# HPAM 7660 Data Assignment 2

## Grace Arkfeld

### 2024-02-19

```r
options(repos = "https://cloud.r-project.org/")
```

```r
install.packages("dplyr")
```

```
## Installing package into 'C:/Users/graci/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)

## package 'dplyr' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##   C:\Users\graci\AppData\Local\Temp\RtmpcfDVoJ\downloaded_packages
```

```r
install.packages("nycflights13")
```

```
## Installing package into 'C:/Users/graci/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)

## package 'nycflights13' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##   C:\Users\graci\AppData\Local\Temp\RtmpcfDVoJ\downloaded_packages
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(nycflights13)
```

```r
united_flights <- flights %>% filter(carrier == "UA")
```

```r
united_flights_lga <- flights %>% filter(carrier == "UA", origin == "LGA")
```

```r
united_flights_dest <- flights %>%
                  filter(carrier == "UA",
                          origin == "LGA",
                          dest %in% c("ORD", "DEN", "IAH", "CLE"))
```

```r
delay_summary <- united_flights_dest %>%
              summarize(mean_delay = mean(arr_delay, na.rm = TRUE),
                      sd_delay = sd(arr_delay, na.rm = TRUE))
```

```r
knitr::kable(delay_summary)
```

| mean_delay | sd_delay |
|-----------:|---------:|
| 4.642189 | 47.74975 |

```r
delay_summary <- united_flights_dest %>%
              summarize(mean_delay = mean(arr_delay, na.rm = TRUE),
                      sd_delay = sd(arr_delay, na.rm = TRUE),
                      total_obs = n())
```

```r
delay_by_month <- united_flights_dest %>%
                group_by(month) %>%
                summarize(mean_delay = mean(arr_delay, na.rm = TRUE),
                        sd_delay = sd(arr_delay, na.rm = TRUE),
                        total_obs = n())
```

```r
knitr::kable(delay_summary)
```

| mean_delay | sd_delay | total_obs |
|-----------:|---------:|----------:|
| 4.642189 | 47.74975 | 8044 |