

**TUGAS *OUTLIER***

**MATA KULIAH MANAJEMEN DAN ANALISIS DATA DENGAN R**



**ASTI OKTOVIANTI SUNMAYA ANANDA PUTRI**

**NPM. 131520220005**

**PROGRAM STUDI EPIDEMIOLOGI**

**FAKULTAS KEDOKTERAN**

**UNIVERSITAS PADJADJARAN**

**2023**

```
#reading data
```

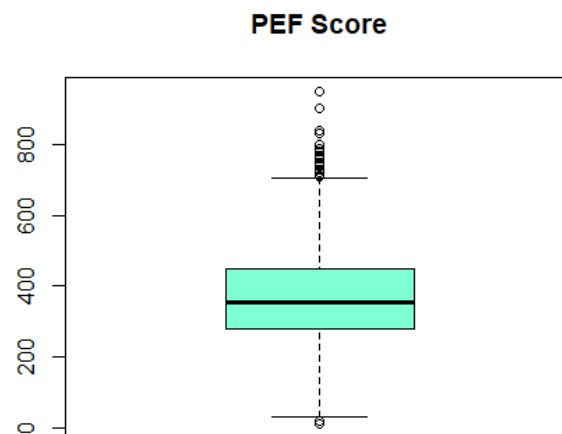
```
pef_final
```

```
#1. Mengidentifikasi outlier berdasarkan kriteria dan visualisasi grafik boxplot
```

```
summary(pef_final$pef)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
10.0	280.0	355.0	367.4	450.0	951.0

```
boxplot(pef_final$pef,col="aquamarine",main ="PEF Score")
```



```
#2. Menentukan cut off outlier (batas atas dan bawah) dari grafik boxplot tersebut.
```

```
min(boxplot(pef_final$pef, plot = FALSE)$out)
```

```
max(boxplot(pef_final$pef, plot = FALSE)$out)
```

Min	Max
10	951

```
quartiles <- quantile(pef_final$pef, probs=c(.25, .75), na.rm=FALSE)
```

```
quartiles
```

25%	75%
280	450

```
IQR <- IQR(pef_final$pef)
```

IQR

<b>170</b>
------------

```
Lower <- quartiles[1] - 1.5*IQR
```

Lower

<b>25%</b>
------------

25
----

```
Upper <- quartiles[2] + 1.5*IQR
```

Upper

<b>75%</b>
------------

705
-----

#3. Membuat dataset yang tidak berisi outlier sesuai cut off no 2.

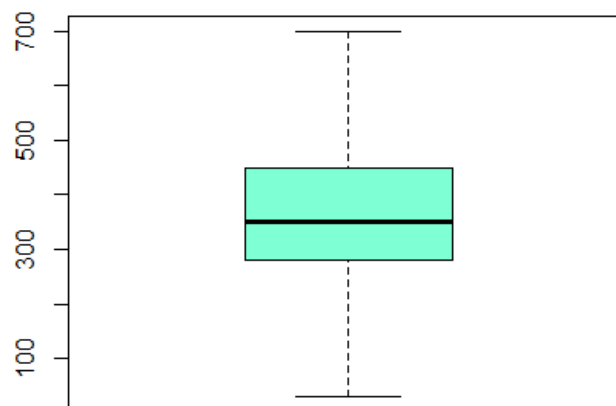
```
pef_final_no <- subset(pef_final, pef_final$pef > Lower
```

```
& pef_final$pef < Upper)
```

```
summary(pef_final_no$pef)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
30.0	280.0	350.0	366.2	450.0	700.0

```
boxplot(pef_final_no$pef, col="aquamarine")
```



#4. Melakukan tes normalitas pada dataset dengan outlier dan tanpa outlier.

#For Big Sample (Kolmogorov-Smirnov)

```
library(nortest)
```

```
lillie.test(pef_final$pef)
```

```
Lilliefors (kolmogorov-smirnov) normality test  
data:  pef_final$pef  
D = 0.055561, p-value < 2.2e-16
```

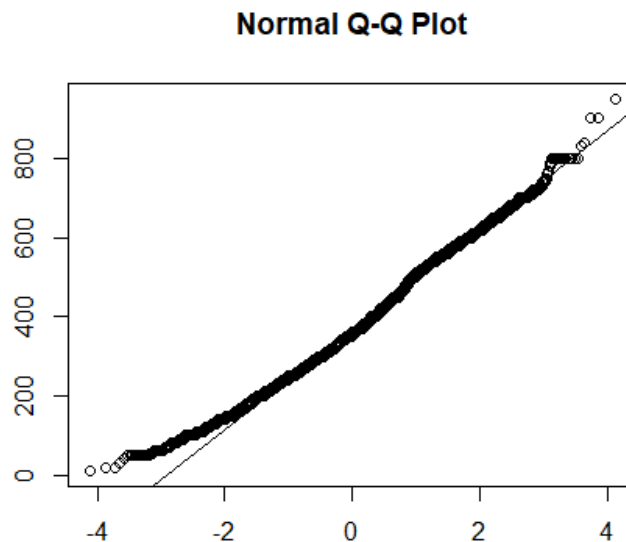
```
lillie.test(pef_final_no$pef)
```

```
Lilliefors (kolmogorov-smirnov) normality test  
data:  pef_final_no$pef  
D = 0.053763, p-value < 2.2e-16
```

#5. Membuat grafik QQ line untuk membandingkan visualisasi nilai pef pada dataset dengan outlier dan tanpa outlier.

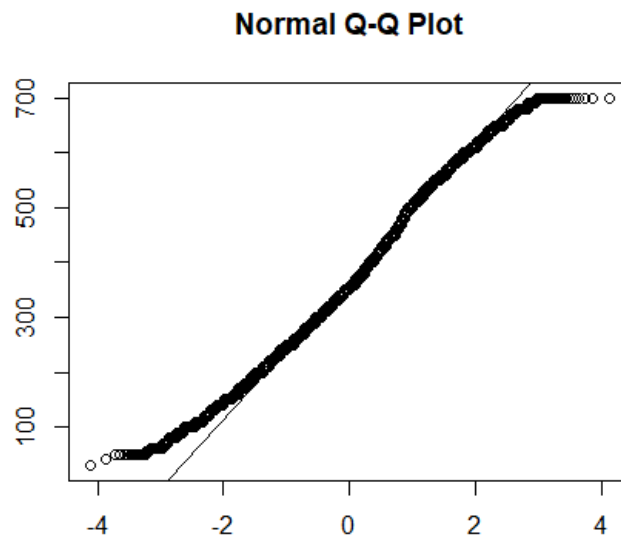
```
qqnorm(pef_final$pef); qqline(pef_final$pef)
```

(DENGAN OUTLIERS)



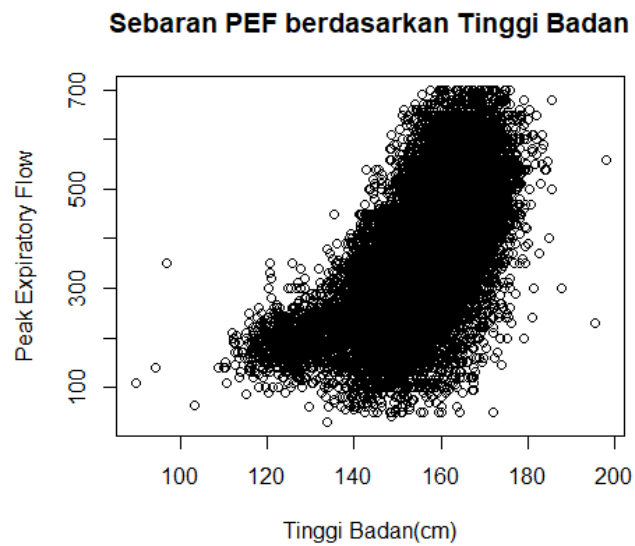
```
qqnorm(pef_final_no$pef); qqline(pef_final_no$pef)
```

(TANPA OUTLIERS)

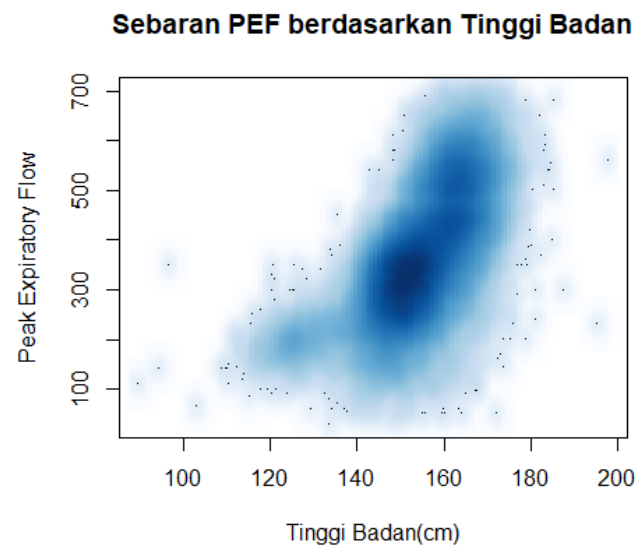


#6. Membuat scatterplot yang memperlihatkan hubungan antara pef dengan height, dengan penambahan garis linear/regresi dan smoothed dengan loes (local regression smoothing).

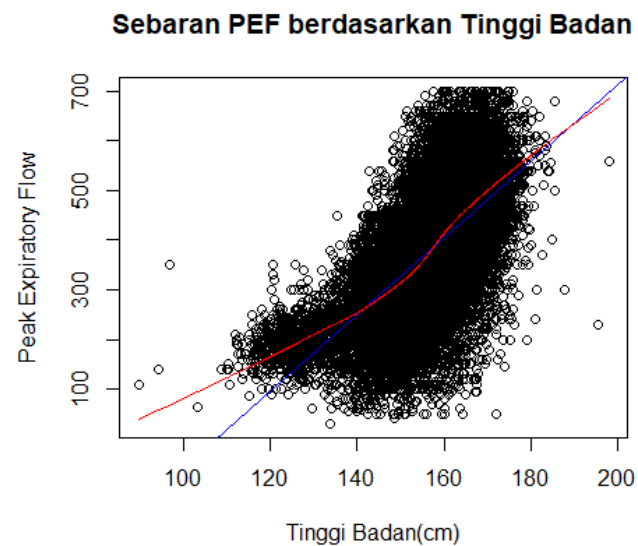
```
plot(pef_final_no$pef~pef_final_no$height, xlab = "Tinggi Badan(cm)",  
     ylab="Peak Expiratory Flow", main="Sebaran PEF berdasarkan Tinggi Badan")
```



```
smoothScatter(pef_final_no$pef~pef_final_no$height, xlab = "Tinggi Badan(cm)",  
  ylab="Peak Expiratory Flow", main="Sebaran PEF berdasarkan Tinggi Badan")
```

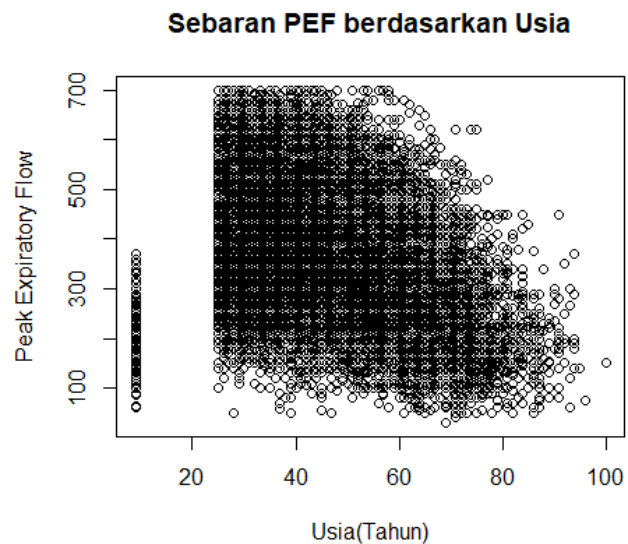


```
abline(lm(pef_final_no$pef~pef_final_no$height, data = pef_final_no), col = "blue")  
lines(lowess(pef_final_no$height, pef_final_no$pef), col = "red")
```

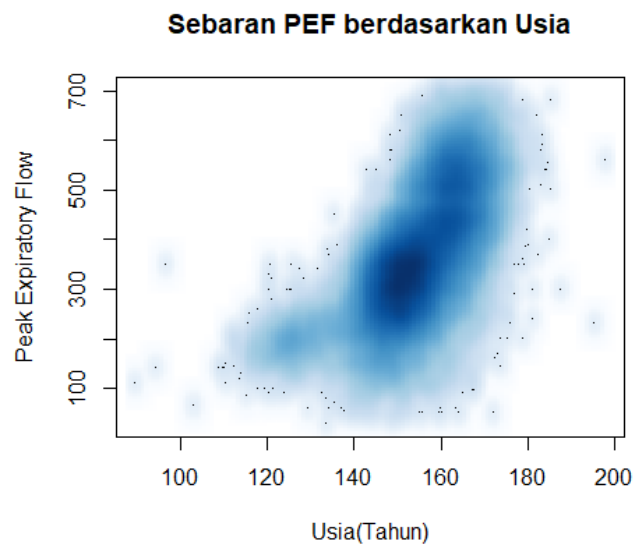


#7. Membuat scatterplot yang memperlihatkan hubungan antara pef dengan umur, dengan penambahan garis linear/regresi dan smoothed dengan loes (local regression smoothing).

```
plot(pef_final_no$pef~pef_final_no$age, xlab = "Usia(Tahun)",  
     ylab="Peak Expiratory Flow", main="Sebaran PEF berdasarkan Usia")
```



```
smoothScatter(pef_final_no$pef~pef_final_no$height, xlab = "Usia(Tahun)",  
              ylab="Peak Expiratory Flow", main="Sebaran PEF berdasarkan Usia")
```



```
abline(lm(pef_final_no$pef~pef_final_no$age, data = pef_final_no), col = "blue")  
lines(lowess(pef_final_no$age, pef_final_no$pef), col = "red")
```

