrier =ifelse(carrier =='UA','UnitedAir', carrier))
```