

Obesity

```
set.seed(1)
```

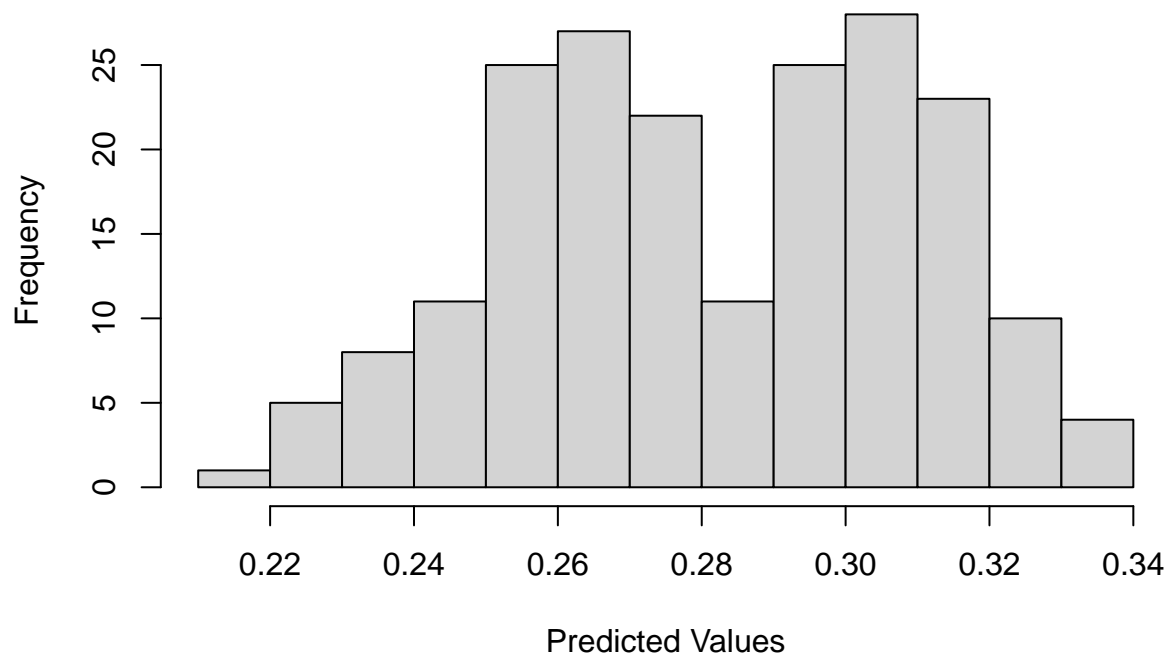
R Markdown

Dependent Variable - Obesity Rate Independent Variables - Region, Income, Poverty Rate, Real.GDP.Growth

```
obesity <- read.csv("Obesity_GDP_PanelData.csv")
model_obesity <- lm(formula = Adult.Obesity ~ Region + Average.Income + Poverty.Rate + Real.GDP.Growth,
                     data = obesity)
summary(model_obesity)
```

```
##
## Call:
## lm(formula = Adult.Obesity ~ Region + Average.Income + Poverty.Rate +
##     Real.GDP.Growth, data = obesity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.053968 -0.013476  0.001056  0.013598  0.053682
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.231e-01  3.367e-02  12.567  < 2e-16 ***
## RegionGreat Lakes Region  3.497e-02  6.427e-03   5.442 1.62e-07 ***
## RegionMideast Region    2.621e-02  6.536e-03   4.010 8.73e-05 ***
## RegionNew England Region  9.602e-03  6.245e-03   1.538 0.125833
## RegionPlains Region     2.958e-02  6.096e-03   4.852 2.54e-06 ***
## RegionRocky Mountain Region -2.106e-02  6.705e-03  -3.141 0.001953 **
## RegionSoutheast Region   4.323e-02  6.136e-03   7.045 3.35e-11 ***
## RegionSouthwest Region   3.316e-03  7.450e-03   0.445 0.656788
## Average.Income          -4.275e-06  6.870e-07  -6.222 3.08e-09 ***
## Poverty.Rate            -1.993e-01  1.078e-01  -1.848 0.066116 .
## Real.GDP.Growth         -3.045e-01  8.293e-02  -3.672 0.000313 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02065 on 189 degrees of freedom
## Multiple R-squared:  0.6557, Adjusted R-squared:  0.6375
## F-statistic: 35.99 on 10 and 189 DF,  p-value: < 2.2e-16
# Histogram of Outcome
predicted_values <- predict(model_obesity)
hist(predicted_values, main = "Histogram of Predicted Values", xlab = "Predicted Values")
```

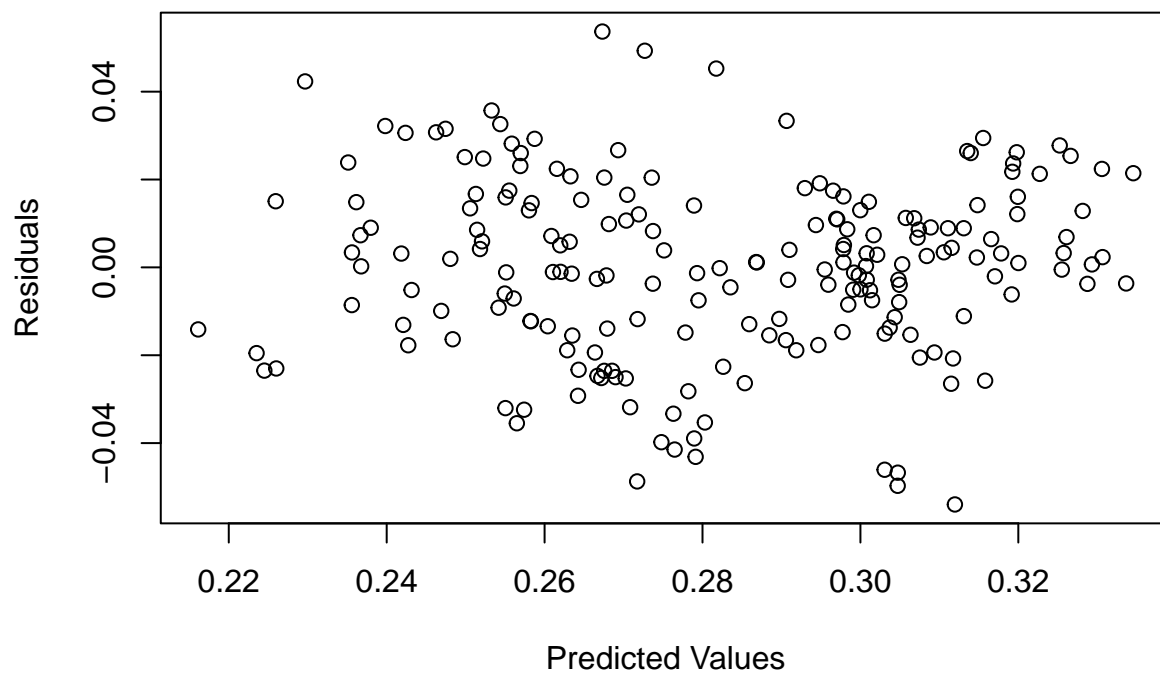
Histogram of Predicted Values



Measure of Model Fit

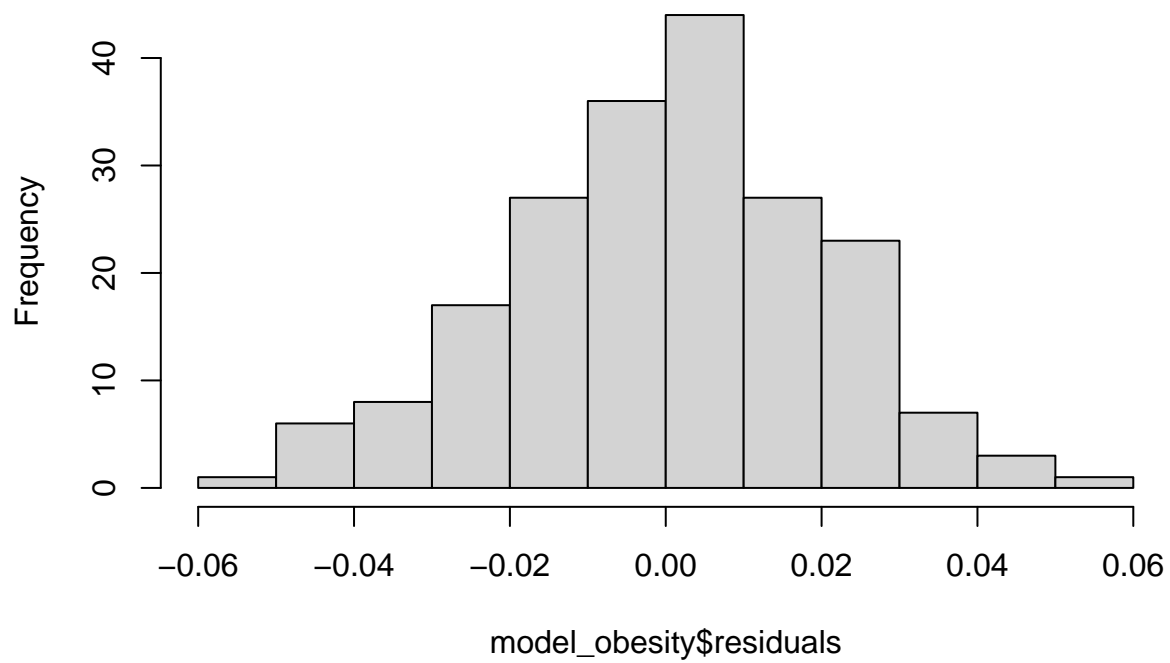
```
plot(x = predicted_values, y = model_obesity$residuals, main = "Fitted vs. Observed", xlab = "Predicted
```

Fitted vs. Observed



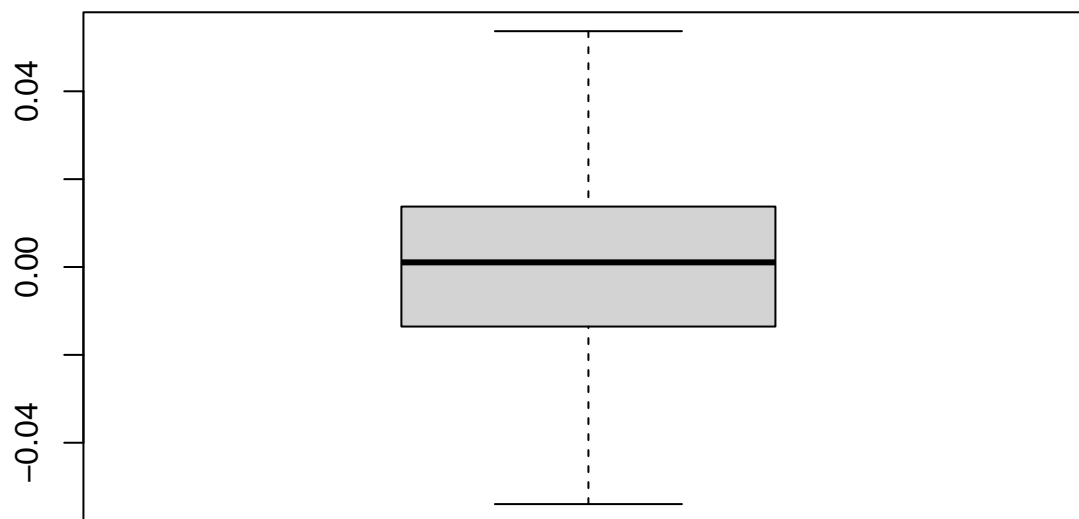
```
hist(model_obesity$residuals)
```

Histogram of model_obesity\$residuals



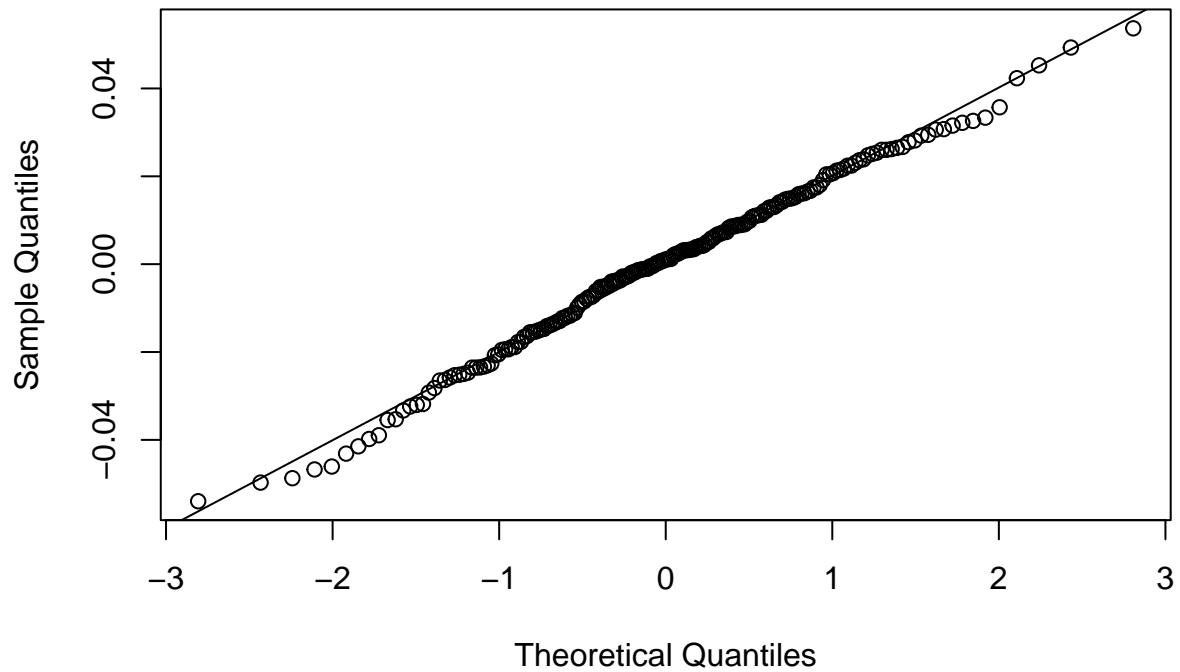
```
boxplot(model_obesity$residuals, main = "Boxplot For Residuals")
```

Boxplot For Residuals

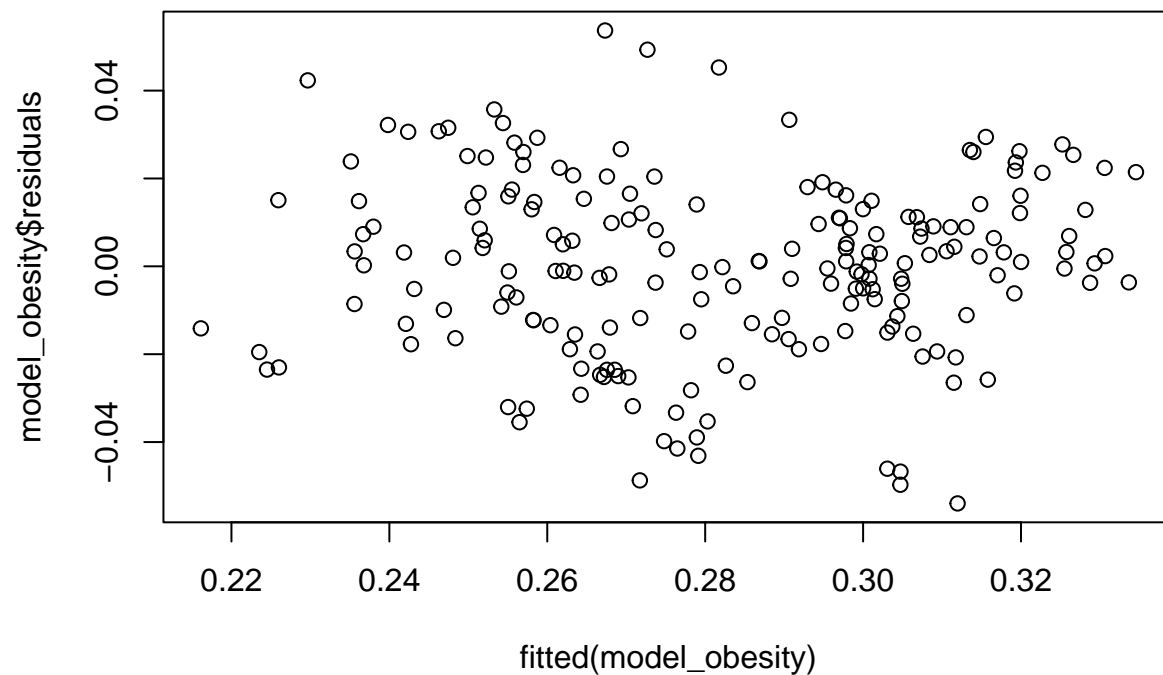


```
qqnorm(model_obesity$residuals)  
qqline(model_obesity$residuals)
```

Normal Q-Q Plot



```
plot(fitted(model_obesity), model_obesity$residuals)
```



```
cat("Average Income Mean:", mean(obesity$Average.Income), "\n")
```

```
## Average Income Mean: 30040.02
```

```
cat("Average Income Median:", median(obesity$Average.Income), "\n")
```

```
## Average Income Median: 29444.9
```

```

cat("Average Income Range:" ,range(obesity$Average.Income), "\n")

## Average Income Range: 20921.02 41961.14
cat("Average Income Standard Deviation:", sd(obesity$Average.Income), "\n")

## Average Income Standard Deviation: 4454.894
print("Average Income Quantile Ranges:")

## [1] "Average Income Quantile Ranges:"
quantile(obesity$Average.Income)

##          0%          25%          50%          75%          100%
## 20921.02 26614.34 29444.90 32952.50 41961.14
cat("Poverty Rate Mean:", mean(obesity$Poverty.Rate), "\n")

## Poverty Rate Mean: 0.1389031
cat("Poverty Rate Median:", median(obesity$Poverty.Rate), "\n")

## Poverty Rate Median: 0.1356303
cat("Poverty Rate Range:" ,range(obesity$Poverty.Rate), "\n")

## Poverty Rate Range: 0.0729647 0.2199588
cat("Poverty Rate Standard Deviation:", sd(obesity$Poverty.Rate), "\n")

## Poverty Rate Standard Deviation: 0.030797
print("Poverty Rate Quantile Ranges:")

## [1] "Poverty Rate Quantile Ranges:"
quantile(obesity$Poverty.Rate)

##          0%          25%          50%          75%          100%
## 0.0729647 0.1133144 0.1356303 0.1584019 0.2199588
cat("Real GDP Growth Mean:", mean(obesity$Real.GDP.Growth), "\n")

## Real GDP Growth Mean: 0.0159575
cat("Real GDP Growth Median:", median(obesity$Real.GDP.Growth), "\n")

## Real GDP Growth Median: 0.01655
cat("Real GDP Growth Range:" ,range(obesity$Real.GDP.Growth), "\n")

## Real GDP Growth Range: -0.0707 0.0767
cat("Real GDP Growth Standard Deviation:", sd(obesity$Real.GDP.Growth), "\n")

## Real GDP Growth Standard Deviation: 0.01834917
print("Real GDP Growth Quantile Ranges:")

## [1] "Real GDP Growth Quantile Ranges:"
quantile(obesity$Real.GDP.Growth)

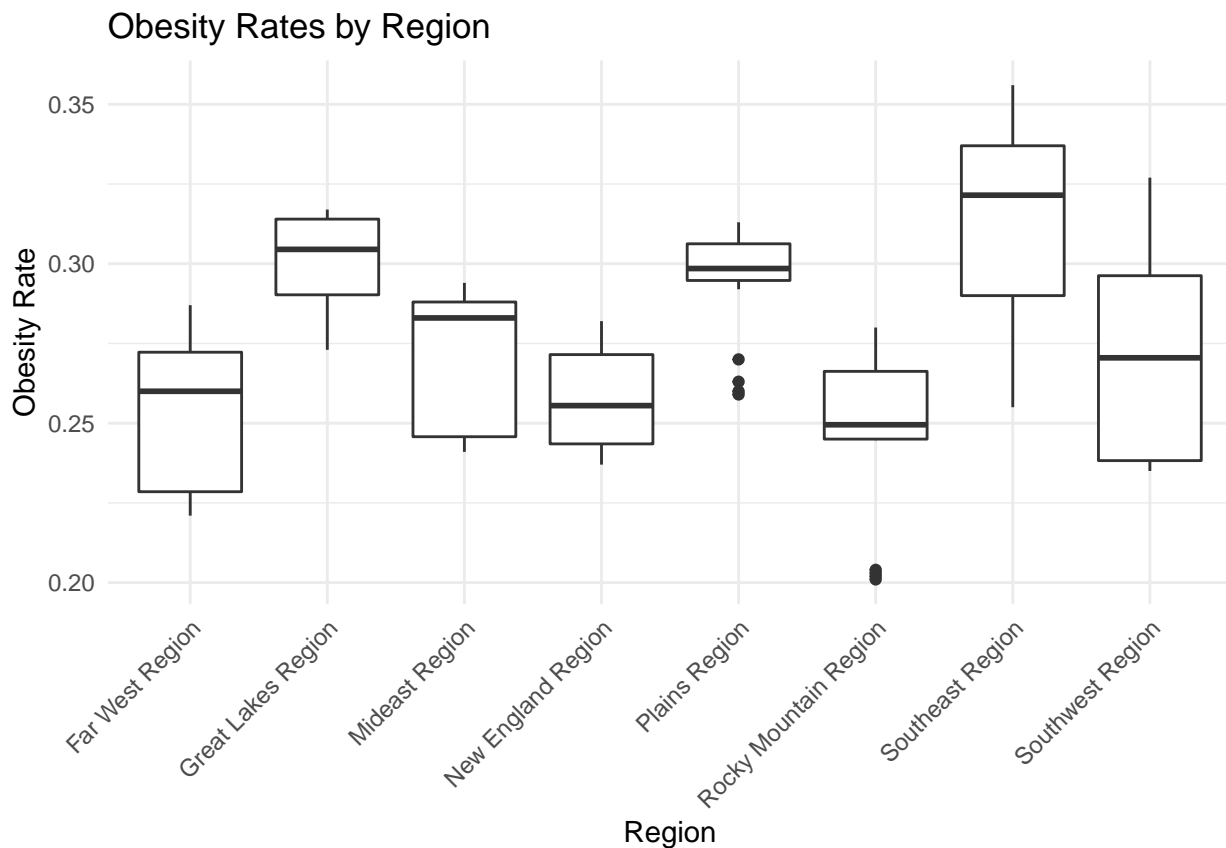
```

```
##           0%      25%      50%      75%      100%
## -0.070700  0.004775  0.016550  0.026175  0.076700
```

```
region_count <- obesity %>%
  group_by(Region) %>%
  summarise(Count = n(), Proportion = n()/200)
print(region_count)
```

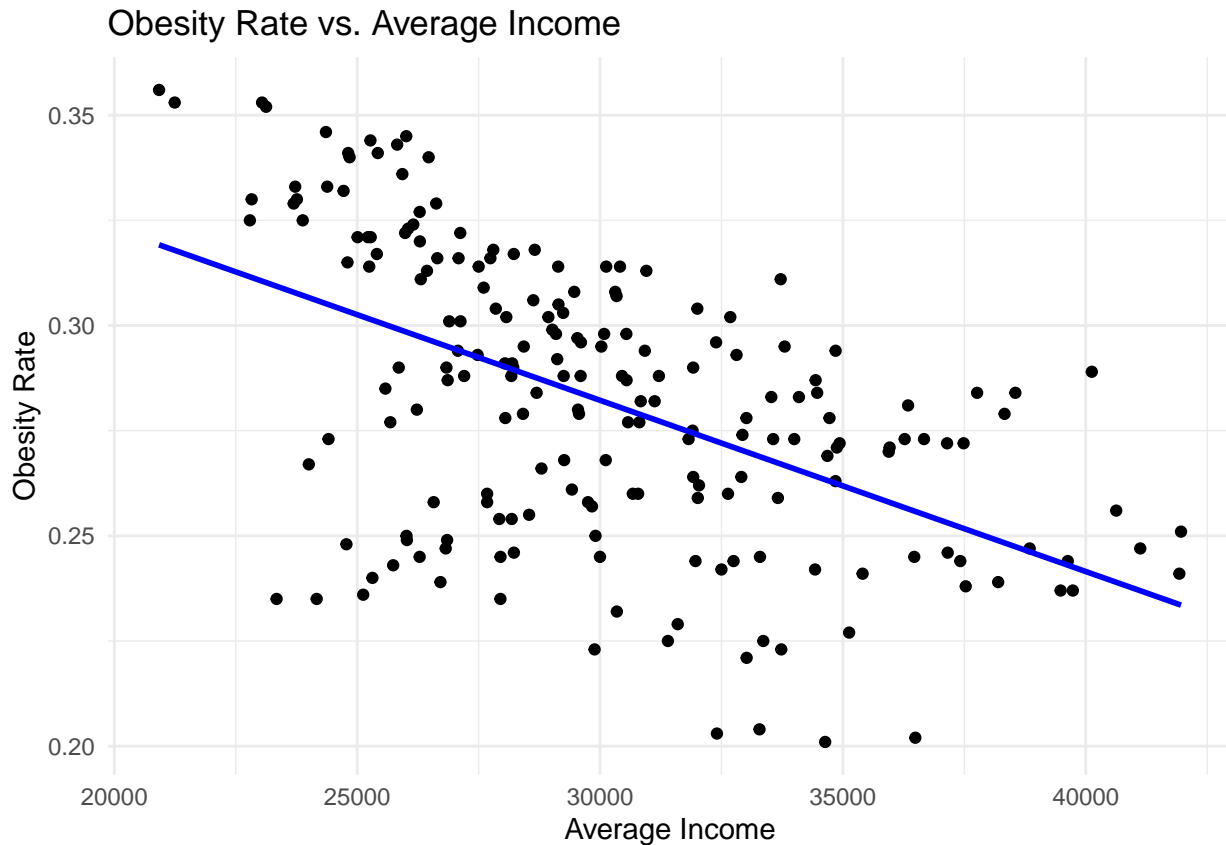
```
## # A tibble: 8 x 3
##   Region          Count Proportion
##   <chr>          <int>     <dbl>
## 1 Far West Region      24      0.12
## 2 Great Lakes Region   20      0.1
## 3 Mideast Region       20      0.1
## 4 New England Region   24      0.12
## 5 Plains Region        28      0.14
## 6 Rocky Mountain Region 20      0.1
## 7 Southeast Region     48      0.24
## 8 Southwest Region     16      0.08
```

```
ggplot(obesity, aes(x = Region, y = Adult.Obesity)) +
  geom_boxplot() +
  labs(title = "Obesity Rates by Region",
       x = "Region",
       y = "Obesity Rate") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



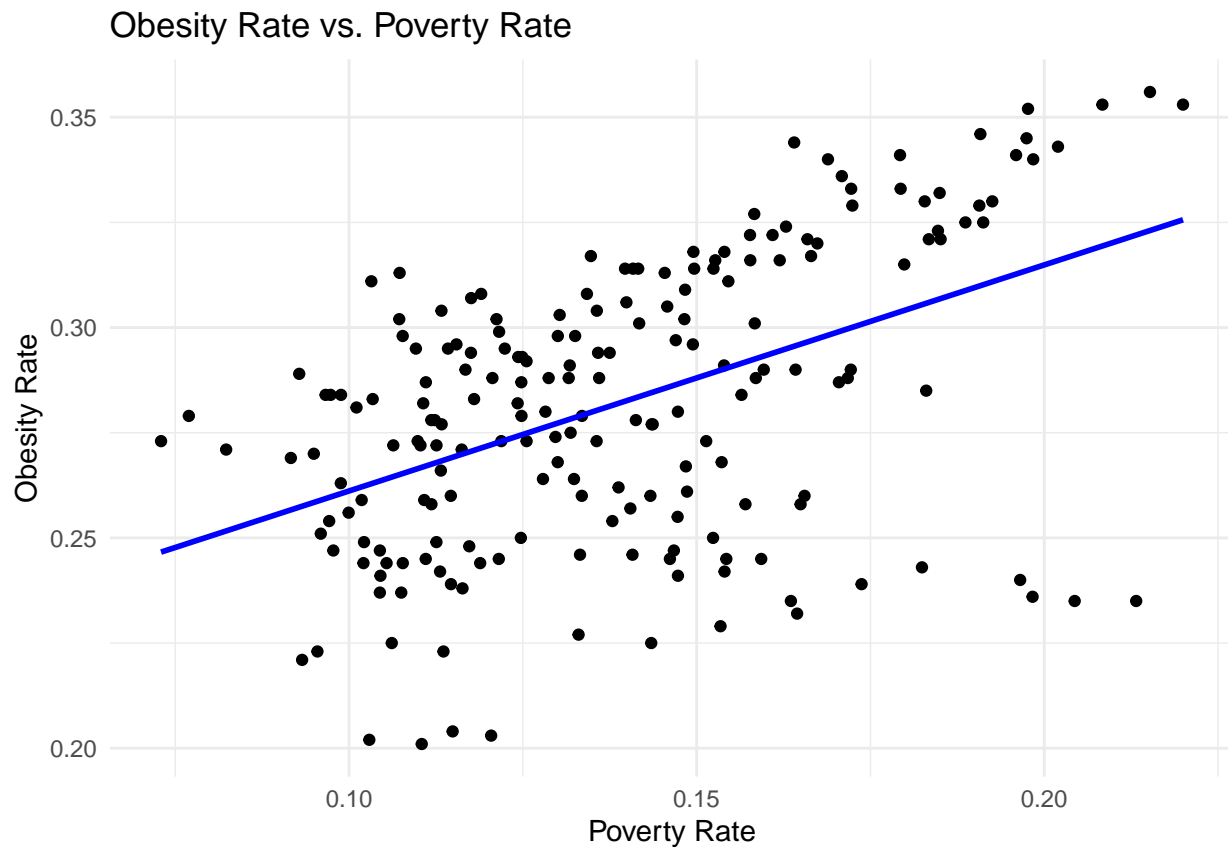
```
ggplot(obesity, aes(x = Average.Income, y = Adult.Obesity)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE, color = "blue") +
  labs(title = "Obesity Rate vs. Average Income",
        x = "Average Income",
        y = "Obesity Rate"
  ) +
  theme_minimal()
```

```
## `geom_smooth()` using formula 'y ~ x'
```



```
ggplot(obesity, aes(x = Poverty.Rate, y = Adult.Obesity)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE, color = "blue") +
  labs(title = "Obesity Rate vs. Poverty Rate",
        x = "Poverty Rate",
        y = "Obesity Rate") +
  theme_minimal()
```

```
## `geom_smooth()` using formula 'y ~ x'
```



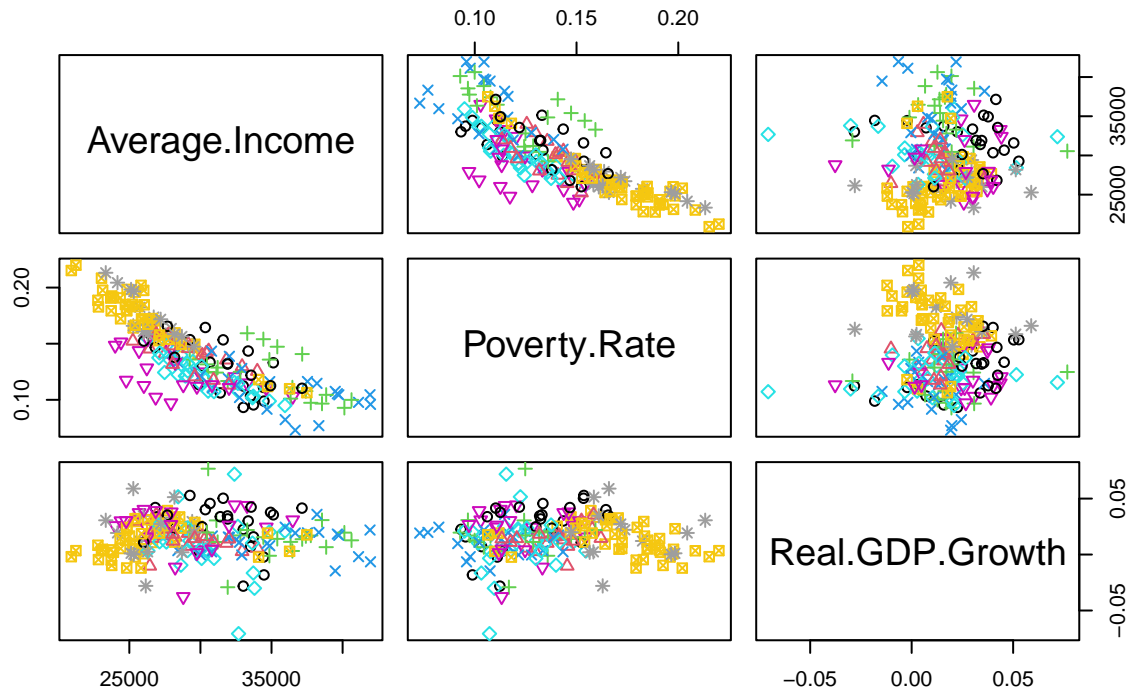
```
ggplot(obesity, aes(x = Real.GDP.Growth, y = Adult.Obesity)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "blue") +  
  labs(title = "Obesity Rate vs. Real GDP Growth",  
        x = "Real GDP Growth",  
        y = "Obesity Rate") +  
  theme_minimal()
```

```
## `geom_smooth()` using formula 'y ~ x'
```

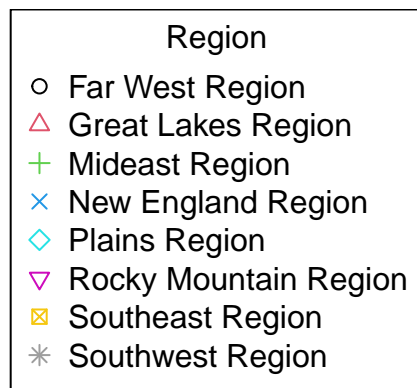



```
pairs(obesity[, c("Average.Income", "Poverty.Rate", "Real.GDP.Growth")],  
      col = as.numeric(factor(obesity$Region)),  
      pch = as.numeric(factor(obesity$Region)),  
      main = "Scatterplot Matrix with Color-Coded 'Region'")
```

Scatterplot Matrix with Color-Coded 'Region'



```
legend_plot <- plot(1, type = "n", axes = FALSE, xlab = "", ylab = "")
legend("topright", legend = levels(factor(obesity$Region)), col = 1:8, pch = 1:8, title = "Region")
```



```
legend_plot
```

```
## NULL
```