

Cleaning

```
# Load necessary libraries  
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(ggplot2)  
library(caret)
```

Loading required package: lattice

```
library(readxl)  
  
Air_Quality <- read_excel("/Users/yegireddimounika/Downloads/Air_Quality.xlsx")  
  
#View  
View(Air_Quality)  
  
#Summary  
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:122.9375	1st Qu.: 75.28857	1st Qu.: 49.41759
Median :2906	Median :249.1752	Median :147.66485	Median :100.41926
Mean :2906	Mean :248.5003	Mean :148.67229	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.7322	3rd Qu.:222.48613	3rd Qu.:151.31922
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497
	NA's :9	NA's :8	NA's :11

N02	S02	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.53854	1st Qu.:24.89137	1st Qu.: 74.13083	1st Qu.: 2.482
Median :102.96266	Median :49.54601	Median :149.76082	Median :14.944
Mean :102.29337	Mean :49.47704	Mean :149.42024	Mean :14.981
3rd Qu.:151.82816	3rd Qu.:73.38877	3rd Qu.:223.45188	3rd Qu.:27.492
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963
NA's :12	NA's :18	NA's :19	NA's :8

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.00	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.54	Median :10.052435	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.61	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
NA's :16	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
#Converting all the NA values in Air Quality Index as the mean
mean_AQI <- mean(Air_Quality$AQI, na.rm = TRUE)
Air_Quality$AQI[is.na(Air_Quality$AQI)] <- mean_AQI
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.28857	1st Qu.: 49.41759
Median :2906	Median :248.9063	Median :147.66485	Median :100.41926

Mean :2906	Mean :248.5003	Mean :148.67229	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.48613	3rd Qu.:151.31922
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497
	NA's :8	NA's :11	

N02	S02	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.53854	1st Qu.:24.89137	1st Qu.: 74.13083	1st Qu.: 2.482
Median :102.96266	Median :49.54601	Median :149.76082	Median :14.944
Mean :102.29337	Mean :49.47704	Mean :149.42024	Mean :14.981
3rd Qu.:151.82816	3rd Qu.:73.38877	3rd Qu.:223.45188	3rd Qu.:27.492
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963
NA's :12	NA's :18	NA's :19	NA's :8

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.00	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.54	Median :10.052435	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.61	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
NA's :16	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
#Converting all the NA values in particulate matter less than 10 micrometers as the median
median_PM10 <- median(Air_Quality$PM10, na.rm = TRUE)
Air_Quality$PM10[is.na(Air_Quality$PM10)] <- median_PM10
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.41068	1st Qu.: 49.41759
Median :2906	Median :248.9063	Median :147.66485	Median :100.41926
Mean :2906	Mean :248.5003	Mean :148.67091	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.32879	3rd Qu.:151.31922
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497
			NA's :11

NO2		SO2		O3		Temperature	
Min.	: 0.00962	Min.	: 0.01102	Min.	: 0.00166	Min.	: -9.991
1st Qu.:	53.53854	1st Qu.:	24.89137	1st Qu.:	74.13083	1st Qu.:	2.482
Median	:102.96266	Median	:49.54601	Median	:149.76082	Median	:14.944
Mean	:102.29337	Mean	:49.47704	Mean	:149.42024	Mean	:14.981
3rd Qu.:	151.82816	3rd Qu.:	73.38877	3rd Qu.:	223.45188	3rd Qu.:	27.492
Max.	:199.98020	Max.	:99.96956	Max.	:299.93681	Max.	:39.963
NA's	:12	NA's	:18	NA's	:19	NA's	:8

Humidity		WindSpeed		RespiratoryCases		CardiovascularCases	
Min.	: 10.00	Min.	: 0.002094	Min.	: 1.000	Min.	: 0.000
1st Qu.:	32.00	1st Qu.:	4.955375	1st Qu.:	8.000	1st Qu.:	3.000
Median	: 54.54	Median	:10.052435	Median	:10.000	Median	: 5.000
Mean	: 54.77	Mean	: 9.992814	Mean	: 9.974	Mean	: 4.989
3rd Qu.:	77.61	3rd Qu.:	14.968804	3rd Qu.:	12.000	3rd Qu.:	6.000
Max.	:100.00	Max.	:19.999139	Max.	:23.000	Max.	:14.000
NA's	:16	NA's	:11				

HospitalAdmissions		HealthImpactScore		HealthImpactClass	
Min.	: 0.000	Min.	: 22.45	Min.	:0.0000
1st Qu.:	1.000	1st Qu.:	98.20	1st Qu.:	0.0000
Median	: 2.000	Median	:100.00	Median	:0.0000
Mean	: 2.001	Mean	: 93.79	Mean	:0.2812
3rd Qu.:	3.000	3rd Qu.:	100.00	3rd Qu.:	0.0000
Max.	:12.000	Max.	:100.00	Max.	:4.0000

```
#Converting all the NA values in particulate matter less than 2.5 micrometers as the mean
mean_PM2_5 <- mean(Air_Quality$PM2_5, na.rm = TRUE)
Air_Quality$PM2_5[is.na(Air_Quality$PM2_5)] <- mean_PM2_5
summary(Air_Quality)
```

RecordID		AQI		PM10		PM2_5	
Min.	: 1	Min.	: 0.0058	Min.	: 0.01585	Min.	: 0.03155
1st Qu.:	1454	1st Qu.:	123.2514	1st Qu.:	75.41068	1st Qu.:	49.46447
Median	:2906	Median	:248.9063	Median	:147.66485	Median	:100.19937
Mean	:2906	Mean	:248.5003	Mean	:148.67091	Mean	:100.16235
3rd Qu.:	4358	3rd Qu.:	373.3924	3rd Qu.:	222.32879	3rd Qu.:	151.21691
Max.	:5811	Max.	:499.8588	Max.	:299.90196	Max.	:199.98497

NO2		SO2		O3		Temperature	
Min.	: 0.00962	Min.	: 0.01102	Min.	: 0.00166	Min.	: -9.991
1st Qu.:	53.53854	1st Qu.:	24.89137	1st Qu.:	74.13083	1st Qu.:	2.482
Median	:102.96266	Median	:49.54601	Median	:149.76082	Median	:14.944

Mean :102.29337	Mean :49.47704	Mean :149.42024	Mean :14.981
3rd Qu.:151.82816	3rd Qu.:73.38877	3rd Qu.:223.45188	3rd Qu.:27.492
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963
NA's :12	NA's :18	NA's :19	NA's :8

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.00	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.54	Median :10.052435	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.61	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
NA's :16	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
#Converting all the NA values in Concentration of nitrogen dioxide as the median
median_NO2 <- median(Air_Quality$NO2, na.rm = TRUE)
Air_Quality$NO2[is.na(Air_Quality$NO2)] <- median_NO2
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.41068	1st Qu.: 49.46447
Median :2906	Median :248.9063	Median :147.66485	Median :100.19937
Mean :2906	Mean :248.5003	Mean :148.67091	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.32879	3rd Qu.:151.21691
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497

N02	S02	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.80846	1st Qu.:24.89137	1st Qu.: 74.13083	1st Qu.: 2.482
Median :102.96266	Median :49.54601	Median :149.76082	Median :14.944
Mean :102.29475	Mean :49.47704	Mean :149.42024	Mean :14.981
3rd Qu.:151.59518	3rd Qu.:73.38877	3rd Qu.:223.45188	3rd Qu.:27.492
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963
	NA's :18	NA's :19	NA's :8

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.00	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.54	Median :10.052435	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.61	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
NA's :16	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
#Converting all the NA values in Concentration of sulfur dioxide as the median
median_S02 <- median(Air_Quality$S02, na.rm = TRUE)
Air_Quality$S02[is.na(Air_Quality$S02)] <- median_S02
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.41068	1st Qu.: 49.46447
Median :2906	Median :248.9063	Median :147.66485	Median :100.19937
Mean :2906	Mean :248.5003	Mean :148.67091	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.32879	3rd Qu.:151.21691
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497

N02	S02	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.80846	1st Qu.:24.94479	1st Qu.: 74.13083	1st Qu.: 2.482
Median :102.96266	Median :49.54601	Median :149.76082	Median :14.944
Mean :102.29475	Mean :49.47725	Mean :149.42024	Mean :14.981
3rd Qu.:151.59518	3rd Qu.:73.34662	3rd Qu.:223.45188	3rd Qu.:27.492
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963
		NA's :19	NA's :8

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.00	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.54	Median :10.052435	Median :10.000	Median : 5.000

Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.61	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
NA's :16	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
#Converting all the NA values in Concentration of ozone as the mean
mean_O3 <- mean(Air_Quality$O3, na.rm = TRUE)
Air_Quality$O3[is.na(Air_Quality$O3)] <- mean_O3
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.41068	1st Qu.: 49.46447
Median :2906	Median :248.9063	Median :147.66485	Median :100.19937
Mean :2906	Mean :248.5003	Mean :148.67091	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.32879	3rd Qu.:151.21691
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497

N02	S02	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.80846	1st Qu.:24.94479	1st Qu.: 74.28772	1st Qu.: 2.482
Median :102.96266	Median :49.54601	Median :149.42024	Median :14.944
Mean :102.29475	Mean :49.47725	Mean :149.42024	Mean :14.981
3rd Qu.:151.59518	3rd Qu.:73.34662	3rd Qu.:223.17835	3rd Qu.:27.492
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963
			NA's :8

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.00	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.54	Median :10.052435	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.61	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
NA's :16	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
#Converting all the NA values in temperature as the mean
mean_Temperature<- mean(Air_Quality$Temperature, na.rm = TRUE)
Air_Quality$Temperature[is.na(Air_Quality$Temperature)] <- mean_Temperature
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.41068	1st Qu.: 49.46447
Median :2906	Median :248.9063	Median :147.66485	Median :100.19937
Mean :2906	Mean :248.5003	Mean :148.67091	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.32879	3rd Qu.:151.21691
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497

NO2	S02	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.80846	1st Qu.:24.94479	1st Qu.: 74.28772	1st Qu.: 2.488
Median :102.96266	Median :49.54601	Median :149.42024	Median :14.981
Mean :102.29475	Mean :49.47725	Mean :149.42024	Mean :14.981
3rd Qu.:151.59518	3rd Qu.:73.34662	3rd Qu.:223.17835	3rd Qu.:27.465
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.00	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.54	Median :10.052435	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.61	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
NA's :16	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000

Mean	: 2.001	Mean	: 93.79	Mean	:0.2812
3rd Qu.:	3.000	3rd Qu.:	100.00	3rd Qu.:	0.0000
Max.	:12.000	Max.	:100.00	Max.	:4.0000

```
#Converting all the NA values in Humidity as the mean
mean_Humidity<- mean(Air_Quality$Humidity, na.rm = TRUE)
Air_Quality$Humidity[is.na(Air_Quality$Humidity)] <- mean_Humidity
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.41068	1st Qu.: 49.46447
Median :2906	Median :248.9063	Median :147.66485	Median :100.19937
Mean :2906	Mean :248.5003	Mean :148.67091	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.32879	3rd Qu.:151.21691
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497

NO2	S02	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.80846	1st Qu.:24.94479	1st Qu.: 74.28772	1st Qu.: 2.488
Median :102.96266	Median :49.54601	Median :149.42024	Median :14.981
Mean :102.29475	Mean :49.47725	Mean :149.42024	Mean :14.981
3rd Qu.:151.59518	3rd Qu.:73.34662	3rd Qu.:223.17835	3rd Qu.:27.465
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.06	1st Qu.: 4.955375	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.63	Median :10.052435	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.54	3rd Qu.:14.968804	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000
	NA's :11		

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
#Converting all the NA values in Wind Speed as the mean
mean_WindSpeed<- mean(Air_Quality$WindSpeed, na.rm = TRUE)
Air_Quality$WindSpeed[is.na(Air_Quality$WindSpeed)] <- mean_WindSpeed
summary(Air_Quality)
```

RecordID	AQI	PM10	PM2_5
Min. : 1	Min. : 0.0058	Min. : 0.01585	Min. : 0.03155
1st Qu.:1454	1st Qu.:123.2514	1st Qu.: 75.41068	1st Qu.: 49.46447
Median :2906	Median :248.9063	Median :147.66485	Median :100.19937
Mean :2906	Mean :248.5003	Mean :148.67091	Mean :100.16235
3rd Qu.:4358	3rd Qu.:373.3924	3rd Qu.:222.32879	3rd Qu.:151.21691
Max. :5811	Max. :499.8588	Max. :299.90196	Max. :199.98497

NO2	SO2	O3	Temperature
Min. : 0.00962	Min. : 0.01102	Min. : 0.00166	Min. : -9.991
1st Qu.: 53.80846	1st Qu.:24.94479	1st Qu.: 74.28772	1st Qu.: 2.488
Median :102.96266	Median :49.54601	Median :149.42024	Median :14.981
Mean :102.29475	Mean :49.47725	Mean :149.42024	Mean :14.981
3rd Qu.:151.59518	3rd Qu.:73.34662	3rd Qu.:223.17835	3rd Qu.:27.465
Max. :199.98020	Max. :99.96956	Max. :299.93681	Max. :39.963

Humidity	WindSpeed	RespiratoryCases	CardiovascularCases
Min. : 10.00	Min. : 0.002094	Min. : 1.000	Min. : 0.000
1st Qu.: 32.06	1st Qu.: 4.958763	1st Qu.: 8.000	1st Qu.: 3.000
Median : 54.63	Median :10.032817	Median :10.000	Median : 5.000
Mean : 54.77	Mean : 9.992814	Mean : 9.974	Mean : 4.989
3rd Qu.: 77.54	3rd Qu.:14.959056	3rd Qu.:12.000	3rd Qu.: 6.000
Max. :100.00	Max. :19.999139	Max. :23.000	Max. :14.000

HospitalAdmissions	HealthImpactScore	HealthImpactClass
Min. : 0.000	Min. : 22.45	Min. :0.0000
1st Qu.: 1.000	1st Qu.: 98.20	1st Qu.:0.0000
Median : 2.000	Median :100.00	Median :0.0000
Mean : 2.001	Mean : 93.79	Mean :0.2812
3rd Qu.: 3.000	3rd Qu.:100.00	3rd Qu.:0.0000
Max. :12.000	Max. :100.00	Max. :4.0000

```
# Check for duplicate rows
duplicates <- Air_Quality[duplicated(Air_Quality), ]
print(paste("Number of duplicate rows:", nrow(duplicates)))
```

```
[1] "Number of duplicate rows: 0"
```

```
# Save the cleaned dataset to a new CSV file  
write.csv(Air_Quality, "Air_Quality_CleanedData.csv", row.names = FALSE)
```