

# Assignment 1

Anushka Rampay

2024-02-04

```
# Source from: https://www.kaggle.com/datasets/mirichoi0218/insurance/data

# Importing the data set
library(readxl)
Insurance_data <- read_excel("C:/Users/rampa/OneDrive/Desktop/Insurance_data.xlsx")
head(Insurance_data)
```

```
## # A tibble: 6 x 7
##   age sex    bmi children smoker region    charges
##   <dbl> <chr> <dbl>    <dbl> <chr>   <chr>    <dbl>
## 1    19 female  27.9         0 yes    southwest 16885.
## 2    18 male   33.8         1 no     southeast  1726.
## 3    28 male   33          3 no     southeast  4449.
## 4    33 male   22.7         0 no     northwest 21984.
## 5    32 male   28.9         0 no     northwest  3867.
## 6    31 female 25.7         0 no     southeast  3757.
```

```
# Display descriptive statistics for quantitative variables
summary(Insurance_data$age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    18.00   27.00   39.00   39.21   51.00   64.00
```

```
summary(Insurance_data$bmi)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    15.96   26.30   30.40   30.66   34.69   53.13
```

```
# Display descriptive statistics for categorical variables
table(Insurance_data$sex)
```

```
##
## female    male
##     662     676
```

```
table(Insurance_data$smoker)
```

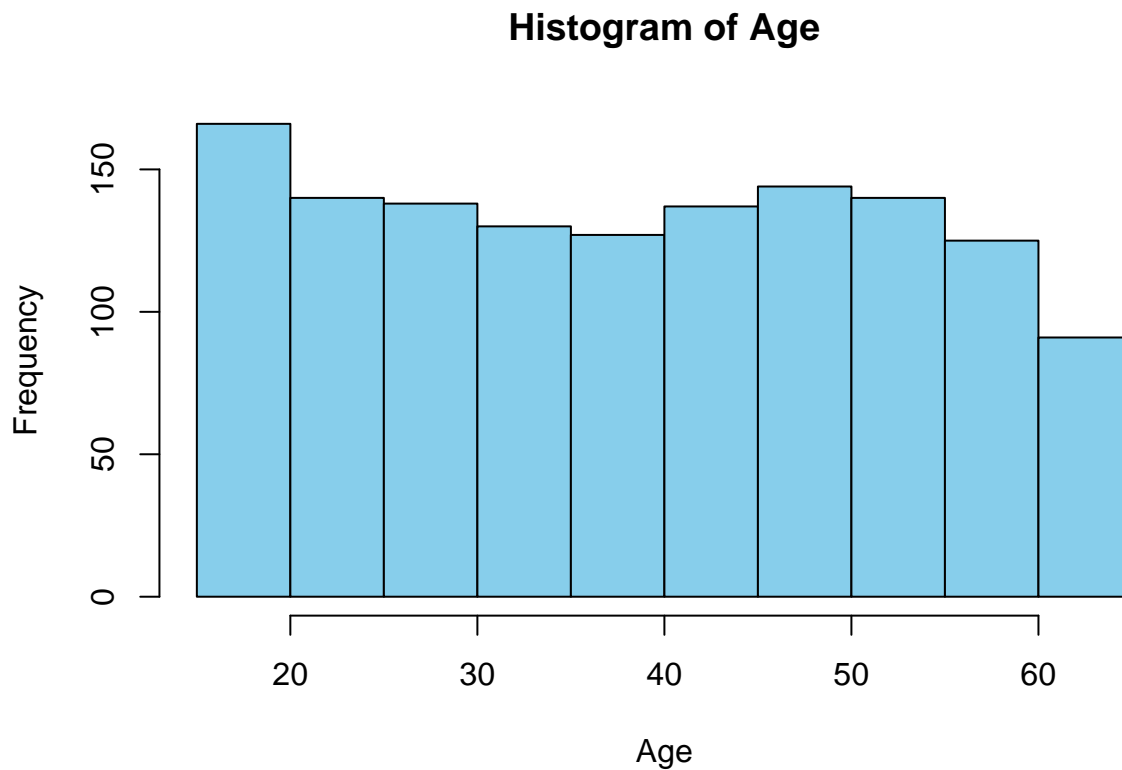
```
##
## no  yes
## 1064 274
```

```
#Transformation of one variable
```

```
Insurance_data$age_squared <- Insurance_data$age^2
```

```
# Plot a histogram for the 'age' variable
```

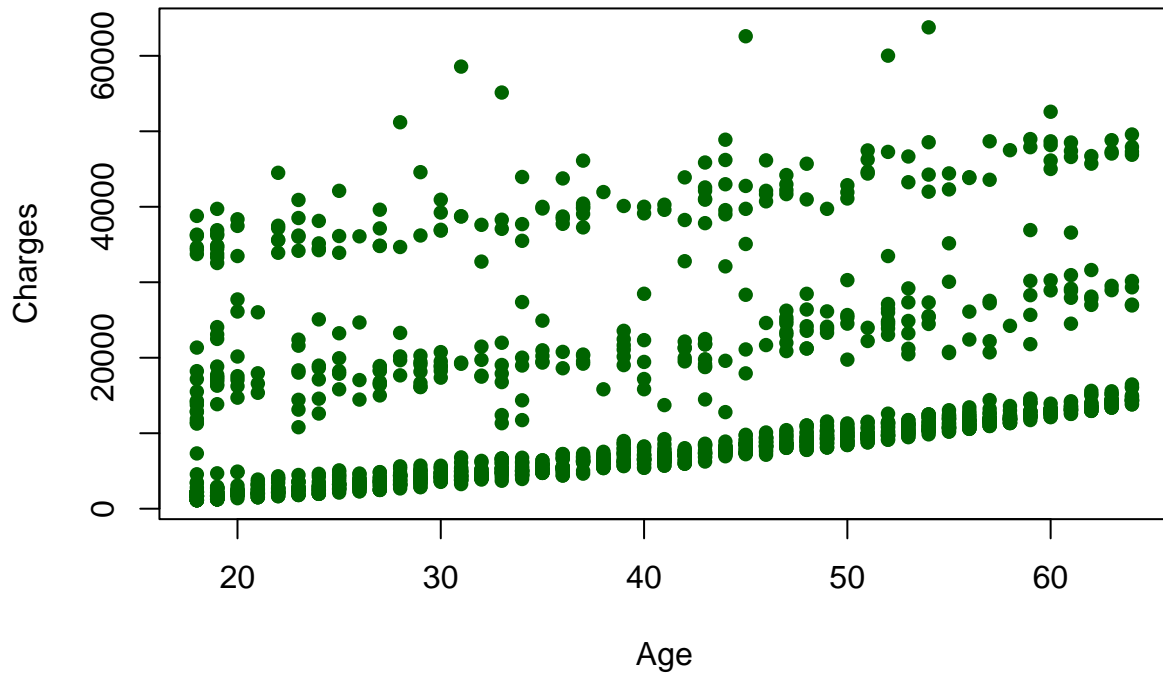
```
hist(Insurance_data$age, main = "Histogram of Age", xlab = "Age", col = "skyblue", border = "black")
```



```
# Scatterplot of age against charges
```

```
plot(Insurance_data$age, Insurance_data$charges,  
     main = "Scatterplot: Age vs Charges",  
     xlab = "Age", ylab = "Charges",  
     col = "darkgreen", pch = 16)
```

**Scatterplot: Age vs Charges**



### Summary

As you get older, insurance costs go up. Different regions, like the Northwest, might have higher charges. Having kids or a higher BMI also makes insurance more expensive. Sometimes, younger people pay a lot due to factors like high BMI, having children, and living in the Northwest.