



Flights Dataset

5 Questions ask about flights dataset

```
# Load the library
library(nycflights13)

View(flights)
glimpse(flights)
```

5 Question

1.) Which airlines had the top 5 most on-time flights?

```
ontime_flights <- flights %>%
  filter(dep_delay <= 0) %>%
  count(carrier) %>%
  arrange(-n) %>%
  head(5)

print(ontime_flights)
```

Result

```
# A tibble: 5 × 2
  carrier     n
```

```
<chr> <int>
1 B6      32724
2 DL      32520
3 UA      30718
4 EV      28217
5 AA      21931
```

2.) What are the top 5 most-flown destinations?

```
top_destinations <- flights %>%
  count(dest) %>%
  arrange(-n) %>%
  head(5)

print(top_destinations)
```

Result

```
# A tibble: 5 × 2
  dest      n
  <chr> <int>
1 ORD     17283
2 ATL     17215
3 LAX     16174
4 BOS     15508
5 MCO     14082
```

3.) Which airlines had the top 5 most delayed flights (dep_delay > 60)?

```
delayed_flights_by_airline <- flights %>%
  filter(dep_delay > 60) %>%
  count(carrier) %>%
  arrange(-n) %>%
  head(5)

print(delayed_flights_by_airline)
```

Result

```
# A tibble: 5 × 2
  carrier     n
  <chr>   <int>
1 EV        6861
2 B6        4571
3 UA        3824
4 DL        2651
5 AA        2003
```

4.) What are the top 5 airlines with the longest average departure delay?

```
average_delay_by_airline <- flights %>%
  group_by(carrier) %>%
  summarize(average_dep_delay = mean(dep_delay, na.rm = TRUE)) %>%
  arrange(-average_dep_delay) %>%
  head(5)

print(average_delay_by_airline)
```

Result

```
# A tibble: 5 × 2
  carrier average_dep_delay
  <chr>           <dbl>
1 F9              20.2
2 EV              20.0
3 YV              19.0
4 FL              18.7
5 WN              17.7
```

5.) Which airlines operated the top 5 most flights on July 1st?

```
flights_on_july1 <- flights %>%
  filter(month == 7, day == 1) %>%
  count(carrier) %>%
  arrange(-n) %>%
  head(5)
```

Result

```
# A tibble: 5 × 2
  carrier     n
  <chr>   <int>
1 UA        168
2 B6        161
3 EV        148
4 DL        136
5 AA         96
```