Homework #1.1-Example A

A. Student

1/24/2025

```
# Setting Document Options
knitr::opts_chunk$set(
  echo = FALSE,
  warning = FALSE,
  message = FALSE,
  fig.align = "center"
)

# Add additional packages by name to the following list
packages <- c("tidyverse", "knitr", "kableExtra", "psych")
invisible(lapply(
  X = packages,
  FUN = library,
  character.only = TRUE,
  warn.conflicts = FALSE,
  quietly = TRUE
))
```

```
## -- 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 become errors
```

```
oreoData <- read.table(
  file = "https://raw.github.com/neilhatfield/STAT461/master/dataFiles/oreo1.dat",
  header = TRUE,
  sep = ","
)
```

For this assignment, I was assigned the Oreo 1 data set.

Here are the values of statistics:

```
summary(
  oreoData[which(oreoData$Type == "Regular"), "Filling.Mass"],
```

```
  na.rm = TRUE
)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.786   2.936   3.039   3.178   3.227   6.766
```

```
var(oreoData[which(oreoData$Type == "Regular"), "Filling.Mass"])
```

```
## [1] 0.418189
```

```
psych::skew(oreoData[which(oreoData$Type == "Regular"), "Filling.Mass"])
```

```
## [1] 4.715976
```

```
summary(
  oreoData[which(oreoData$Type == "Double Stuf"), "Filling.Mass"],
  na.rm = TRUE
)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.771   5.590   5.846   5.915   6.407   6.935
```

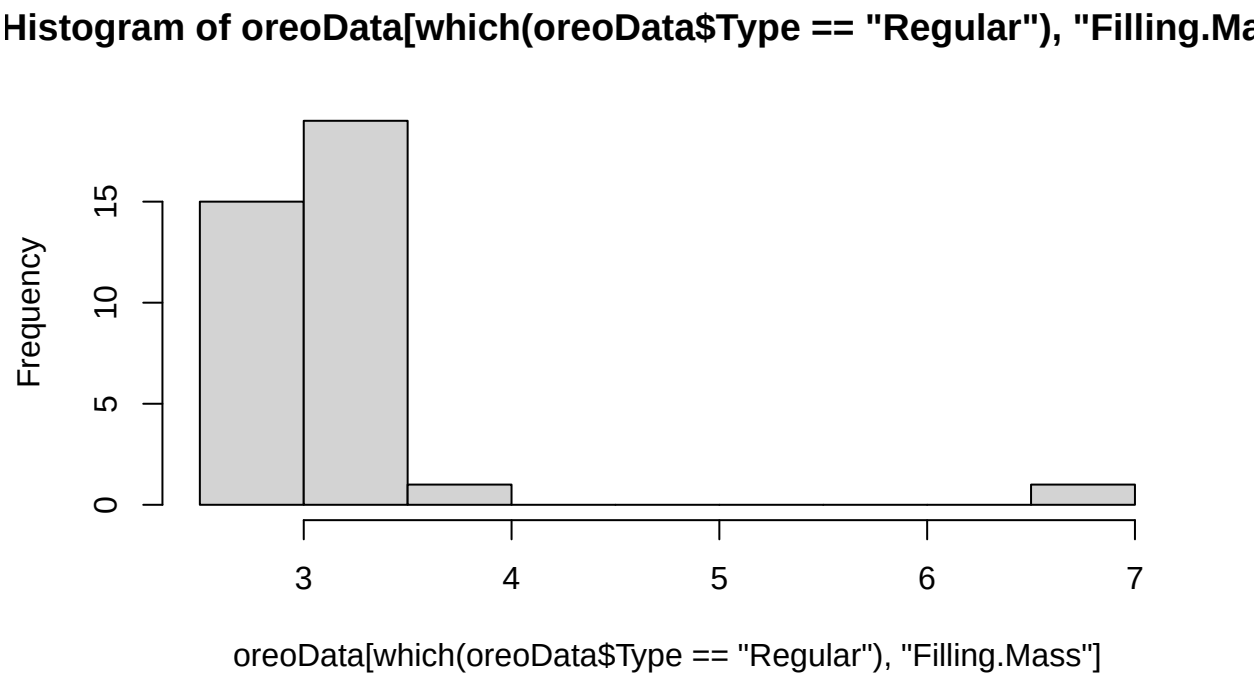
```
var(oreoData[which(oreoData$Type == "Double Stuf"), "Filling.Mass"])
```

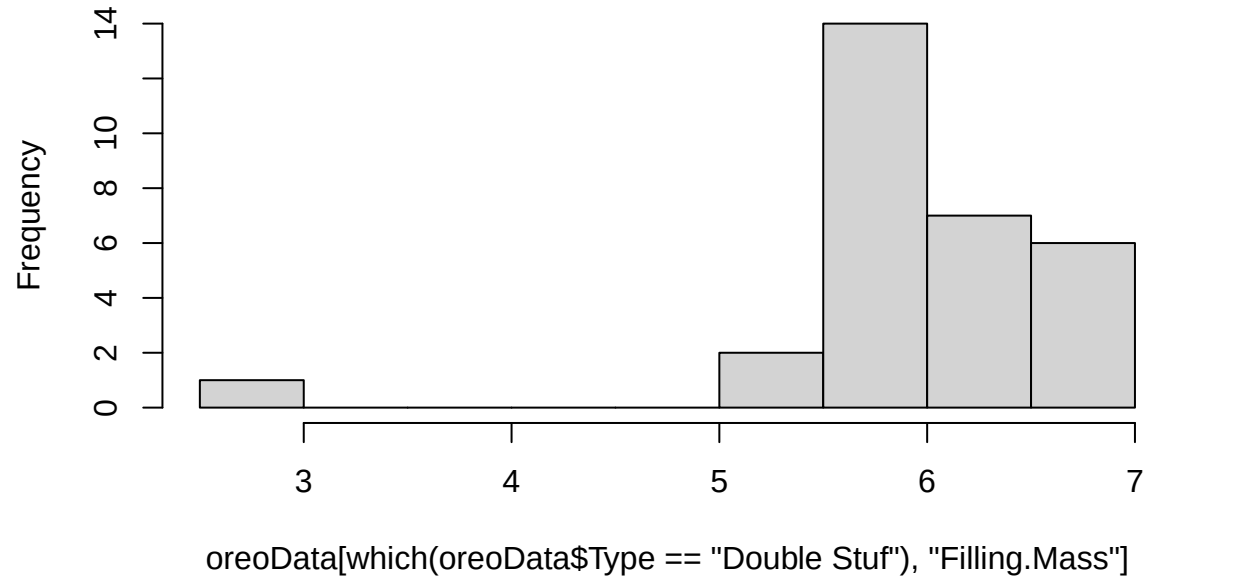
```
## [1] 0.5674834
```

```
psych::skew(oreoData[which(oreoData$Type == "Double Stuf"), "Filling.Mass"])
```

```
## [1] -2.181845
```

The sample min is the smallest value, the sample max is the biggest value. Twenty-five percent of the cookies have 2.936 grams of filling (or 5.59). Half of the cookies have less than 3.04 grams (5.85 grams) of créme filling. The average of regular oreos is 3.178 grams; double stuf are 5.915 grams. Q3 for the oreos are 3.23 and 6.41. The variances are the square of the standard deviation, so they are the squared typical deviation from the average. Skewness tells us that the regular oreos are left skewed while double stuff are right skewed as shown in the histograms.





The question of whether double stuf oreos are actually double the stuff is a two-sample location problem.

The null hypothesis would be $D = 2R$; while the alternative hypothesis would be D not equal 2R. I can treat this as $D - 2R = 0$ vs $D - 2R \neq 0$. This would be a two-sample t test.

```r
t.test(
  x = oreoData[which(oreoData$Type == "Double Stuf"), "Filling.Mass"],
  y = oreoData[which(oreoData$Type == "Regular"), "Filling.Mass"],
  conf.level = 0.97
)
```

```
##
##  Welch Two Sample t-test
##
## data:  oreoData[which(oreoData$Type == "Double Stuf"), "Filling.Mass"] and oreoData[which(oreoData$Ty
## t = 15.661, df = 57.566, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 97 percent confidence interval:
##  2.347738 3.125384
## sample estimates:
## mean of x mean of y
##  5.914533  3.177972
```

My t value is 15.661; degrees of freedom is 57.566. The p-value is really small ($< 2.2e\text{-}16$). This would mean that there is essentially no chance that the null hypothesis is true. The 97% confidence interval is (2.348, 3.125); we're 97% confident that the true difference between double stuf and regular oreos is between 2.348 and 3.125 grams. The mean of regular oreos is 3.178, which is not in this interval; so clearly double stuf can't have double the filling; they have less.