Project

Zachary Nash

2024-12-12

```
library(tidyverse)
```

```
## -- 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 1.9.3     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
```

```
library(tidyr)
library(tinytex)
library(plotly)
```

```
##
## 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)
library(dplyr)
library(tm)
```

```
## Loading required package: NLP
##
## Attaching package: 'NLP'
##
```

```
## The following object is masked from 'package:ggplot2':
##
##     annotate
```

```
library(wordcloud)
```

```
## Loading required package: RColorBrewer
```

```
library(readxl)
```

Reasoning - there are different circumstances in my step count, I might go on a two day study spree where I barely move. In comparison, when i work the conditions are typically consistent

```
steps <- read_excel("C:/Users/zknas/Downloads/StepsWorkVSOff.xlsx")
```

$H_0 : \mu_{host} = \mu_{off}$ $H_a : \mu_{host} > \mu_{off}$

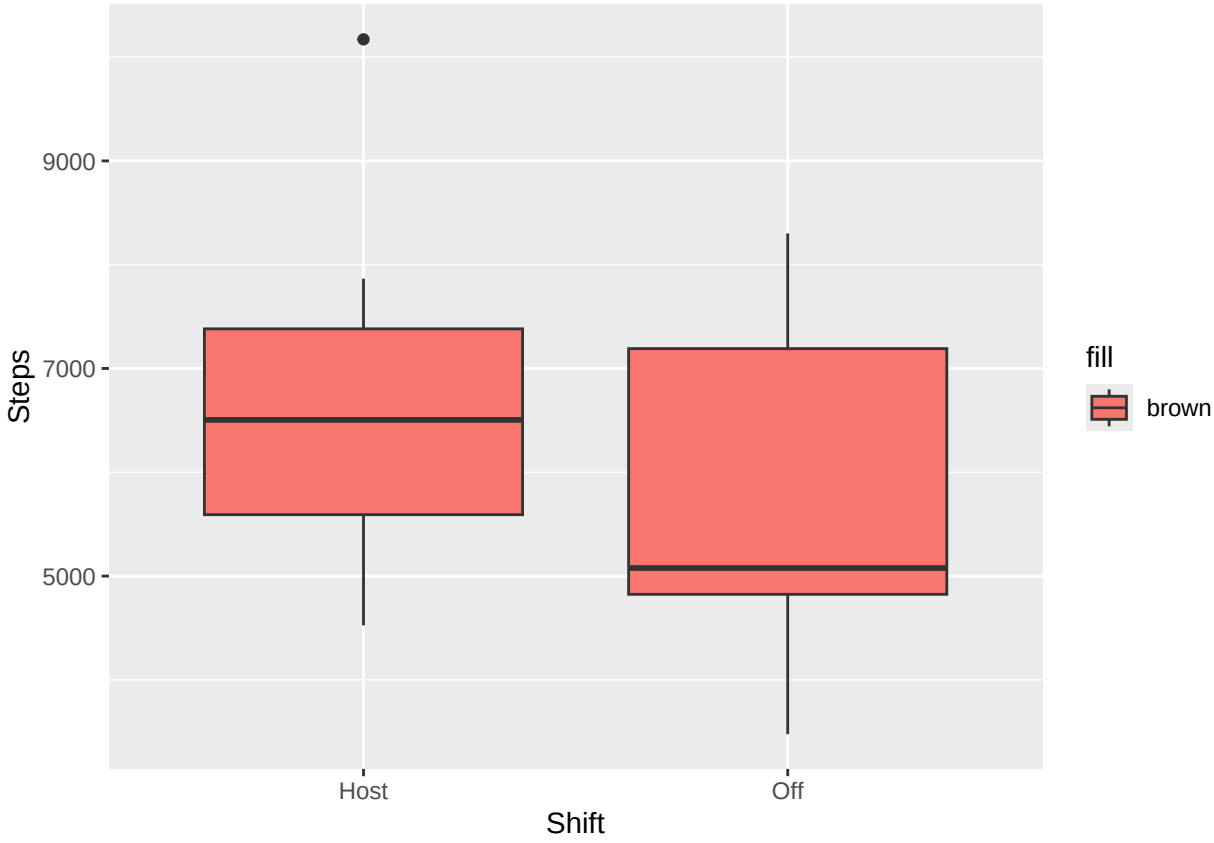
```
stepsBO<-steps%>%
  filter(Shift!="Busser")%>%
  select(Steps,Shift)

t.test(Steps~Shift,data=stepsBO,alternative='greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  Steps by Shift
## t = 1.2016, df = 17.961, p-value = 0.1226
## alternative hypothesis: true difference in means between group Host and group Off is greater than 0
## 95 percent confidence interval:
##  -388.7731       Inf
## sample estimates:
## mean in group Host  mean in group Off
##             6632.5             5755.6
```

Since the p-value ended up being .1226, greater than 0.05, we reject the null hypothesis. There is not significant evidence to back the claim that the mean number of steps taken on days when I work as a restaurant host is greater than the mean number of steps taken on days when I do not work altogether.

```
mapsteps<-stepsBO%>%
  ggplot(aes(x=Shift,y=Steps))+geom_boxplot(aes(fill = 'brown'))
mapsteps
```


9000
7000
5000
Steps
Host
Off
Shift
fill
brown