

# Second\_Assessment\_2080.R

arpan

2023-07-25

## ## SECOND ASSESSMENT

6 Do the following in R using ggplot2 package a. Create a dataset with the following variables: age (18-99 years), sex (male/female), educational levels (No education/Primary/ Secondary/Beyond Secondary), socioeconomic status (Low, Middle,High) and body mass index (14-38) with random 100 cases of each variable. b. Create a line chart of age variable using ggplot2 and interpret. c. Create scatter plot of age and body mass index variable using ggplot and interpret d. Create classes of body mass index variable as: <18, 18-24, 25-30, 30+ and show it as pie chart using ggplot e. Create classes of age variable as <15, 15-59 and 60+ and show it as bar diagram using ggplot.

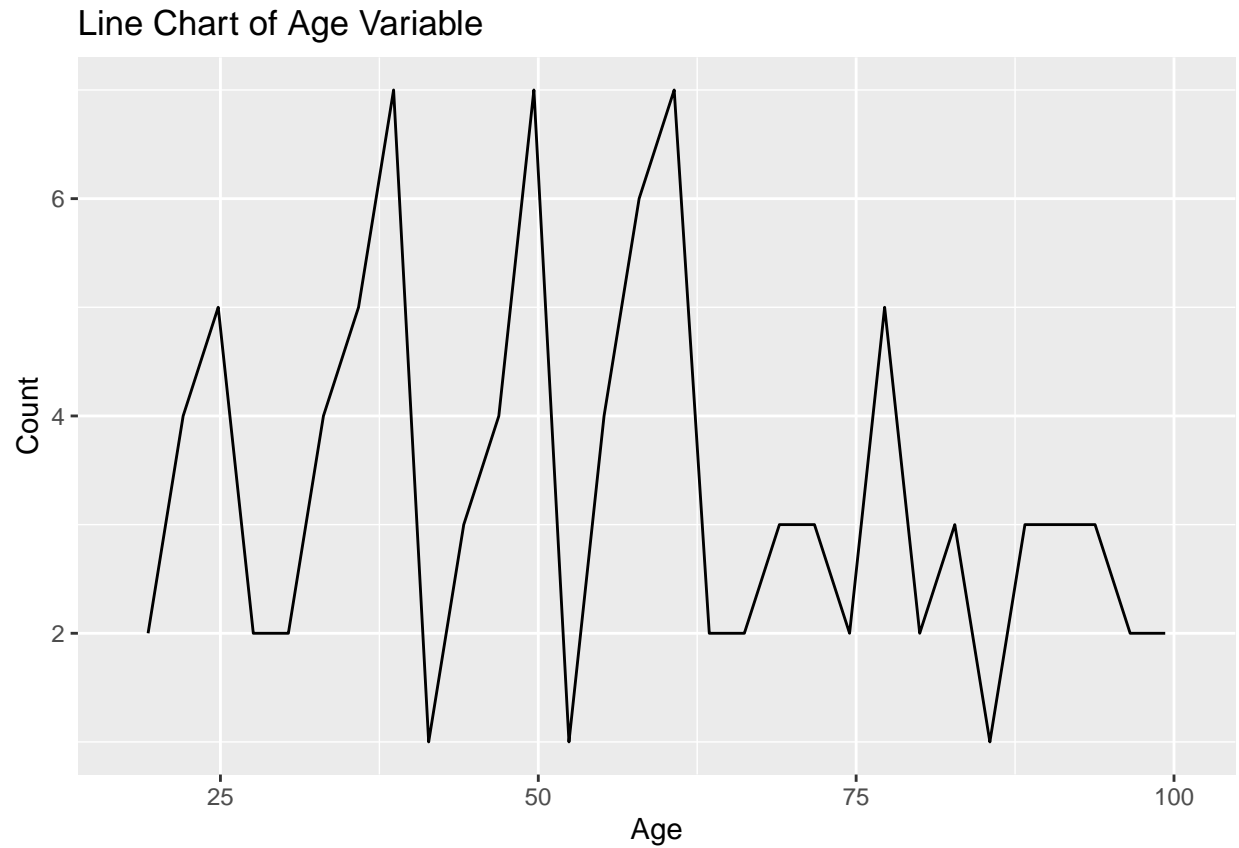
```
#Create the dataset
set.seed(07)
age <- sample(18:99, 100, replace = TRUE)
sex <- sample(c("male", "female"), 100, replace = TRUE)
educational_levels <- sample(c("No education", "Primary", "Secondary", "Beyond Secondary"), 100, replace = TRUE)
socioeconomic_status <- sample(c("Low", "Middle", "High"), 100, replace = TRUE)
body_mass_index <- runif(100, 14, 38)
data <- data.frame(age, sex, educational_levels, socioeconomic_status, body_mass_index)
head(data)
```

```
##   age    sex educational_levels socioeconomic_status body_mass_index
## 1  59 female      Beyond Secondary             Middle          21.03819
## 2  48  male           Primary                   High           14.80905
## 3  83  male      Beyond Secondary             Middle          18.91093
## 4  32  male      Beyond Secondary             Middle          32.95656
## 5  25 female           Secondary             Middle          20.11518
## 6  84  male      Beyond Secondary             High           15.67420
```

```
library(ggplot2)
# Line chart of age variable
ggplot(data, aes(x = age)) +
  geom_line(aes(y = ..count..), stat = "bin") +
  labs(title = "Line Chart of Age Variable",
       x = "Age",
       y = "Count")
```

```
## Warning: The dot-dot notation ('..count..') was deprecated in ggplot2 3.4.0.
## i Please use 'after_stat(count)' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

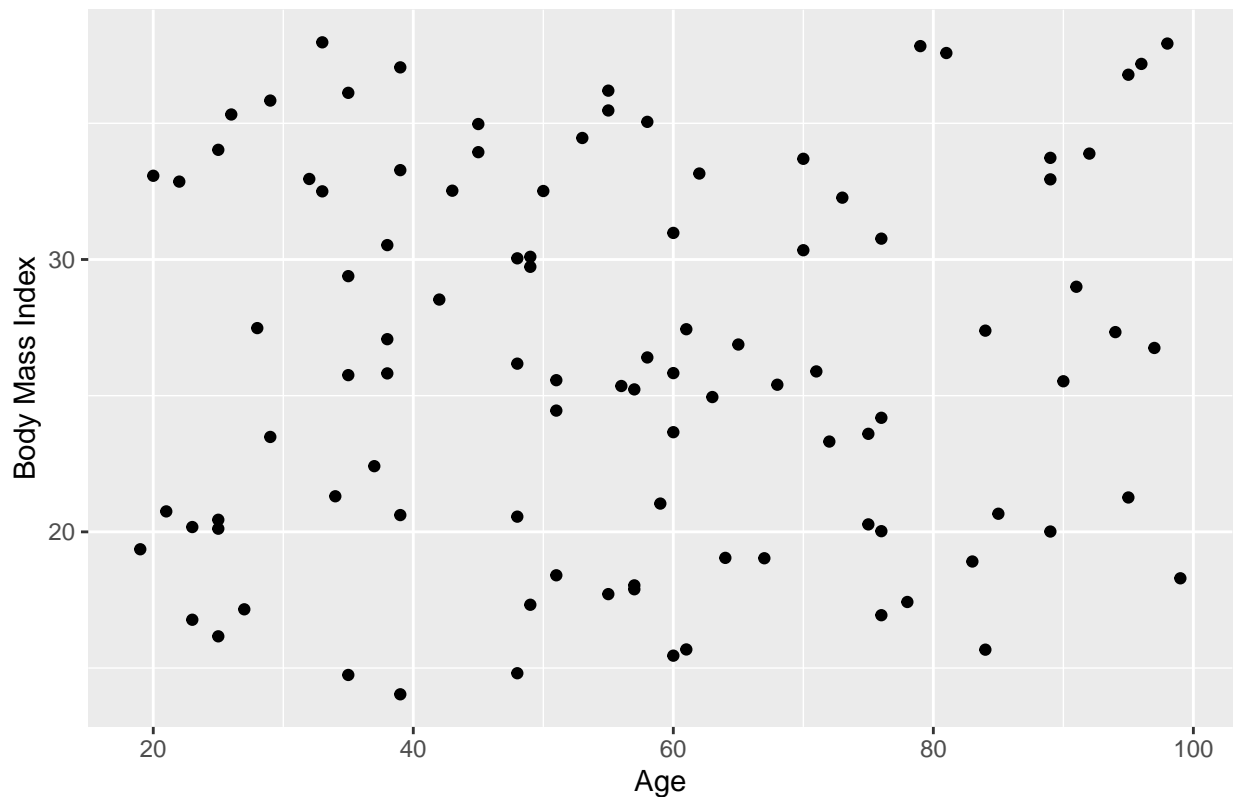
```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



*# Interpretation of the line chart: The line chart displays the trend of age across the dataset. The x-axis represents age, and the y-axis represents the count.*

```
# Scatter plot of age and body mass index variable
ggplot(data, aes(x = age, y = body_mass_index)) +
  geom_point() +
  labs(title = "Scatter Plot of Age and Body Mass Index",
       x = "Age",
       y = "Body Mass Index")
```

### Scatter Plot of Age and Body Mass Index

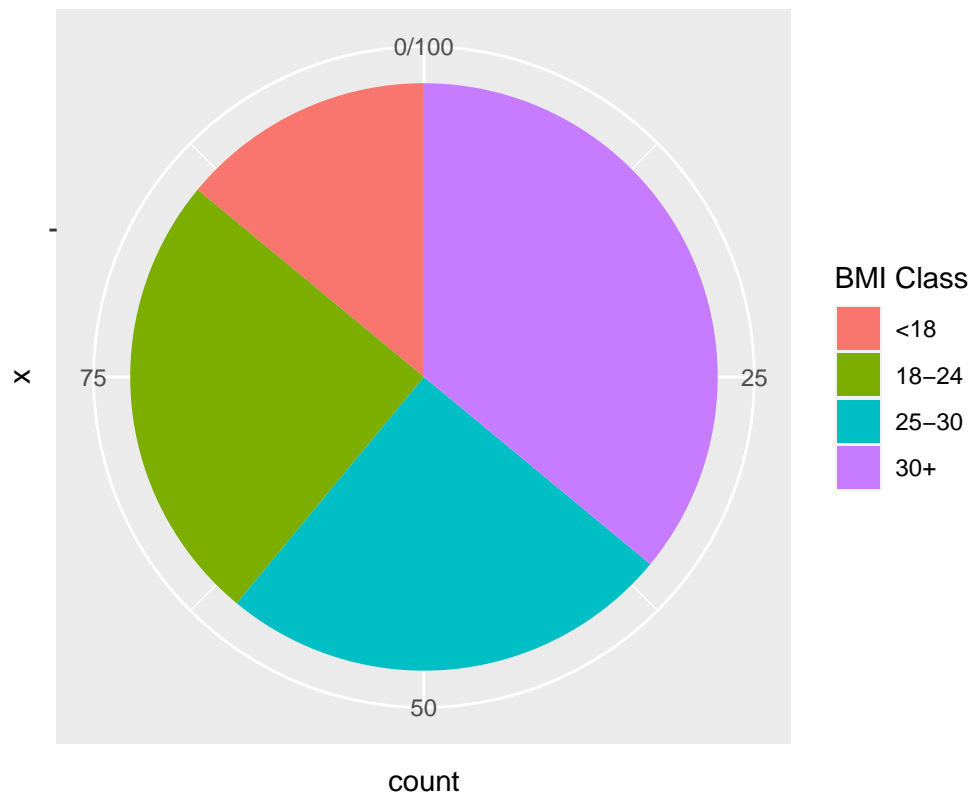


*#Interpretation of the scatter plot: The scatter plot shows the relationship between age and body mass*

```
# Create classes of body mass index variable
data$BMI_class <- cut(data$body_mass_index, breaks = c(0, 18, 24, 30, Inf), labels = c("<18", "18-24", "25-30", "31-40", "41-50", "51-60", "61-70", "71-80", "81-90", "91-100"))

# Pie chart of body mass index classes
ggplot(data, aes(x = "", fill = BMI_class)) +
  geom_bar(width = 1) +
  coord_polar("y", start = 0) +
  labs(title = "Pie Chart of Body Mass Index Classes",
       fill = "BMI Class")
```

### Pie Chart of Body Mass Index Classes

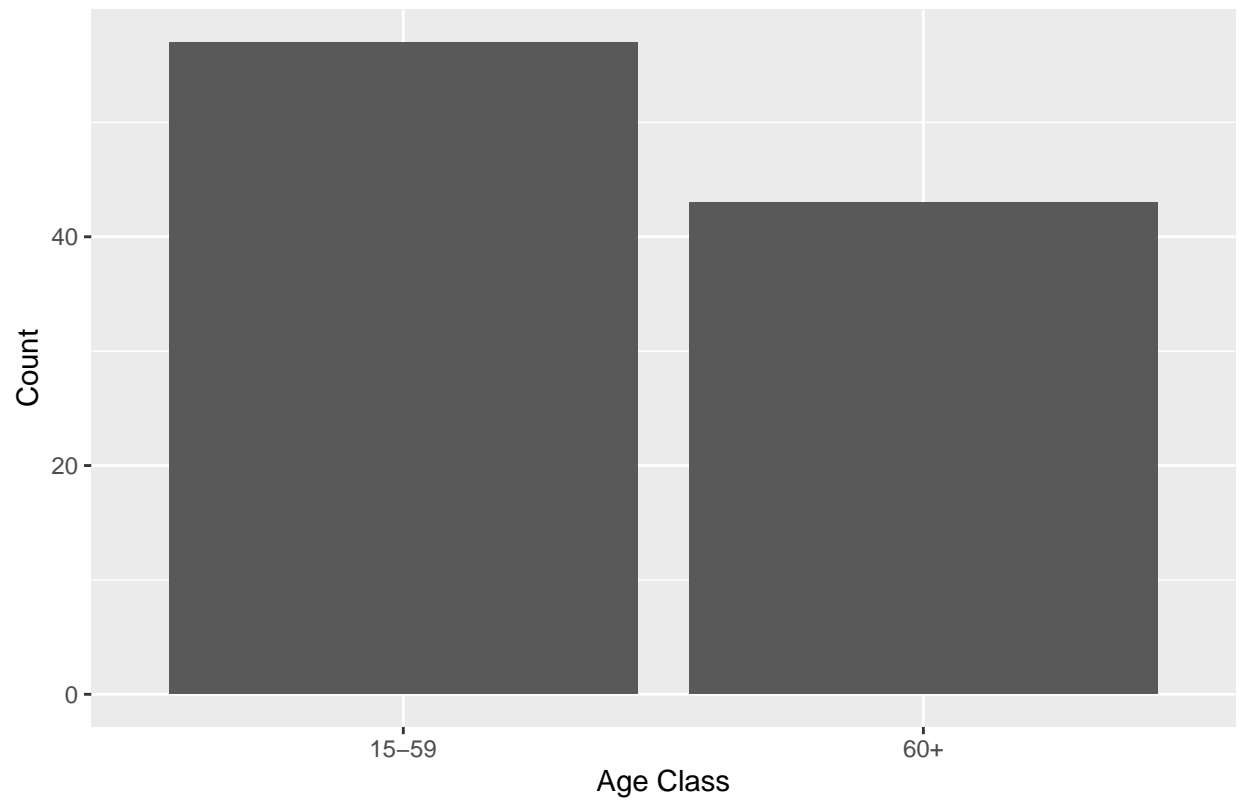


*#Interpretation of the pie chart: The pie chart represents the distribution of individuals across different BMI classes.*

```
# Create classes of age variable
data$age_class <- cut(data$age, breaks = c(0, 15, 59, Inf), labels = c("<15", "15-59", "60+"))

# Bar diagram of age classes
ggplot(data, aes(x = age_class)) +
  geom_bar() +
  labs(title = "Bar Diagram of Age Classes",
       x = "Age Class",
       y = "Count")
```

Bar Diagram of Age Classes



*#Interpretation of the bar diagram: The bar diagram shows the distribution of individuals across differ*