Visualization

```
library(ggplot2)
library(ggcorrplot)
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
library(GGally)
```

```
Registered S3 method overwritten by 'GGally':
  method from
  +.gg   ggplot2
```

```
library(readr)

Air_Quality <- read_csv("/Users/yegireddimounika/Desktop/AirQuality/Air_Quality_CleanedData.c
```

```
Rows: 5811 Columns: 15
```

```
-- Column specification --------------------------------------------------------
Delimiter: ","
dbl (15): RecordID, AQI, PM10, PM2_5, NO2, SO2, O3, Temperature, Humidity, W...

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
cor_matrix <- cor(Air_Quality %>% select(where(is.numeric)))
print("Correlation matrix:")
```

```
[1] "Correlation matrix:"
```

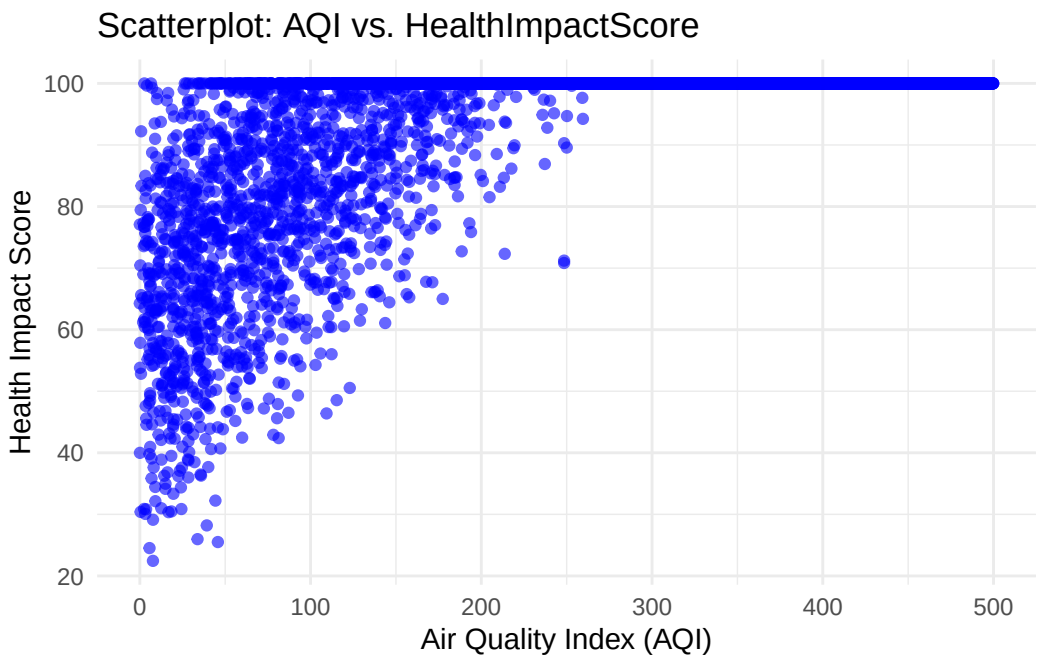
```r
print(cor_matrix)
```

```
                      RecordID          AQI         PM10        PM2_5
RecordID           1.000000000 -0.004489943  0.021973749  0.024451129
AQI               -0.004489943  1.000000000  0.023013622  0.005685949
PM10               0.021973749  0.023013622  1.000000000 -0.011864553
PM2_5              0.024451129  0.005685949 -0.011864553  1.000000000
NO2                0.010630472  0.007928765  0.008244873  0.004768957
SO2               -0.022459670 -0.006566274 -0.007535980  0.015298137
O3                -0.002148044  0.002438408 -0.003676096  0.006542829
Temperature        0.003766024  0.003817746 -0.017929193 -0.001621353
Humidity          -0.001313263 -0.004910000 -0.017923236  0.008076406
WindSpeed          0.018098854 -0.018521595 -0.018886610  0.006612453
RespiratoryCases  -0.014544781  0.009699325 -0.001157492  0.025900188
CardiovascularCases -0.030537083  0.008742064  0.015392986  0.017305732
HospitalAdmissions  -0.008298656 -0.013775159 -0.009303021 -0.023402070
HealthImpactScore   0.015785839  0.614070657  0.182695553  0.220068353
HealthImpactClass   0.198229480 -0.347836491 -0.108376186 -0.151839387
                             NO2           SO2            O3   Temperature
RecordID              0.010630472 -0.0224596699 -0.0021480443  0.0037660242
AQI                   0.007928765 -0.0065662739  0.0024384080  0.0038177462
PM10                  0.008244873 -0.0075359802 -0.0036760957 -0.0179291931
PM2_5                 0.004768957  0.0152981371  0.0065428287 -0.0016213526
NO2                   1.000000000 -0.0204700369 -0.0155040266  0.0068505624
SO2                  -0.020470037  1.0000000000 -0.0050778599 -0.0230526898
O3                   -0.015504027 -0.0050778599  1.0000000000 -0.0006949235
Temperature           0.006850562 -0.0230526898 -0.0006949235  1.0000000000
Humidity             -0.011084329  0.0023339281  0.0058063172  0.0018707525
WindSpeed            -0.002129010 -0.0001418513 -0.0037970762  0.0019169481
```

```
RespiratoryCases     0.019460342  0.0104201197 -0.0093085643  0.0091640545
CardiovascularCases -0.005325073 -0.0028767099  0.0034171483  0.0086916721
HospitalAdmissions  -0.003270290 -0.0101642014 -0.0132012040  0.0081815174
HealthImpactScore    0.124513760  0.0160319559  0.1575436976 -0.0103988463
HealthImpactClass   -0.090551014 -0.0088077065 -0.1098868509  0.0058355727
                         Humidity     WindSpeed RespiratoryCases
RecordID            -0.0013132627  0.0180988545     -0.014544781
AQI                 -0.0049100004 -0.0185215950      0.009699325
PM10                -0.0179232355 -0.0188866104     -0.001157492
PM2_5                0.0080764059  0.0066124530      0.025900188
NO2                 -0.0110843286 -0.0021290096      0.019460342
SO2                  0.0023339281 -0.0001418513      0.010420120
O3                   0.0058063172 -0.0037970762     -0.009308564
Temperature          0.0018707525  0.0019169481      0.009164055
Humidity             1.0000000000  0.0217115995      0.009634135
WindSpeed            0.0217115995  1.0000000000     -0.009379147
RespiratoryCases     0.0096341354 -0.0093791471      1.000000000
CardiovascularCases -0.0252992281  0.0040929108     -0.002621434
HospitalAdmissions  -0.0039658683 -0.0059408872      0.009679222
HealthImpactScore    0.0005699451 -0.0252221576      0.019267742
HealthImpactClass   -0.0066464217  0.0057598636     -0.008536223
                    CardiovascularCases HospitalAdmissions HealthImpactScore
RecordID                    -0.030537083       -0.008298656      0.0157858385
AQI                          0.008742064       -0.013775159      0.6140706575
PM10                         0.015392986       -0.009303021      0.1826955533
PM2_5                        0.017305732       -0.023402070      0.2200683528
NO2                         -0.005325073       -0.003270290      0.1245137597
SO2                         -0.002876710       -0.010164201      0.0160319559
O3                           0.003417148       -0.013201204      0.1575436976
Temperature                  0.008691672        0.008181517     -0.0103988463
Humidity                    -0.025299228       -0.003965868      0.0005699451
WindSpeed                    0.004092911       -0.005940887     -0.0252221576
RespiratoryCases            -0.002621434        0.009679222      0.0192677425
CardiovascularCases          1.000000000       -0.035631615     -0.0088574331
HospitalAdmissions          -0.035631615        1.000000000     -0.0293041240
HealthImpactScore           -0.008857433       -0.029304124      1.0000000000
HealthImpactClass            0.001848024        0.014873113     -0.6556114838
                    HealthImpactClass
RecordID                    0.198229480
AQI                        -0.347836491
PM10                       -0.108376186
PM2_5                      -0.151839387
NO2                        -0.090551014
```

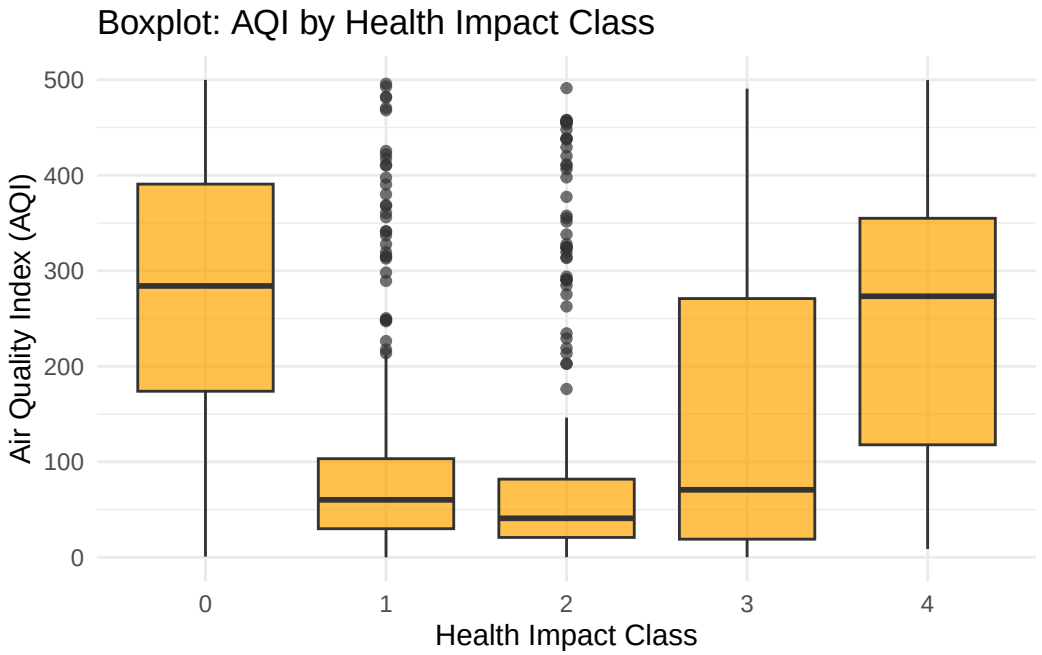
```
SO2                      -0.008807706
O3                       -0.109886851
Temperature               0.005835573
Humidity                 -0.006646422
WindSpeed                 0.005759864
RespiratoryCases         -0.008536223
CardiovascularCases       0.001848024
HospitalAdmissions        0.014873113
HealthImpactScore        -0.655611484
HealthImpactClass         1.000000000
```

```r
#Scatterplot: AQI vs. HealthImpactScore
ggplot(Air_Quality, aes(x = AQI, y = HealthImpactScore)) +
  geom_point(color = "blue", alpha = 0.6) +
  ggtitle("Scatterplot: AQI vs. HealthImpactScore") +
  xlab("Air Quality Index (AQI)") +
  ylab("Health Impact Score") +
  theme_minimal()
```

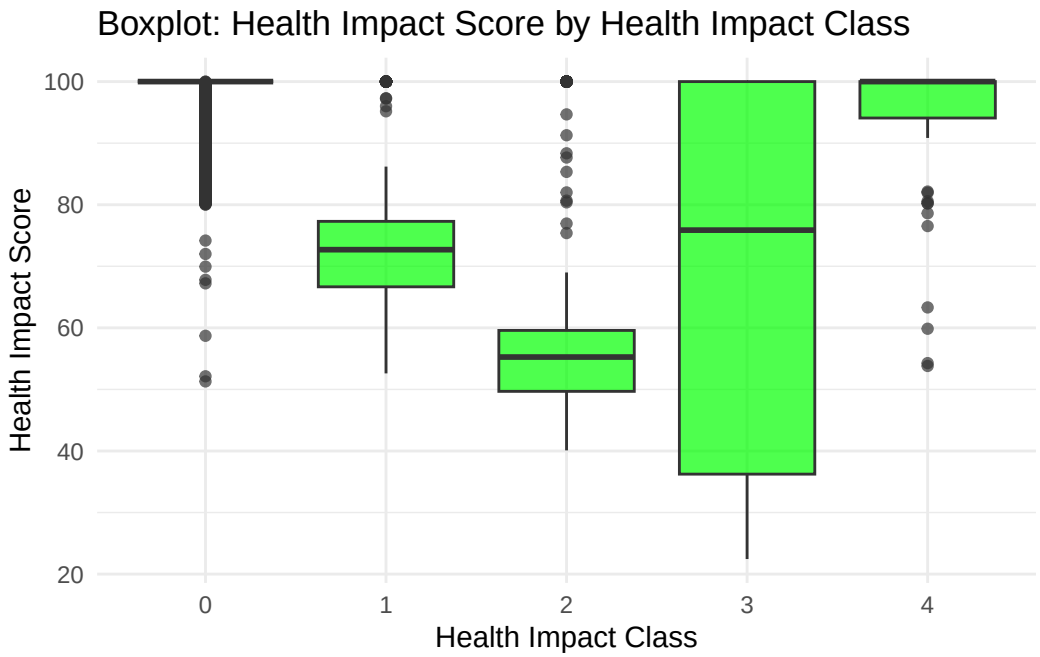



```r
# Boxplot: AQI by HealthImpactClass
ggplot(Air_Quality, aes(x = as.factor(HealthImpactClass), y = AQI)) +
  geom_boxplot(fill = "orange", alpha = 0.7) +
```

```
  ggtitle("Boxplot: AQI by Health Impact Class") +
  xlab("Health Impact Class") +
  ylab("Air Quality Index (AQI)") +
  theme_minimal()
```




```
# Boxplot: HealthImpactScore by HealthImpactClass
ggplot(Air_Quality, aes(x = as.factor(HealthImpactClass), y = HealthImpactScore)) +
  geom_boxplot(fill = "green", alpha = 0.7) +
  ggtitle("Boxplot: Health Impact Score by Health Impact Class") +
  xlab("Health Impact Class") +
  ylab("Health Impact Score") +
  theme_minimal()
```


Boxplot: Health Impact Score by Health Impact Class
Health Impact Score
100
80
60
40
20
0
1
2
3
4
Health Impact Class