EPI 590R Surveillance Assignment 3

Victoria Ngo

2024-10-28

## Tables and Figures for Dataset B

# Filtering out unknown or NA values for gender  
datab\_filtered <- dataB %>%  
 filter(!is.na(case\_gender), case\_gender != "Unknown")  
  
# Calculating total number of patients for each gender  
total\_by\_gender <- datab\_filtered %>%  
 group\_by(case\_gender) %>%  
 summarise(Total = n())  
  
# Summarizing the count of each symptom by gender  
symptom\_counts <- datab\_filtered %>%  
 group\_by(case\_gender) %>%  
 summarise(  
 Fever = sum(sym\_fever == "Yes", na.rm = TRUE),  
 Subjective\_Fever = sum(sym\_subjfever == "Yes", na.rm = TRUE),  
 Myalgia = sum(sym\_myalgia == "Yes", na.rm = TRUE),  
 Loss\_Taste\_Smell = sum(sym\_losstastesmell == "Yes", na.rm = TRUE),  
 Sore\_Throat = sum(sym\_sorethroat == "Yes", na.rm = TRUE),  
 Cough = sum(sym\_cough == "Yes", na.rm = TRUE),  
 Headache = sum(sym\_headache == "Yes", na.rm = TRUE)  
 ) %>%  
 ungroup()  
  
# Merging the total counts and calculate percentages for each gender  
symptom\_counts\_with\_percent <- symptom\_counts %>%  
 left\_join(total\_by\_gender, by = "case\_gender") %>%  
 mutate(  
 Fever\_Percentage = round((Fever / sum(Fever)) \* 100, 2),  
 Subjective\_Fever\_Percentage = round((Subjective\_Fever / sum(Subjective\_Fever)) \* 100, 2),  
 Myalgia\_Percentage = round((Myalgia / sum(Myalgia)) \* 100, 2),  
 Loss\_Taste\_Smell\_Percentage = round((Loss\_Taste\_Smell / sum(Loss\_Taste\_Smell)) \* 100, 2),  
 Sore\_Throat\_Percentage = round((Sore\_Throat / sum(Sore\_Throat)) \* 100, 2),  
 Cough\_Percentage = round((Cough / sum(Cough)) \* 100, 2),  
 Headache\_Percentage = round((Headache / sum(Headache)) \* 100, 2)  
 )  
  
# Reshaping the table to have symptoms as rows and gender as columns  
transposed\_table <- symptom\_counts\_with\_percent %>%  
 select(case\_gender, Fever\_Percentage:Headache\_Percentage) %>%  
 pivot\_longer(cols = -case\_gender, names\_to = "Symptom", values\_to = "Percentage") %>%  
 pivot\_wider(names\_from = case\_gender, values\_from = Percentage)  
  
# View the final transposed table  
print(transposed\_table)

## # A tibble: 7 × 3  
## Symptom Female Male  
## <chr> <dbl> <dbl>  
## 1 Fever\_Percentage 47.7 52.3  
## 2 Subjective\_Fever\_Percentage 52.1 47.9  
## 3 Myalgia\_Percentage 54.4 45.6  
## 4 Loss\_Taste\_Smell\_Percentage 64.2 35.8  
## 5 Sore\_Throat\_Percentage 55.6 44.4  
## 6 Cough\_Percentage 52.4 47.6  
## 7 Headache\_Percentage 59.3 40.7

|  |  |  |
| --- | --- | --- |
| Symptom | Female (%) | Male (%) |
| Fever | 638 (47.7) | 699 (52.3) |
| Subjective Fever | 560 (52.1) | 514 (47.9) |
| Myalgia | 684 (54.4) | 574 (45.6) |
| Loss of Taste and Smell | 68 (64.2) | 38 (35.8) |
| Sore Throat | 397 (55.6) | 317 (44.4) |
| Cough | 820 (52.4) | 745 (47.6) |
| Headache | 678 (59.3) | 465 (40.7) |

Table 1. Gender and percentages by symptom

# Summarizing the data by city, gender, and race, excluding "Unknown" and NA values for gender  
city\_summary <- dataB %>%  
 filter(confirmed\_case == "Yes",   
 !is.na(case\_gender),   
 case\_gender != "Unknown") %>%  
 group\_by(case\_city, case\_gender, case\_race) %>%  
 summarise(Case\_Count = n()) %>%  
 ungroup()

## `summarise()` has grouped output by 'case\_city', 'case\_gender'. You can  
## override using the `.groups` argument.

# Plotting a bar chart to visualize the disease burden by city, gender, and race  
ggplot(city\_summary, aes(x = case\_gender, y = Case\_Count, fill = case\_race)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 facet\_wrap(~ case\_city) +  
 labs(title = "Disease Burden by City, Gender, and Race",  
 x = "City",  
 y = "Number of Confirmed Cases",  
 fill = "Race") +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1))

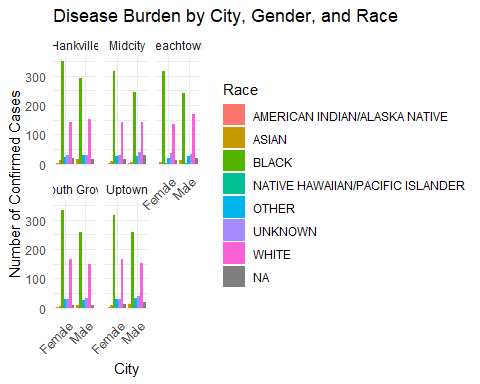


Figure 1. Disease Burden by Case City, Gender, and Race