

# Advanced Regression Analysis (6301)

Sai Krishna Nandyala

```
library(readxl)

## Warning: package 'readxl' was built under R version 4.3.3

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(ggplot2)

data <- read_excel("dataset.xlsx", sheet = "All cases")
df <- data[,c('Age', 'Sex', 'Migratory_Pain',
              'Lower_Right_Abd_Pain', 'Coughing_Pain',
              'Loss_of_Appetite', 'Nausea',
              'Body_Temperature', 'Diagnosis')]
names(df) <- c('Age', 'Gender', 'Migratory Pain',
              'Lower Right Abd Pain', 'Coughing Pain',
              'Anorexia', 'Nausea', 'Body Temperature', 'Diagnosis')

head(df)

## # A tibble: 6 × 9
##   Age Gender `Migratory Pain` `Lower Right Abd Pain` `Coughing Pain`
##   <dbl> <chr>   <chr>                <chr>                <chr>
## 1  5.11 female no                yes                no                no
## 2  5.03 male   no                yes                no                no
## 3  5.72 female no                no                 no                yes
## 4 11.0  female no                yes                no                yes
## 5 15.6  male   no                yes                no                no
## 6 14.7  female yes               yes                no                yes
## # i 3 more variables: Nausea <chr>, `Body Temperature` <dbl>, Diagnosis
## <chr>

dim(df)
```

```
## [1] 782 9

df$Age <- floor(df$Age)

colnames(df)

## [1] "Age" "Gender" "Migratory Pain"
## [4] "Lower Right Abd Pain" "Coughing Pain" "Anorexia"
## [7] "Nausea" "Body Temperature" "Diagnosis"

colSums(is.na(df))

##           Age           Gender           Migratory Pain
##           1             2             9
## Lower Right Abd Pain      Coughing Pain      Anorexia
##           8             16             10
##           Nausea      Body Temperature      Diagnosis
##           8             7             2

summary(df)

##           Age           Gender           Migratory Pain           Lower Right Abd
Pain
## Min.      : 0.00      Length:782      Length:782      Length:782
## 1st Qu.: 9.00      Class :character      Class :character      Class :character
## Median :11.00      Mode  :character      Mode  :character      Mode  :character
## Mean      :10.85
## 3rd Qu.:14.00
## Max.      :18.00
## NA's      :1
## Coughing Pain           Anorexia           Nausea           Body Temperature
## Length:782      Length:782      Length:782      Min.      :26.9
## Class :character      Class :character      Class :character      1st Qu.:36.8
## Mode  :character      Mode  :character      Mode  :character      Median :37.2
##                                     Mean      :37.4
##                                     3rd Qu.:37.9
##                                     Max.      :40.2
##                                     NA's      :7
## Diagnosis
## Length:782
## Class :character
## Mode  :character
##
##
##
##

catColumns <- names(df)[sapply(df, is.character)]
numColumns <- names(df)[sapply(df, is.numeric)]

Mode <- function(x) {
  ux <- unique(x)
```

```

    ux[which.max(tabulate(match(x, ux)))]
  }
  for (col in catColumns) {
    df[[col]] <- ifelse(is.na(df[[col]]), Mode(df[[col]]), df[[col]])
  }
  for (col in numColumns) {
    df[[col]] <- ifelse(is.na(df[[col]]), mean(df[[col]], na.rm = TRUE),
df[[col]])
  }

```

```
df$Age <- floor(df$Age)
```

```
colSums(is.na(df))
```

```
##           Age           Gender           Migratory Pain
##           0           0           0
## Lower Right Abd Pain    Coughing Pain    Anorexia
##           0           0           0
##           Nausea    Body Temperature    Diagnosis
##           0           0           0
```

```

total_count <- df %>%
  group_by(Age) %>%
  summarise(TotalCount = n())
filtered_data <- df %>%
  filter(Diagnosis == "appendicitis") %>%
  group_by(Age) %>%
  summarise(Count = n())
filtered_data

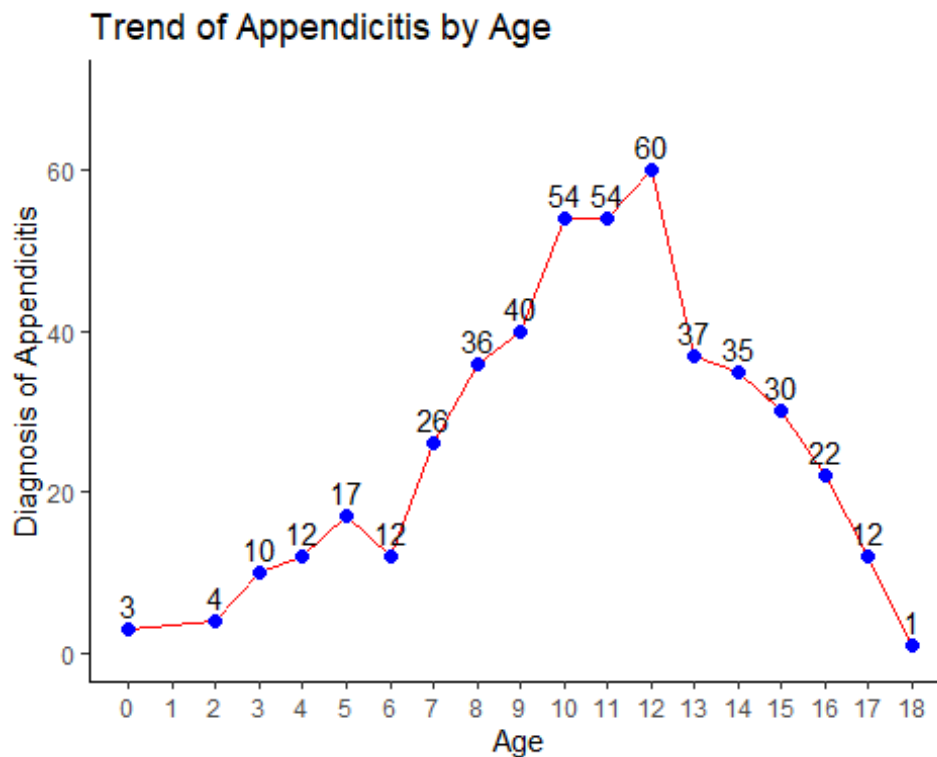
```

```

## # A tibble: 18 × 2
##   Age Count
##   <dbl> <int>
## 1     0     3
## 2     2     4
## 3     3    10
## 4     4    12
## 5     5    17
## 6     6    12
## 7     7    26
## 8     8    36
## 9     9    40
## 10    10    54
## 11    11    54
## 12    12    60
## 13    13    37
## 14    14    35
## 15    15    30
## 16    16    22
## 17    17    12
## 18    18     1

```

```
ggplot(filtered_data, aes(x = Age, y = Count)) +
  geom_line(color="red") + # Line plot
  geom_point(color="blue",size=2) + # Points on the line
  geom_text(aes(label = Count), vjust = -0.5, color = "black") +
  scale_y_continuous(limits = c(0, 70))+
  scale_x_continuous(breaks = seq(min(filtered_data$Age),
max(filtered_data$Age), by = 1)) + # Adding text labels above each point
  labs(x = "Age", y = "Diagnosis of Appendicitis", title = "Trend of
Appendicitis by Age") +
  theme_classic()
```

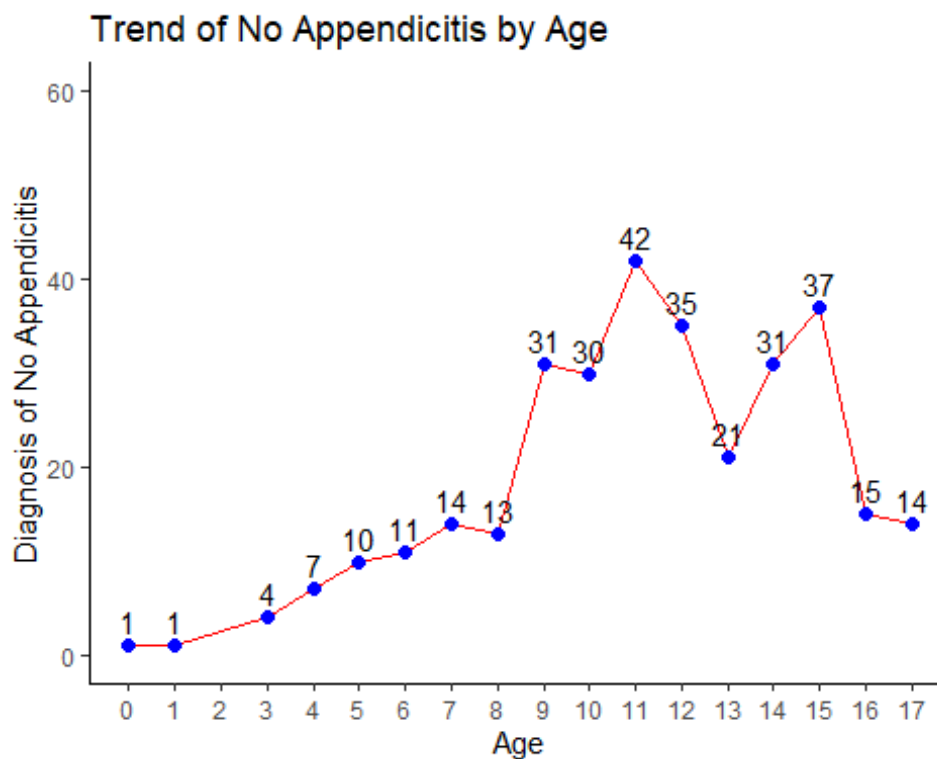


```
filtered_data <- df %>%
  filter(Diagnosis == "no appendicitis") %>%
  group_by(Age) %>%
  summarise(Count = n())
filtered_data
```

```
## # A tibble: 17 × 2
##   Age Count
##   <dbl> <int>
## 1     0     1
## 2     1     1
## 3     3     4
## 4     4     7
## 5     5    10
## 6     6    11
## 7     7    14
```

```
## 8      8      13
## 9      9      31
## 10     10     30
## 11     11     42
## 12     12     35
## 13     13     21
## 14     14     31
## 15     15     37
## 16     16     15
## 17     17     14
```

```
ggplot(filtered_data, aes(x = Age, y = Count)) +
  geom_line(color="red") + # Line plot
  geom_point(color="blue",size=2) + # Points on the line
  geom_text(aes(label = Count), vjust = -0.5, color = "black") +
  scale_x_continuous(breaks = seq(min(filtered_data$Age),
max(filtered_data$Age), by = 1)) + # Adding text labels above each point
  scale_y_continuous(limits = c(0, 60))+
  labs(x = "Age", y = "Diagnosis of No Appendicitis", title = "Trend of No
Appendicitis by Age") +
  theme_classic()
```



```
filtered_data <- merge(filtered_data, total_count, by = "Age")
filtered_data <- filtered_data %>%
  mutate(Percentage = (Count / TotalCount) * 100)
ggplot(filtered_data, aes(x = Age, y = Percentage)) +
  geom_line(color = "red") + # Line plot
```

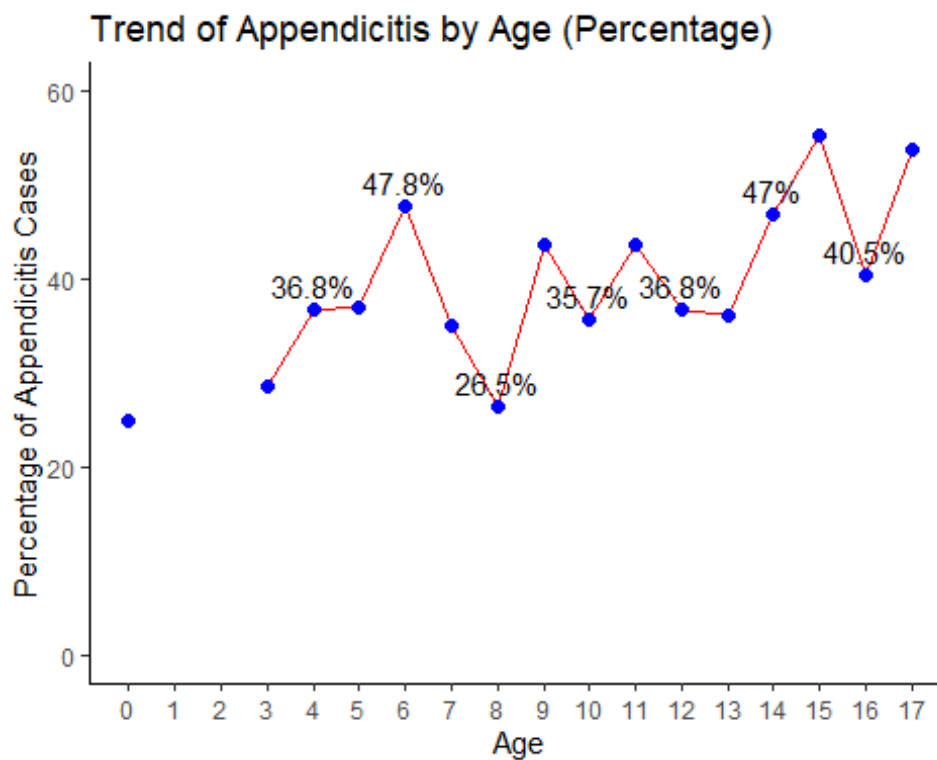
```

geom_point(color = "blue", size = 2) + # Points on the Line
geom_text(data = filtered_data %>% filter(row_number() %% 2 == 0),
aes(label = paste0(round(Percentage, 1), "%"), vjust = -0.5, color =
"black")) +
scale_x_continuous(breaks = seq(min(filtered_data$Age),
max(filtered_data$Age), by = 1)) + # Adding text labels above each point
scale_y_continuous(limits = c(0, 60)) +
labs(x = "Age", y = "Percentage of Appendicitis Cases", title = "Trend of
Appendicitis by Age (Percentage)") +
theme_classic()

```

## Warning: Removed 1 rows containing missing values (`geom\_point()`).

## Warning: Removed 1 rows containing missing values (`geom\_text()`).



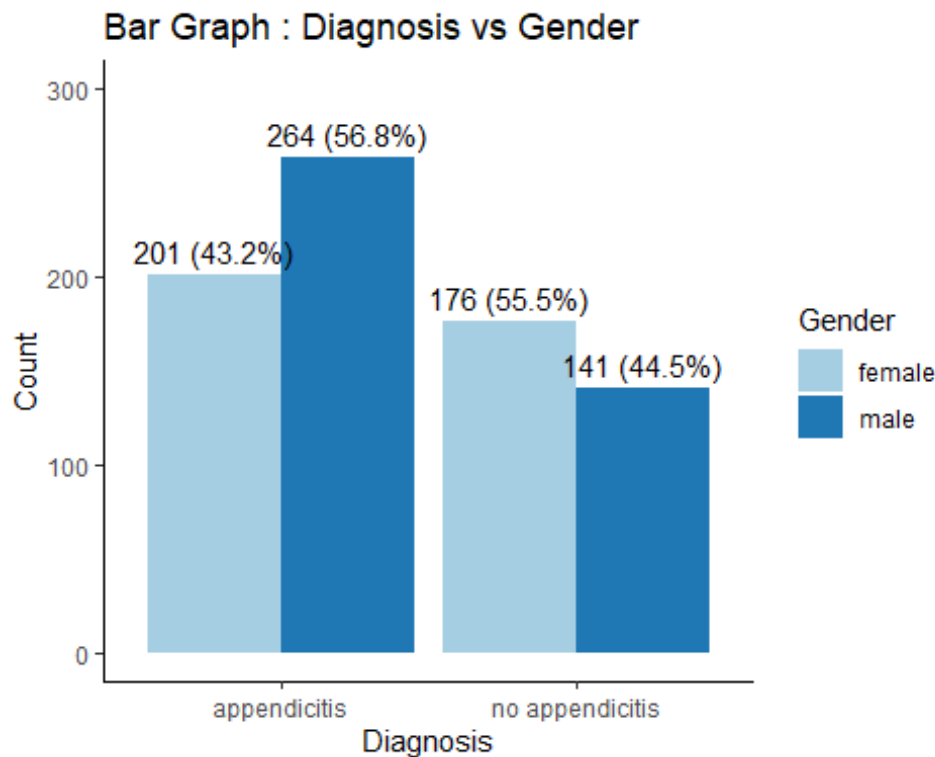
```

total_counts <- df %>%
  group_by(Diagnosis) %>%
  summarise(total_count = n())
gender_counts <- df %>%
  group_by(Diagnosis, Gender) %>%
  summarise(count = n()) %>%
  mutate(percentage = count / sum(count) * 100)

## `summarise()` has grouped output by 'Diagnosis'. You can override using
the
## `.groups` argument.

```

```
ggplot(df, aes(x = Diagnosis, fill = Gender)) +
  geom_bar(position = "dodge") +
  labs(x = "Diagnosis", y = "Count", title = "Bar Graph : Diagnosis vs
Gender") +
  theme_classic() +
  scale_fill_brewer(palette = "Paired") +
  scale_y_continuous(limits = c(0, 300)) +
  geom_text(data = gender_counts, aes(label = paste0(count, " (",
round(percentage, 1), "%)"), y = count),
  position = position_dodge(width = 0.9), vjust = -0.5)
```

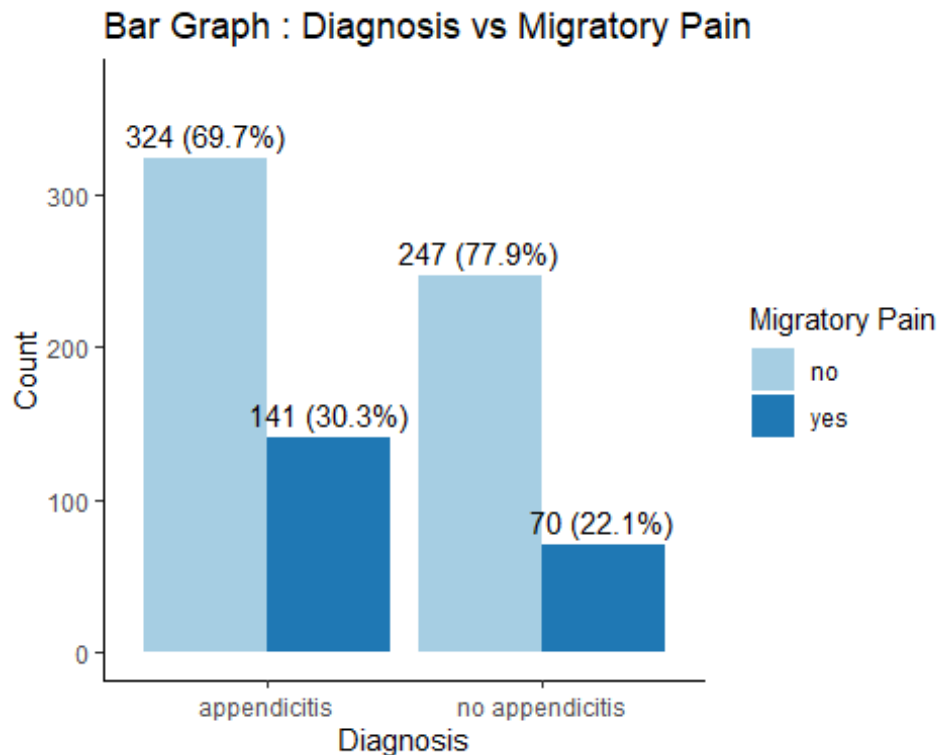


```
total_counts <- df %>%
  group_by(Diagnosis) %>%
  summarise(total_count = n())
migratory_pain_counts <- df %>%
  group_by(Diagnosis, `Migratory Pain`) %>%
  summarise(count = n()) %>%
  mutate(percentage = count / sum(count) * 100)

## `summarise()` has grouped output by 'Diagnosis'. You can override using
the
## `.groups` argument.

ggplot(df, aes(x = Diagnosis, fill = `Migratory Pain`)) +
  geom_bar(position = "dodge") +
  labs(x = "Diagnosis", y = "Count", title = "Bar Graph : Diagnosis vs
Migratory Pain") +
```

```
theme_classic() +
scale_fill_brewer(palette = "Paired") +
scale_y_continuous(limits = c(0, 370)) +
geom_text(data = migratory_pain_counts, aes(label = paste0(count, " (",
round(percentage, 1), "%)"), y = count),
position = position_dodge(width = 0.9), vjust = -0.5)
```



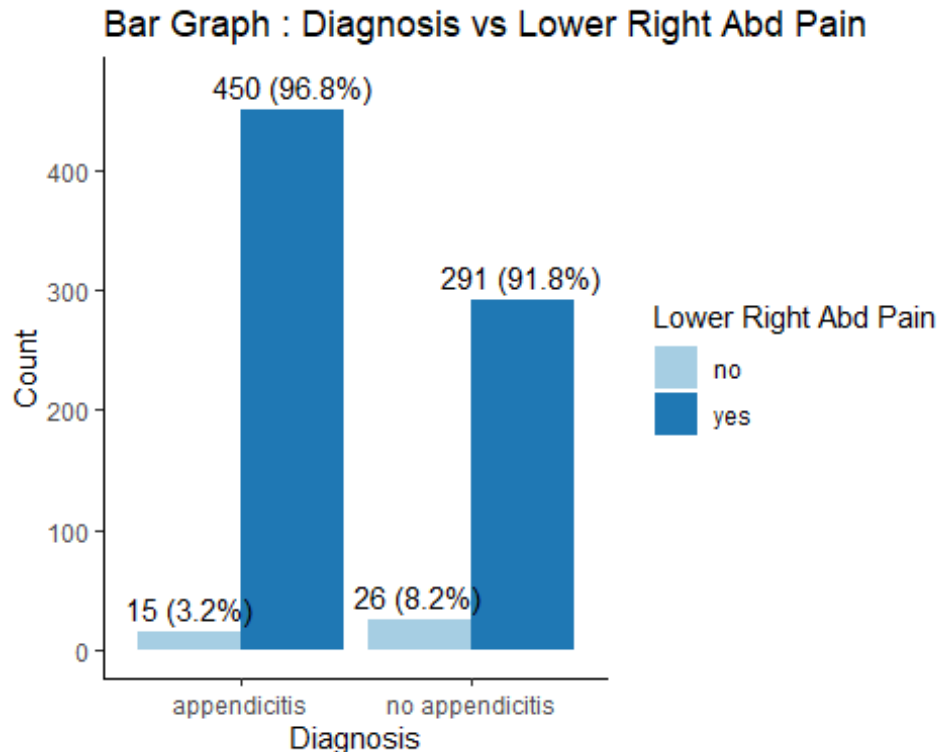
```
total_counts <- df %>%
  group_by(Diagnosis) %>%
  summarise(total_count = n())
lower_right_abd_pain_counts <- df %>%
  group_by(Diagnosis, `Lower Right Abd Pain`) %>%
  summarise(count = n()) %>%
  mutate(percentage = count / sum(count) * 100)

## `summarise()` has grouped output by 'Diagnosis'. You can override using
the
## `.groups` argument.

ggplot(df, aes(x = Diagnosis, fill = `Lower Right Abd Pain`)) +
  geom_bar(position = "dodge") +
  labs(x = "Diagnosis", y = "Count", title = "Bar Graph : Diagnosis vs Lower
Right Abd Pain") +
  theme_classic() +
  scale_fill_brewer(palette = "Paired") +
  scale_y_continuous(limits = c(0, 470)) +
  geom_text(data = lower_right_abd_pain_counts, aes(label =
```



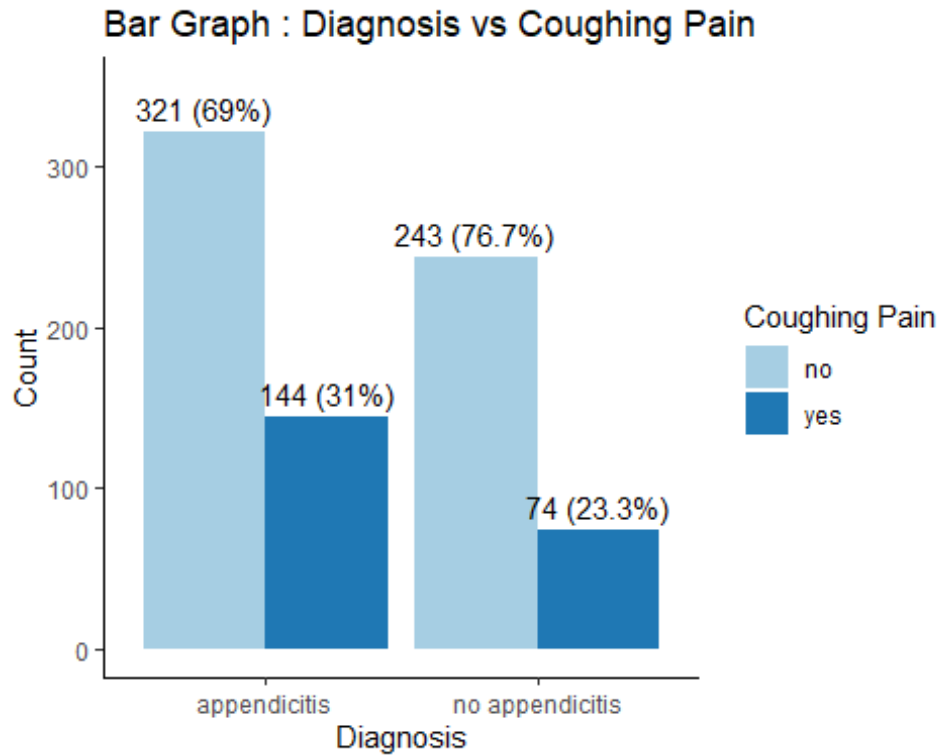
```
paste0(round(count, 1), " (", round(percentage, 1), "%)"), y = count),
  position = position_dodge(width = 0.9), vjust = -0.5)
```



```
total_counts <- df %>%
  group_by(Diagnosis) %>%
  summarise(total_count = n())
coughing_pain_counts <- df %>%
  group_by(Diagnosis, `Coughing Pain`) %>%
  summarise(count = n()) %>%
  mutate(percentage = count / sum(count) * 100)

## `summarise()` has grouped output by 'Diagnosis'. You can override using
the
## `.groups` argument.

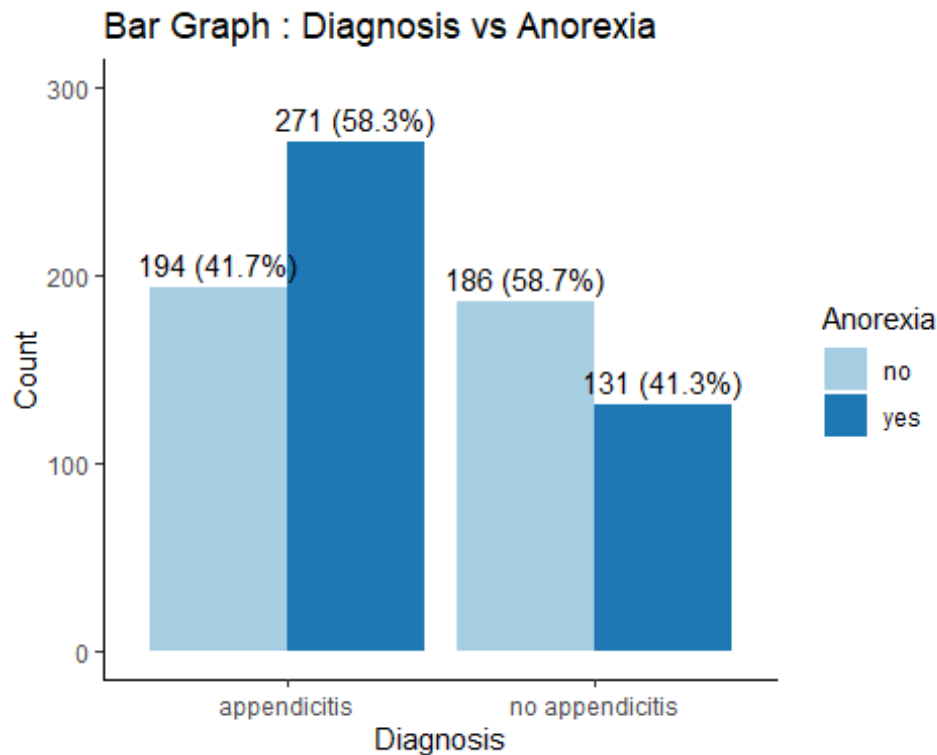
ggplot(df, aes(x = Diagnosis, fill = `Coughing Pain`)) +
  geom_bar(position = "dodge") +
  labs(x = "Diagnosis", y = "Count", title = "Bar Graph : Diagnosis vs
Coughing Pain") +
  theme_classic() +
  scale_fill_brewer(palette = "Paired") +
  scale_y_continuous(limits = c(0, 350)) +
  geom_text(data = coughing_pain_counts, aes(label = paste0(count, " (",
round(percentage, 1), "%)"), y = count),
  position = position_dodge(width = 0.9), vjust = -0.5)
```



```
total_counts <- df %>%
  group_by(Diagnosis) %>%
  summarise(total_count = n())
anorexia_counts <- df %>%
  group_by(Diagnosis, Anorexia) %>%
  summarise(count = n()) %>%
  mutate(percentage = count / sum(count) * 100)

## `summarise()` has grouped output by 'Diagnosis'. You can override using
the
## `.groups` argument.

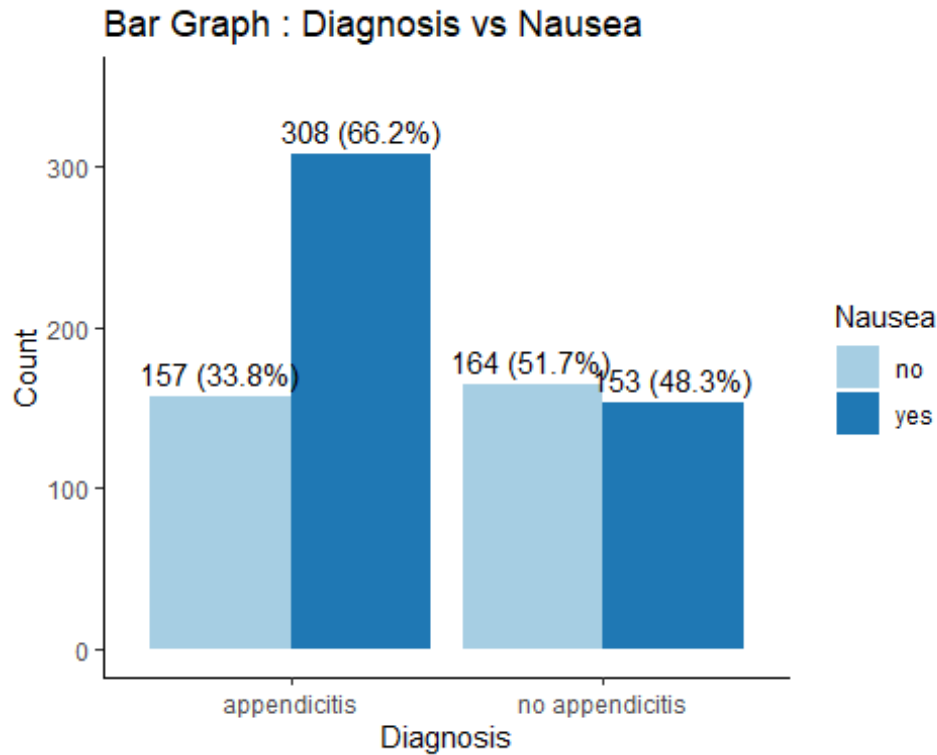
ggplot(df, aes(x = Diagnosis, fill = Anorexia)) +
  geom_bar(position = "dodge") +
  labs(x = "Diagnosis", y = "Count", title = "Bar Graph : Diagnosis vs
Anorexia") +
  theme_classic() +
  scale_fill_brewer(palette = "Paired") +
  scale_y_continuous(limits = c(0, 300)) +
  geom_text(data = anorexia_counts, aes(label = paste0(count, " (",
round(percentage, 1), "%)"), y = count),
            position = position_dodge(width = 0.9), vjust = -0.5)
```



```
total_counts <- df %>%
  group_by(Diagnosis) %>%
  summarise(total_count = n())
nausea_counts <- df %>%
  group_by(Diagnosis, Nausea) %>%
  summarise(count = n()) %>%
  mutate(percentage = count / sum(count) * 100)

## `summarise()` has grouped output by 'Diagnosis'. You can override using
the
## `.groups` argument.

ggplot(df, aes(x = Diagnosis, fill = Nausea)) +
  geom_bar(position = "dodge") +
  labs(x = "Diagnosis", y = "Count", title = "Bar Graph : Diagnosis vs
Nausea") +
  theme_classic() +
  scale_fill_brewer(palette = "Paired") +
  scale_y_continuous(limits = c(0, 350)) +
  geom_text(data = nausea_counts, aes(label = paste0(count, " (",
round(percentage, 1), "%)"), y = count),
            position = position_dodge(width = 0.9), vjust = -0.5)
```



```
cor(df[,numColumns])
```

```
##
##           Age Body Temperature
## Age           1.0000000      -0.2905143
## Body Temperature -0.2905143          1.0000000
```

```
#Encoding
```

```
df$Gender<- ifelse(df$Gender == "male", 1, 0)
df$`Migratory Pain` <- ifelse(df$`Migratory Pain` == "yes", 1, 0)
df$`Lower Right Abd Pain` <- ifelse(df$`Lower Right Abd Pain` == "yes", 1, 0)
df$`Coughing Pain` <- ifelse(df$`Coughing Pain` == "yes", 1, 0)
df$Anorexia <- ifelse(df$Anorexia == "yes", 1, 0)
df$Nausea <- ifelse(df$Nausea == "yes", 1, 0)
df$Diagnosis<- ifelse(df$Diagnosis == "appendicitis", 1, 0)
```

```
head(df)
```

```
## # A tibble: 6 × 9
##   Age Gender `Migratory Pain` `Lower Right Abd Pain` `Coughing Pain`
##   <dbl> <dbl>          <dbl>          <dbl>          <dbl>
## 1     5     0              0              1              0
## 2     5     1              0              1              0
## 3     5     0              0              0              0
```

```

1
## 4      11      0              0              1              0
1
## 5      15      1              0              1              0
0
## 6      14      0              1              1              0
1
## # i 3 more variables: Nausea <dbl>, `Body Temperature` <dbl>, Diagnosis
<dbl>

#Scaling
df[-which(names(df) == "Diagnosis")] <- scale(df[-which(names(df) ==
"Diagnosis")])

#Modeling
#Full Model
fullModel<- glm(Diagnosis ~ Age+Gender+`Migratory Pain`+`Lower Right Abd
Pain`+
                  `Coughing Pain`+Anorexia+Nausea+`Body
Temperature`,
                  data = df, family = binomial(link = "logit"))
summary(fullModel)

##
## Call:
## glm(formula = Diagnosis ~ Age + Gender + `Migratory Pain` + `Lower Right
Abd Pain` +
##      `Coughing Pain` + Anorexia + Nausea + `Body Temperature`,
##      family = binomial(link = "logit"), data = df)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.41858    0.07719   5.423 5.87e-08 ***
## Age           -0.07253    0.08356  -0.868 0.385344
## Gender         0.23350    0.07827   2.983 0.002852 **
## `Migratory Pain` 0.13569    0.07979   1.701 0.089030 .
## `Lower Right Abd Pain` 0.24506    0.08037   3.049 0.002297 **
## `Coughing Pain`  0.11883    0.07895   1.505 0.132310
## Anorexia        0.25957    0.08372   3.100 0.001933 **
## Nausea          0.24948    0.08255   3.022 0.002510 **
## `Body Temperature` 0.32770    0.09192   3.565 0.000364 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1055.90  on 781  degrees of freedom
## Residual deviance:  975.67  on 773  degrees of freedom
## AIC: 993.67

```

```
##
## Number of Fisher Scoring iterations: 4

#Simple Model
simpleModel<- glm(Diagnosis ~ Gender+`Lower Right Abd Pain`+Anorexia+
                  Nausea+`Body Temperature`,
                  data = df, family = binomial(link = "logit"))

summary(simpleModel)

##
## Call:
## glm(formula = Diagnosis ~ Gender + `Lower Right Abd Pain` + Anorexia +
##      Nausea + `Body Temperature`, family = binomial(link = "logit"),
##      data = df)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.41349    0.07678   5.386 7.22e-08 ***
## Gender            0.24133    0.07651   3.154 0.001610 **
## `Lower Right Abd Pain` 0.26995    0.07873   3.429 0.000606 ***
## Anorexia          0.28471    0.08268   3.444 0.000574 ***
## Nausea            0.25584    0.08204   3.118 0.001819 **
## `Body Temperature`  0.33954    0.08754   3.879 0.000105 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1055.90  on 781  degrees of freedom
## Residual deviance:  981.75  on 776  degrees of freedom
## AIC: 993.75
##
## Number of Fisher Scoring iterations: 4

#Interaction Model 1
IM1 <- glm(Diagnosis ~Gender*`Lower Right Abd Pain` +
            Gender*Anorexia +
            Gender*Nausea +
            Gender*`Body Temperature` +
            `Lower Right Abd Pain`*Anorexia+
            `Lower Right Abd Pain`*Nausea+
            `Lower Right Abd Pain`*`Body Temperature`+
            Anorexia*Nausea+
            Anorexia*`Body Temperature`+
            Nausea*`Body Temperature`,
            data = df, family = binomial(link = "logit"))

summary(IM1)

##
## Call:
## glm(formula = Diagnosis ~ Gender * `Lower Right Abd Pain` + Gender *
```

```

##      Anorexia + Gender * Nausea + Gender * `Body Temperature` +
##      `Lower Right Abd Pain` * Anorexia + `Lower Right Abd Pain` *
##      Nausea + `Lower Right Abd Pain` * `Body Temperature` + Anorexia *
##      Nausea + Anorexia * `Body Temperature` + Nausea * `Body Temperature`,
##      family = binomial(link = "logit"), data = df)
##
## Coefficients:
##
##                                     Estimate Std. Error z value
Pr(>|z|)
## (Intercept)                        0.43094      0.08682   4.964
6.92e-07
## Gender                            0.22606      0.07931   2.850
0.004366
## `Lower Right Abd Pain`            0.30130      0.12121   2.486
0.012925
## Anorexia                          0.28928      0.08485   3.409
0.000651
## Nausea                            0.25734      0.08805   2.923
0.003470
## `Body Temperature`               0.34999      0.09172   3.816
0.000136
## Gender:`Lower Right Abd Pain`     -0.05959      0.10846  -0.549
0.582719
## Gender:Anorexia                   -0.10445      0.08496  -1.229
0.218917
## Gender:Nausea                     0.09522      0.08412   1.132
0.257660
## Gender:`Body Temperature`        -0.27643      0.09184  -3.010
0.002613
## `Lower Right Abd Pain`:Anorexia   0.08214      0.10019   0.820
0.412311
## `Lower Right Abd Pain`:Nausea     -0.08342      0.14533  -0.574
0.565972
## `Lower Right Abd Pain`:`Body Temperature` -0.15206      0.09968  -1.526
0.127119
## Anorexia:Nausea                   -0.02394      0.08580  -0.279
0.780209
## Anorexia:`Body Temperature`       0.06393      0.09689   0.660
0.509327
## Nausea:`Body Temperature`        -0.04949      0.09607  -0.515
0.606461
##
## (Intercept)                       ***
## Gender                            **
## `Lower Right Abd Pain`            *
## Anorexia                          ***
## Nausea                            **
## `Body Temperature`                ***
## Gender:`Lower Right Abd Pain`
## Gender:Anorexia

```

```

## Gender:Nausea
## Gender:`Body Temperature` **
## `Lower Right Abd Pain`:Anorexia
## `Lower Right Abd Pain`:Nausea
## `Lower Right Abd Pain`:`Body Temperature`
## Anorexia:Nausea
## Anorexia:`Body Temperature`
## Nausea:`Body Temperature`
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1055.90 on 781 degrees of freedom
## Residual deviance: 965.58 on 766 degrees of freedom
## AIC: 997.58
##
## Number of Fisher Scoring iterations: 5

#Interaction Model 2
IM2 <- glm(Diagnosis ~ Gender*`Body Temperature` + `Lower Right Abd Pain` +
  Anorexia+Nausea ,
  data = df, family = binomial(link = "logit"))
summary(IM2)

##
## Call:
## glm(formula = Diagnosis ~ Gender * `Body Temperature` + `Lower Right Abd
Pain` +
## Anorexia + Nausea, family = binomial(link = "logit"), data = df)
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 0.41999 0.07720 5.440 5.33e-08 ***
## Gender 0.23217 0.07712 3.010 0.002610 **
## `Body Temperature` 0.34898 0.08953 3.898 9.70e-05 ***
## `Lower Right Abd Pain` 0.26523 0.07957 3.333 0.000859 ***
## Anorexia 0.29326 0.08323 3.523 0.000426 ***
## Nausea 0.25387 0.08246 3.079 0.002079 **
## Gender:`Body Temperature` -0.24191 0.08918 -2.713 0.006677 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1055.90 on 781 degrees of freedom
## Residual deviance: 974.23 on 775 degrees of freedom
## AIC: 988.23
##
## Number of Fisher Scoring iterations: 4

```



```

lrt_result1 <- anova(simpleModel, fullModel, test = "Chisq")
lrt_result1

## Analysis of Deviance Table
##
## Model 1: Diagnosis ~ Gender + `Lower Right Abd Pain` + Anorexia + Nausea +
##   `Body Temperature`
## Model 2: Diagnosis ~ Age + Gender + `Migratory Pain` + `Lower Right Abd
Pain` +
##   `Coughing Pain` + Anorexia + Nausea + `Body Temperature`
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      776      981.75
## 2      773      975.67  3    6.0828  0.1076

lrt_result2 <- anova(simpleModel, IM1, test = "Chisq")
lrt_result2

## Analysis of Deviance Table
##
## Model 1: Diagnosis ~ Gender + `Lower Right Abd Pain` + Anorexia + Nausea +
##   `Body Temperature`
## Model 2: Diagnosis ~ Gender * `Lower Right Abd Pain` + Gender * Anorexia +
##   Gender * Nausea + Gender * `Body Temperature` + `Lower Right Abd Pain`
*
##   Anorexia + `Lower Right Abd Pain` * Nausea + `Lower Right Abd Pain` *
##   `Body Temperature` + Anorexia * Nausea + Anorexia * `Body Temperature`
+
##   Nausea * `Body Temperature`
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      776      981.75
## 2      766      965.58 10    16.169  0.0949 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lrt_result3 <- anova(simpleModel, IM2, test = "Chisq")
lrt_result3

## Analysis of Deviance Table
##
## Model 1: Diagnosis ~ Gender + `Lower Right Abd Pain` + Anorexia + Nausea +
##   `Body Temperature`
## Model 2: Diagnosis ~ Gender * `Body Temperature` + `Lower Right Abd Pain`
+
##   Anorexia + Nausea
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      776      981.75
## 2      775      974.23  1    7.5258 0.006082 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

pvals <- data.frame(
  Comparision = c( "Simple Model vs Full Model", "Simple Model vs Interaction
Model 1", "Simple Model vs Interaction Model 2"),
  `p-values` =
c(round(lrt_result1$`Pr(>Chi)`[2],2), round(lrt_result2$`Pr(>Chi)`[2],2),
round(lrt_result3$`Pr(>Chi)`[2],2))
)
print(pvals)

##                               Comparision p.values
## 1           Simple Model vs Full Model      0.11
## 2 Simple Model vs Interaction Model 1      0.09
## 3 Simple Model vs Interaction Model 2      0.01

results <- data.frame(
  Model = c("Full Model", "Simple Model", "Interaction Model 1", "Interaction
Model 2"),
  AIC = c(AIC(fullModel), AIC(simpleModel), AIC(IM1), AIC(IM2)),
  BIC = c(BIC(fullModel), BIC(simpleModel), BIC(IM1), BIC(IM2))
)
print(results)

##           Model      AIC      BIC
## 1      Full Model 993.6690 1035.626
## 2      Simple Model 993.7518 1021.723
## 3 Interaction Model 1 997.5830 1072.173
## 4 Interaction Model 2 988.2260 1020.859

coefficients <- coef(IM2)
odds_ratios <- exp(coefficients)
odds_ratio_simple <- data.frame(
  Coefficients = coefficients,
  OddsRatio = odds_ratios
)
odds_ratio_simple <- round(odds_ratio_simple, 2)
print(odds_ratio_simple)

##                               Coefficients OddsRatio
## (Intercept)                0.42      1.52
## Gender                    0.23      1.26
## `Body Temperature`        0.35      1.42
## `Lower Right Abd Pain`    0.27      1.30
## Anorexia                  0.29      1.34
## Nausea                    0.25      1.29
## Gender:`Body Temperature` -0.24      0.79

```