



LAMPIRAN

Lampiran 1. Data Mahasiswa UNTIRTA 2021/2022

No	NIM	Jenis Kelamin	Jalur Masuk	Kategori Provinsi	Kategori Pulau	Pekerjaan Ayah	Tahun Lulus SMTA	Kategori SMTA	kWh Listrik	Bidik misi	IPK
1	1111210001	P	SNM PTN	Prop. Banten	Pulau Jawa	Karyawan Swasta	2021	SMAN	1300	Tidak	3.98
2	1111210002	P	SNM PTN	Prop. Banten	Pulau Jawa	Karyawan Swasta	2021	SMAN	900	Tidak	3.76
3	1111210003	P	SNM PTN	Prop. Banten	Pulau Jawa	Karyawan Swasta	2021	SMAS	2200	Tidak	3.92
4	1111210005	P	SNM PTN	Prop. Banten	Pulau Jawa	Pedagang/Wiraswasta	2021	MAN	450	Ya	3.92
...
4096	8884210039	P	SMM PTN	Prop. Banten	Luar Pulau Jawa	Petani/Nelayan	2021	SMAN	450	Tidak	2.92

**Data tidak dilampirkan semua karena bersifat rahasia*

Lampiran 2. Hasil One-Way ANOVA pada Variabel Jenis Kelamin dengan Variabel IPK

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Jenis Kelamin	1	64.83	64.8331	330.68	0.000
Error	4094	802.67	0.1961		
Total	4095	867.50			

Lampiran 3. Hasil One-Way ANOVA pada Variabel Jalur Masuk dengan Variabel IPK

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Jalur Masuk	2	13.37	6.6873	32.05	0.000
Error	4093	854.13	0.2087		
Total	4095	867.50			

Lampiran 4. Hasil *One-Way* ANOVA pada Variabel Kategori Provinsi dengan Variabel IPK

Analysis of Variance						
Source	DF	Adj SS	Adj MS	F-Value	P-Value	
Kategori Provinsi	1	3.281	3.2812	15.54	0.000	
Error	4094	864.221	0.2111			
Total	4095	867.502				

Lampiran 5. Hasil *One-Way* ANOVA pada Variabel Kategori Pulau dengan Variabel IPK

Analysis of Variance						
Source	DF	Adj SS	Adj MS	F-Value	P-Value	
Kategori Pulau	1	0.080	0.07960	0.38	0.540	
Error	4094	867.422	0.21188			
Total	4095	867.502				

Lampiran 6. Hasil *One-Way* ANOVA pada Variabel Pekerjaan Ayah dengan Variabel IPK

Analysis of Variance						
Source	DF	Adj SS	Adj MS	F-Value	P-Value	
Pekerjaan Ayah	16	5.497	0.3436	1.63	0.054	
Error	4079	862.004	0.2113			
Total	4095	867.502				

Lampiran 7. Hasil *One-Way* ANOVA pada Variabel Tahun Lulus SMTA dengan Variabel IPK

Analysis of Variance						
Source	DF	Adj SS	Adj MS	F-Value	P-Value	
Tahun Lulus SMTA	2	0.429	0.2144	1.01	0.364	
Error	4093	867.073	0.2118			
Total	4095	867.502				

Lampiran 8. Hasil *One-Way* ANOVA pada Variabel kWh Listrik dengan Variabel IPK

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Kwh Listrik	4	3.385	0.8463	4.01	0.003
Error	4091	864.117	0.2112		
Total	4095	867.502			

Lampiran 9. Hasil *One-Way* ANOVA pada Variabel Bidikmisi dengan Variabel IPK

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Bidikmisi	1	5.927	5.9269	28.16	0.000
Error	4094	861.575	0.2104		
Total	4095	867.502			

Lampiran 10. Hasil *One-Way* ANOVA pada Variabel Kategori SMTA dengan Variabel IPK

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Kategori SMTA	6	3.421	0.5702	2.70	0.013
Error	4089	864.081	0.2113		
Total	4095	867.502			

Lampiran 11. Syntax K-Prototype

```

#====input data====
data_awal<-read.csv("DATA VARIABEL SIGNIFIKAN ANOVA.csv", header = T, sep = ",")

# Simpan nama baris sebelum menghapus kolom "NIM"
nama_baris <- data_awal$NIM

# Menghapus kolom "NIM"
data_klaster <- subset(data_awal, select = -NIM)

# Konversi data menjadi data frame dan tetapkan nama baris
data_klaster <- as.data.frame(data_klaster)
rownames(data_klaster) <- nama_baris

# Menampilkan data dengan perubahan
print(data_klaster)

# Menentukan jenis variabel
data_klaster$Jenis_kelamin <- as.factor(data_klaster$Jenis_kelamin)
data_klaster$Jalur_Masuk <- as.factor(data_klaster$Jalur_Masuk)
data_klaster$Kategori_Provinsi <- as.factor(data_klaster$Kategori_Provinsi)
data_klaster$Kategori_SMTA <- as.factor(data_klaster$Kategori_SMTA)
data_klaster$KWH_Listrik <- as.factor(data_klaster$KWH_Listrik)
data_klaster$Bidikmisi <- as.factor(data_klaster$Bidikmisi)

# Cek tipe data
str(data_klaster)

#====ALGORITMA K-PROTOTYPES====
library(clustMixType)
library(data.table)
library(dplyr)

# Menentukan pemilihan alternatif
nilai.seed <- sample(1:9999,1)
set.seed(nilai.seed)

```

Lampiran 11. Syntax K-Prototype (Lanjutan)

```

# Menentukan pemilihan alternatif
nilai.seed <- sample(1:9999,1)
set.seed(nilai.seed)

# Menghitung jarak euclidean
jarak_euclidean <- dist(data_klaster, method = "euclidean")

# Menentukan klaster
#2 klaster
dua_klaster <- kproto(data_klaster, k=2)
#3 klaster
tiga_klaster <- kproto(data_klaster, k=3)
#4 klaster
empat_klaster <- kproto(data_klaster, k=4)
#5 klaster
lima_klaster <- kproto(data_klaster, k=5)
#6 klaster
enam_klaster <- kproto(data_klaster, k=6)

hasil_klaster <- cbind(data_awal,
                      "dua klaster" = dua_klaster$cluster,
                      "tiga klaster" = tiga_klaster$cluster,
                      "empat klaster" = empat_klaster$cluster,
                      "lima klaster" = lima_klaster$cluster,
                      "enam klaster" = enam_klaster$cluster)

# statistik setiap klaster
library(fpc)
stat_dua_klaster <- cluster.stats(jarak_euclidean, dua_klaster$cluster)
stat_tiga_klaster <- cluster.stats(jarak_euclidean, tiga_klaster$cluster)
stat_empat_klaster <- cluster.stats(jarak_euclidean, empat_klaster$cluster)
stat_lima_klaster <- cluster.stats(jarak_euclidean, lima_klaster$cluster)
stat_enam_klaster <- cluster.stats(jarak_euclidean, enam_klaster$cluster)

```

Lampiran 11. *Syntax K-Prototype (Lanjutan)*

```

====VISUALISASI====
library(factoextra)
library(cluster)
library(NbClust)

#2 klaster
sil_dua_klaster <- silhouette(dua_klaster$cluster, jarak_euclidean)
grafik_dua_klaster <- fviz_silhouette(sil_dua_klaster)

#3 klaster
sil_tiga_klaster <- silhouette(tiga_klaster$cluster, jarak_euclidean)
grafik_tiga_klaster <- fviz_silhouette(sil_tiga_klaster)

#4 klaster
sil_empat_klaster <- silhouette(empat_klaster$cluster, jarak_euclidean)
grafik_empat_klaster <- fviz_silhouette(sil_empat_klaster)

#5 klaster
sil_lima_klaster <- silhouette(lima_klaster$cluster, jarak_euclidean)
grafik_lima_klaster <- fviz_silhouette(sil_lima_klaster)

#6 klaster
sil_enam_klaster <- silhouette(enam_klaster$cluster, jarak_euclidean)
grafik_enam_klaster <- fviz_silhouette(sil_enam_klaster)

====Validasi Internal====

#average silhouette width
silhouette_avg <- cbind.data.frame("2 klaster" = stat_dua_klaster$avg.silwidth,
                                   "3 klaster" = stat_tiga_klaster$avg.silwidth,
                                   "4 klaster" = stat_empat_klaster$avg.silwidth,
                                   "5 klaster" = stat_lima_klaster$avg.silwidth,
                                   "6 klaster" = stat_enam_klaster$avg.silwidth)
rownames(silhouette_avg) <- "Rata-rata Koefisien Silhouette"

```

Lampiran 11. *Syntax K-Prototype (Lanjutan)*

```

#indeks Dunn
indeks_dunn <- cbind.data.frame("2 klaster" = stat_dua_klaster$dunn,
                                "3 klaster" = stat_tiga_klaster$dunn,
                                "4 klaster" = stat_empat_klaster$dunn,
                                "5 klaster" = stat_lima_klaster$dunn,
                                "6 klaster" = stat_enam_klaster$dunn)
rownames(indeks_dunn) <- "Indeks Dunn"

#resume validasi internal
validasi_internal <- rbind.data.frame(silhouette_avg, indeks_dunn)

====validasi eksternal====
entropy_klaster <- cbind.data.frame("2 klaster" = stat_dua_klaster$entropy,
                                    "3 klaster" = stat_tiga_klaster$entropy,
                                    "4 klaster" = stat_empat_klaster$entropy,
                                    "5 klaster" = stat_lima_klaster$entropy,
                                    "6 klaster" = stat_enam_klaster$entropy)

write.csv(validasi_internal, "validasi internal fix.csv")
write.csv(entropy_klaster, "validasi eksternal fix.csv")
write.csv(hasil_klaster, "hasil klaster fix.csv")

getwd()

```

Lampiran 12. Hasil Klaster

No	NIM	Klaster				
		2	3	4	5	6
1	1111210001	2	2	1	2	6
2	1111210002	1	3	4	2	6
3	1111210003	2	2	4	2	6
4	1111210005	2	2	3	4	6
...
4096	8884210039	2	2	2	3	5

Lampiran 13. Hasil *General Linear Model*

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Jenis Kelamin	1	0.00	0.001	0.01	0.907
Jalur Masuk	2	0.12	0.058	1.19	0.304
Kategori Provinsi	1	0.03	0.034	0.71	0.401
Kategori SMTA	6	0.64	0.106	2.18	0.042
Kwh Listrik	4	650.89	162.722	3348.76	0.000
Bidikmisi	1	0.28	0.281	5.78	0.016
IPK	239	157.18	0.658	13.53	0.000
Error	3841	186.64	0.049		
Lack-of-Fit	2611	186.64	0.071	*	*
Pure Error	1230	0.00	0.000		
Total	4095	1021.84			