

LAMPIRAN

1. Fungsi R

```
library(readxl)
library(factoextra)
library(knitr)
library(summarytools)
library(gridExtra)
library(vtable)
```

2. Data Awal

```
#Data
Data.Awal <- read_excel("E:/DATA/KLASIFIKASI BUKU.xlsx",
                        sheet = "Sheet6")
view(Data.Awal)
Data.Awal = data.frame(Data.Awal)
head(Data.Awal)
```

	NO.	KLASIFIKASI	TOTAL.PEMINJAM	STOK.BUKU
1	1	STATISTIK	103	106
2	2	ALJABAR	31	28
3	3	KALKULUS	41	29
4	4	MATEMATIKA UMUM	56	62
5	5	BAHASA INDONESIA	79	118
6	6	BAHASA INGGRIS	59	81
7	7	ILMU POLITIK	167	146
8	8	ILMU ADMINISTRASI	161	72
9	9	ILMU SOSIAL	142	133
10	10	KAMUS	8	17
11	11	AKUNTANSI	27	37
12	12	KEUANGAN	84	104
13	13	KIMIA	72	84
14	14	BIOLOGI	65	90
15	15	FILSAFAT	71	44
16	16	MANAJEMEN	149	138
17	17	EKONOMI	534	301
18	18	PENDIIKAN	127	143
19	19	METODE	143	116
20	20	ILMU HUKUM	60	75
21	21	AGAMA	32	160

▲	NO.	KLASIFIKASI	TOTAL.PEMINJAM	STOK.BUKU
	22	PSIKOLOGI	15	15
	23	SASTRA	62	68
	24	TEKNIK	35	26
	25	KOMUNIKASI	98	89
	26	SENI DAN DESAIN	23	20
	27	PERTANIAN	275	262
	28	BELAJAR DAN PEMBELAJARAN	29	48
	29	SOSIOLOGI	43	33
	30	PETERNAKAN	43	60
	31	SEJARAH	53	29
	32	FILOSOFI	21	10
	33	TOEFL	16	22
	34	ANALISIS	42	51
	35	PENULISAN	32	34
	36	LOGIKA	62	13
	37	ARGUMENTASI	8	9
	38	PANDUAN	13	15
	39	FIKSI	211	147
	40	INDONESIA	99	68
	41	LINGUISTIK	37	49
	42	ORGANISASI	82	89

▲	NO.	KLASIFIKASI	TOTAL.PEMINJAM	STOK.BUKU
	43	PEMERINTAHAN	38	21
	44	ILMU TANAH	14	28
	45	BAHASA	24	71
	46	CERPEN DAN PUISI	26	9
	47	SDM	17	14
	48	ETIKA/ KARAKTER	43	77
	49	WIRUSAHA	4	4
	50	LINGKUNGAN DAN BUDAYA	23	57
	51	GURU	19	26
	52	JURNAL	18	15
	53	LAIN-LAIN	113	149

3. Mencari data hilang

```
dfsummary(Data.Awal$TOTAL.PEMINJAM)
```

No	variable	stats / values	Freqs (% of valid)	Graph	valid	Missing
1	TOTAL.PEMINJAM [numeric]	Mean (sd) : 72.6 (85.4) min < med < max: 4 < 43 < 534 IQR (cv) : 60 (1.2)	47 distinct values	:	53 (100.0%)	0 (0.0%)

4. Pemusatan dan Penyebaran data

```
st(Data.Awal)
```

Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
NO.	53	27	15	1	14	40	53
TOTAL.PEMINJAM	53	73	85	4	24	84	534
STOK.BUKU	53	70	61	4	26	90	301

```
summary(Data.Awal)
```

```
> st(Data.Awal)
> summary(Data.Awal)
      NO.      KLASIFIKASI      TOTAL.PEMINJAM      STOK.BUKU
Min.   : 1      Length:53      Min.   : 4.00      Min.   : 4.00
1st Qu.:14      Class :character      1st Qu.: 24.00      1st Qu.: 26.00
Median :27      Mode  :character      Median : 43.00      Median : 57.00
Mean   :27                                     Mean   : 72.62      Mean   : 70.04
3rd Qu.:40                                     3rd Qu.: 84.00      3rd Qu.: 90.00
Max.   :53                                     Max.   :534.00      Max.   :301.00
```

5. Hasil Ragam/Varian dan Standar Deviansi

```
var(Data.Awal$TOTAL.PEMINJAM)
sqrt(var(Data.Awal$TOTAL.PEMINJAM))
```

```
> var(Data.Awal$TOTAL.PEMINJAM)
[1] 7285.047
> sqrt(var(Data.Awal$TOTAL.PEMINJAM))
[1] 85.35249
```

6. Boxplot

```
#Boxplot
boxplot(Data.Awal$TOTAL.PEMINJAM,las = 1,col = "blue",notch=T,
        boxwex = 0.2,frame=F)
# menambahkan garis rata-rata
abline(h = mean(Data.Awal$TOTAL.PEMINJAM),col = "red",lwd = 3)
# menambahkan data rata-rata
text(y=mean(Data.Awal$TOTAL.PEMINJAM)+0.5,x=0.6,paste("Mean:",
        round(mean(Data.Awal$TOTAL.PEMINJAM),1)),col = "black",cex = 0.8,pos = 3)
# menambahkan data kuartil
text(y=boxplot.stats(Data.Awal$TOTAL.PEMINJAM)$stats,
        labels = round(boxplot.stats(Data.Awal$TOTAL.PEMINJAM)$stats,3),
        x = 1.1, cex = 0.8,col = "black")
# menambahkan data pencilan
text(y=boxplot.stats(Data.Awal$TOTAL.PEMINJAM)$out, labels =
        round(boxplot.stats(Data.Awal$TOTAL.PEMINJAM)$out,1),x = 1.1, cex = 0.8,col = "red")
```

7. Hasil Tranformasi Data Plot dan Tabel sebelum penaganan asumsi

```
#Transformasi+ plot normal
data.trans=as.data.frame(scale(Data.Awal$TOTAL.PEMINJAM))
ggplot(data.trans, aes(sample=V1)) +
  stat_qq() +
  geom_abline(color="red", intercept=mean(data.trans$V1), slope=sd(data.trans$V1)) +
  labs(title="Gaussian (normal) Data", subtitle="Q-Q Plot Data") +
  theme_light()
```

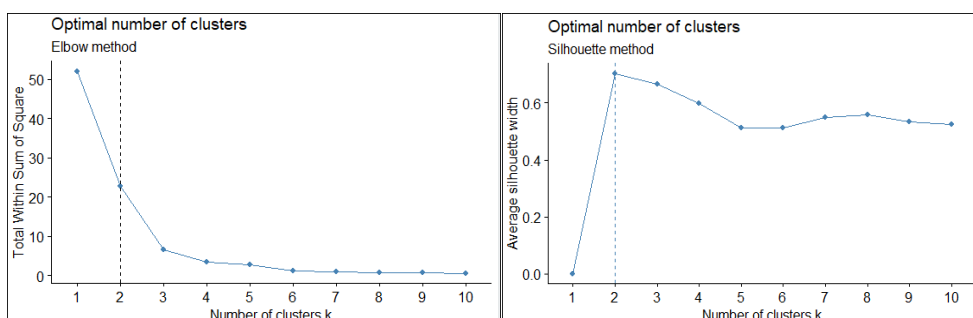
	V1
1	0.355904781
2	-0.487655868
3	-0.370494667
4	-0.194752865
5	0.074717898
6	-0.159604504
7	1.105736469
8	1.035439749
9	0.812833466
10	-0.757126631
11	-0.534520348
12	0.133298499
13	-0.007294943
14	-0.089307784
15	-0.019011063
16	0.894846307
17	5.405552556
18	0.637091664
19	0.824549586
20	-0.147888384
21	-0.475939748

	V1
22	-0.675113790
23	-0.124456144
24	-0.440791387
25	0.297324181
26	-0.581384829
27	2.371077443
28	-0.511088108
29	-0.347062426
30	-0.347062426
31	-0.229901225
32	-0.604817069
33	-0.663397670
34	-0.358778547
35	-0.475939748
36	-0.124456144
37	-0.757126631
38	-0.698546030
39	1.621245755
40	0.309040301
41	-0.417359147
42	0.109866259

	V1
43	-0.405643027
44	-0.686829910
45	-0.569668709
46	-0.546236469
47	-0.651681550
48	-0.347062426
49	-0.803991111
50	-0.581384829
51	-0.628249309
52	-0.639965430
53	0.473065983

8. WSS dan Silhouette

```
#cluster optimal
fviz_nbclust(data.trans, kmeans, method = "wss")+geom_vline(xintercept=2,
  linetype=2)+labs(subtitle="Elbow method")
fviz_nbclust(data.trans, kmeans, method = "silhouette")+labs(subtitle="Silhouette method")
```



9. *Index Dunn*

```
library(c1valid)
set.seed(1234)
validation <- c1valid(data.frame(Data.Awal$TOTAL.PEMINJAM),
2:10, c("kmeans"), validation = "internal", length(Data.Awal), verbose = T)
summary(validation)
```

```
The memory and time required may be excessive, do you wish to continue?
(y to continue, any other character to exit)
y
[1] "Finished internal validation, kmeans 2 clusters"
[1] "Finished internal validation, kmeans 3 clusters"
[1] "Finished internal validation, kmeans 4 clusters"
[1] "Finished internal validation, kmeans 5 clusters"
[1] "Finished internal validation, kmeans 6 clusters"
[1] "Finished internal validation, kmeans 7 clusters"
[1] "Finished internal validation, kmeans 8 clusters"
[1] "Finished internal validation, kmeans 9 clusters"
[1] "Finished internal validation, kmeans 10 clusters"
Warning message:
In c1valid(data.frame(Data.Awal$TOTAL.PEMINJAM), 2:10, c("kmeans"), :
rownames for data not specified, using 1:nrow(data)
> summary(validation)

Clustering Methods:
kmeans

Cluster sizes:
 2 3 4 5 6 7 8 9 10

Validation Measures:
           2      3      4      5      6      7      8      9     10
kmeans Connectivity  5.1079  7.7004 14.0718 15.9548 17.2048 23.9175 28.2036 34.5091 37.9710
      Dunn          0.2471  0.0791  0.0795  0.1562  0.1852  0.0750  0.0769  0.0323  0.0455
      Silhouette    0.8093  0.6668  0.6255  0.6052  0.5850  0.5533  0.5488  0.5292  0.5549

Optimal Scores:
           Score Method Clusters
Connectivity 5.1079 kmeans 2
Dunn         0.2471 kmeans 2
Silhouette   0.8093 kmeans 2
```

10. Hasil *cluster* optimal sebelum penanganan asumsi

```
#c1cluster
c1<- kmeans(data.trans,2,nstart = 25)
data.new=data.frame(Data.Awal$TOTAL.PEMINJAM,c1$cluster)
colnames(data.new)=c("x1","c")
rownames(data.new)=Data.Awal$KLASIFIKASI
head(data.new)
```

	X1	C		X1	C
STATISTIK	103	2	PSIKOLOGI	15	2
ALJABAR	31	2	SASTRA	62	2
KALKULUS	41	2	TEKNIK	35	2
MATEMATIKA UMUM	56	2	KOMUNIKASI	98	2
BAHASA INDONESIA	79	2	SENI DAN DESAIN	23	2
BAHASA INGGRIS	59	2	PERTANIAN	275	1
ILMU POLITIK	167	2	BELAJAR DAN PEMBELAJARAN	29	2
ILMU ADMINISTRASI	161	2	SOSIOLOGI	43	2
ILMU SOSIAL	142	2	PETERNAKAN	43	2
KAMUS	8	2	SEJARAH	53	2
AKUNTANSI	27	2	FILOSOFI	21	2
KEUANGAN	84	2	TOEFL	16	2
KIMIA	72	2	ANALISIS	42	2
BIOLOGI	65	2	PENULISAN	32	2
FILSAFAT	71	2	LOGIKA	62	2
MANAJEMEN	149	2	ARGUMENTASI	8	2
EKONOMI	534	1	PANDUAN	13	2
PENDIIKAN	127	2	FIKSI	211	2
METODE	143	2	INDONESIA	99	2
ILMU HUKUM	60	2	LINGUISTIK	37	2
AGAMA	32	2	ORGANISASI	82	2

	X1	C
PEMERINTAHAN	38	2
ILMU TANAH	14	2
BAHASA	24	2
CERPEN DAN PUISI	26	2
SDM	17	2
ETIKA/ KARAKTER	43	2
WIRUSAHA	4	2
LINGKUNGAN DAN BUDAYA	23	2
GURU	19	2
JURNAL	18	2
LAIN-LAIN	113	2

11. Hasil Transformasi Data Plot dan Tabel sebelum penanganan asumsi

▲	V1	▼
1	4.634729	
2	3.433987	
3	3.713572	
4	4.025352	
5	4.369448	
6	4.077537	
7	5.117994	
8	5.081404	
9	4.955827	
10	2.079442	
11	3.295837	
12	4.430817	
13	4.276666	
14	4.174387	
15	4.262680	
16	5.003946	
17	6.280396	
18	4.844187	
19	4.962845	
20	4.094345	
21	3.465736	

▲	V1	▼
22	2.708050	
23	4.127134	
24	3.555348	
25	4.584967	
26	3.135494	
27	5.616771	
28	3.367296	
29	3.761200	
30	3.761200	
31	3.970292	
32	3.044522	
33	2.772589	
34	3.737670	
35	3.465736	
36	4.127134	
37	2.079442	
38	2.564949	
39	5.351858	
40	4.595120	
41	3.610918	
42	4.406719	

▲	V1	▼
43	3.637586	
44	2.639057	
45	3.178054	
46	3.258097	
47	2.833213	
48	3.761200	
49	1.386294	
50	3.135494	
51	2.944439	
52	2.890372	
53	4.727388	

12. Hasil *cluster* optimal sesudah penanganan asumsi

	X1	C		X1	C
STATISTIK	103	1	PSIKOLOGI	15	2
ALJABAR	31	2	SASTRA	62	1
KALKULUS	41	2	TEKNIK	35	2
MATEMATIKA UMUM	56	1	KOMUNIKASI	98	1
BAHASA INDONESIA	79	1	SENI DAN DESAIN	23	2
BAHASA INGGRIS	59	1	PERTANIAN	275	1
ILMU POLITIK	167	1	BELAJAR DAN PEMBELAJARAN	29	2
ILMU ADMINISTRASI	161	1	SOSIOLOGI	43	2
ILMU SOSIAL	142	1	PETERNAKAN	43	2
KAMUS	8	2	SEJARAH	53	1
AKUNTANSI	27	2	FILOSOFI	21	2
KEUANGAN	84	1	TOEFL	16	2
KIMIA	72	1	ANALISIS	42	2
BIOLOGI	65	1	PENULISAN	32	2
FILSAFAT	71	1	LOGIKA	62	1
MANAJEMEN	149	1	ARGUMENTASI	8	2
EKONOMI	534	1	PANDUAN	13	2
PENDIIKAN	127	1	FIKSI	211	1
METODE	143	1	INDONESIA	99	1
ILMU HUKUM	60	1	LINGUISTIK	37	2
AGAMA	32	2	ORGANISASI	82	1

	X1	C
PEMERINTAHAN	38	2
ILMU TANAH	14	2
BAHASA	24	2
CERPEN DAN PUISI	26	2
SDM	17	2
ETIKA/ KARAKTER	43	2
WIRUSAHA	4	2
LINGKUNGAN DAN BUDAYA	23	2
GURU	19	2
JURNAL	18	2
LAIN-LAIN	113	1

DAFTAR RIWAYAT HIDUP



Penulis dilahirkan di Atambua Provinsi Nusa Tenggara Timur pada tanggal 24 April 2000, sebagai anak kedua dari empat bersaudara dari pasangan Bapak Oktovinus Bouk dan Ibu Imelda Funan Hane. Penulis berkebangsaan Indonesia dan beragama katolik. Berikut riwayat pendidikan penulis, yaitu pada tahun 2013 lulus dari SDI Haeneno, Kecamatan Io Kufeu, Kabupaten Malaka. Kemudian melanjutkan pendidikan di SMP Negeri 1 Atambua dan lulus pada tahun 2016. Pada tahun 2019 penulis lulus dari SMK Katholik St. Willibrodus Betun dan melanjutkan studi di Universitas Timor Program Studi Matematika melalui jalur MANDIRI hingga penulis selesai menyusun skripsi ini dengan motto “Keluar dari zona nyaman, bertekad menyelesaikan misi”.

Akhir kata penulis mengucapkan rasa syukur yang sebesar – besarnya atas selesainya menyusun skripsi dengan judul “IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING DALAM MENETUKAN PENAMBAHAN KOLEKSI BUKU DI PERPUSTAKAAN UNIVERSITAS TIMOR”.

Kefamenanu, Desember 2023

Paskalia Yunita Mau