

Jawaban Bagian B (2) UTS STA1342

Faisal Arkan

10/10/2022

Library

```
library(factoextra)
library(ggcorrplot)
library(readxl)
```

Input Data

```
IPM <- read_excel("D:/Kuliah/Semester 5/STA1342 Teknik Peubah Ganda/UTS/Data IPM KotaKab.xlsx")
```

Data Managing

```
IPM <- IPM[,-1]
colnames(IPM)[1] <- "X1" #Harapan Lama Sekolah
colnames(IPM)[2] <- "X2" #Pengeluaran Per Kapita
colnames(IPM)[3] <- "X3" #Rata-rata Lama Sekolah
colnames(IPM)[4] <- "X4" #Umur Harapan Hidup
head(IPM)
```

```
## # A tibble: 6 x 4
##       X1      X2      X3      X4
##   <dbl> <dbl> <dbl> <dbl>
## 1  13.9  7148   9.48  65.3
## 2  14.3  8776   8.68  67.4
## 3  14.6  8180   8.88  64.4
## 4  14.0  8030   9.67  68.2
## 5  13.0  8577   8.21  68.7
## 6  14.3 10780   9.86  68.9
```

Korelasi

Matriks Korelasi

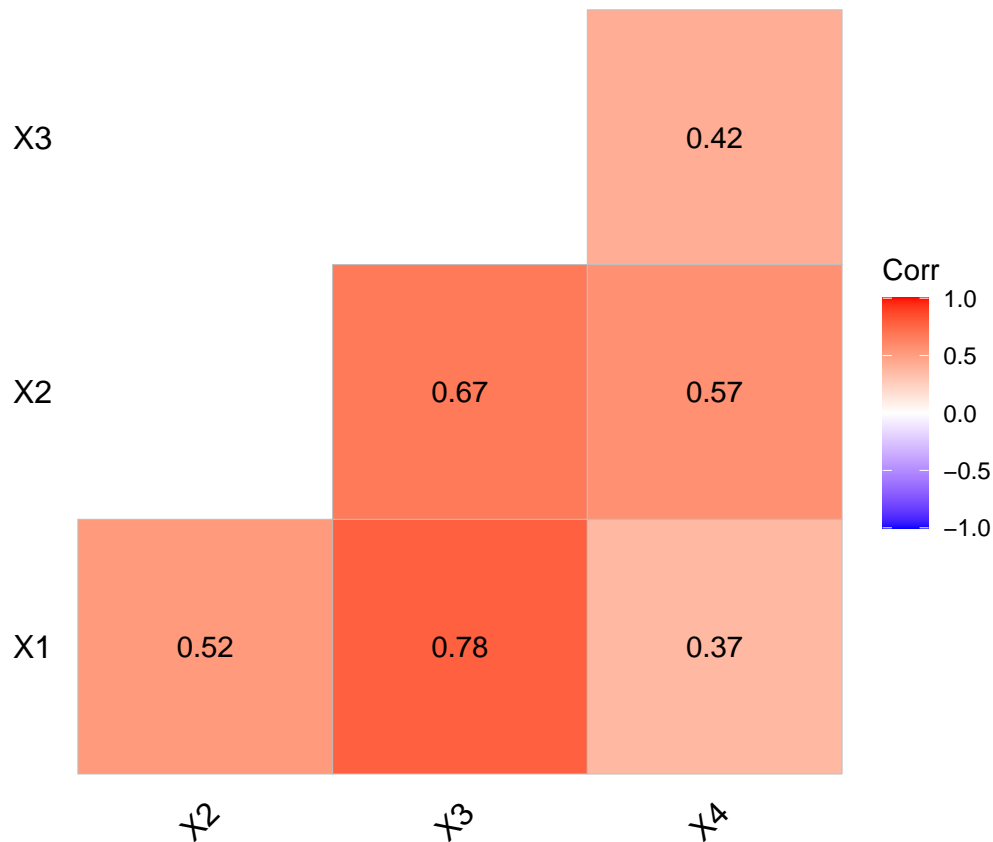
```
cor_IPM <- cor(IPM)
cor_IPM
```

```
##           X1           X2           X3           X4
## X1 1.0000000 0.5243790 0.7777449 0.3713379
## X2 0.5243790 1.0000000 0.6688473 0.5668087
## X3 0.7777449 0.6688473 1.0000000 0.4160388
## X4 0.3713379 0.5668087 0.4160388 1.0000000
```

Plot Korelasi

Selain itu, untuk memudahkan dalam menganalisis secara visual, matriks korelasi dapat dibentuk grafik sebagai berikut

```
ggcorrplot(cor_IPM, type = "lower", lab = T, ggtheme = ggplot2::theme_void)
```



Jawab 2A : Dari plot korelasi di atas terlihat bahwa peubah X1 dengan X3 memiliki korelasi yang paling tinggi yakni sekitar 0.78. Sedangkan untuk korelasi paling rendah dimiliki oleh peubah X1 dengan X4 yang memiliki nilai korelasi sebesar 0.37.

Jawab 2B : Matriks korelasi yang berisi nilai-nilai korelasi pada dasarnya digunakan untuk mendapatkan keeratan hubungan antar peubah yang diteliti. Nilai korelasi inilah yang nantinya akan dibandingkan dengan nilai korelasi yang diperoleh dari Analisis Komponen Utama.

Analisis Komponen Utama

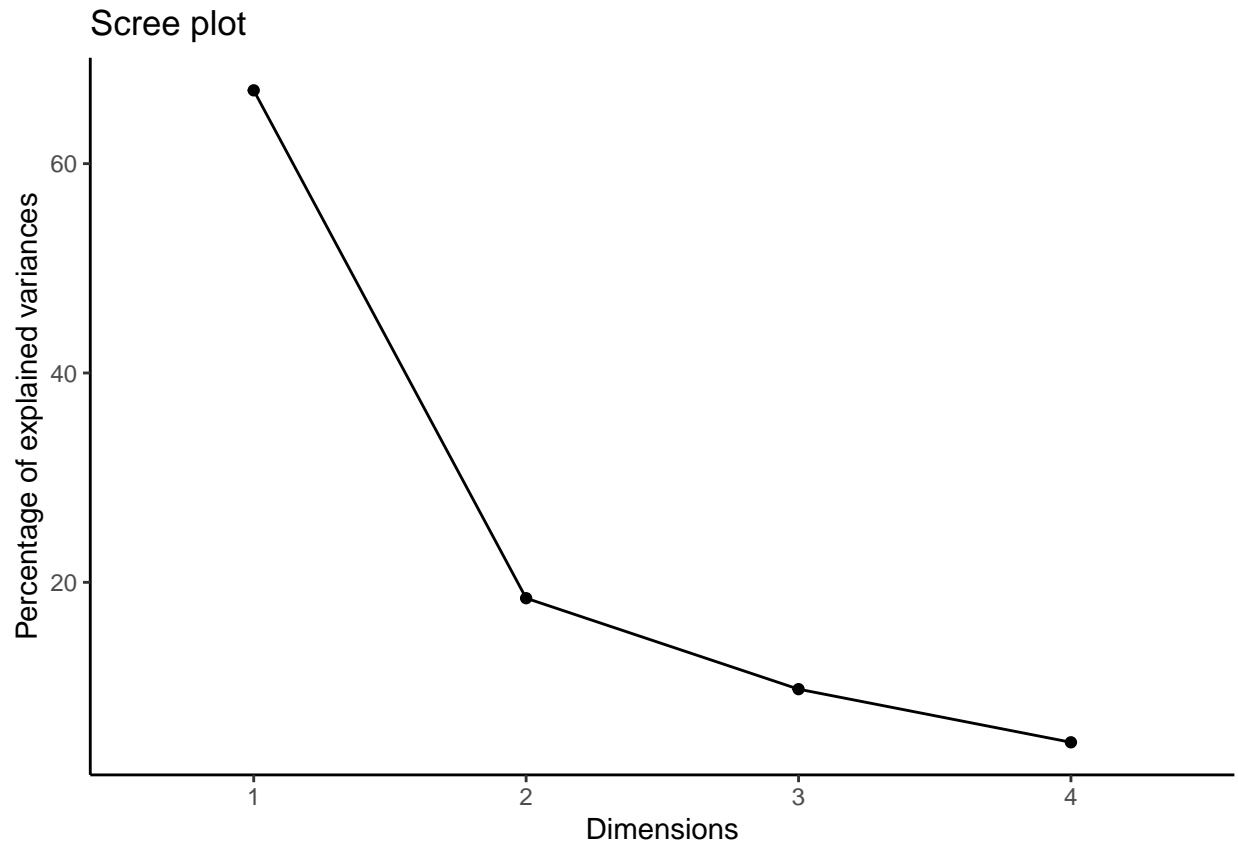
Analisis Komponen Utama (Principal Component Analysis) adalah analisis multivariat yang digunakan untuk mereduksi peubah. Pada umumnya, analisis ini bukan merupakan analisis akhir. AKU atau PCA menghasilkan peubah baru yang disebut sebagai **Komponen Utama** yang mempunyai karakteristik yakni memuat hampir seluruh informasi dan Saling bebas.

```
aku_IPM <- prcomp(IPM, scale.=TRUE, center=TRUE)
summary(aku_IPM)
```

```
## Importance of components:
##              PC1      PC2      PC3      PC4
## Standard deviation    1.637 0.8599 0.62592 0.43441
## Proportion of Variance 0.670 0.1848 0.09794 0.04718
## Cumulative Proportion 0.670 0.8549 0.95282 1.00000
```

Jawab 2C : Terdapat 3 hasil keluaran yang didapatkan dari `summary` di atas, yakni *Standard Deviation*, *Proportion of Variance*, dan *Cumulative Proportion*. *Standard Deviation* merupakan akar dari eigenvalue-nya. Akar ciri (eigenvalue) berperan sebagai ragam dari komponen utama. *Proportion of Variance* memberikan informasi seberapa besar keragaman peubah asal yang dapat dijelaskan oleh masing-masing komponen utama. Semakin besar nilai dari *Proportion of Variance* maka semakin baik pula komponen utama tersebut untuk merepresentasikan peubah asal. *Cumulative Proportion* menjelaskan seberapa besar keragaman yang dapat dijelaskan oleh komponen utama secara kumulatif. Dua komponen utama yang dipilih (PC1 dan PC2) memberikan kontribusi sebanyak 85.49%. Hal ini mengindikasikan bahwa hanya dengan menggunakan PC1 dan PC2, sebanyak 85.49% data sudah dapat dijelaskan.

```
fviz_screplot(aku_IPM, geom="line")+theme_classic()
```



Banyaknya komponen utama bisa ditentukan dengan melihat komponen utama yang garisnya berbentuk seperti siku (elbow). Berdasarkan *screeplot* di atas, garis membentuk siku saat berada di komponen utama kedua (dimension = 2). Sehingga banyaknya komponen utama yang digunakan adalah sebanyak dua (PC1 dan PC2).

Interpretasi

Interpretasi metode PCA dapat dilakukan dengan menggunakan vektor ciri pada masing-masing komponen utama. Semakin besar vektor ciri pada komponen utama tertentu, semakin besar pula kontribusi dari peubah asal untuk membangun komponen utama tersebut. Nilai negatif pada vektor ciri menandakan peubah asal memberikan kontribusi yang berkebalikan pada pembentukan komponen utama. Dalam konteks vektor ciri negatif, semakin besar nilai peubah asal semakin kecil nilai pada komponen utama.

```
aku_IPM$rotation
```

```
##          PC1          PC2          PC3          PC4
## X1 -0.5070841  0.4859192 -0.44734046  0.55374614
## X2 -0.5175424 -0.2280017  0.75363468  0.33496251
## X3 -0.5451726  0.3561005  0.06925768 -0.75576627
## X4 -0.4216663 -0.7649119 -0.47657598 -0.09991335
```

Jawab 2D : Interpretasi hanya dilakukan pada dua komponen utama sehingga vektor ciri yang akan dideskripsikan hanya pada PC1 dan PC2. Pada PC1, vektor ciri memiliki nilai yang relatif sama walaupun pada peubah X4 (Umur Harapan Hidup Saat Lahir) nilainya cenderung jauh dari -0.5. Hal ini mengindikasikan bahwa nilai-nilai pada PC1 dapat menggambarkan Indeks Pembangunan Manusia

(IPM) untuk semua peubah. Oleh karena itu, PC1 dapat digunakan untuk menentukan kota/kabupaten mana yang memiliki IPM tertinggi untuk semua indikator yang digunakan. Vektor ciri PC2 memiliki nilai negatif dan positif. Nilai negatif mengindikasikan bahwa untuk indikator X2 dan X4, semakin besar PC2 maka semakin kecil nilai untuk indikator Pengeluaran per Kapita Disesuaikan dan Umur Harapan Hidup Saat Lahir. Sedangkan, nilai positif mengindikasikan bahwa untuk indikator X1 dan X3, semakin besar PC2 maka semakin besar pula nilai untuk indikator Harapan Lama Sekolah dan Rata-rata Lama Sekolah. Berdasarkan hal tersebut, PC2 dapat digunakan untuk menentukan kota/kabupaten yang memiliki nilai Harapan Lama Sekolah dan Rata-rata Lama Sekolah dengan Pengeluaran Perkapita dan Umur Harapan Hidup Saat Lahir.