PROGRAM STUDI STATISTIKA
SK BAN-PT No. 1765/SK/BAN-PT/AK-PPJ/S/III/2022
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PGRI ADI BUANA SURABAYA

FORM F.SK05
BUKTI BIMBINGAN SKRIPSI

Nama Mahasiswa : Dika Mhardika Sari

NIM : 212409005

Judul Skripsi : Pengelompokkan Kab/Kota di Jawa Timur Berdasarkan Indikator Tenaga Kerja Dengan Metode K-Means dan K-Medoids Clustering

Dosen Pembimbing : 1. Artanti Indrasetianingsih, S. Si., M. Si

Materi Pembimbingan Proposal	Tanda Tangan Dosen Pembimbing
1. Bimbingan variabel	
2. Bimbingan abstrak	
3. Bimbingan bab 1, 2	
4. Bimbingan bab 3	
5. Bimbingan bab 4	
6. Bimbingan bab 5	
7. Bimbingan hasil artikel skripsi	
8. Bimbingan akhir hasil artikel skripsi	

Catatan: *) Coret yang tidak sesuai

PROGRAM STUDI STATISTIKA
SK BAN-PT No. 1765/SK/BAN-PT/AK-PPJ/S/III/2022
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PGRI ADI BUANA SURABAYA

FORM F.SK08
PERBAIKAN/REVISI SEMINAR DAN UJIAN SKRIPSI

Nama Mahasiswa : Dika Mhardika Sari

NIM : 212409005

Judul Skripsi : Pengelompokkan Kab/Kota di Jawa Timur Berdasarkan Indikator Tenaga Kerja Dengan Metode K-Means dan K-Medoids Clustering

Dosen Pembimbing : 1. Artanti Indrasetianingsih, M.Si

Materi Revisi Seminar dan Ujian Skripsi	Tanda Tangan Dosen Penguji
1. Revisi bab 2 dan bab 4	
2. Revisi bab 1, 3, dan 5	
3. Revisi silhouette plot	
4. Revisi cluster plot	
5. Revisi bab 4 faktor 1, 2, 3	

Surabaya, 25 Februari 2023
Dosen Pembimbing,

Artanti Indrasetianingsih, S.Si., M.Si
NIP/NPP : 0609466 / DY

*Catatan: *) Coret yang tidak sesuai*

Lembar ini digunakan untuk bukti perbaikan makalah/jurnal dan hasil ujian skripsi
Batas waktu revisi proposal dua minggu terhitung dari waktu ujian proposal

LAMPIRAN

Lampiran 1. Hasil *Output Uji Bartlett*

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.777
Bartlett's Test of Sphericity	Approx. Chi-Square	412.684
	Df	66
	Sig.	.000

Lampiran 2. Hasil *Output Total Variance Explained*

Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	6.530	54.416	54.416
2	1.987	16.554	70.970
3	1.395	11.625	82.595
4	.567	4.721	87.316
5	.385	3.209	90.525
6	.312	2.601	93.126
7	.280	2.329	95.456
8	.208	1.732	97.188
9	.130	1.083	98.270
10	.088	.737	99.007
11	.077	.638	99.645
12	.043	.355	100.000

Extraction Method: Principal Component Analysis.

Lampiran 3. Hasil *Output Rotated Component Matrix*[a]

Rotated Component Matrix[a]

	Component		
	1	2	3
X1	.010	.292	-.885
X2	.263	.897	-.020
X3	.209	.929	.098
X4	.372	.311	.750
X5	.254	.189	.868
X6	.737	-.263	.219
X7	.794	.353	.124
X8	.885	.199	.285
X9	.677	.385	.329
X10	.864	.272	.100
X11	.928	.219	-.015
X12	.819	.302	.220

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 5 iterations.

Lampiran 4. *Syntax R Studio Average Silhouette Width Method "K-Means"*

```
library(readxl)
```

```
datakmeansfaktor3 <- read_excel("D:/RUNN PENDAPATAN DAN
KETENAGAKERJAAN/KMEANS FAKTOR3/DATA KMEANS
FAKTOR3 2 38.xlsx")
View(datakmeansfaktor3)

library(ggplot2)
library(dplyr)
library(tidyr)
library(readr)
library(purrr)
library(tibble)
library(stringr)
library(forcats)
install.packages("tidyverse")
library(tidyverse)

library(cluster)
library(factoextra)

datakmeansfaktor3
View(datakmeansfaktor3)

str(datakmeansfaktor3)
```

```
is.na(datakmeansfaktor3)
summary(datakmeansfaktor3)

datacluster<-datakmeansfaktor3[2:3]
datacluster
View(datacluster)

data_norm<-scale(datacluster)
data_norm
View(data_norm)

(kmeans.result <- kmeans(data_norm, 4))

fviz_nbclust(data_norm, kmeans, method = "silhouette")

fviz_nbclust(data_norm, kmeans, method = "wss")

gapstat <- clusGap(data_norm, FUN = kmeans, K.max = 10, B = 76)
gapstat
fviz_nbclust(data_norm, kmeans, method = "gap")

final <- kmeans(data_norm, 2)
```

final

fviz_cluster(final, data = data_norm)

Lampiran 5. *Syntax R Studio* Nilai *Silhouette* 1-10 *Method "K-Means"*

```
library(readxl)
datakmeansfaktor3 <- read_excel("D:/RUNN PENDAPATAN DAN KETENAGAKERJAAN/KMEANS FAKTOR3/SILHOUETTE KMEANS FAKTOR3 2 38/DATA KMEANS FAKTOR3 2 38.xlsx")
View(datakmeansfaktor3)

library(ggplot2)
library(dplyr)
library(tidyr)
library(readr)
library(purrr)
library(tibble)
library(stringr)
library(forcats)
install.packages("tidyverse")
library(tidyverse)
```

```
names(datakmeansfaktor3)
head(datakmeansfaktor3)
str(datakmeansfaktor3)
summary(datakmeansfaktor3)
sd(datakmeansfaktor3$X4)
sd(datakmeansfaktor3$X5)

#install
#package:ggplot2
#package:stats
#package:graphics
#package:httr
#package:plotly

library(ggplot2)
library(plotly)
ggplot(datakmeansfaktor3, aes(x = X4)) +
  geom_histogram(aes(y = ..count..),
  binwidth = 5, colour = "#1F3552", fill = "#4271AE")
,+
  scale_x_continuous(name = "X4", breaks = seq(20, 70, 5)) +
  scale_y_continuous(name = "Frekuensi",breaks = seq(0, 40, 5)) +
  ggtitle("Historgram Frekuensi DATA Berdasarkan X4") +
```

```
  geom_vline(xintercept = mean(datakmeansfaktor3$X4), size = 1,
  colour = "#FF3721",linetype = "dashed")

boxplot(datakmeansfaktor3$X4,col="#ff0066",
        main="Boxplot untuk Analisis Deskriptif X4")

library(corrplot)       # muat library untuk menggunakannya
corM<-cor(datakmeansfaktor3[2:3]) # menghitung matriks korelasi
corM

library(purrr)
set.seed(123)
# fungsi untuk menghitung jumlah total intra-cluster optimumkuadrat
(iss)
iss <- function(k) {
kmeans(datakmeansfaktor3[,2:3],k,iter.max=100,nstart=100,algorithm=
"Lloyd")$tot.withinss
}
k.values <- 1:10
iss_values <- map_dbl(k.values, iss)
plot(k.values, iss_values,
     type="b", pch = 19, frame = FALSE,
     xlab="Jumlah cluster optimumK",
```

```
    ylab="Jumlah Total Intra-cluster optimumKuadrat")

library(cluster)
library(gridExtra)
library(grid)

datacluster<-datakmeansfaktor3[2:3]
datacluster
View(datacluster)

data_norm<-scale(datacluster)
data_norm
View(data_norm)

k2<-kmeans(data_norm[,1:2],2,iter.max=100,nstart=25,
algorithm="Lloyd")
s2<-plot(silhouette(k2$cluster,dist(data_norm[,1:2],
"euclidean")))

k3<-kmeans(data_norm[,1:2],3,iter.max=100,nstart=25,
algorithm="Lloyd")
s3<-plot(silhouette(k3$cluster,dist(data_norm[,1:2],
"euclidean")))
```

```
k4<-kmeans(data_norm[,1:2],4,iter.max=100,nstart=25,
algorithm="Lloyd")
s4<-plot(silhouette(k4$cluster,dist(data_norm[,1:2],
"euclidean")))

k5<-kmeans(data_norm[,1:2],5,iter.max=100,nstart=25,
algorithm="Lloyd")
s5<-plot(silhouette(k5$cluster,dist(data_norm[,1:2],
"euclidean")))

k6<-kmeans(data_norm[,1:2],6,iter.max=100,nstart=25,
algorithm="Lloyd")
s6<-plot(silhouette(k6$cluster,dist(data_norm[,1:2],
"euclidean")))

k7<-kmeans(data_norm[,1:2],7,iter.max=100,nstart=25,
algorithm="Lloyd")
s7<-plot(silhouette(k7$cluster,dist(data_norm[,1:2],
"euclidean")))

k8<-kmeans(data_norm[,1:2],8,iter.max=100,nstart=25,
algorithm="Lloyd")
```

```
s8<-plot(silhouette(k8$cluster,dist(data_norm[,1:2],
"euclidean")))

k9<-kmeans(data_norm[,1:2],9,iter.max=100,nstart=25,
algorithm="Lloyd")
s9<-plot(silhouette(k9$cluster,dist(data_norm[,1:2],
"euclidean")))

k10<-kmeans(data_norm[,1:2],10,iter.max=100,nstart=25,
algorithm="Lloyd")
s10<-plot(silhouette(k10$cluster,dist(data_norm[,1:2],
"euclidean")))
```

Lampiran 6. ***Syntax R Studio Method "K-Medoids"***

Keteranggan : Nilai *Silhouette* (crit)

```
library(readxl)
datakmedoidsfaktor3 <- read_excel("D:/RUNN PENDAPATAN DAN
KETENAGAKERJAAN/KMEDOIDS FAKTOR3/DATA KMEDOIDS
FAKTOR3 2 38.xlsx")
View(datakmedoidsfaktor3)

library(ggplot2)
library(dplyr)
```

```
library(tidyr)
library(readr)
library(purrr)
library(tibble)
library(stringr)
library(forcats)
install.packages("tidyverse")
library(tidyverse)

install.packages(fpc)
install.packages(cluster)
install.packages(factoextra)
library(fpc)
library(cluster)
library(factoextra)

str(datakmedoidsfaktor3)
datakmedoidsfaktor3$KABKOTA<-NULL
head(datakmedoidsfaktor3)

summary(is.na(datakmedoidsfaktor3))
na.omit(datakmedoidsfaktor3)
boxplot(datakmedoidsfaktor3)
```

```
datacluster<-datakmedoidsfaktor3[1:2]
datacluster
View(datacluster)

data_norm<-scale(datacluster)
data_norm
View(data_norm)

pamk.hasil <-pamk(data_norm)
pamk.hasil
pamk.hasil$nc

fviz_nbclust(data_norm, pam, method = "silhouette")

pam.hasil  <- pam(data_norm, 2)
summary(pam.hasil)

pam.hasil$medoids
pam.hasil$diss

data.frame(data_norm,pam.hasil$clustering)

fviz_cluster(pam.hasil)
```

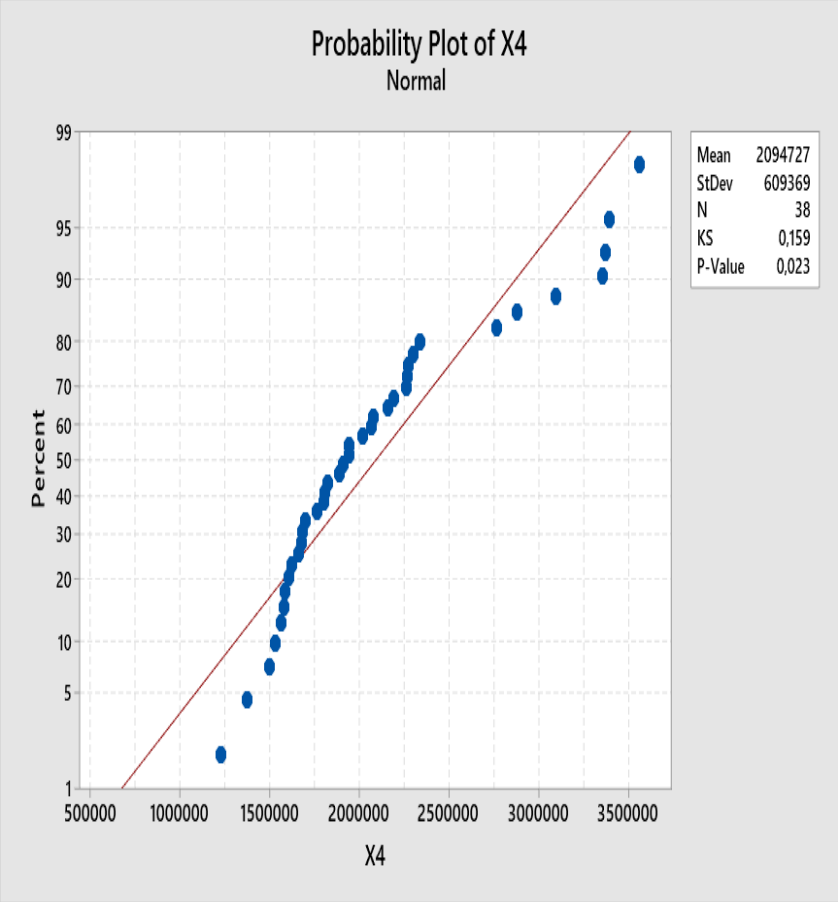
data_norm%>%

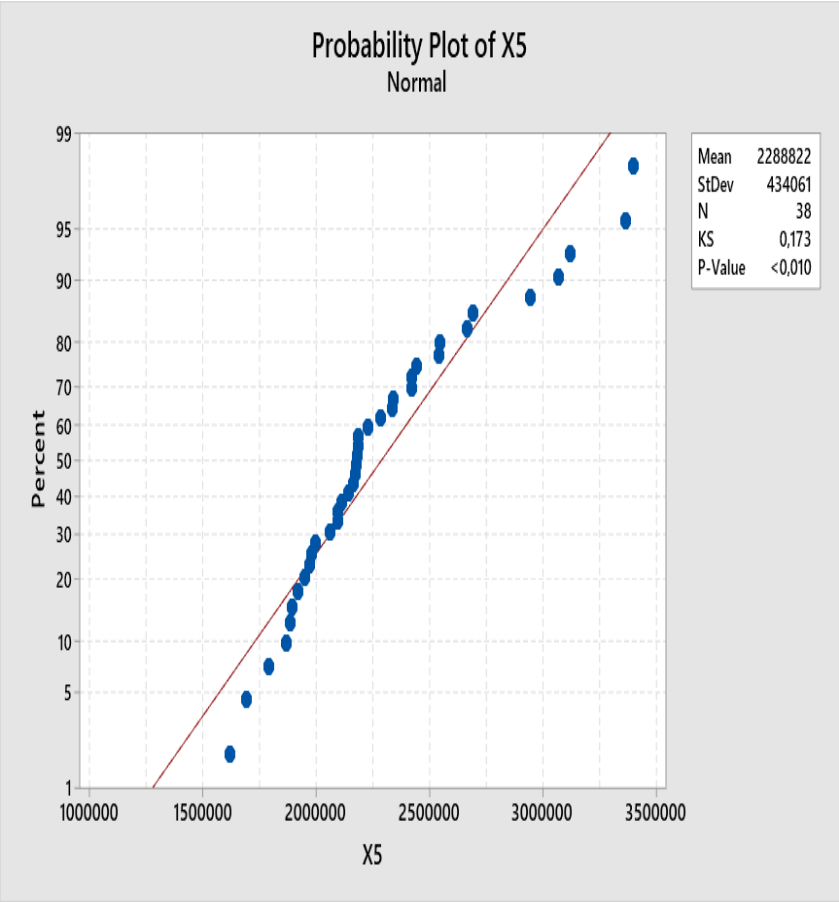
mutate(Cluster optimum= pam.hasil$cluster) %>%

group_by(Cluster) %>%

summarise_all("mean")

Lampiran 7. Hasil *Output* Uji Distribusi Normal *Kolmogorov-Smirnov*





Lampiran 8. Hasil *Output* Deskriptif *Mean* dan *Median* Variabel X4 CLUSTER OPTIMUM1

Statistics

Variable	N	N*	Mean	Median
X4 CLUSTER OPTIMUM1	32	0	1876299	1815962

Variabel X4 CLUSTER OPTIMUM2

Statistics

Variable	N	N*	Mean	Median
X4 CLUSTER OPTIMUM2	6	0	3259673	3367380

Variabel X5 CLUSTER OPTIMUM1

Statistics

Variable	N	N*	Mean	Median
X5 CLUSTER OPTIMUM1	32	0	2157840	2154043

Variabel X5 CLUSTER OPTIMUM2

Statistics

Variable	N	N*	Mean	Median
X5 CLUSTER OPTIMUM2	6	0	2987394	3033031