United Airlines Flight Insights

Group 12: Aria, Leah, Helen, Nana Kofi

2024-09-25

Key Findings

Our team wanted to analyze what airport United Airlines was experiencing the most delays in and build our analysis on ways to improve the structure and flow for travelers. We found that Denver Airport (DEN) experienced the most United flight delays. Further investigation revealed that flights from Denver to Chicago O'Hare International Airport (ORD) suffer from the highest number of delays. After researching possible causes for this correlation we found that weather delays and operational staffing issues are likely contributing factors.

```
library(tidyverse)
library(dplyr)
library(ggplot2)

June1<-read_csv("June_1.csv")
June2<-read_csv("JUne_2.csv")

June1 %>% filter(OP_UNIQUE_CARRIER=="UA")->United
```

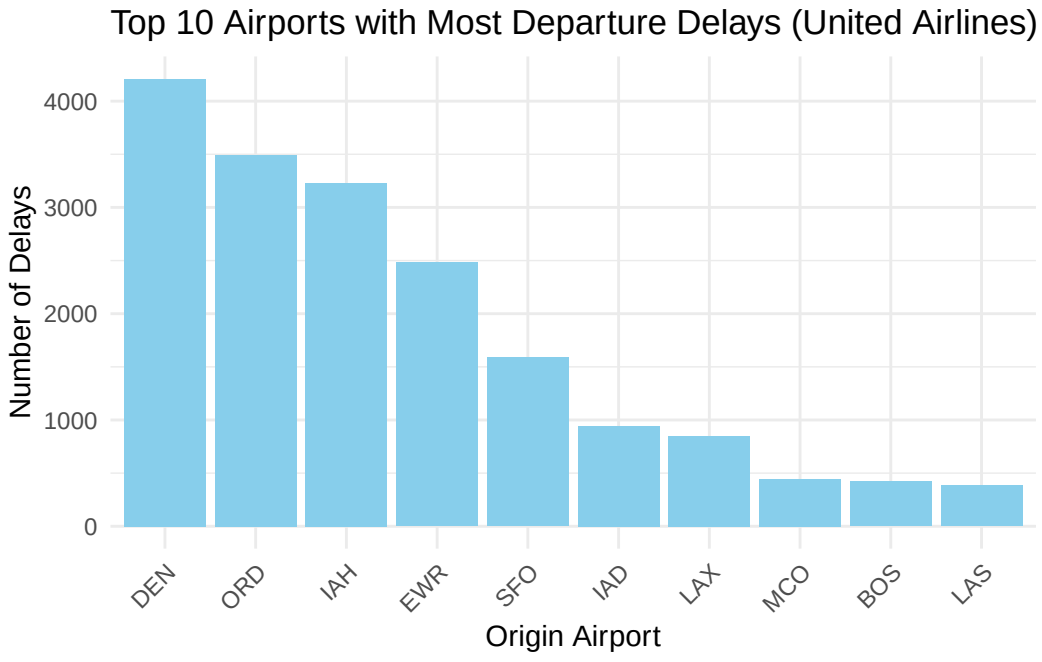
Top 10 Airports with Most Departure Delays (United Airlines)

```
#Filter for positive departure delays,
#count # of delays per airport and arrange in descending order
delay_count_by_airport <- United %>%
  filter(DEP_DELAY > 0) %>%
  group_by(ORIGIN) %>%
  summarise(number_of_delays = n()) %>%
  arrange(desc(number_of_delays)) %>%
  slice(1:10)  # Select the top 10 most delayed airports

# Print the result
print(delay_count_by_airport)
```

```
# A tibble: 10 x 2
   ORIGIN number_of_delays
   <chr>             <int>
 1 DEN                4210
 2 ORD                3493
 3 IAH                3228
 4 EWR                2480
 5 SFO                1586
 6 IAD                 941
 7 LAX                 851
 8 MCO                 443
 9 BOS                 421
10 LAS                 384
```

```r
#Create the bar chart for top 10 most delayed airports by count
ggplot(delay_count_by_airport, aes(x = reorder(ORIGIN, -number_of_delays),
                                   y = number_of_delays)) +
  geom_bar(stat = "identity", fill = "skyblue") +
  labs(title = "Top 10 Airports with Most Departure Delays (United Airlines)",
       x = "Origin Airport",
       y = "Number of Delays") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



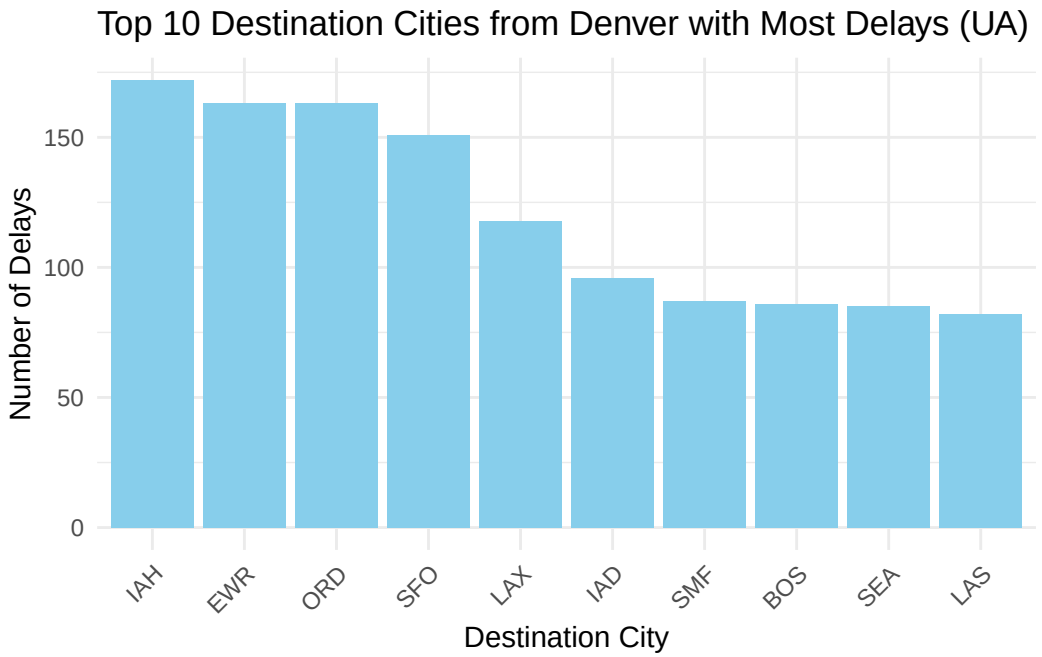
Top 10 Destination Cities from Denver with Most Delays (United Airlines)

```r
denver_delays <- United %>%
  filter(ORIGIN == "DEN", DEP_DELAY > 0) %>%
  group_by(DEST) %>%
  summarise(number_of_delays = n()) %>%
  arrange(desc(number_of_delays)) %>%
  slice(1:10)  # Select the top 10 destination cities with the most delays

# Print the result
print(denver_delays)
```

```
# A tibble: 10 x 2
   DEST  number_of_delays
   <chr>            <int>
 1 IAH                172
 2 EWR                163
 3 ORD                163
 4 SFO                151
 5 LAX                118
 6 IAD                 96
 7 SMF                 87
 8 BOS                 86
 9 SEA                 85
10 LAS                 82
```

```r
#Create the bar chart for top 10 destinations from Denver with the most delays
ggplot(denver_delays, aes(x = reorder(DEST, -number_of_delays),
                          y = number_of_delays)) +
  geom_bar(stat = "identity", fill = "skyblue") +
  labs(title = "Top 10 Destination Cities from Denver with Most Delays (UA)",
       x = "Destination City",
       y = "Number of Delays") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Business Insights

From the graph of the "Top 10 Airports with Most Departure Delays (United Airlines)" it highlights how the Denver airport has the most departure delays at 4210 compared to the 2nd most delayed airport which has 3493 departure delays. We wanted to do more research into why the Denver airport had so many delays and decided to look into the destination cities of delayed flights from Denver airport. From this analysis we found that the destination city of Chicago had the most delayed flights from the origin airport of Denver. We decided to analyze the destination location as flights get delayed because of issues at the origin and destination location so it is vital to look at both places of issue. We did this to narrow down the worst route in terms of amount of delayed flights to tackle United airlines biggest route issues.

Frequent delays for this flight from DEN-ORD could lead to increased customer frustration and possible adverse effects to United Airlines' brand reputation. As a solution to this problem, United Airlines should ensure that their crews for these locations are always well-staffed, as this could improve operational efficiency to prevent delays due to inclement weather which is prevalent at those locations. Additionally, they could invest more into building a strong relationship with air traffic control to prevent congestion.

Functions Report

- **Group_by():** This function was used to group the data by the origin airport to do analysis on each airport to figure out the count of delays for each airport. We also did a group by for the destination to do an analysis on each destination to figure out the count of delays for flights out of Denver.
- **Arrange():** We used the arrange function to format our tables from highest to lowest. We wanted to focus on airports and destinations with the most delays and used the arrange function to have better readability of the data we wanted to look at.
- **Filter():** We used the filter function to count the number of delays at each airport and destination as a flight counts as a delay if the del_dep > 0.