

Tugas Mata Kuliah Manajemen dan Analisis Data dengan R

Disusun oleh:

Syndi Nurmawati NPM 131520220002



Program Studi Magister Epidemiologi
Fakultas Kedokteran
Universitas Padjadjaran
2023

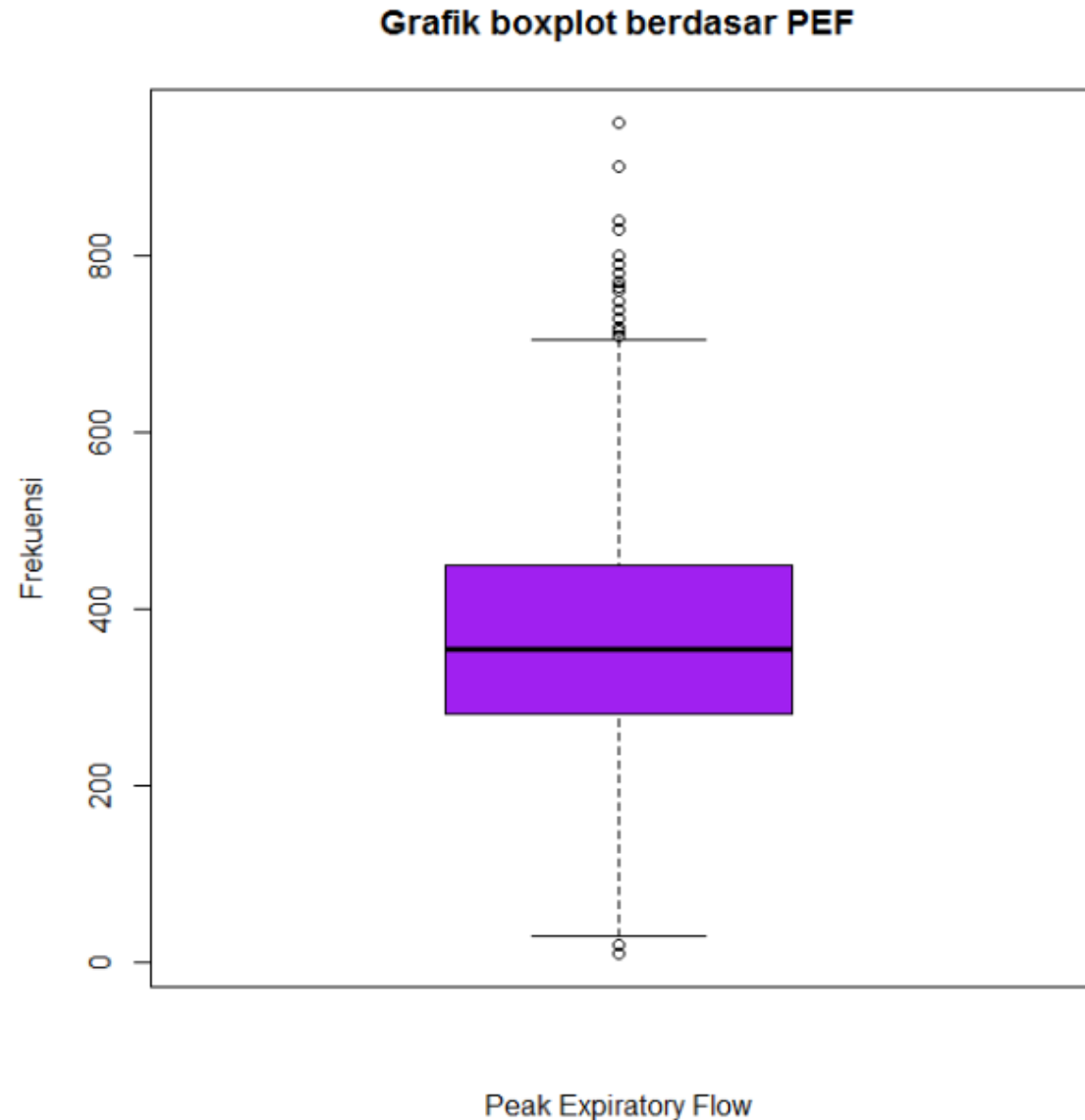
1. Mengidentifikasi outlier berdasarkan kriteria dan visualisasi grafik boxplot

Boxplot berdasarkan pef

Kode: `boxplot(w5_ahsp$pef)`

Hasil:

Pada grafik ini terlihat outlier pada bagian atas dan bawah dari boxplot.



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

a. Menentukan cut-off outlier dengan mempergunakan kriteria mean

Kode: `mean = mean(w5_ahsp$pef)`

Hasil: `mean` | 367.366534924375

Kode: `std = sd(w5_ahsp$pef)`

Hasil: `std` | 123.452358919085

Kode batas bawah: `Tmin1 = mean-(3*std)`

Hasil: `Tmin1` | -2.99054183288121

Kode batas atas: `Tmax1 = mean+(3*std)`

Hasil: `Tmax1` | 737.723611681631

b. Menentukan cut-off outlier dengan mempergunakan kriteria IQR

Kode: `Q1 <- quantile(w5_ahsp$pef, .25)`

`Q3 <- quantile(w5_ahsp$pef, .75)`

Hasil:

<code>Q1</code>	Named num 280
<code>Q3</code>	Named num 450

Kode: `IQR <- IQR(w5_ahsp$pef)`

Hasil: `IQR` | 170

Kode batas bawah: `Lower <- Q1 - 1.5*IQR`

Hasil: `Lower` | Named num 25

Kode batas atas: `Upper <- Q3 + 1.5*IQR`

Hasil: `Upper` | Named num 705

3. Membuat dataset yang tidak berisi outlier sesuai cut-off nomor 2

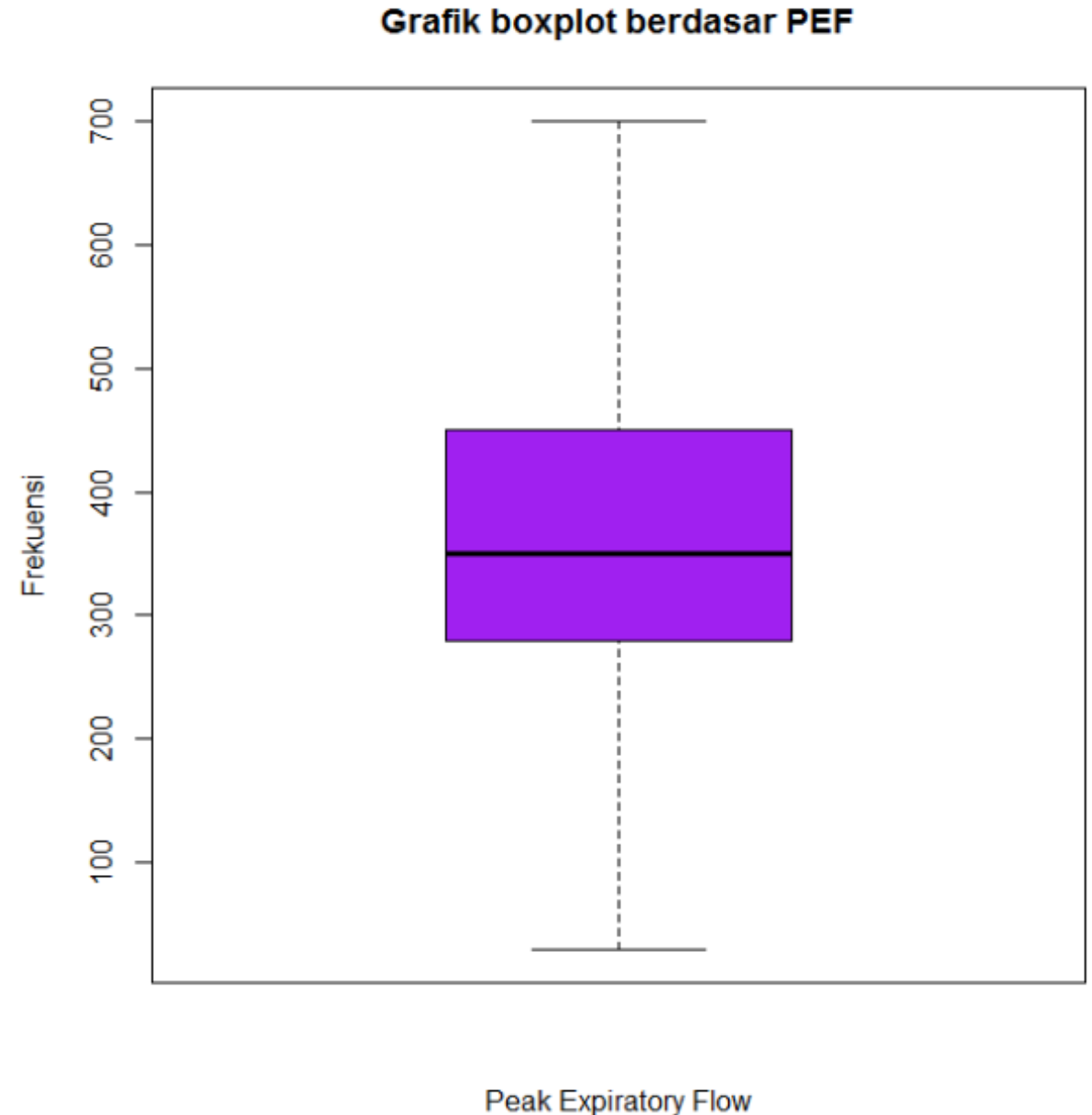
Cut-off outlier yang dipergunakan adalah berdasarkan IQR.

```
Kode: pef_no2 <- subset(w5_ahsp, w5_ahsp$pef > Lower & w5_ahsp$pef < Upper)
```

Hasil: Pada dataset lama terdapat 26.314 observasi, sedangkan pada dataset baru terdapat 26.232 observasi (selisih 82 observasi).

```
w5_ahsp      | 26314 obs. of 31 variables  
pef_no2      | 26232 obs. of 31 variables
```

Pengecekan pada boxplot (gambar kanan) menunjukkan tidak lagi terdapat outlier.



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

a. Tes normalitas dengan Kolmogorov Smirnov pada dataset dengan outlier

Kode: `lillie.test(w5_ahsp$pef)`

Hasil: `Lilliefors (Kolmogorov-Smirnov) normality test`
`data: w5_ahsp$pef`
`D = 0.055561, p-value < 2.2e-16`

b. Tes normalitas dengan Kolmogorov Smirnov pada dataset tanpa outlier

Kode: `lillie.test(pef_no2$pef)`

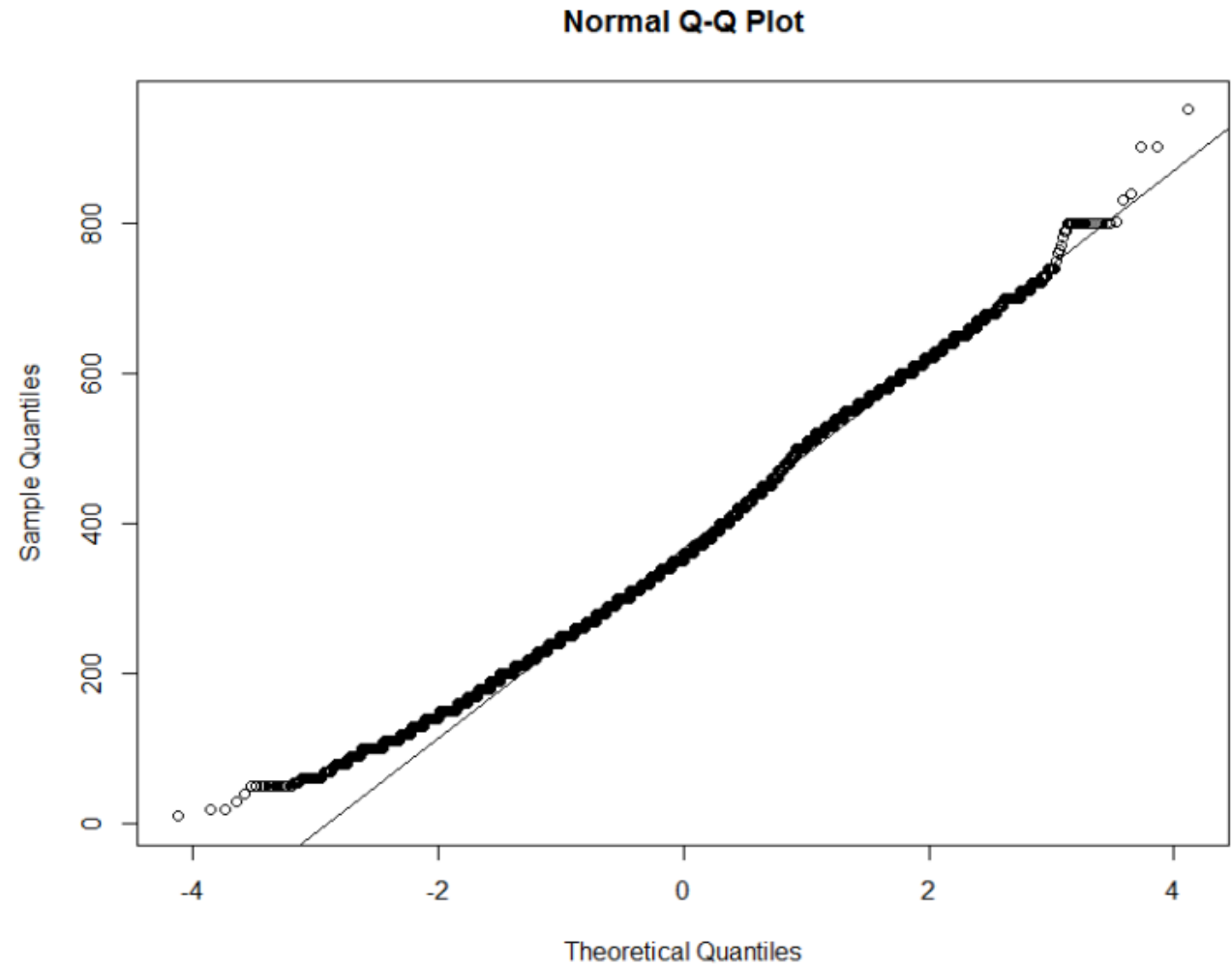
Hasil: `Lilliefors (Kolmogorov-Smirnov) normality test`
`data: pef_no2$pef`
`D = 0.053763, p-value < 2.2e-16`

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

a. Membuat grafik QQ line pada dataset dengan outlier

Kode: `qqnorm(w5_ahsp$pef);
qqline(w5_ahsp$pef)`

Hasil pada gambar di sebelah kanan

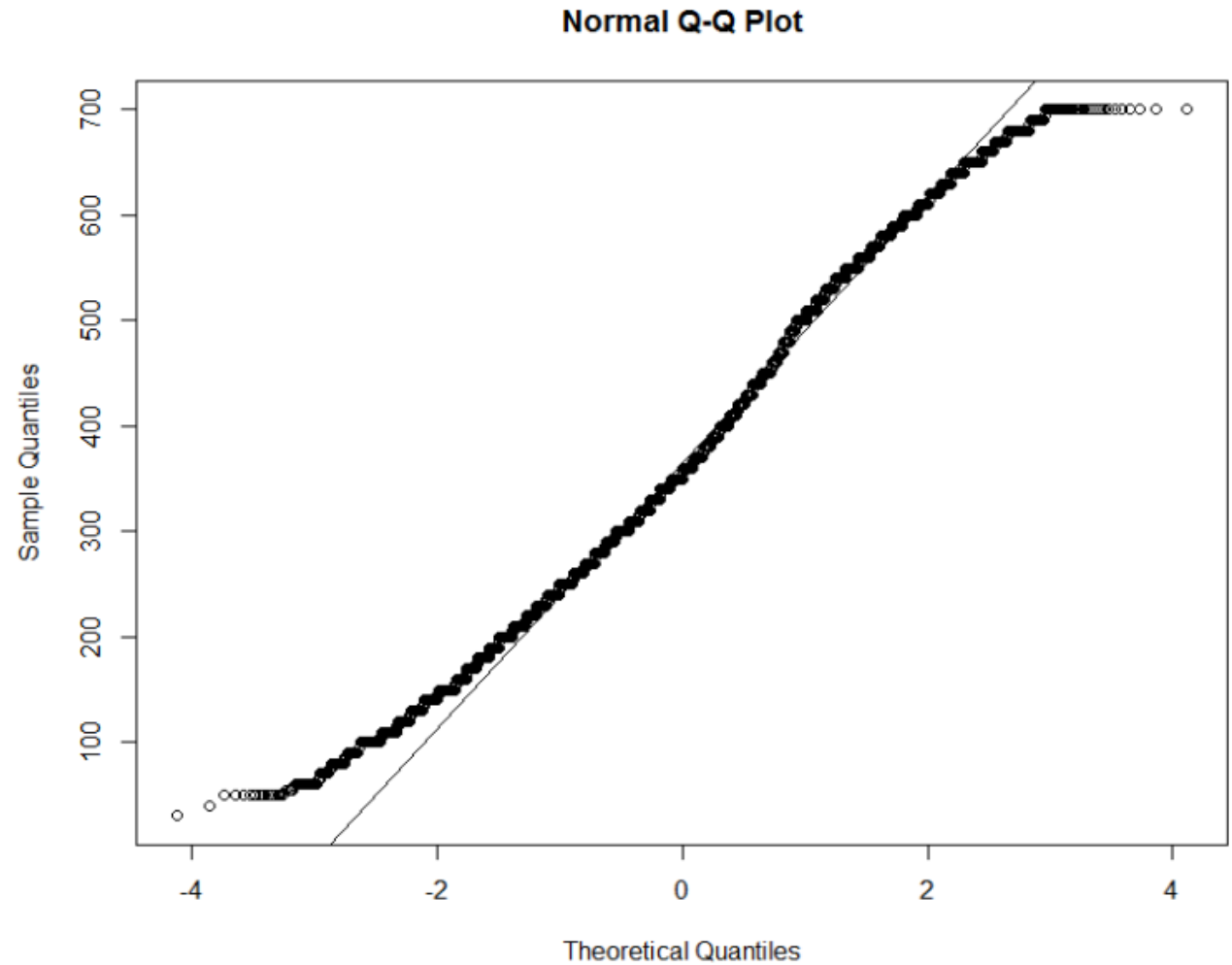


5. Membuat grafik QQ line untuk membandingkan visualisasi nilai pef pada dataset dengan dan tanpa outlier

b. Membuat grafik QQ line pada dataset tanpa outlier

Kode: `qqnorm(pef_no2$pef);`
`qqline(pef_no2$pef)`

Hasil pada gambar di sebelah kanan



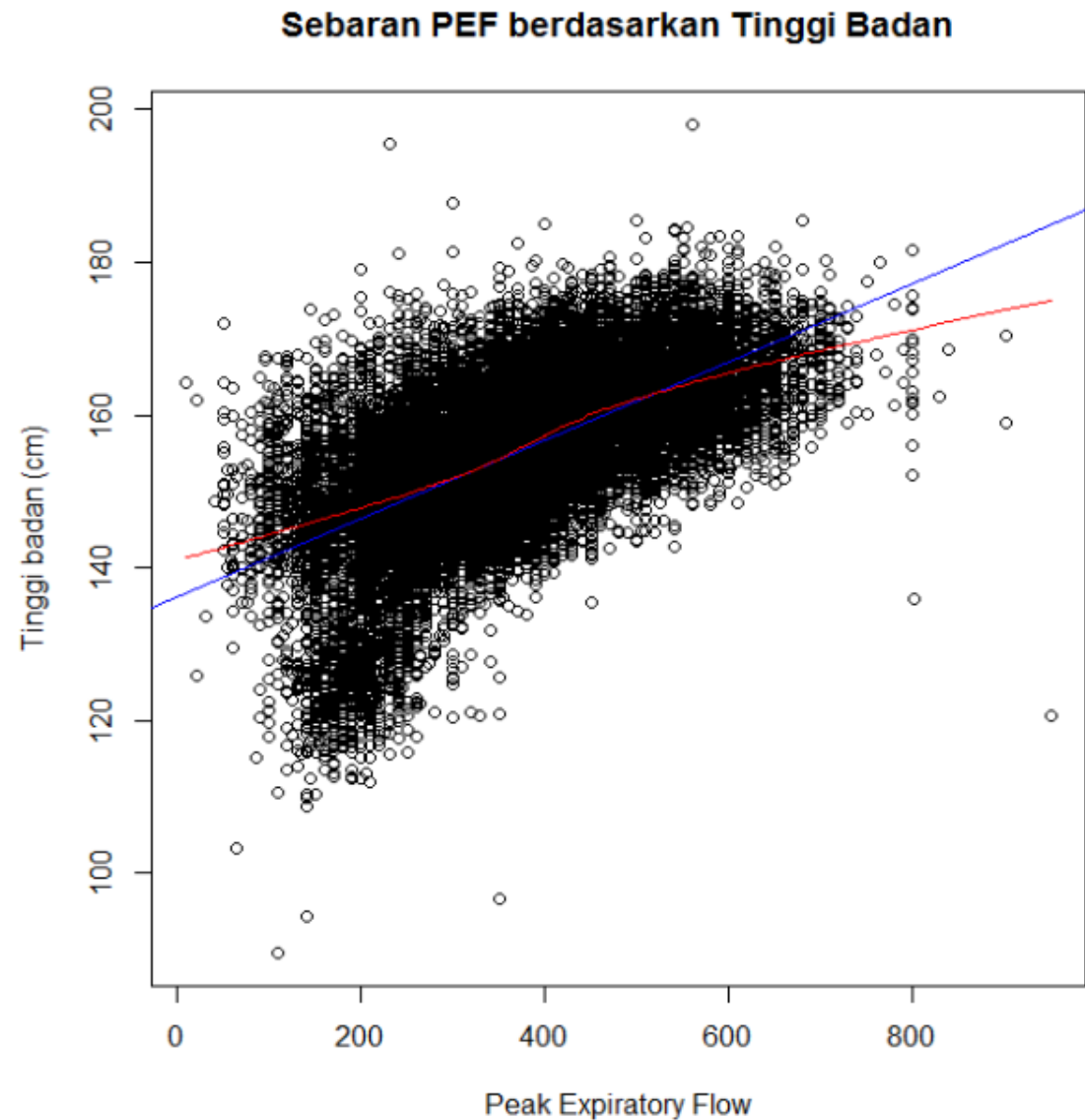
6. Membuat scatterplot yang memperlihatkan hubungan antara pef dengan height, dengan penambahan garis linear/regresi dan smoothed dengan loes (local regression smoothing)

a. Pada dataset dengan outlier

Kode:

```
plot(w5_ahsp$height~w5_ahsp$pef, xlab =  
"Peak Expiratory Flow", ylab = "Tinggi badan  
(cm)", main = "Sebaran PEF berdasarkan  
Tinggi Badan")  
abline(lm(w5_ahsp$height~w5_ahsp$pef, data  
= w5_ahsp), col = "blue")  
lines(lowess(w5_ahsp$pef, w5_ahsp$height),  
col = "red")
```

Hasil pada gambar di sebelah kanan



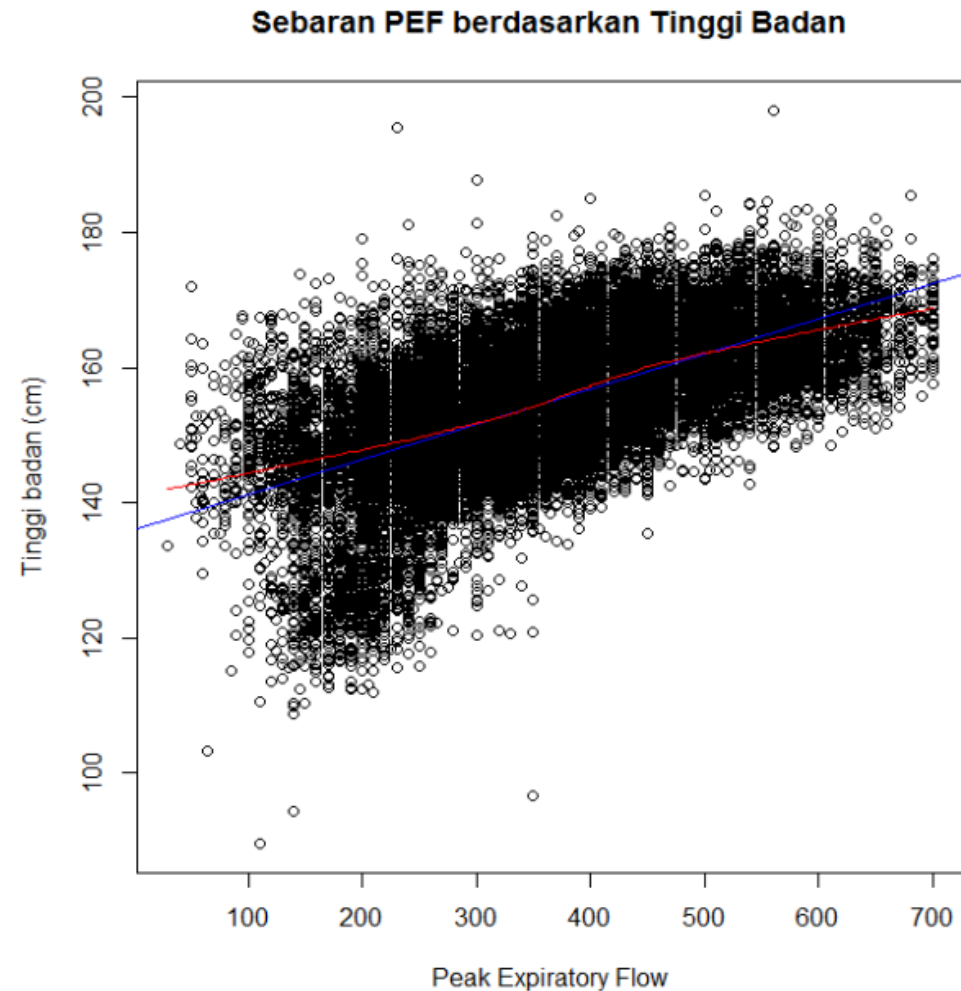
6. Membuat scatterplot yang memperlihatkan hubungan antara pef dengan height, dengan penambahan garis linear/regresi dan smoothed dengan loes (local regression smoothing)

b. Pada dataset tanpa outlier

Kode:

```
plot(pef_no2$height~ pef_no2 $pef, xlab =  
"Peak Expiratory Flow", ylab = "Tinggi badan  
(cm)", main = "Sebaran PEF berdasarkan  
Tinggi Badan")  
abline(lm(pef_no2$height~ pef_no2$pef, data  
= pef_no2), col = "blue")  
lines(lowess(pef_no2$pef, pef_no2$height), col  
= "red")
```

Hasil pada gambar di sebelah kanan



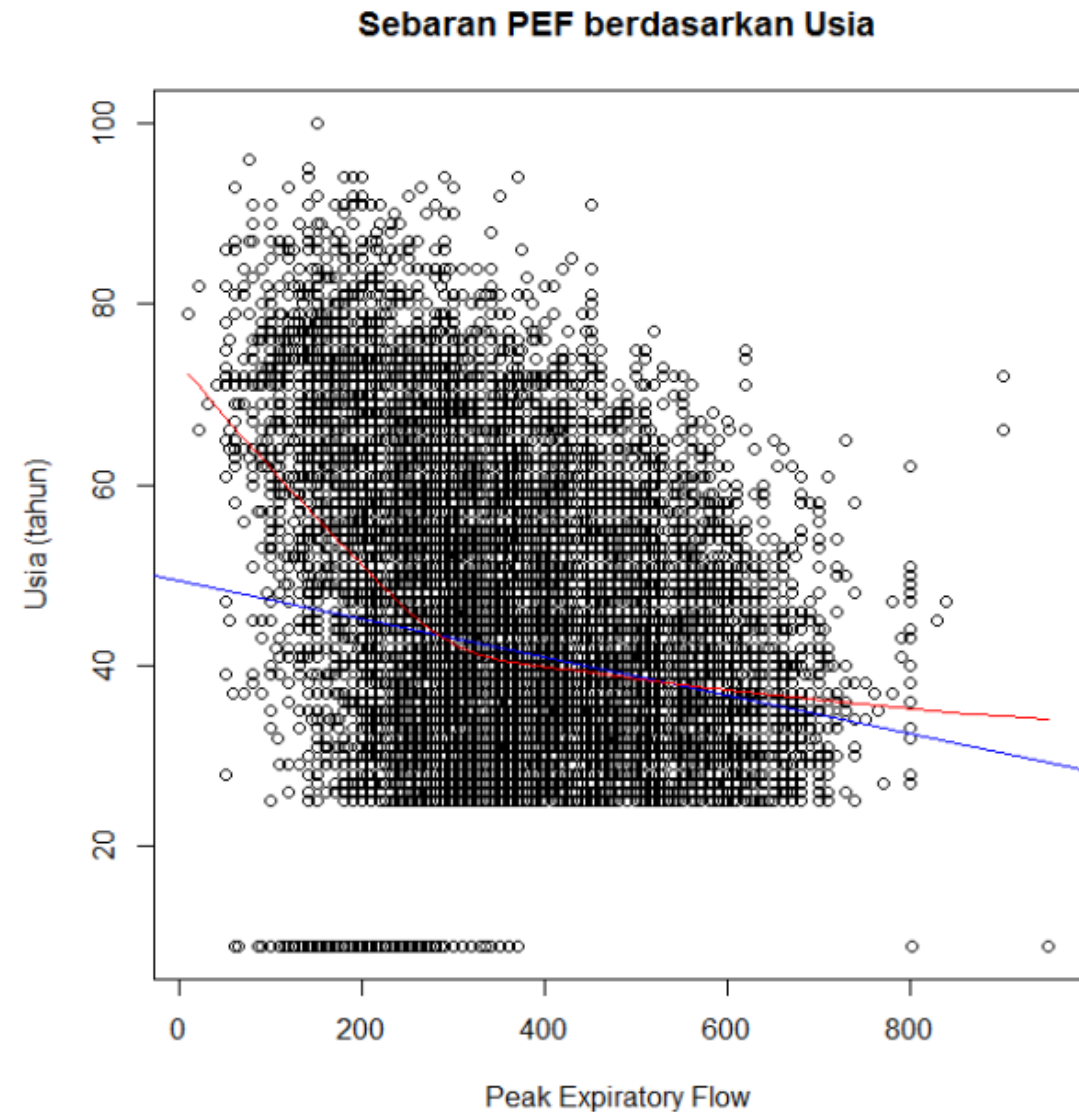
7. Membuat scatterplot yang memperlihatkan hubungan antara pef dengan umur, dengan penambahan garis linear/regresi dan smoothed dengan loes (local regression smoothing)

a. Pada dataset dengan outlier

Kode:

```
plot(w5_ahsp$age~w5_ahsp$pef, xlab = "Peak  
Expiratory Flow", ylab = "Usia (tahun)",  
main = "Sebaran PEF berdasarkan Usia")  
abline(lm(w5_ahsp$age~w5_ahsp$pef, data =  
w5_ahsp), col = "blue")  
lines(lowess(w5_ahsp$pef, w5_ahsp$age), col =  
"red")
```

Hasil pada gambar di sebelah kanan



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

b. Pada dataset tanpa outlier

Kode:

```
plot(pef_no2$age~ pef_no2$pef, xlab = "Peak  
Expiratory Flow", ylab = "Usia (tahun)", main  
= "Sebaran PEF berdasarkan Usia")  
abline(lm(pef_no2$age~ pef_no2$pef, data =  
pef_no2), col = "blue")  
lines(lowess(pef_no2$pef, pef_no2$age), col =  
"red")
```

Hasil pada gambar di sebelah kanan

