

Final Report

due November 16, 2021 by 11:59 PM

Shelby Brown, Katie Lam, Kaeden Hill

10/31/21

Load Packages

```
library(tidyverse)

## Warning in system("timedatectl", intern = TRUE): running command 'timedatectl'
## had status 1

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.5      v dplyr  1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.0.2      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(janitor)

##
## Attaching package: 'janitor'

## The following objects are masked from 'package:stats':
##
##   chisq.test, fisher.test

knitr::opts_chunk$set(warning = FALSE, message = FALSE)
```

Load Data

```
heart <- readr::read_csv("heart.csv")
```

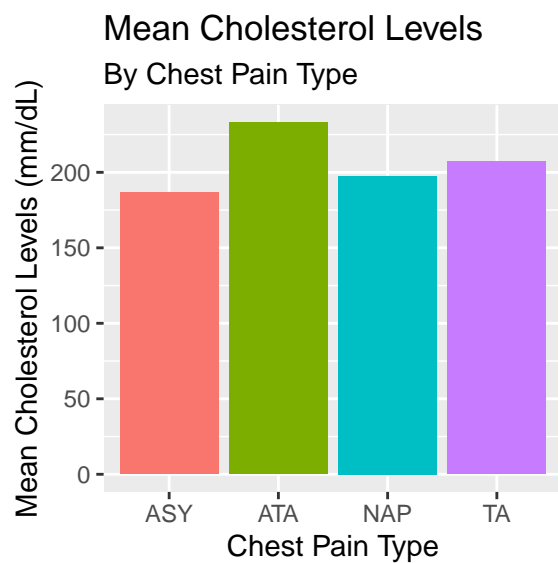
Data Analysis Plan

```
mean_cholesterol <- heart %>%
  group_by(ChestPainType) %>%
  summarize(mean_cholesterol = mean(Cholesterol))%>%
  print()
```

```
## # A tibble: 4 x 2
```

```
## ChestPainType mean_cholesterol
## <chr> <dbl>
## 1 ASY 187.
## 2 ATA 233.
## 3 NAP 197.
## 4 TA 207.

mean_cholesterol %>%
  ggplot()+
  geom_col(mapping = aes(x = ChestPainType, y = mean_cholesterol, fill = ChestPainType), position = "dodge")
  theme(legend.position = "none")+
  labs(title = "Mean Cholesterol Levels",
        subtitle = "By Chest Pain Type",
        x = "Chest Pain Type",
        y = "Mean Cholesterol Levels (mm/dL)")
```



Grouping Variables

```
heart_grouped <- heart %>%
  mutate(chol_level = cut(Cholesterol,
                          breaks = c(-Inf, 120, 200, 239, Inf),
                          labels = c("Low", "Normal", "Intermediate", "High"),
                          right=FALSE))

heart_grouped <- heart_grouped %>%
  mutate(press_level = cut(RestingBP,
                           breaks = c(-Inf, 120, 130, 140, Inf),
                           labels = c("Normal", "Elevated", "Hypertension 1", "Hypertension 2"),
                           right=FALSE))
```

Counting Variables

```
heart_grouped <- heart_grouped %>%
  mutate(Sex, sex_factor=ifelse(Sex=="M", 0,1)) %>%
  mutate(ExerciseAngina, exer_factor=ifelse(ExerciseAngina=="N", 0,1))
```

```
sex_grouped <- heart_grouped %>%
  group_by(ChestPainType, sex_factor) %>%
  summarize(count = n())
```

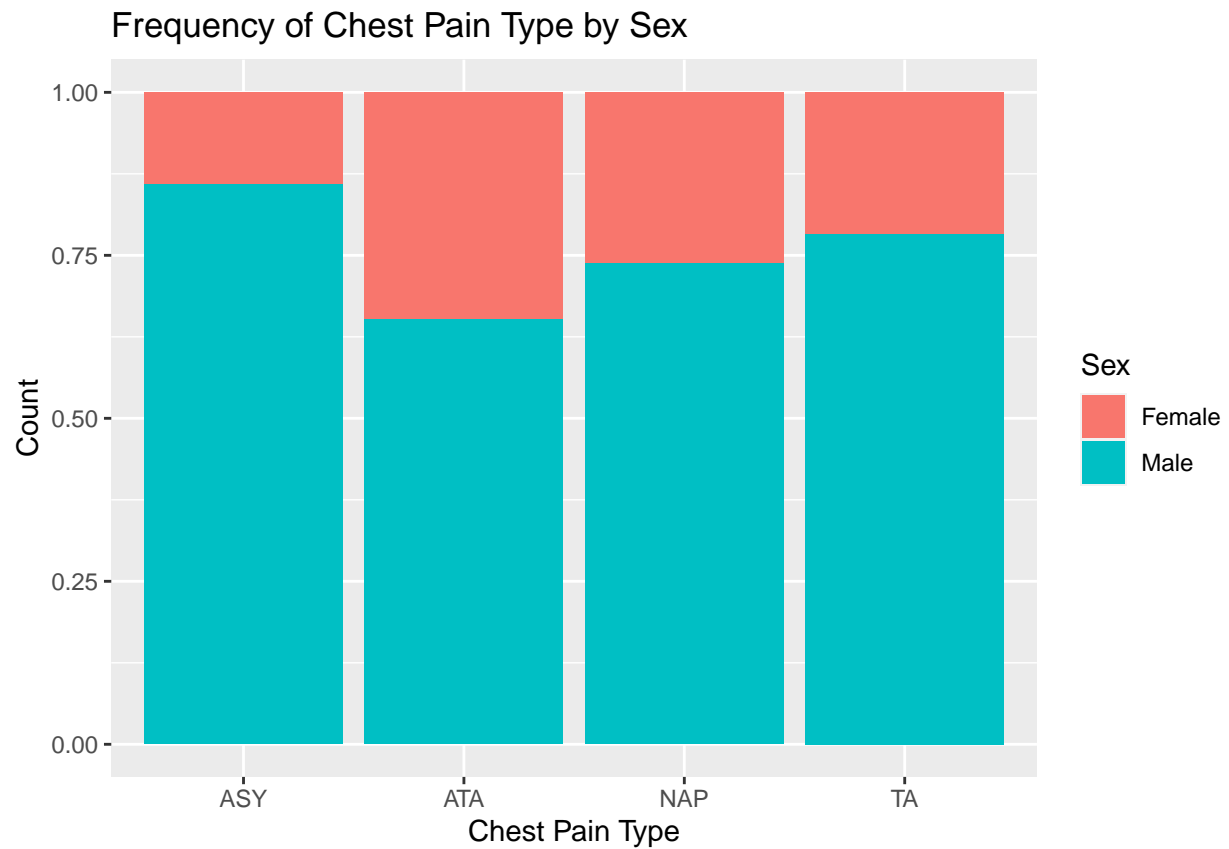
```
exer_grouped <- heart_grouped %>%
  group_by(ChestPainType, exer_factor) %>%
  summarize(count = n())
```

```
RBP_grouped <- heart_grouped %>%
  group_by(ChestPainType, press_level) %>%
  summarize(count = n())
```

```
chol_grouped <- heart_grouped %>%
  group_by(ChestPainType, chol_level)%>%
  summarize(count = n())
```

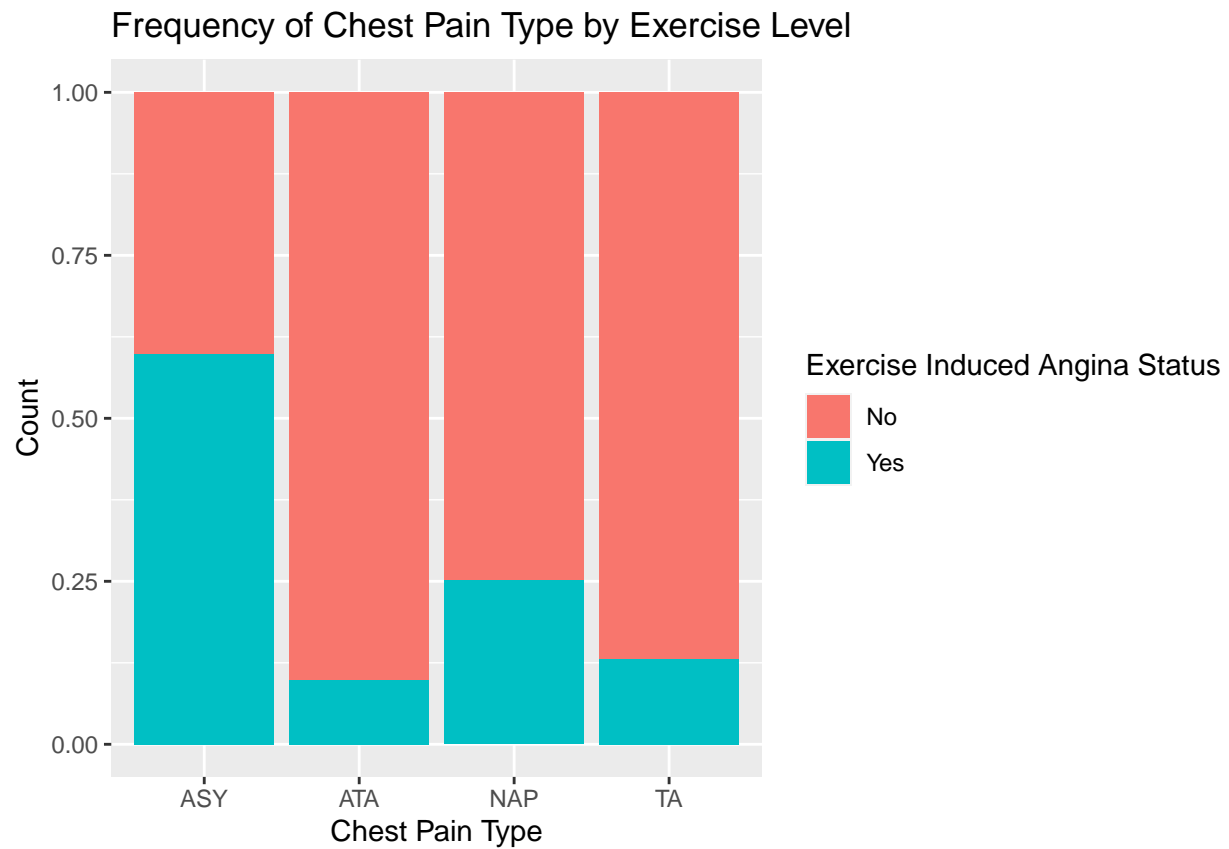
Visualizing Data

```
sex_grouped %>%
  mutate(sex_factor, sex_name=ifelse(sex_factor==0, "Male", "Female")) %>%
  ggplot()+
  geom_col(aes(x = ChestPainType, y = count, fill = as.factor(sex_name)), position = "fill") +
  labs(title = "Frequency of Chest Pain Type by Sex",
       x = "Chest Pain Type",
       y = "Count",
       fill = "Sex")
```

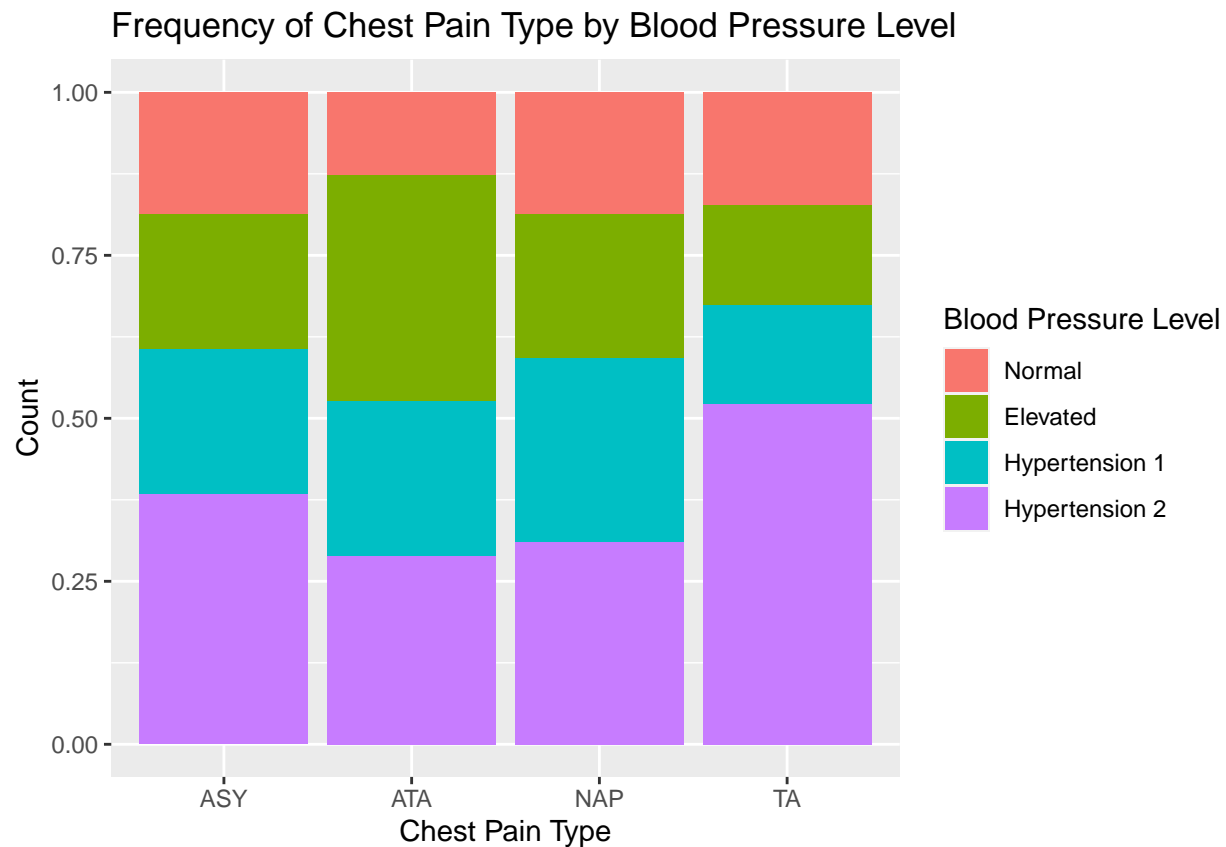


```

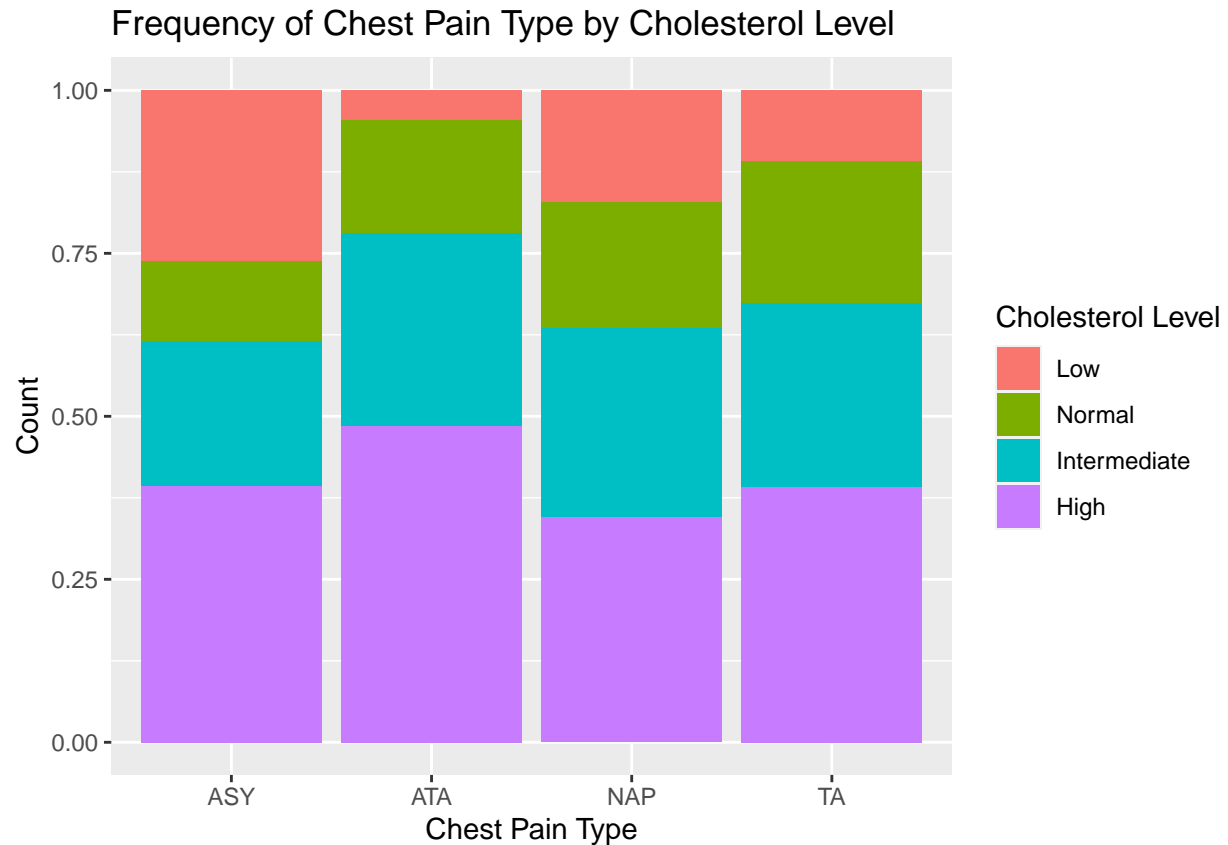
exer_grouped %>%
  mutate(exer_factor, exer_cat=ifelse(exer_factor==0, "No", "Yes")) %>%
  ggplot()+
  geom_col(aes(x = ChestPainType, y = count, fill = as.factor(exer_cat)), position = "fill") +
  labs(title = "Frequency of Chest Pain Type by Exercise Level",
       x = "Chest Pain Type",
       y = "Count",
       fill = "Exercise Induced Angina Status")
  
```



```
RBP_grouped %>%
  ggplot()+
  geom_col(aes(x = ChestPainType, y = count, fill = as.factor(press_level)), position = "fill") +
  labs(title = "Frequency of Chest Pain Type by Blood Pressure Level",
       x = "Chest Pain Type",
       y = "Count",
       fill = "Blood Pressure Level")
```



```
chol_grouped %>%
  ggplot()+
  geom_col(aes(x = ChestPainType, y = count, fill = as.factor(chol_level)), position = "fill") +
  labs(title = "Frequency of Chest Pain Type by Cholesterol Level",
       x = "Chest Pain Type",
       y = "Count",
       fill = "Cholesterol Level")
```



Statistical Tests

```
sex_table <- heart_grouped %>%
  tabyl(Sex, ChestPainType)

exer_table <- heart_grouped %>%
  tabyl(ExerciseAngina, ChestPainType)

RBP_table <- heart_grouped %>%
  tabyl(press_level, ChestPainType)

chol_table <- heart_grouped %>%
  tabyl(chol_level, ChestPainType)
```

```
fisher.test(sex_table)
```

```
##
## Fisher's Exact Test for Count Data
##
## data: sex_table
## p-value = 5.982e-08
## alternative hypothesis: two.sided
```

```
fisher.test(exer_table)
```

```
##
```

```

## Fisher's Exact Test for Count Data
##
## data:  exer_table
## p-value < 2.2e-16
## alternative hypothesis: two.sided
chisq.test(chol_table)

##
## Pearson's Chi-squared test
##
## data:  chol_table
## X-squared = 49.409, df = 9, p-value = 1.391e-07
chisq.test(RBP_table)

##
## Pearson's Chi-squared test
##
## data:  RBP_table
## X-squared = 26.829, df = 9, p-value = 0.001493
heart_grouped2 <- heart_grouped %>%
  filter(ChestPainType %in% c("ATA", "ASY"))

chol_step <- heart_grouped2 %>%
  tabyl(chol_level, ChestPainType)
chisq.test(chol_step)

##
## Pearson's Chi-squared test
##
## data:  chol_step
## X-squared = 36.837, df = 3, p-value = 4.981e-08
heart_grouped3 <- heart_grouped %>%
  filter(ChestPainType %in% c("ASY", "NAP"))

chol_step2 <- heart_grouped3 %>%
  tabyl(chol_level, ChestPainType)
chisq.test(chol_step2)

##
## Pearson's Chi-squared test
##
## data:  chol_step2
## X-squared = 13.433, df = 3, p-value = 0.003788
heart_grouped4 <- heart_grouped %>%
  filter(ChestPainType %in% c("TA", "ASY"))

chol_step3 <- heart_grouped4 %>% ##NOT SIGNIFICANT
  tabyl(chol_level, ChestPainType)
chisq.test(chol_step3)

##
## Pearson's Chi-squared test
##

```



```

## data: chol_step3
## X-squared = 7.5281, df = 3, p-value = 0.05684
heart_grouped5 <- heart_grouped %>%
  filter(ChestPainType %in% c("ATA", "NAP"))

chol_step4 <- heart_grouped5 %>%
  tabyl(chol_level, ChestPainType)
chisq.test(chol_step4)

##
## Pearson's Chi-squared test
##
## data: chol_step4
## X-squared = 17.701, df = 3, p-value = 0.0005069
heart_grouped6 <- heart_grouped %>%
  filter(ChestPainType %in% c("ATA", "TA"))

chol_step5 <- heart_grouped6 %>% #NOT SIGNIFICANT; used fisher bc count was <10 for low:TA
  tabyl(chol_level, ChestPainType)
fisher.test(chol_step5)

##
## Fisher's Exact Test for Count Data
##
## data: chol_step5
## p-value = 0.3106
## alternative hypothesis: two.sided
heart_grouped7 <- heart_grouped %>%
  filter(ChestPainType %in% c("NAP", "TA"))

chol_step6 <- heart_grouped7 %>% #NOT SIGNIFICANT
  tabyl(chol_level, ChestPainType)
chisq.test(chol_step6)

##
## Pearson's Chi-squared test
##
## data: chol_step6
## X-squared = 1.3071, df = 3, p-value = 0.7274

```