Final Project

John Clark

2024-12-11

```
library(readxl)
library(ggplot2)
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
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v forcats   1.0.0     v stringr   1.5.1
## v lubridate 1.9.3     v tibble    3.2.1
## v purrr     1.0.2     v tidyr     1.3.1
## v readr     2.1.5
```

```
## -- 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
```

```
library(plotly)
```

```
## Warning: package 'plotly' was built under R version 4.4.2
```

```
##
## Attaching package: 'plotly'
##
## The following object is masked from 'package:ggplot2':
##
##     last_plot
##
```

```
## The following object is masked from 'package:stats':
##
##     filter
##
## The following object is masked from 'package:graphics':
##
##     layout
```

```
library(DT)
```

```
## Warning: package 'DT' was built under R version 4.4.2
```

```
library(tm)
```

```
## Warning: package 'tm' was built under R version 4.4.2

## Loading required package: NLP
##
## Attaching package: 'NLP'
##
## The following object is masked from 'package:ggplot2':
##
##     annotate
```

```
library(wordcloud)
```

```
## Warning: package 'wordcloud' was built under R version 4.4.2

## Loading required package: RColorBrewer
```

```
library(readxl)
library(lubridate)
```

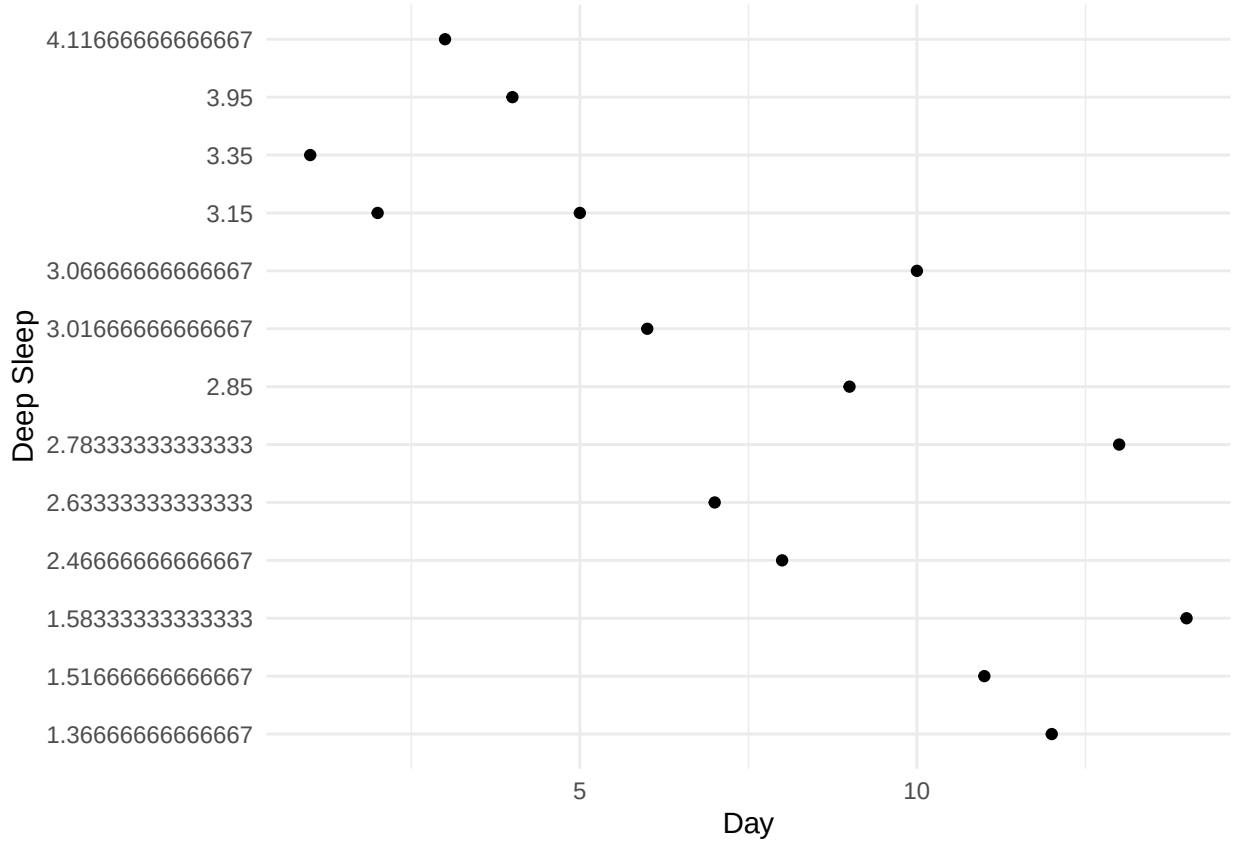
Format the data.

```
df = read_excel("C:/Users/pokej/OneDrive/Data211 Fall 2024/Week Final/Final Project Data.xlsx")

df$Start <- format(df$Start, "%H:%M")
df$Stop <- format(df$Stop, "%H:%M")
df$Awake <- format(df$Awake, "%H:%M")
df$Length <- format(df$Length, "%H:%M")
df$`Deep Sleep` <- format(df$`Deep Sleep`, "%H:%M")
df <- df[-c(15,16),]
df$Day <- as.numeric(df$Day)

my_func <- function(user_str) {
  number <- as.numeric(substr(user_str, 1, 2))
  number
  number <- number + (as.numeric(substr(user_str, 4, 5)) * 5/300)
  return (number)
```

```
}

for (data in 1:length(df$`Deep Sleep`)) {
  df$`Deep Sleep`[data] <-  my_func(df$`Deep Sleep`[data])
}

df$`Deep Sleep`
```

```
##  [1] "3.35"              "3.15"              "4.11666666666667" "3.95"
##  [5] "3.15"              "3.01666666666667" "2.63333333333333" "2.46666666666667"
##  [9] "2.85"              "3.06666666666667" "1.51666666666667" "1.36666666666667"
## [13] "2.78333333333333" "1.58333333333333"
```

```
df <- df %>%
  arrange(Day)

ggplot(df, aes(x = Day, y=`Deep Sleep`)) +
  geom_point() +
  theme_minimal()
```




```
t.test(as.numeric(df$`Deep Sleep`), mu = 2.0, alternative = "greater")
```

```
##
##  One Sample t-test
```

```
##
## data:  as.numeric(df$`Deep Sleep`)
## t = 3.525, df = 13, p-value = 0.001865
## alternative hypothesis: true mean is greater than 2
## 95 percent confidence interval:
##  2.390982      Inf
## sample estimates:
## mean of x
##  2.785714
```