NYCflights13

▼ Q1: Top 5 destinations in December 2013

▼ Code

```
Q1 <- flights %>%
    filter(month == 12, year == 2013) %>%
    count(dest, sort = TRUE, name = "Total_flight") %>%
    arrange(desc(5)) %>%
    head(5) %>%
    left_join(airports,
            by = c("dest" = "faa"))
```

▼ Results

```
> Q1 <- flights %>%
+     filter(month == 12, year == 2013) %>%
+     count(dest, sort = TRUE, name = "Total_flight") %>%
+     arrange(desc(5)) %>%
+     head(5) %>%
+     left_join(airports,
+             by = c("dest" = "faa"))
> Q1
# A tibble: 5 × 9
  dest  Total_flight name         lat    lon   alt    tz dst
  <chr>        <int> <chr>      <dbl>  <dbl> <dbl> <dbl> <chr>
1 ATL           1463 Hartsfie…  33.6  -84.4  1026    -5 A
2 LAX           1408 Los Ange…  33.9 -118.    126    -8 A
3 MCO           1219 Orlando …  28.4  -81.3    96    -5 A
4 ORD           1196 Chicago …  42.0  -87.9   668    -6 A
5 CLT           1193 Charlott…  35.2  -80.9   748    -5 A
```

▼ Q2: Top 5 biggest carriers of JFK origin

▼ Code

```
Q2 <- flights %>%
  left_join(planes,
            by = c("tailnum" = "tailnum"))%>%
  select(flight, origin, dest, model, seats)%>%
  filter(origin == "JFK") %>%
  arrange(desc(seats))%>%
  head(5)
```

▼ Results

```
> Q2 <- flights %>%
+   left_join(planes,
+             by = c("tailnum" = "tailnum"))%>%
+   select(flight, origin, dest, model, seats)%>%
+   filter(origin == "JFK") %>%
+   arrange(desc(seats))%>%
+   head(5)
> Q2
# A tibble: 5 × 5
  flight origin dest  model    seats
   <int> <chr>  <chr> <chr>    <int>
1     27 JFK    ATL   747-451    450
2    595 JFK    SFO   777-222    400
3    389 JFK    SFO   777-222    400
4     17 JFK    LAX   777-232    400
5     27 JFK    PHX   A321-231   379
```

▼ Q3: Top 3 most delayed flights in April 2013

▼ Code

```
Q3 <- flights %>%
  select(carrier, flight, arr_delay, year, month, day)%>%
  filter(year == 2013, month == 4)%>%
  mutate(arr_delay_hr = glue("{round(arr_delay/60, digit = 2)} hrs")) %>%
  arrange(desc(arr_delay))%>%
  head(3)
```

▼ Results

```
> Q3 <- flights %>%
+   select(carrier, flight, arr_delay, year, month, day)%>%
+   filter(year == 2013, month == 4)%>%
+   mutate(arr_delay_hr = glue("{round(arr_delay/60, digit = 2)} hrs")) %>%
+   arrange(desc(arr_delay))%>%
+   head(3)
> Q3
# A tibble: 3 × 7
  carrier flight arr_delay  year month   day arr_delay_hr
  <chr>    <int>     <dbl> <int> <int> <int> <glue>
1 DL        2391       931  2013     4    10 15.52 hrs
2 DL        1435       821  2013     4    19 13.68 hrs
3 AA        1901       783  2013     4    19 13.05 hrs
```

▼ Q4: Number of flights of LGA-IAH route on 1st October 2013

▼ Code

```
Q4 <- flights %>%
  filter(origin == "LGA", dest == "IAH", year == 2013, month == 10, day == 1) %>%
  count()
```

▼ Results

```
> Q4 <- flights %>%
+   filter(origin == "LGA", dest == "IAH", year == 2013, month == 10, day == 1) %>%
+   count()
> Q4
# A tibble: 1 × 1
      n
  <int>
1     9
```

▼ Q5: Top 5 most popular routes in 2013

▼ Code

```
Q5 <- flights %>%
  mutate(route = glue("{origin} - {dest}")) %>%
  count(route, sort = TRUE, name = "Total_flight") %>%
  arrange(desc(5)) %>%
  head(5)
```

▼ Results

```
> Q5 <- flights %>%
+   mutate(route = glue("{origin} - {dest}")) %>%
+   count(route, sort = TRUE, name = "Total_flight") %>%
+   arrange(desc(5)) %>%
+   head(5)
> Q5
# A tibble: 5 × 2
  route     Total_flight
  <glue>           <int>
1 JFK - LAX        11262
2 LGA - ATL        10263
3 LGA - ORD         8857
4 JFK - SFO         8204
5 LGA - CLT         6168
```