Final Report

due November 16, 2021 by 11:59 PM

Shelby Brown, Katie Lam, Kaeden Hill

10/31/21

```
## 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()
```

```
##
## Attaching package: 'janitor'
```

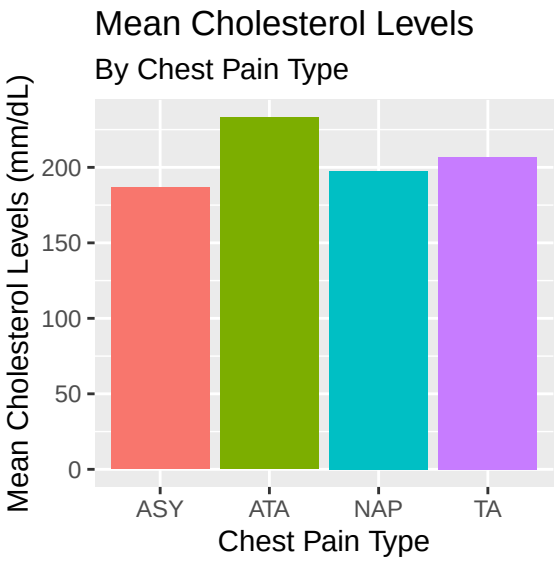
```
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

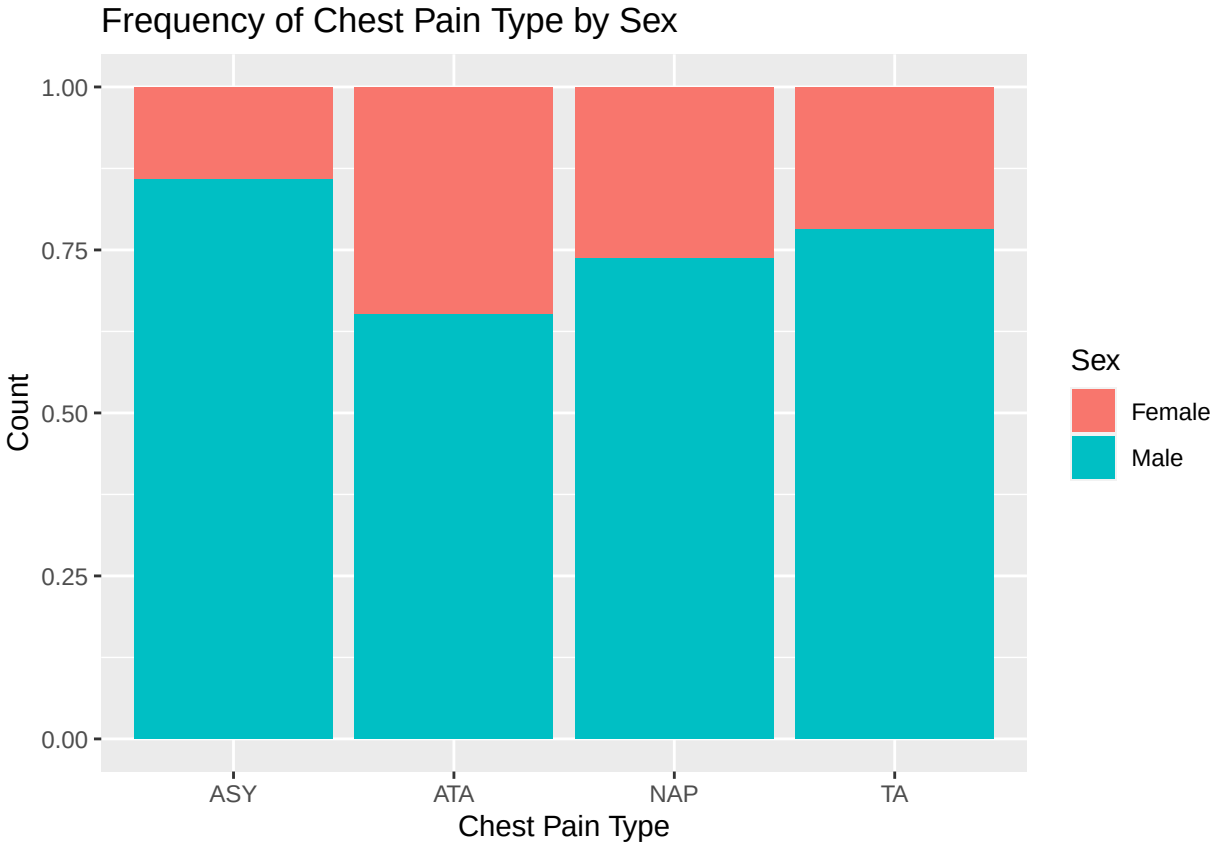
Abstract

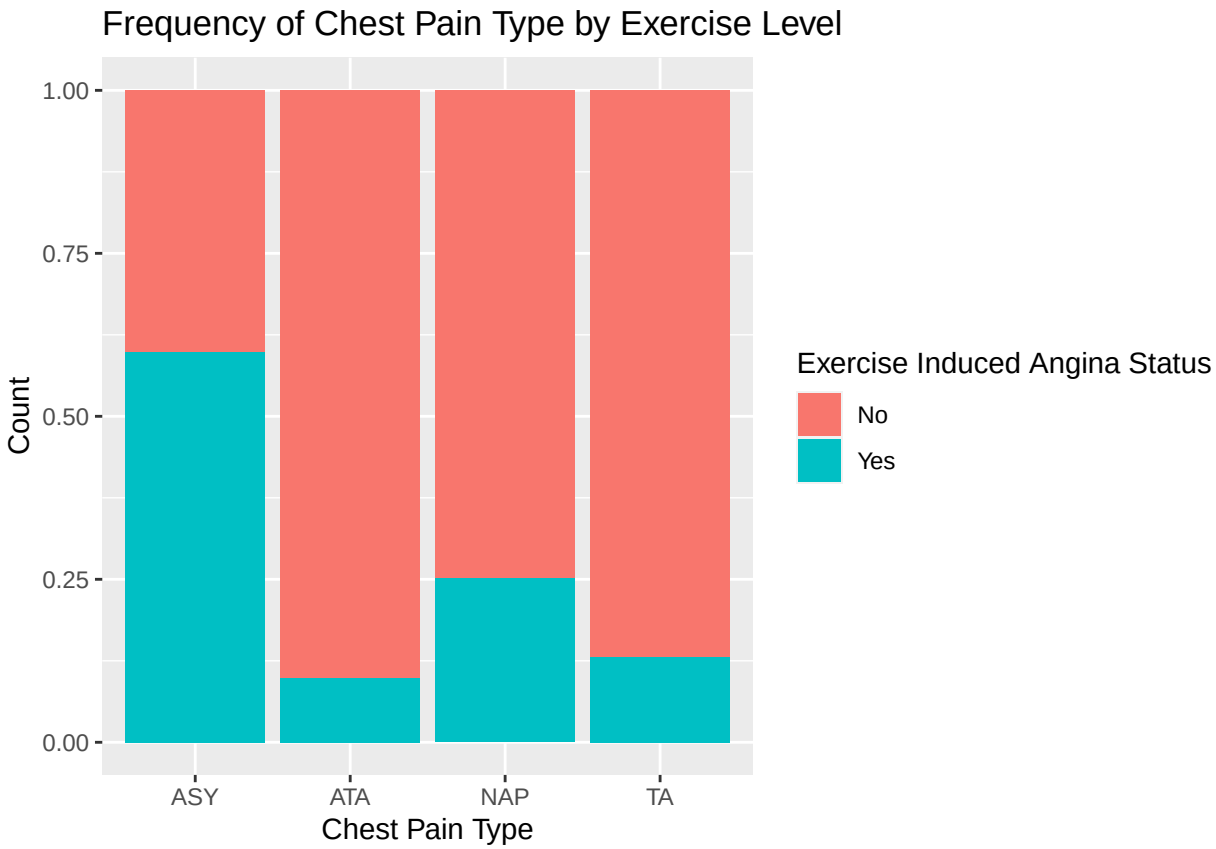
Introduction

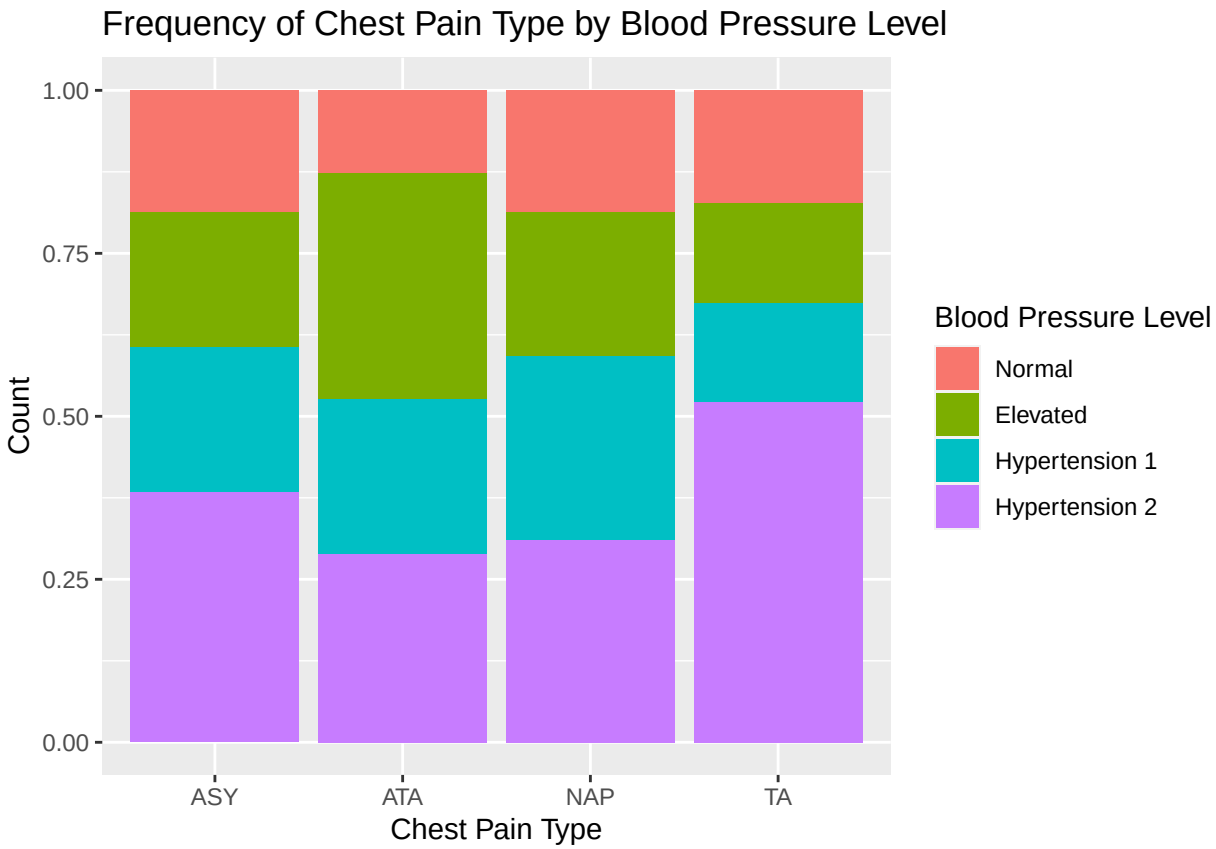
For this data visualization, chest pain types and their relation to other physiological factors were analyzed to search for an association between factors, such as blood pressure, cholesterol, and exercise, and the type of chest pain a patient experiences. The dataset being analyzed is the Heart Failure Prediction Dataset, retrieved from Kaggle, and compiled from five sets with common variables. These sources are the Hungarian Institute of Cardiology. Budapest, University Hospital, Zurich, Switzerland, University Hospital, Basel, Switzerland, the V.A. Medical Center, Long Beach, and the Cleveland Clinic Foundation. The variables of interest include gender, resting blood pressure (mm Hg), serum cholesterol (mm/d), whether or not the angina was exercise induced, and whether or not the patient was diagnosed with heart disease (**diagnosis before or after the angina??). To allow for the use of statistical tests, each variable was transformed into a categorical variable.

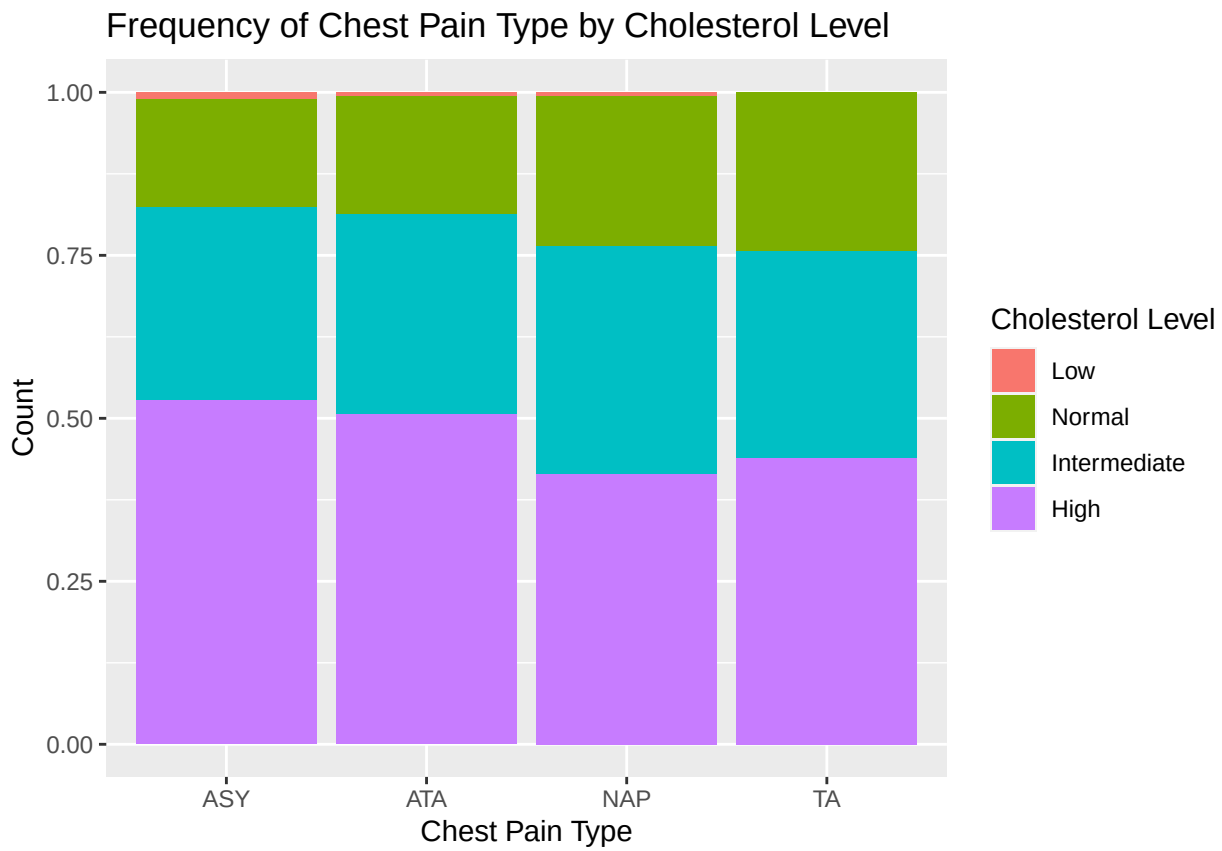
```
## # A tibble: 4 x 2
##   ChestPainType mean_cholesterol
##   <chr>                    <dbl>
## 1 ASY                       187.
## 2 ATA                       233.
## 3 NAP                       197.
## 4 TA                        207.
```


Mean Cholesterol Levels
By Chest Pain Type
Mean Cholesterol Levels (mm/dL)
200
150
100
50
0
ASY
ATA
NAP
TA
Chest Pain Type


Frequency of Chest Pain Type by Sex
1.00
0.75
0.50
0.25
0.00
Count
ASY
ATA
NAP
TA
Chest Pain Type
Sex
Female
Male


Frequency of Chest Pain Type by Exercise Level
1.00
0.75
0.50
0.25
0.00
Count
ASY
ATA
NAP
TA
Chest Pain Type
Exercise Induced Angina Status
No
Yes


Frequency of Chest Pain Type by Blood Pressure Level
1.00
0.75
0.50
0.25
0.00
Count
ASY
ATA
NAP
TA
Chest Pain Type
Blood Pressure Level
Normal
Elevated
Hypertension 1
Hypertension 2



Statistical Tests

```
##
##  Pearson's Chi-squared test
##
## data:  sex_table
## X-squared = 36.879, df = 3, p-value = 4.88e-08
```

```
##
##  Pearson's Chi-squared test
##
## data:  exer_table
## X-squared = 179.27, df = 3, p-value < 2.2e-16
```

```
##
##  Pearson's Chi-squared test
##
## data:  chol_table
## X-squared = 8.6117, df = 9, p-value = 0.4739
```

```
##
##  Pearson's Chi-squared test
##
## data:  RBP_table
## X-squared = 26.829, df = 9, p-value = 0.001493
```

```
##
```

```
##  Pearson's Chi-squared test
##
## data:  disease_table
## X-squared = 268.07, df = 3, p-value < 2.2e-16

##
##  Fisher's Exact Test for Count Data
##
## data:  chol_step
## p-value = 0.9297
## alternative hypothesis: two.sided

##
##  Fisher's Exact Test for Count Data
##
## data:  chol_step2
## p-value = 0.06204
## alternative hypothesis: two.sided

##
##  Fisher's Exact Test for Count Data
##
## data:  chol_step3
## p-value = 0.5284
## alternative hypothesis: two.sided

##
##  Fisher's Exact Test for Count Data
##
## data:  chol_step4
## p-value = 0.3368
## alternative hypothesis: two.sided

##
##  Fisher's Exact Test for Count Data
##
## data:  chol_step5
## p-value = 0.6594
## alternative hypothesis: two.sided

##
##  Fisher's Exact Test for Count Data
##
## data:  chol_step6
## p-value = 0.94
## alternative hypothesis: two.sided

##
##  Fisher's Exact Test for Count Data
##
## data:  sex_step
## p-value = 2.369e-08
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.2031800 0.4733313
## sample estimates:
## odds ratio
```

```
##  0.3100921
##
##  Fisher's Exact Test for Count Data
##
## data:  sex_step2
## p-value = 0.0002766
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.3054468 0.7119609
## sample estimates:
## odds ratio
##  0.4656136
##
##  Fisher's Exact Test for Count Data
##
## data:  sex_step3
## p-value = 0.1896
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.2722762 1.4001345
## sample estimates:
## odds ratio
##  0.5922146
##
##  Fisher's Exact Test for Count Data
##
## data:  sex_step4
## p-value = 0.07251
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.9418551 2.3984968
## sample estimates:
## odds ratio
##   1.501076
##
##  Fisher's Exact Test for Count Data
##
## data:  sex_step5
## p-value = 0.1107
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.8524745 4.6148967
## sample estimates:
## odds ratio
##   1.906248
##
##  Fisher's Exact Test for Count Data
##
## data:  sex_step6
## p-value = 0.7073
## alternative hypothesis: true odds ratio is not equal to 1
```

```
## 95 percent confidence interval:
##  0.5679163 3.0759273
## sample estimates:
## odds ratio
##   1.270772

##
##  Fisher's Exact Test for Count Data
##
## data:  exer_step
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.04032086 0.12583163
## sample estimates:
## odds ratio
## 0.07331366

##
##  Fisher's Exact Test for Count Data
##
## data:  exer_step2
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.1529646 0.3280057
## sample estimates:
## odds ratio
##  0.2253116

##
##  Fisher's Exact Test for Count Data
##
## data:  exer_step3
## p-value = 4.234e-10
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.03430591 0.24552456
## sample estimates:
## odds ratio
##   0.100903

##
##  Fisher's Exact Test for Count Data
##
## data:  exer_step4
## p-value = 0.0001411
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.656297 5.935882
## sample estimates:
## odds ratio
##   3.070166

##
##  Fisher's Exact Test for Count Data
```

```
##
## data:  exer_step5
## p-value = 0.5883
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.4161352 3.9597702
## sample estimates:
## odds ratio
##   1.374316
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  exer_step6
## p-value = 0.0836
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.1466831 1.1500616
## sample estimates:
## odds ratio
##  0.4483045
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  RBP_step
## p-value = 0.001125
## alternative hypothesis: two.sided
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  RBP_step2
## p-value = 0.2331
## alternative hypothesis: two.sided
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  RBP_step3
## p-value = 0.34
## alternative hypothesis: two.sided
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  RBP_step4
## p-value = 0.0436
## alternative hypothesis: two.sided
```

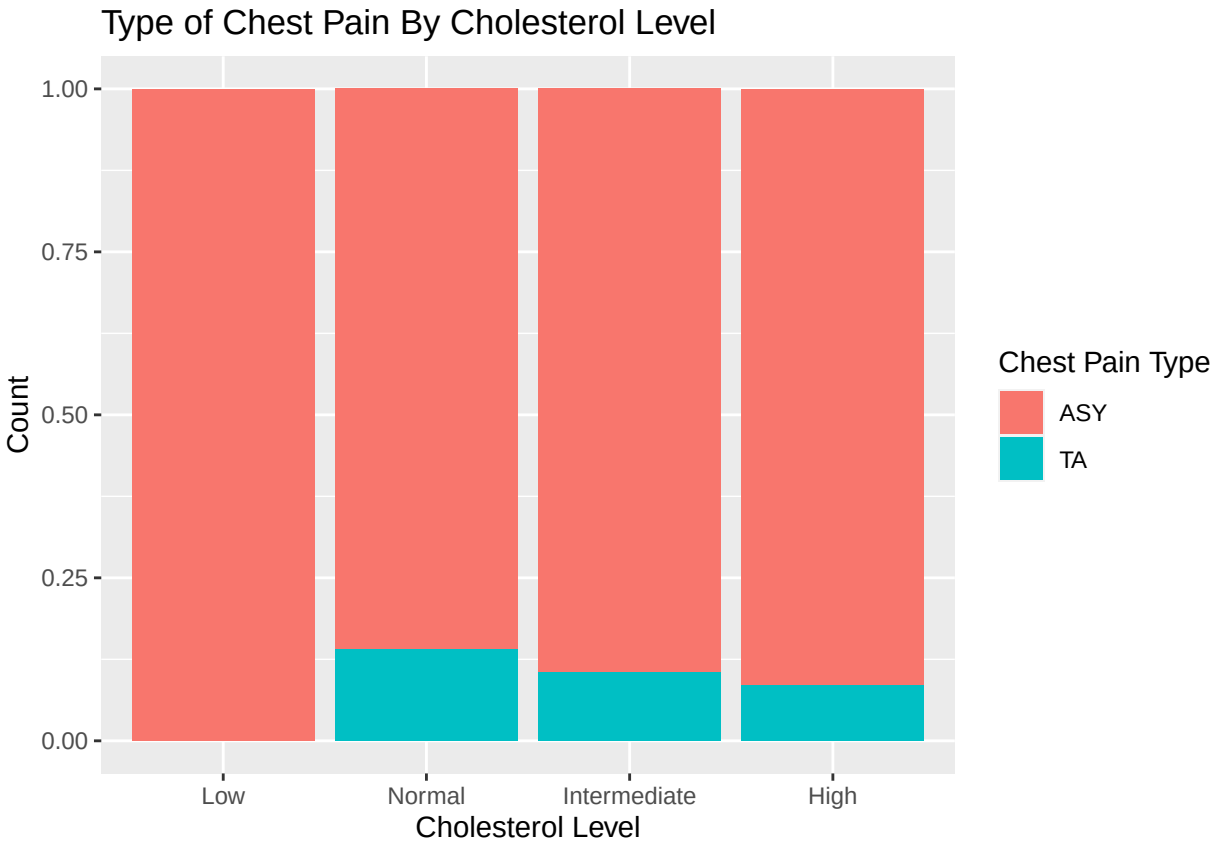
```
##
##  Fisher's Exact Test for Count Data
##
## data:  RBP_step5
## p-value = 0.006839
## alternative hypothesis: two.sided
```

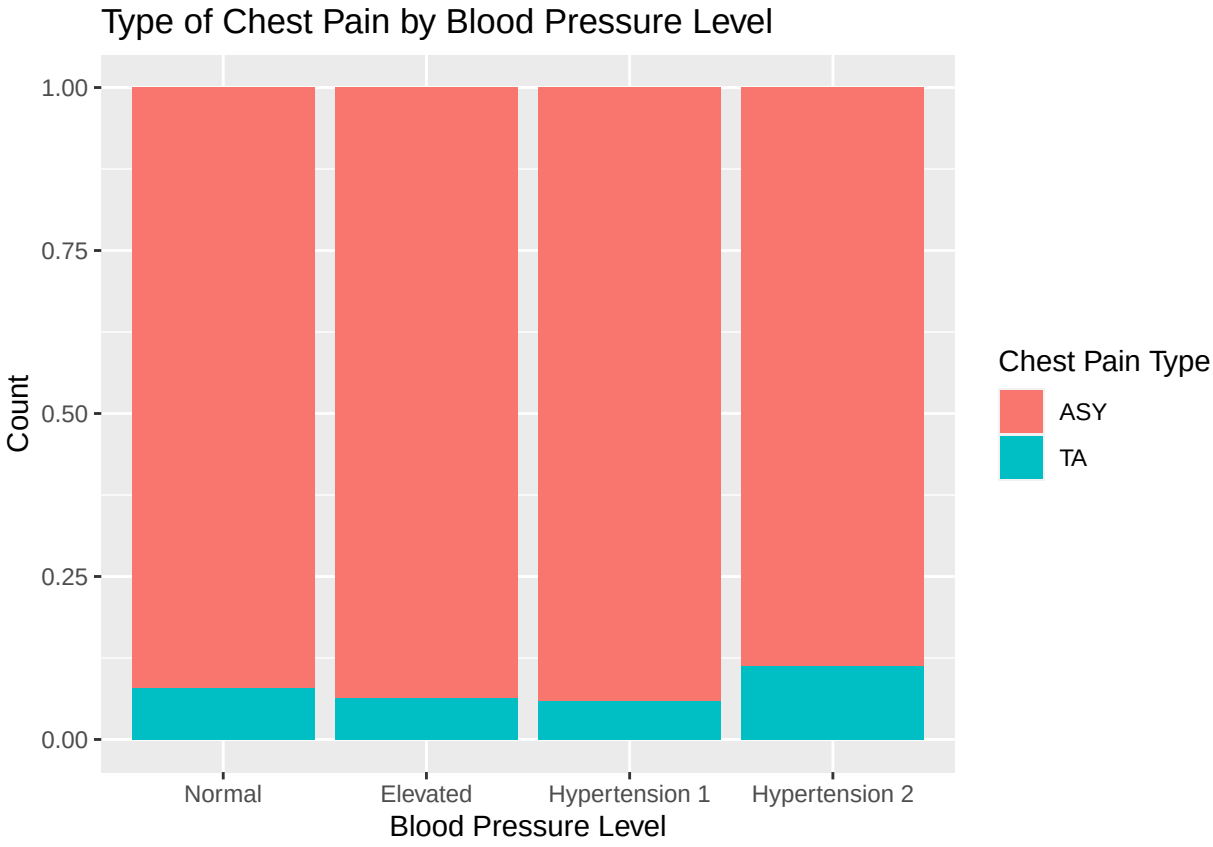
```
##
##  Fisher's Exact Test for Count Data
##
## data:  RBP_step6
## p-value = 0.05117
## alternative hypothesis: two.sided
##
##  Fisher's Exact Test for Count Data
##
## data:  disease_step
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.02532899 0.07052378
## sample estimates:
## odds ratio
## 0.04297732
##
##  Fisher's Exact Test for Count Data
##
## data:  disease_step2
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.1001099 0.2122784
## sample estimates:
## odds ratio
##  0.1463117
##
##  Fisher's Exact Test for Count Data
##
## data:  disease_step3
## p-value = 6.656e-07
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.1039251 0.3983253
## sample estimates:
## odds ratio
##   0.204859
##
##  Fisher's Exact Test for Count Data
##
## data:  disease_step4
## p-value = 1.618e-06
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.984092 5.990659
## sample estimates:
## odds ratio
##   3.401259
##
```

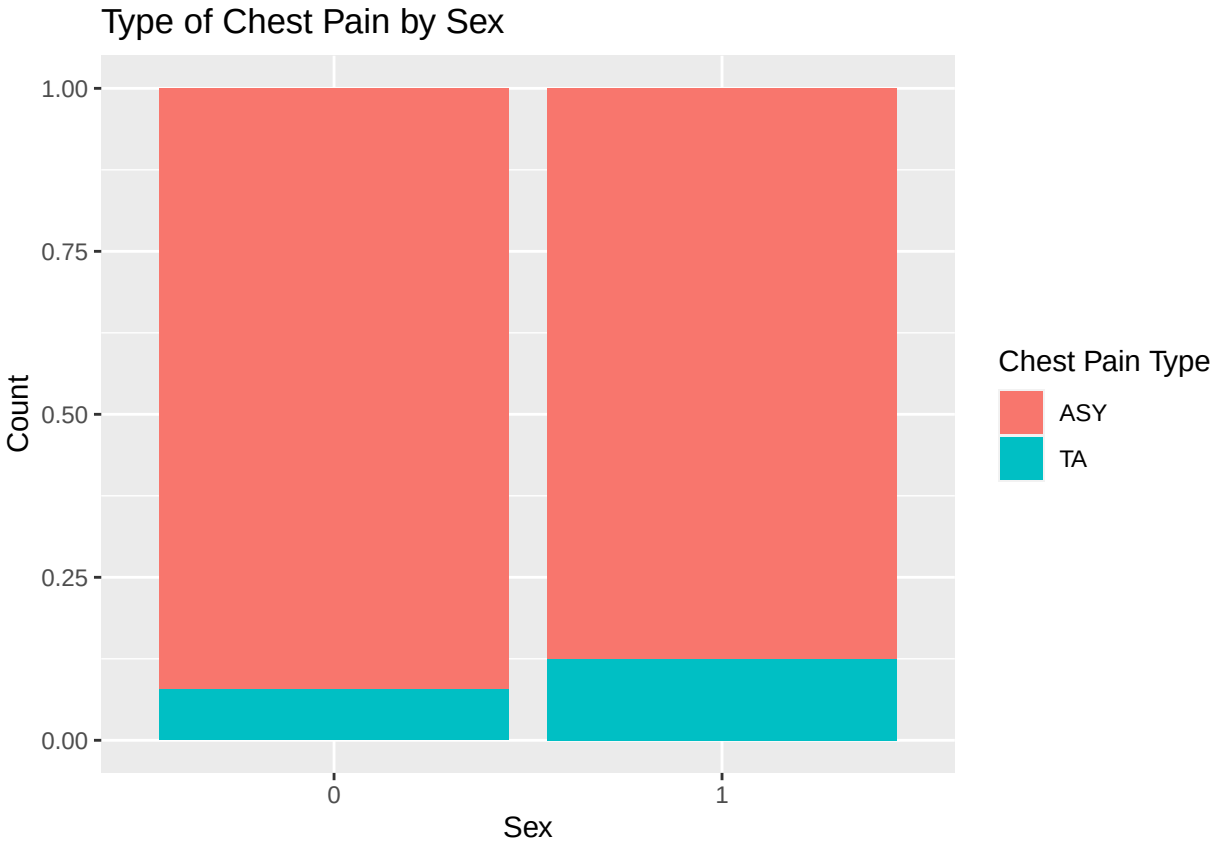
```
##  Fisher's Exact Test for Count Data
##
## data:  disease_step5
## p-value = 3.929e-05
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##   2.15486 10.44700
## sample estimates:
## odds ratio
##    4.73165
```

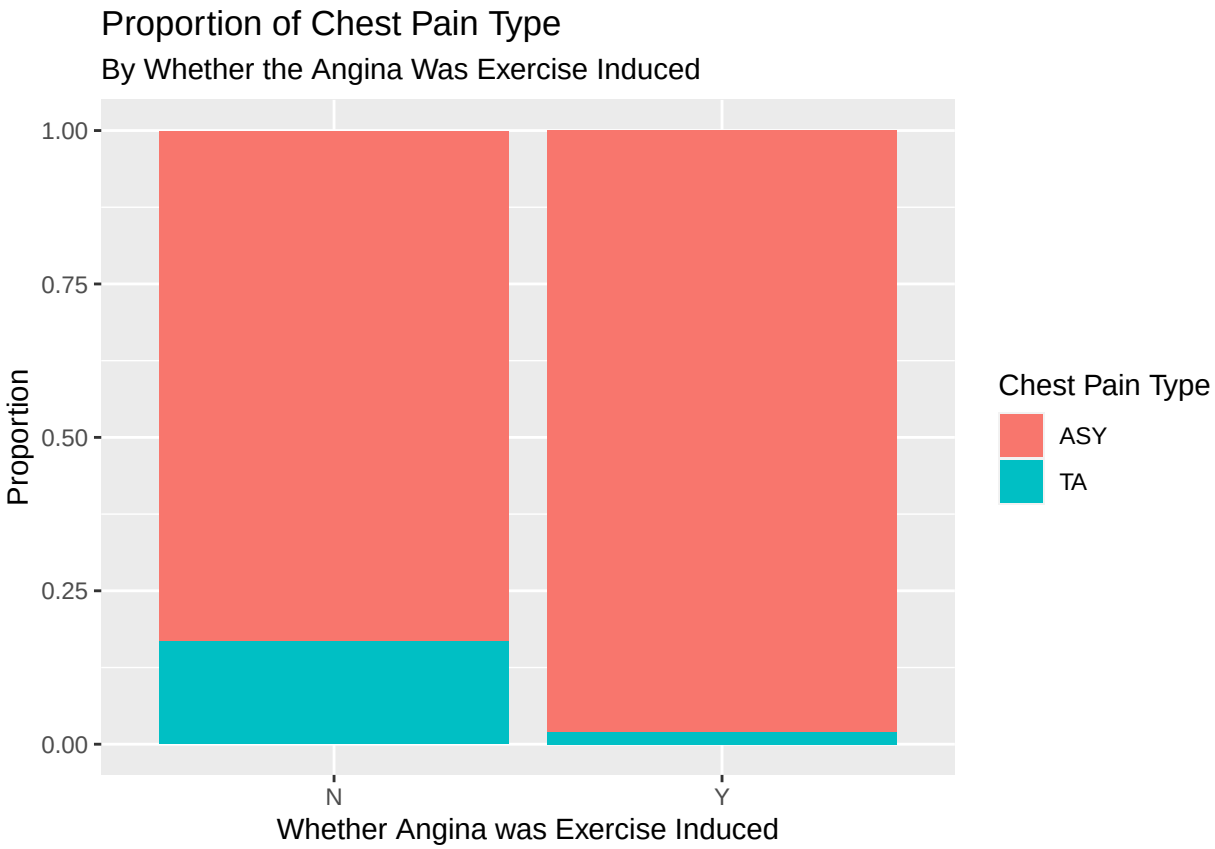
```
##
##  Fisher's Exact Test for Count Data
##
## data:  disease_step
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.02532899 0.07052378
## sample estimates:
## odds ratio
## 0.04297732
```

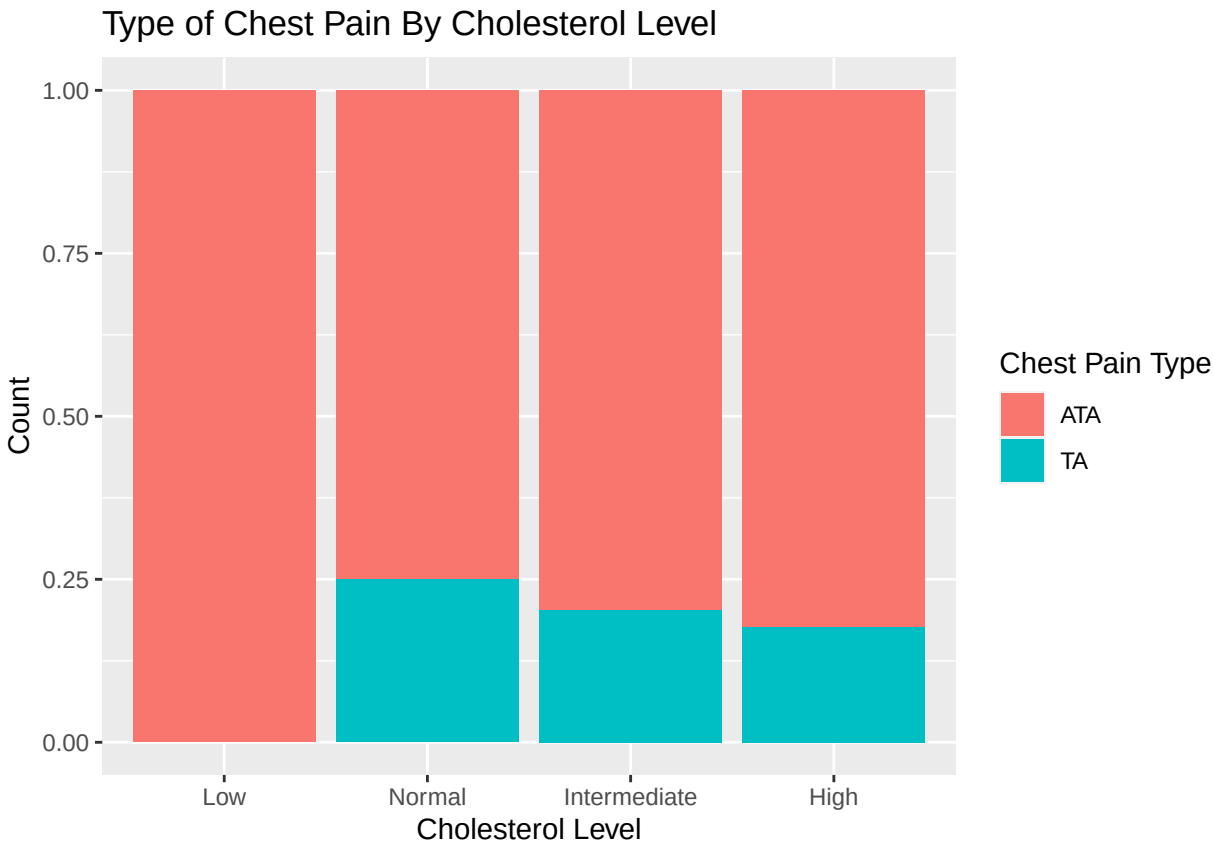
Results

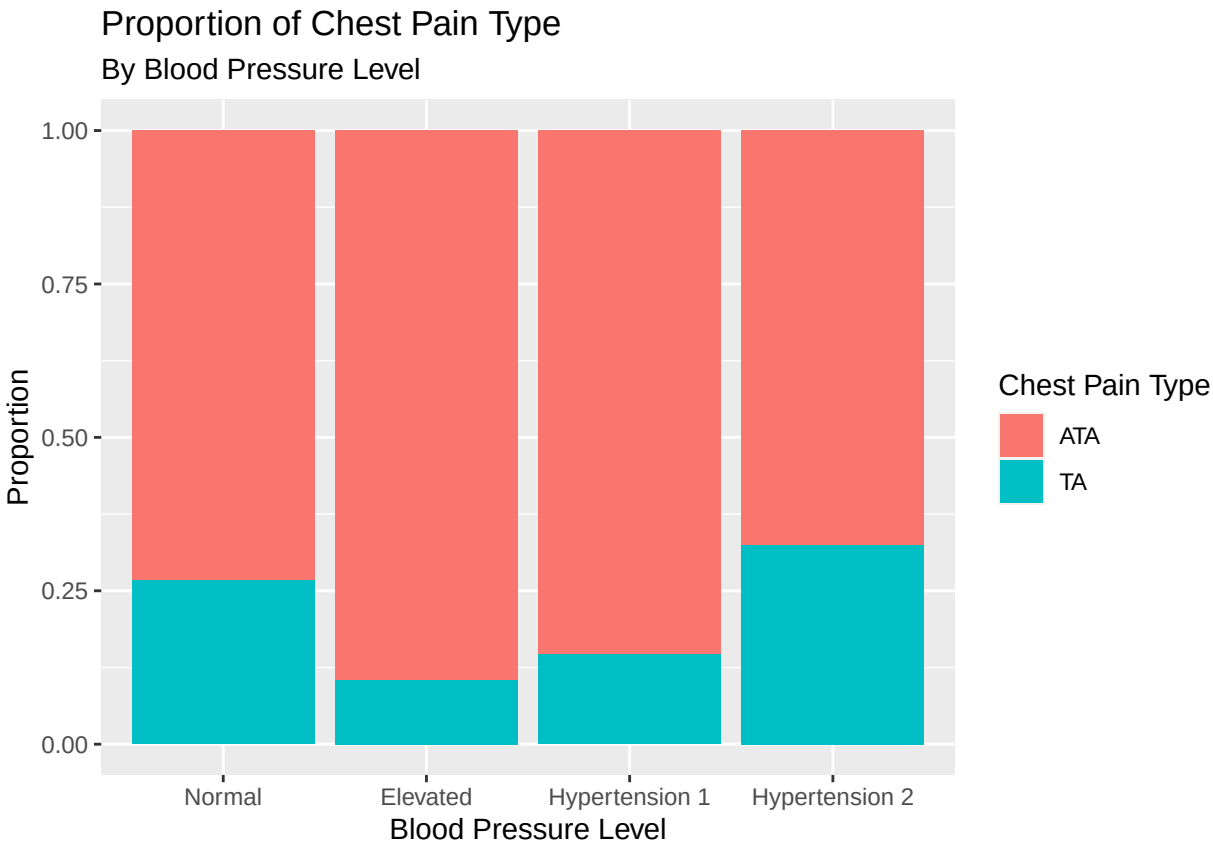

Type of Chest Pain By Cholesterol Level
1.00
0.75
0.50
0.25
0.00
Count
Low
Normal
Intermediate
High
Cholesterol Level
Chest Pain Type
ASY
TA

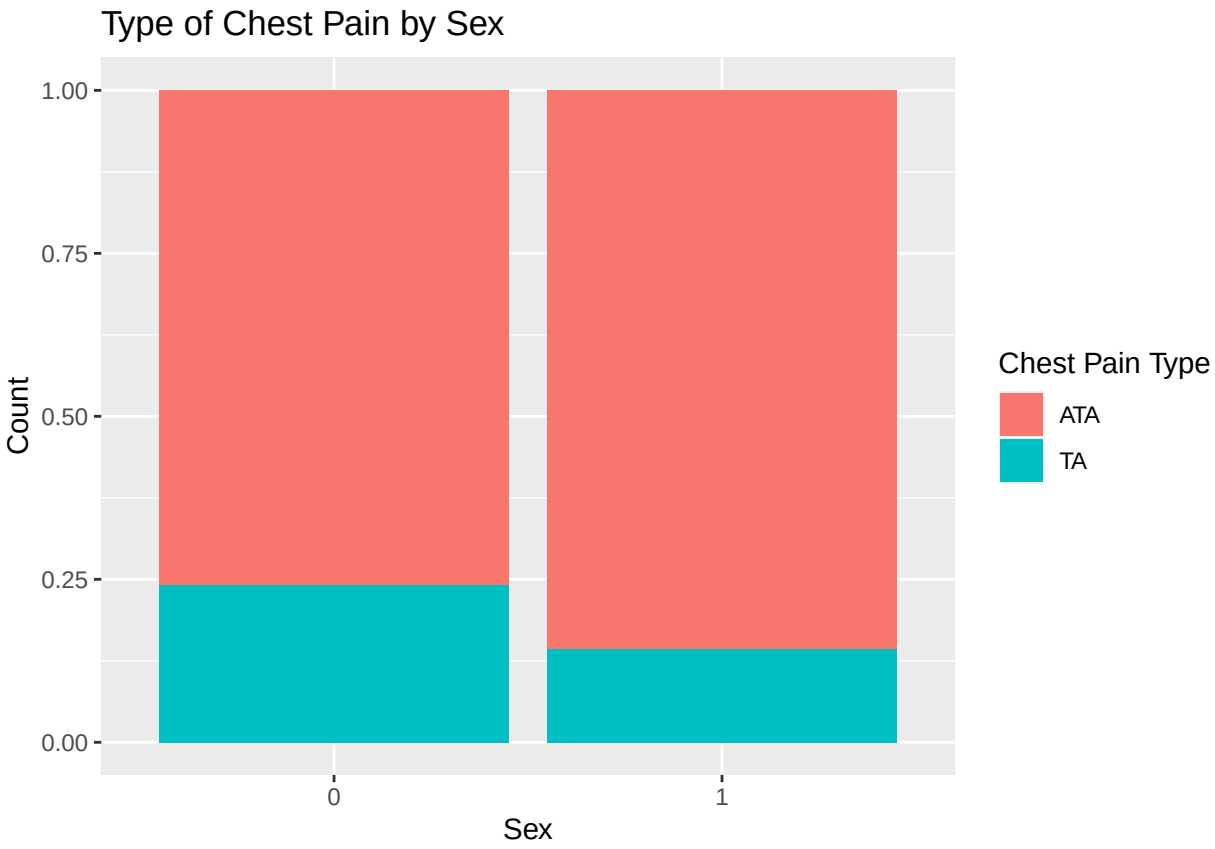

Type of Chest Pain by Blood Pressure Level
1.00
0.75
0.50
0.25
0.00
Count
Normal
Elevated
Hypertension 1
Hypertension 2
Blood Pressure Level
Chest Pain Type
ASY
TA

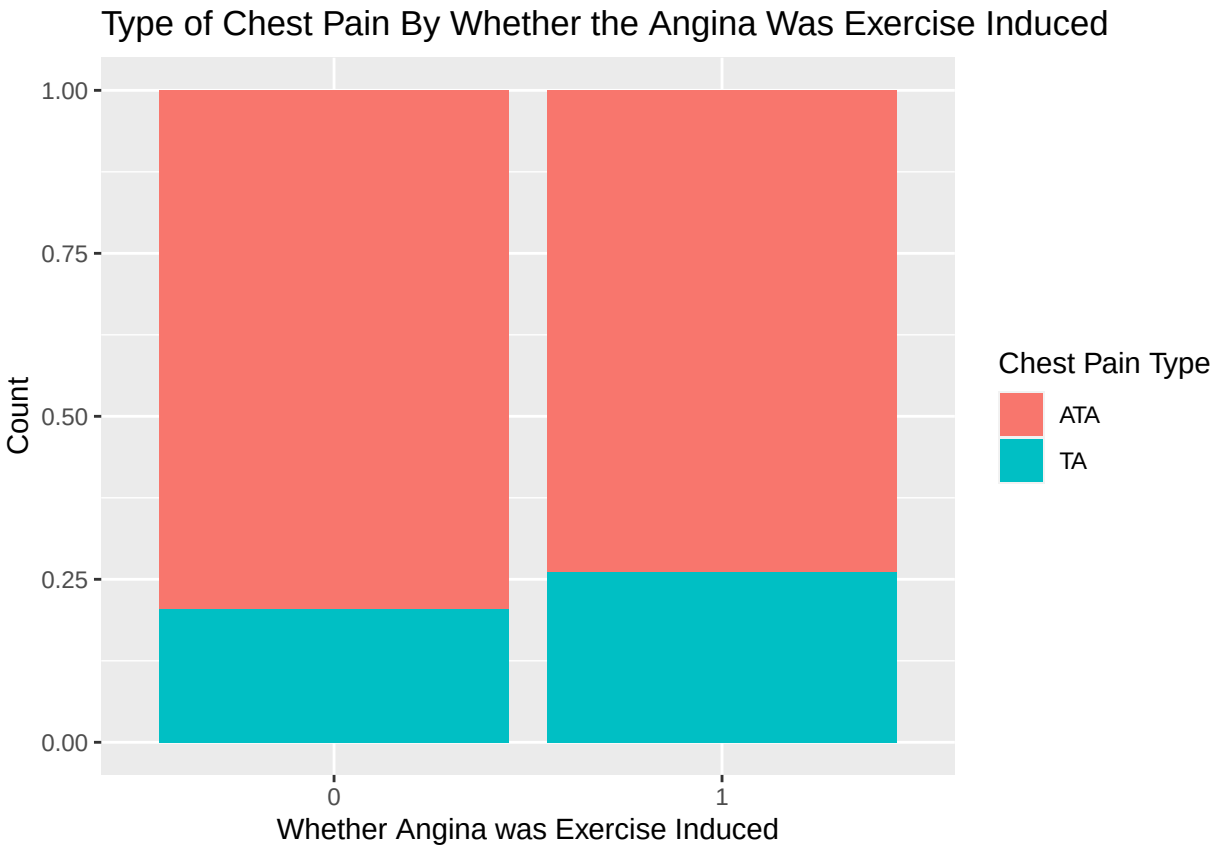

Type of Chest Pain by Sex
1.00
0.75
0.50
0.25
0.00
Count
0
1
Sex
Chest Pain Type
ASY
TA

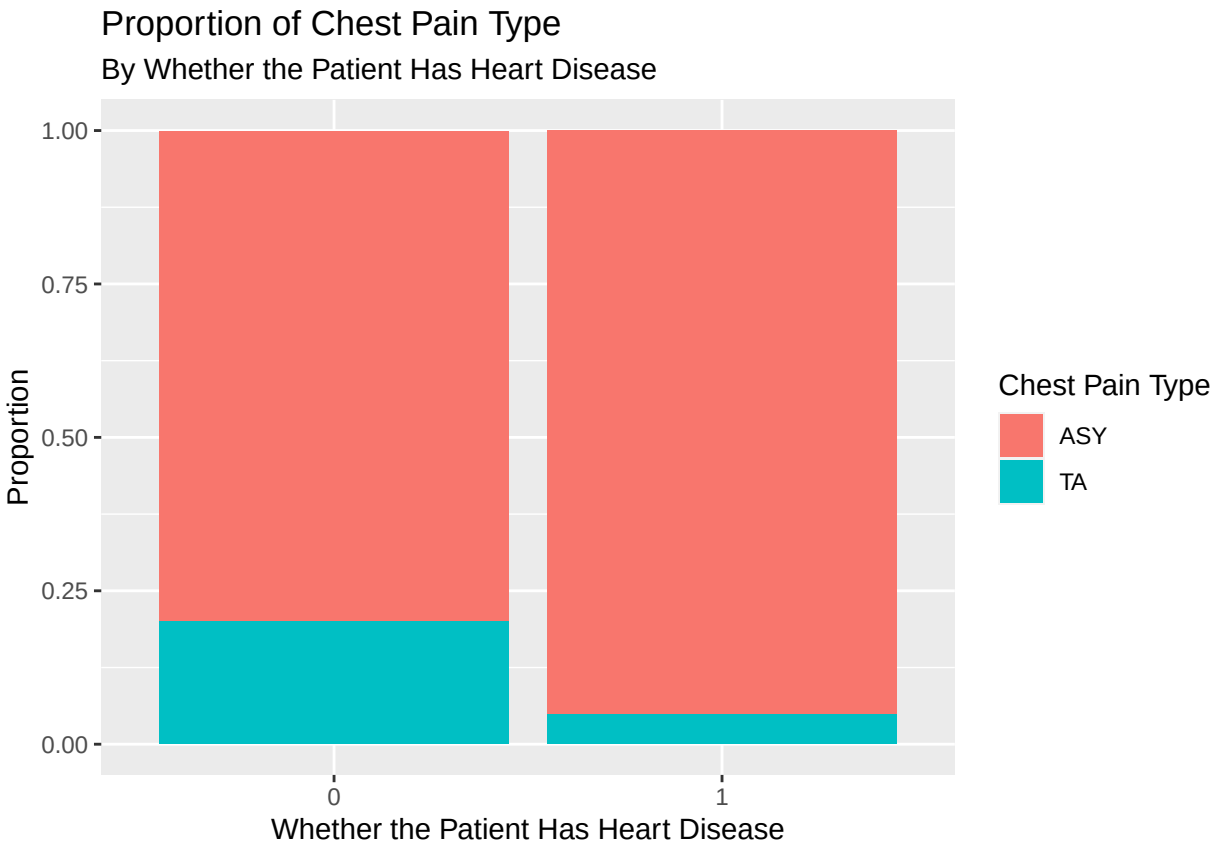

Proportion of Chest Pain Type
By Whether the Angina Was Exercise Induced
1.00
0.75
0.50
0.25
0.00
Proportion
N
Y
Whether Angina was Exercise Induced
Chest Pain Type
ASY
TA

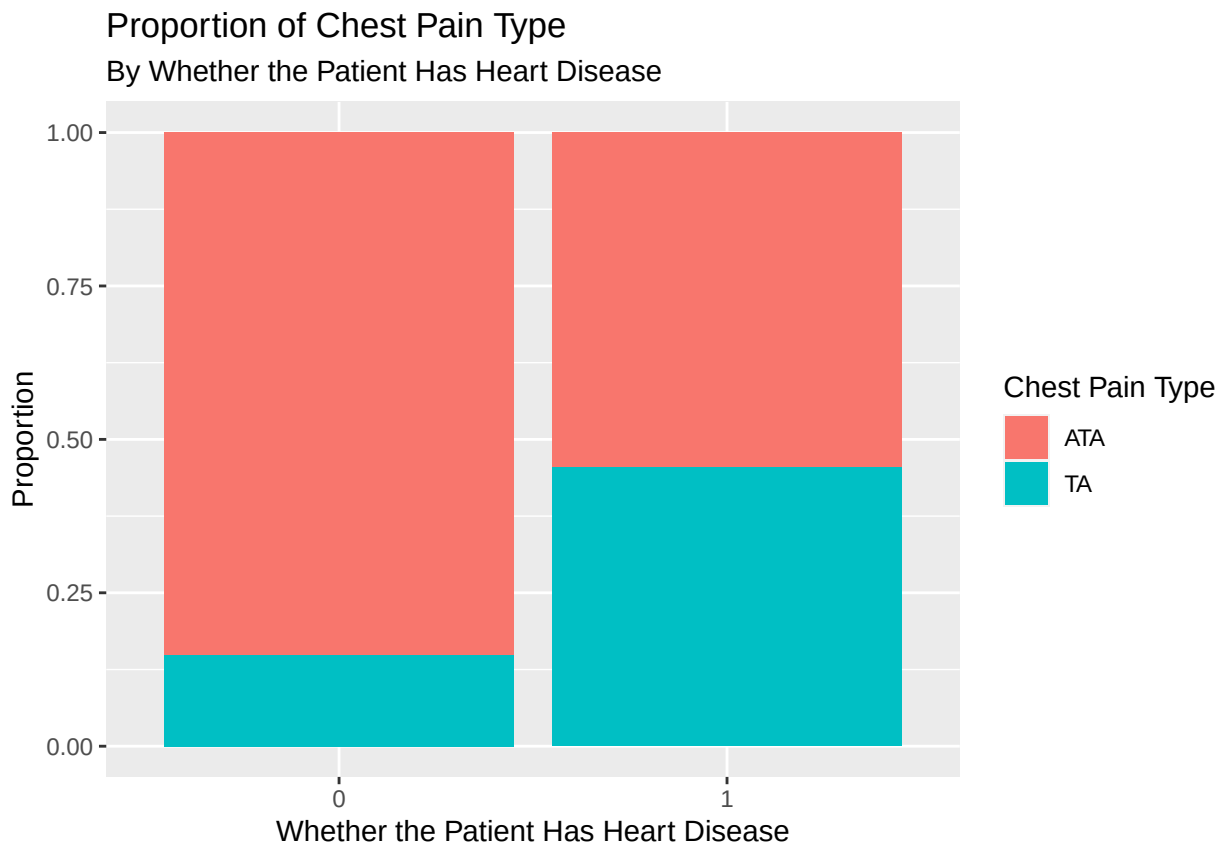

Type of Chest Pain By Cholesterol Level
1.00
0.75
0.50
0.25
0.00
Count
Low
Normal
Intermediate
High
Cholesterol Level
Chest Pain Type
ATA
TA


Proportion of Chest Pain Type
By Blood Pressure Level
1.00
0.75
0.50
0.25
0.00
Proportion
Normal
Elevated
Hypertension 1
Hypertension 2
Blood Pressure Level
Chest Pain Type
ATA
TA


Type of Chest Pain by Sex
1.00
0.75
0.50
0.25
0.00
Count
0
1
Sex
Chest Pain Type
ATA
TA


Type of Chest Pain By Whether the Angina Was Exercise Induced
Count
1.00
0.75
0.50
0.25
0.00
0
1
Whether Angina was Exercise Induced
Chest Pain Type
ATA
TA


Proportion of Chest Pain Type
By Whether the Patient Has Heart Disease
Proportion
1.00
0.75
0.50
0.25
0.00
0
1
Whether the Patient Has Heart Disease
Chest Pain Type
ASY
TA




```
## # A tibble: 2 x 2
##   Sex   count
##   <chr> <int>
## 1 F        80
## 2 M       462
```

```
## # A tibble: 4 x 3
## # Groups:   Sex [2]
##   Sex   ChestPainType count
##   <chr> <chr>         <int>
## 1 F     ASY              70
## 2 F     TA               10
## 3 M     ASY             426
## 4 M     TA               36
```

```
## # A tibble: 2 x 2
##   Sex   count
##   <chr> <int>
## 1 F        70
## 2 M       149
```

```
## # A tibble: 4 x 3
## # Groups:   Sex [2]
##   Sex   ChestPainType count
##   <chr> <chr>         <int>
## 1 F     ATA              60
## 2 F     TA               10
## 3 M     ATA             113
```

```
## 4 M     TA               36
```

```
## # A tibble: 2 x 2
##   ExerciseAngina count
##   <chr>          <int>
## 1 N                239
## 2 Y                303
```

```
## # A tibble: 4 x 3
## # Groups:   ExerciseAngina [2]
##   ExerciseAngina ChestPainType count
##   <chr>          <chr>         <int>
## 1 N              ASY             199
## 2 N              TA               40
## 3 Y              ASY             297
## 4 Y              TA                6
```

```
## # A tibble: 2 x 2
##   ExerciseAngina count
##   <chr>          <int>
## 1 N                196
## 2 Y                 23
```

```
## # A tibble: 4 x 3
## # Groups:   ExerciseAngina [2]
##   ExerciseAngina ChestPainType count
##   <chr>          <chr>         <int>
## 1 N              ATA             156
## 2 N              TA               40
## 3 Y              ATA              17
## 4 Y              TA                6
```

Discussion