AE 02

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.5.1     v tibble    3.2.1
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

```
library(nycflights13)
```

`nycflights` data

1. Find all flights that flew to Portland (destination PWM).

```
flights %>% filter(dest == "PWM")
```

```
# A tibble: 2,352 x 19
    year month   day dep_time sched_dep_time dep_delay arr_time sched_arr_time
   <int> <int> <int>    <int>          <int>     <dbl>    <int>          <int>
 1  2013     1     1      947            953        -6     1053           1110
 2  2013     1     1     1056           1059        -3     1203           1209
 3  2013     1     1     1350           1355        -5     1456           1510
 4  2013     1     1     1454           1458        -4     1554           1615
 5  2013     1     1     1832           1823         9     1948           1940
 6  2013     1     1     2224           2200        24     2324           2316
 7  2013     1     1     2240           2245        -5     2340           2356
```

```
 8  2013     1     2      940            844        56     1055           1003
 9  2013     1     2      952            953        -1     1104           1110
10  2013     1     2     1205           1129        36     1316           1239
# i 2,342 more rows
# i 11 more variables: arr_delay <dbl>, carrier <chr>, flight <int>,
#   tailnum <chr>, origin <chr>, dest <chr>, air_time <dbl>, distance <dbl>,
#   hour <dbl>, minute <dbl>, time_hour <dttm>
```

2. Which carriers flew to Portland?

```
flights %>% filter(dest == "PWM") %>% distinct(carrier)
```

```
# A tibble: 3 x 1
  carrier
  <chr>
1 B6
2 EV
3 DL
```

The carriers are B6, EV, and DL.

3. Find the number of flights, by carrier, that flew to Portland.

```
flights %>% filter(dest == "PWM") %>% count(carrier)
```

```
# A tibble: 3 x 2
  carrier     n
  <chr>   <int>
1 B6       1304
2 DL        235
3 EV        813
```

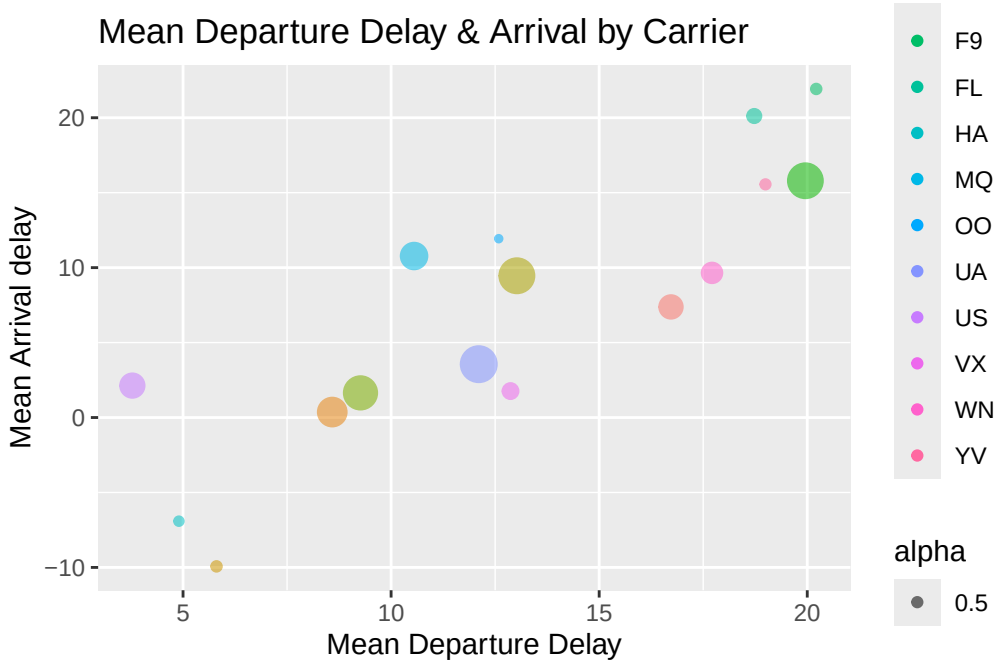
The # of lights by carrier are 334424

Recreate visualizations

Mean departure arrival delays by carrier

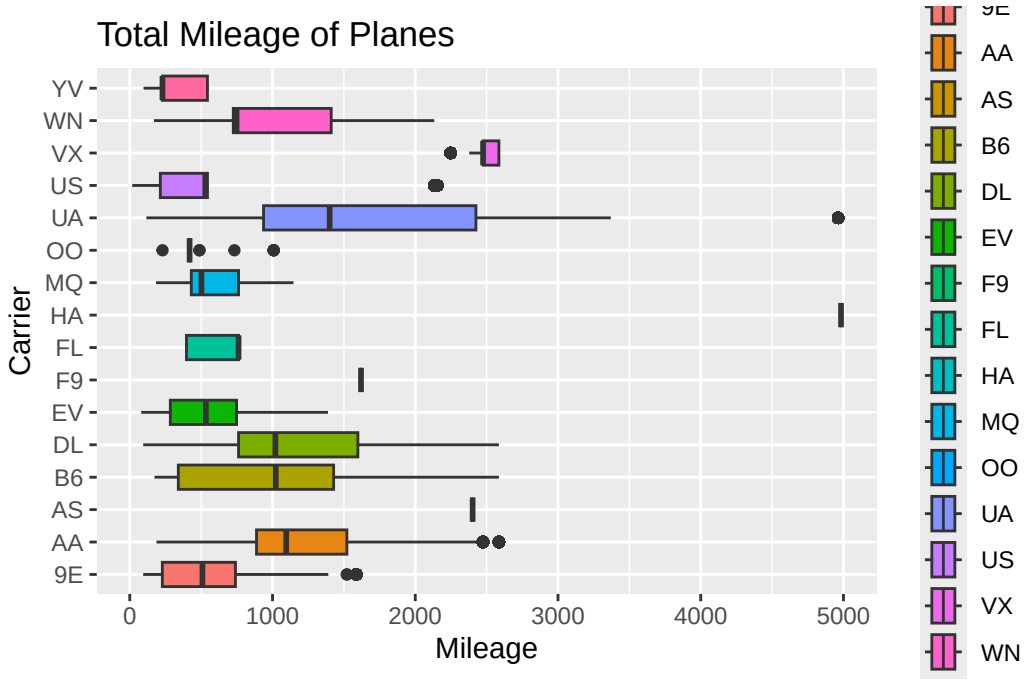
```
flights %>%
  group_by(carrier) %>%
  summarize(
    Mean_Arr_Delay = mean(arr_delay, na.rm = TRUE),
    Mean_Dep_Delay = mean(dep_delay, na.rm = TRUE),
```

```
    count = n()
  ) %>%
ggplot(aes(y = Mean_Arr_Delay, x = Mean_Dep_Delay, color = carrier)) +
  geom_point(aes(size = count, alpha = .5)) + labs(
    title = "Mean Departure Delay & Arrival by Carrier",
    y = "Mean Arrival delay",
    x = "Mean Departure Delay"
  )
```



total Mileage of planes

```
flights %>%
  ggplot(aes(x = distance, y = carrier, fill = carrier)) +
  geom_boxplot() +
  labs(
  title = "Total Mileage of Planes",
  y = "Carrier",
  x = "Mileage"
  )
```


Total Mileage of Planes
Carrier
YV
WN
VX
US
UA
OO
MQ
HA
FL
F9
EV
DL
B6
AS
AA
9E
0
1000
2000
3000
4000
5000
Mileage
9E
AA
AS
B6
DL
EV
F9
FL
HA
MQ
OO
UA
US
VX
WN