Obesity Ghana Analysis

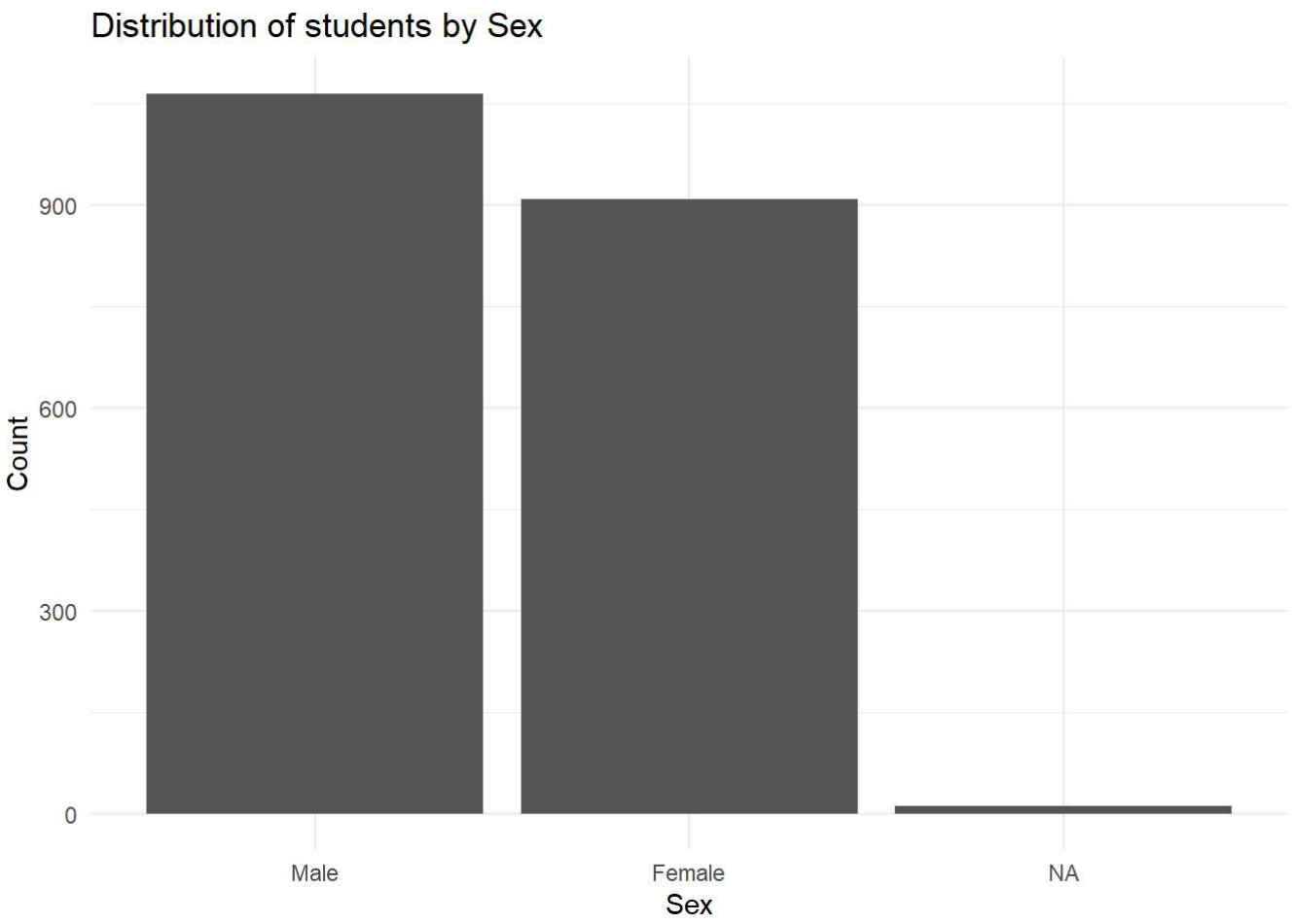
Vasco Ayere Avoka

2024-06-19

The dataset obesity_ghana contains information on various health and lifestyle variables collected from senior high school students in Ghana in 2012 and updated in 2018. It includes variables such as age, sex, weight, height, physical activity, and other related measures.

```
# Distribution of students by sex
age_sex_plot <- ggplot(obesity_ghana, aes(x = factor(Sex, levels = c(1, 2), labels = c("M
  geom_bar() +
  theme_minimal() +
  labs(title = "Distribution of students by Sex",
       x = "Sex",
       y = "Count")

print(age_sex_plot)
```

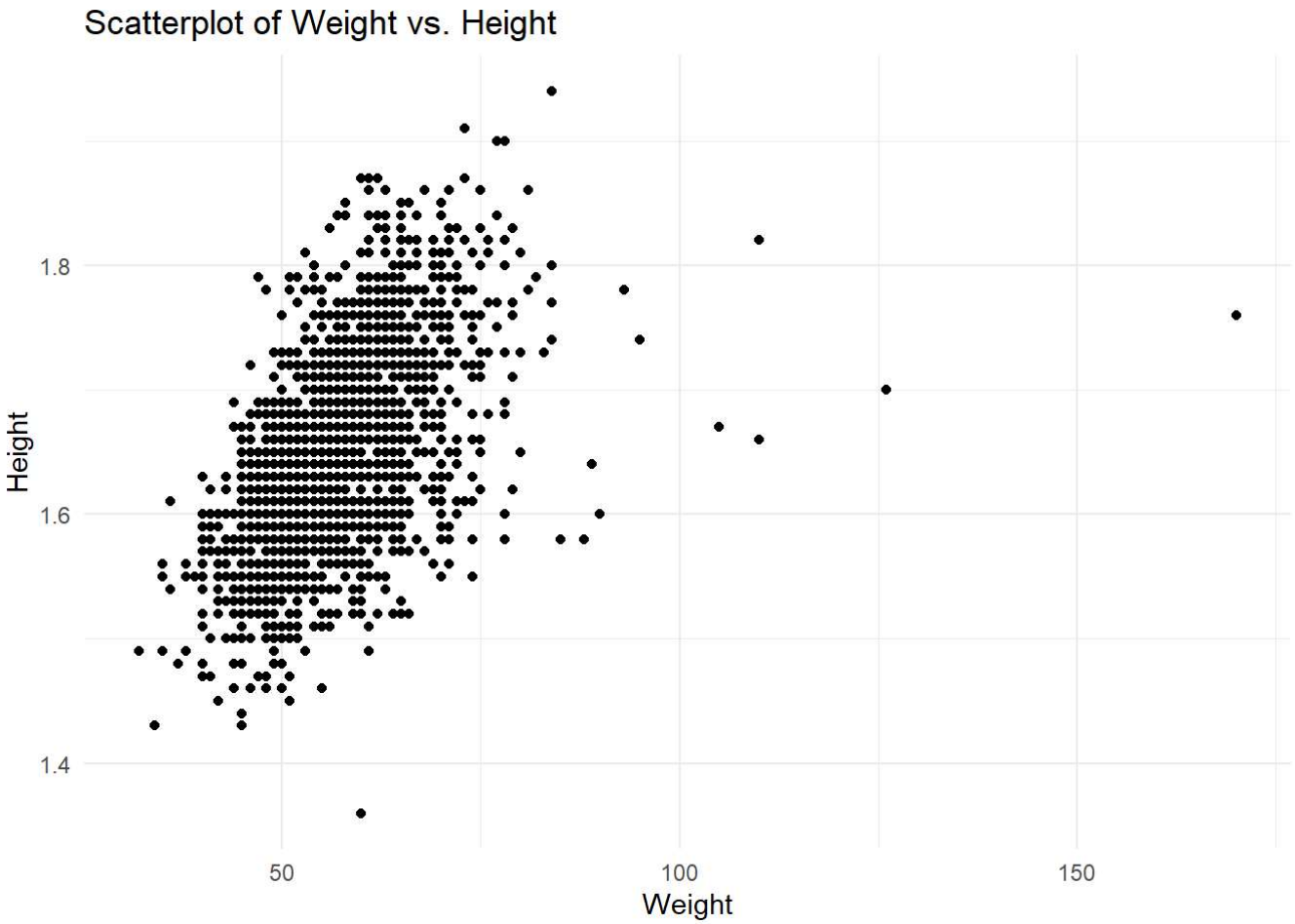


The bar plot above shows the distribution of students age and sex. It provides a visual representation of the number of male and female students in the dataset.

```
# A scatter plot of weight and height
weight_height_scatter <- ggplot(obesity_ghana, aes(x = Weight, y = Height)) +
  geom_point() +
  theme_minimal() +
  labs(title = "Scatterplot of Weight vs. Height",
       x = "Weight",
       y = "Height")

print(weight_height_scatter)
```

```
## Warning: Removed 35 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

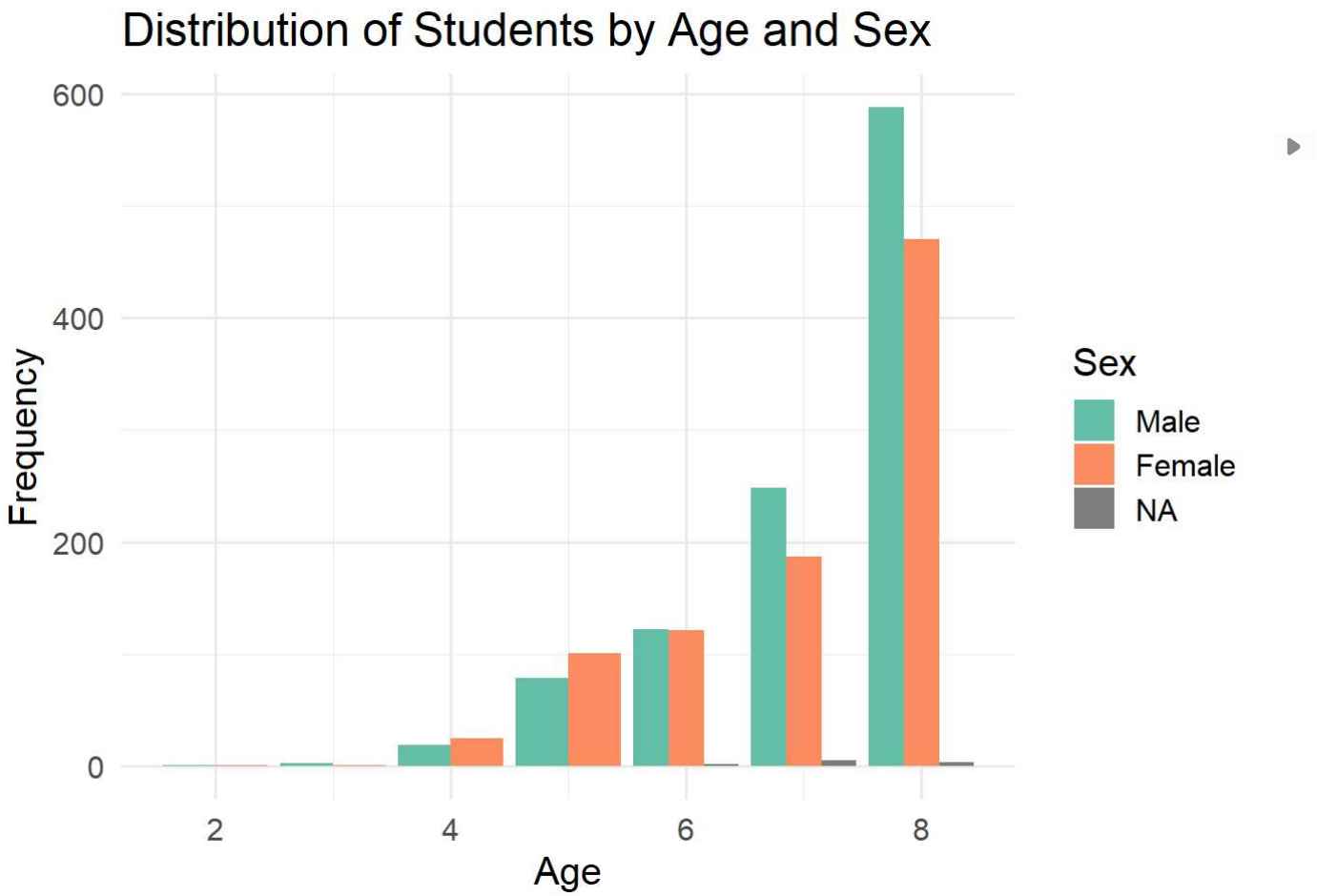


The scatterplot above illustrates the relationship between the weight and height of the students. Each point represents an individual student, showing how weight and height vary among the population.

```
# Distribution of age by sex
age_sex_distribution <- ggplot(obesity_ghana, aes(x = Age, fill = factor(Sex))) +
  geom_bar(position = 'dodge') +
  scale_fill_manual(values = c("#66c2a5", "#fc8d62"), labels = c('Male', 'Female')) +
  labs(title = 'Distribution of Students by Age and Sex', x = 'Age', y = 'Frequency', fil
  theme_minimal(base_size = 15)
```

```
print(age_sex_distribution)
```

```
## Warning: Removed 7 rows containing non-finite outside the scale range
## (`stat_count()`).
```

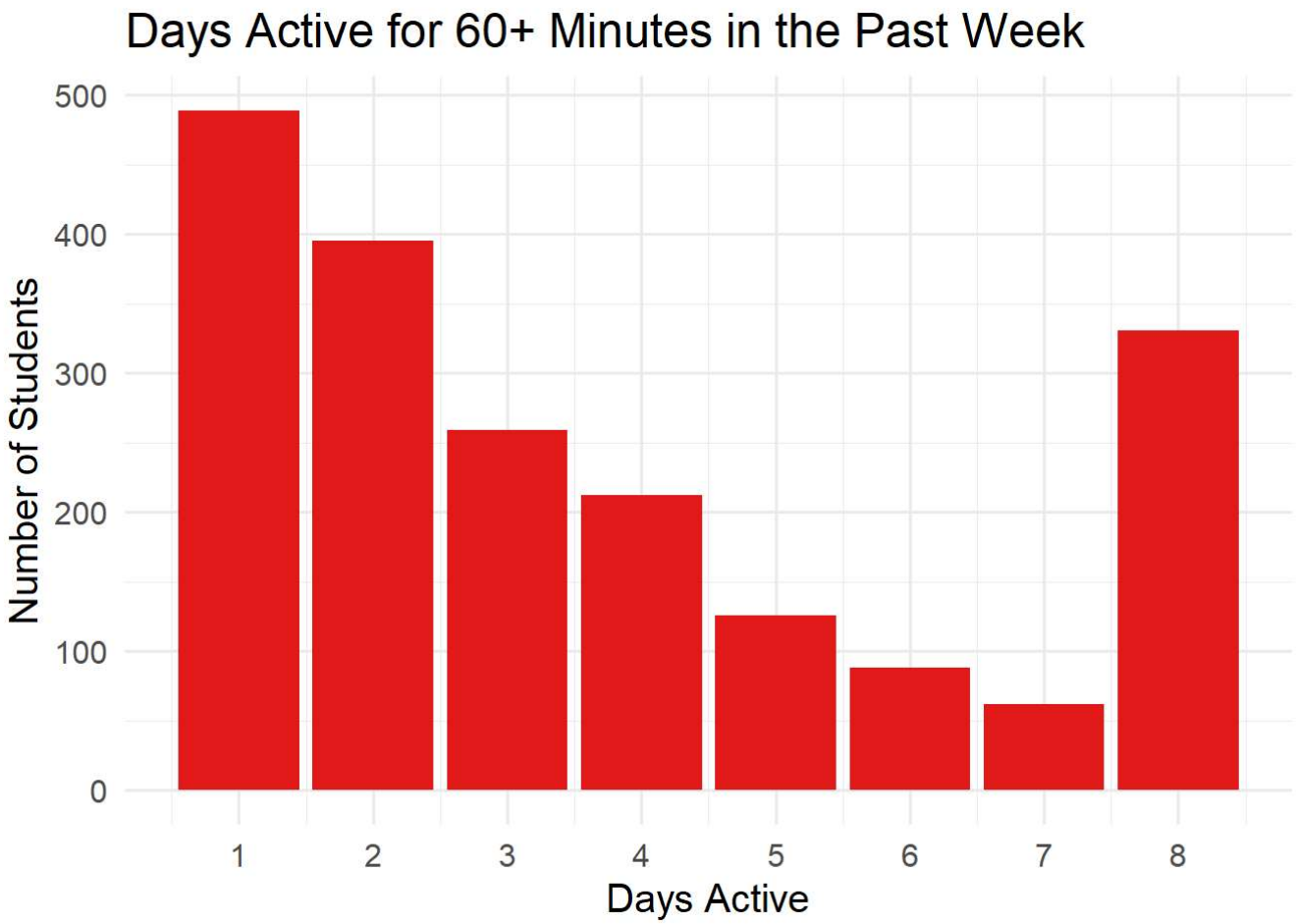


The bar chart shows the distribution of students by age and sex. It provides a comparison between male and female students across different age groups.

```
# Bar charts for active days(60mins+) in the past week
bar_days_active <- ggplot(obesity_ghana, aes(x = `days active 6o mins plus 7 past days`))
  geom_bar(fill = "#e41a1c") +  # 'Set1' color
  labs(title = 'Days Active for 60+ Minutes in the Past Week', x = 'Days Active', y = 'Nu
  scale_x_continuous(breaks = 1:8) +
  theme_minimal(base_size = 15)

print(bar_days_active)
```

```
## Warning: Removed 22 rows containing non-finite outside the scale range
## (`stat_count()`).
```

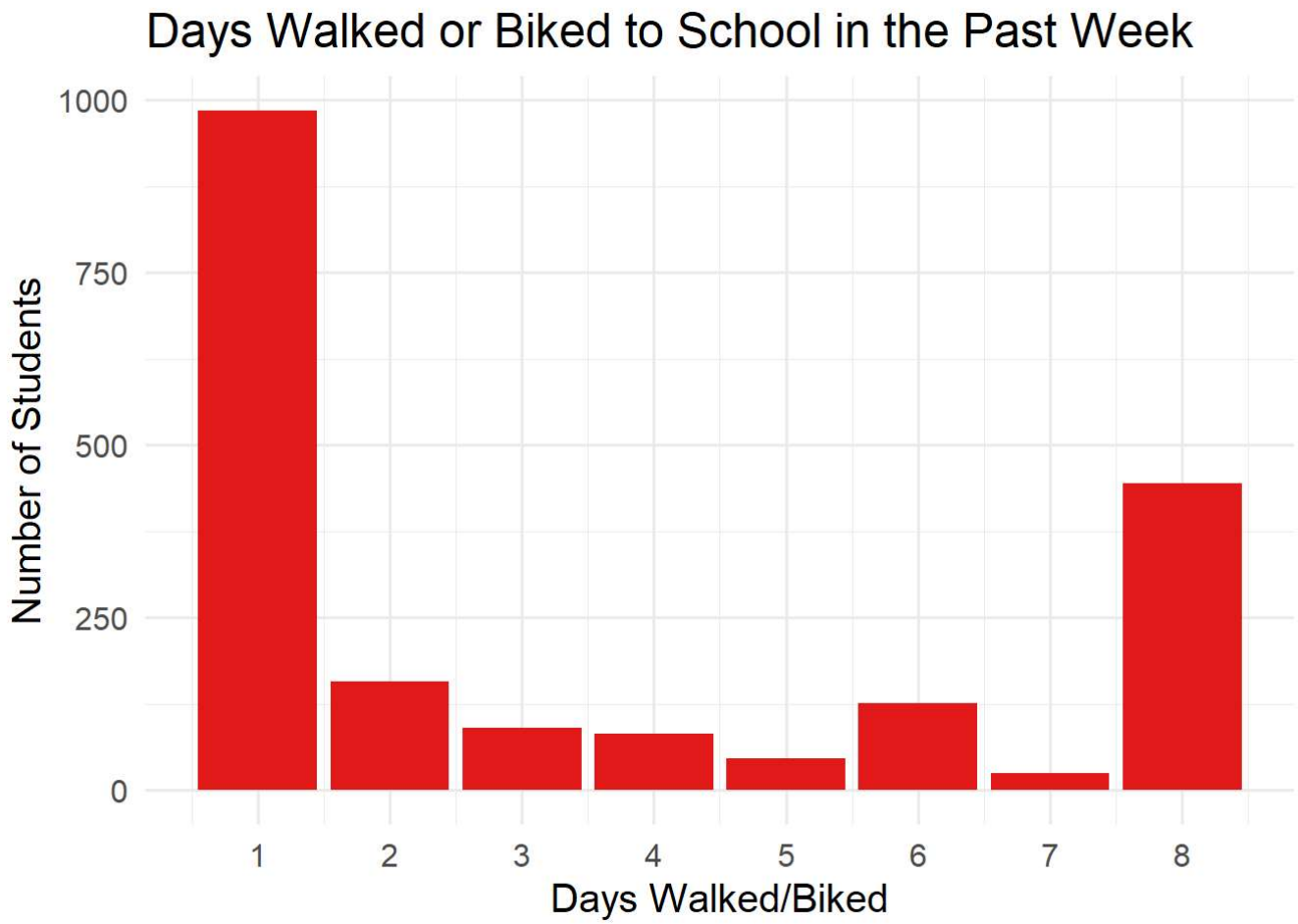


The bar chart above displays the number of days students were active for at least 60 minutes in the past week. It highlights the level of physical activity among the students.

```
# A bar chart for number of days biked/walked to school
bar_walked_biked <- ggplot(obesity_ghana, aes(x = `walk or bike to school past 7 days`))
  geom_bar(fill = "#e41a1c") +  # 'Set1' color
  labs(title = 'Days Walked or Biked to School in the Past Week', x = 'Days Walked/Biked'
  scale_x_continuous(breaks = 1:8) +
  theme_minimal(base_size = 15)

print(bar_walked_biked)
```

```
## Warning: Removed 27 rows containing non-finite outside the scale range
## (`stat_count()`).
```



The bar chart shows the number of days students walked or biked to school in the past week, providing insight into their transportation habits and physical activity.`