

Nama : Andi Illa Erviani Nensi

Nim : 200102502002

Kelas : DS02

1. Menentukan objek yang akan dianalisis
- Objek dalam latihan ini adalah “Belanja” dengan 2 kategori yaitu sedikit (0) dan banyak (1). Variabel bebas lainnya ada 10. Data tersebut diimpor ke dalam R dengsn cara sebagai berikut:

```
dt.disk<-read.csv(file.choose()),header = T,sep = ",",dec = ",")

dt.disk1<-dt.disk[1:40,]

dt.disk2<-dt.disk[41:50,]

dt.disk1
```

ii..Layout Bersih Harga Kasir Lengkap AC Diskon Parkir Belanja Staf Citra

1	3	3	3	4	3	3	4	3	0	3	3
2	3	3	3	3	3	3	4	3	0	3	3
3	3	3	3	3	3	3	4	3	1	3	3
4	4	3	3	2	2	3	4	3	1	4	3
5	3	3	3	2	1	4	3	3	0	4	5
6	4	3	3	3	2	4	3	3	0	4	5
7	2	3	3	4	3	2	3	3	0	2	3
8	3	3	3	2	2	2	4	3	1	2	5
9	4	4	3	5	2	2	4	3	1	2	3
10	4	4	5	2	2	4	2	4	1	2	5
11	4	4	5	5	3	4	2	4	0	2	4
12	5	2	3	5	2	4	2	4	1	3	5
13	5	2	5	2	1	3	2	2	1	4	3
14	3	2	3	5	2	3	5	2	0	3	4
15	4	2	5	3	2	3	5	2	0	4	3
16	4	2	4	5	1	3	3	2	0	3	3
17	5	3	5	5	2	3	3	2	0	3	3
18	4	4	3	5	4	2	3	3	0	3	4
19	3	3	4	3	4	2	4	4	0	2	3
20	3	4	3	3	4	2	4	3	1	4	4
21	3	3	3	4	3	3	5	4	0	4	4
22	3	3	3	4	3	3	3	3	1	4	4
23	3	3	4	4	3	3	2	3	0	3	2
24	2	2	3	2	3	4	3	3	1	3	2
25	4	4	3	2	2	4	3	2	0	3	2
26	4	4	5	2	2	4	2	4	1	3	2
27	4	4	5	5	3	4	2	4	0	5	2
28	5	2	3	5	2	4	2	4	1	5	3
29	5	2	5	2	1	3	2	2	1	3	4
30	3	2	3	5	2	3	5	2	0	5	3
31	4	2	5	3	2	3	5	2	0	3	4

32	4	5	4	5	1	3	3	2	0	5	3
33	5	3	5	5	2	3	3	2	0	4	3
34	4	5	3	5	4	2	3	3	0	5	3
35	3	4	4	3	4	2	4	4	0	3	2
36	4	3	3	2	2	3	4	3	1	4	3
37	3	3	3	2	1	4	3	3	0	4	5
38	4	3	3	3	2	4	3	3	0	4	5
39	2	3	3	4	3	2	3	3	0	2	3
40	3	3	3	2	2	2	4	3	1	2	5

2. Pengujian asumsi multikolenaritas

```
> uji.multi<-rcorr(as.matrix(dt.disk1[,-5],type="pearson"))

> uji.multi
```

> uji.multi

	i..Layout	Bersih	Harga	Kasir	AC	Diskon	Parkir	Belanja	Staf	Citra
i..Layout	1.00	-0.03	0.51	0.19	0.32	-0.40	-0.16	0.16	0.30	0.05
Bersih	-0.03	1.00	0.00	0.11	-0.17	-0.19	0.31	-0.13	0.03	-0.14
Harga	0.51	0.00	1.00	0.01	0.19	-0.32	-0.11	-0.08	-0.05	-0.22
Kasir	0.19	0.11	0.01	1.00	-0.14	-0.04	-0.03	-0.40	0.20	-0.17
AC	0.32	-0.17	0.19	-0.14	1.00	-0.45	0.15	0.02	0.30	0.11
Diskon	-0.40	-0.19	-0.32	-0.04	-0.45	1.00	-0.27	-0.19	-0.01	0.04
Parkir	-0.16	0.31	-0.11	-0.03	0.15	-0.27	1.00	0.20	-0.16	0.07
Belanja	0.16	-0.13	-0.08	-0.40	0.02	-0.19	0.20	1.00	-0.12	0.12
Staf	0.30	0.03	-0.05	0.20	0.30	-0.01	-0.16	-0.12	1.00	-0.12
Citra	0.05	-0.14	-0.22	-0.17	0.11	0.04	0.07	0.12	-0.12	1.00

n= 40

P

	i..Layout	Bersih	Harga	Kasir	AC	Diskon	Parkir	Belanja	Staf	Citra
i..Layout		0.8431	0.0008	0.2530	0.0466	0.0114	0.3220	0.3123	0.0614	0.7392
Bersih	0.8431		0.9782	0.4904	0.2903	0.2428	0.0518	0.4093	0.8494	0.4008
Harga	0.0008	0.9782		0.9419	0.2472	0.0434	0.4847	0.6112	0.7384	0.1780
Kasir	0.2530	0.4904	0.9419		0.3960	0.7932	0.8593	0.0104	0.2252	0.3008
AC	0.0466	0.2903	0.2472	0.3960		0.0039	0.3605	0.9149	0.0606	0.5087
Diskon	0.0114	0.2428	0.0434	0.7932	0.0039		0.0860	0.2416	0.9727	0.7918
Parkir	0.3220	0.0518	0.4847	0.8593	0.3605	0.0860		0.2128	0.3158	0.6714
Belanja	0.3123	0.4093	0.6112	0.0104	0.9149	0.2416	0.2128		0.4457	0.4632
Staf	0.0614	0.8494	0.7384	0.2252	0.0606	0.9727	0.3158	0.4457		0.4687
Citra	0.7392	0.4008	0.1780	0.3008	0.5087	0.7918	0.6714	0.4632	0.4687	

Berdasarkan hasil uji multikolinearitas diperoleh bahwa variabel bebas mempunyai nilai P > 0.05 dan P < 0.05. Hal ini menunjukkan bahwa terjadi kasus multikolinearitas pada data.

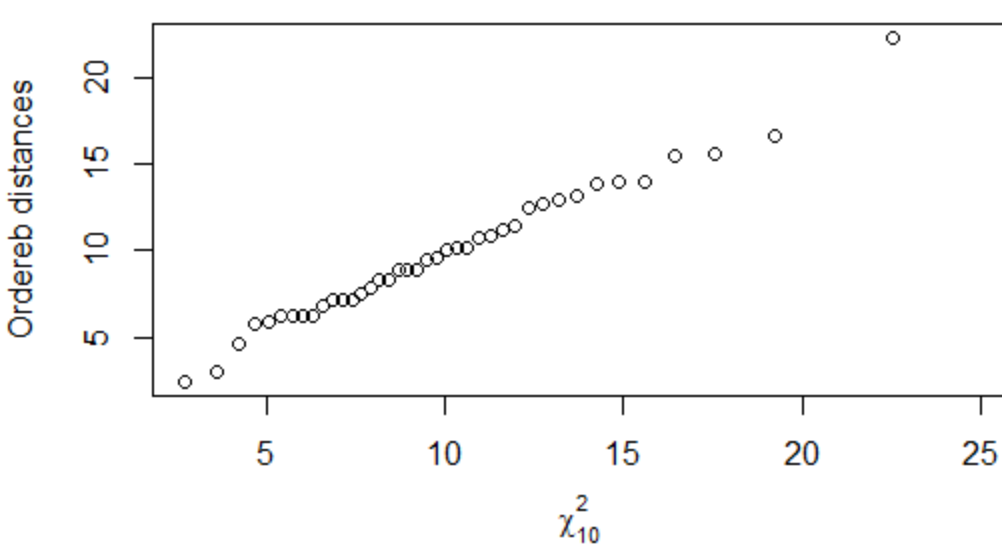
3. Pengujian asumsi distribusi normal multivariate

```
> plot(qc <- qchisq((1:nrow(x) - 1/2) / nrow(x), df = 10),

+ sd <- sort(d),

+ xlab = (expression(paste(chi[10]^2, "Quantile"))),

+ ylab = "Ordereb distances", xlim = range(qc)*c(1,1.1))
```



Bagan 1 Plot Chi-kuadrat

Berdasarkan Bagan 1, terlihat bahwa pola data sepertinya membentuk sebuah garis lurus. Untuk lebih meyakinkan bahwa data mengikuti distribusi normal multivariat, dilakukan pengujian, sebagai berikut:

```
> uji.mvn<-mvShapiro.Test(as.matrix(dt.disk1[, -5]))
> uji.mvn

Generalized shapiro-wilk test for Multivariate Normality by villasenor-Alva and
Gonzalez-Estrada

data:  as.matrix(dt.disk1[, -5])
MVW = 0.93973, p-value = 6.07e-07
```

Generalized Shapiro-Wilk test for Multivariate Normality by Villasenor-Alva and Gonzalez-Estrada

data: as.matrix(dt.disk1[, -5])
MVW = 0.93973, **p-value = 6.07e-07**

Hasil yang diperoleh adalah nilai -p>0.05. Ini berarti bahwa data mengikuti distribusi normal multivariat.

4. Pengujian asumsi kesamaan matriks varians kovarian

```
> uji.KMV<-BoxM(dt.disk1[, -5],dt.disk1[,5])
> uji.KMV

$chisq
[1] Inf

$df
[1] 165

$P.value
[1] 0

$Test
[1] "BoxM"

attr(,"class")
[1] "MVTests" "list"
```

\$p.value
[1] 0

Hasil yang diperoleh adalah nilai-p>0. Ini berarti bahwa data pada kelompok “Belanja sedikit” dan data pada kelompok “Belanja banyak” adalah bukan homogen.

5. Pengujian asumsi kesamaan vektor mean

```
> uji.KVM<-manova(cbind(Bersih,Harga,Kasir,Lengkap,AC,Diskon,Parkir,Staf,Citra)~Belanja,data=dt.disk1)
> summary(uji.KVM)
```

```
> uji.KVM<-manova(cbind(Bersih,Harga,Kasir,Lengkap,AC,Diskon,Parkir,Staf,Citra)~Belanja,data=dt.disk1)
> summary(uji.KVM)
              Df  Pillai approx F num Df den Df  Pr(>F)
Belanja      1 0.36202    1.8915      9    30 0.09219 .
Residuals   38
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pr(>F)

0.09219

Karena nilai-p>0.05, maka diperoleh bahwa tidak terdapat perbedaan vektor mean dari data. Ini menunjukkan bahwa pemisahan kelompok “Belanja sedikit” dan kelompok “Belanja banya” tidak signifikan.

6. Mengestimasi fungsi diskriminan

```
>model.disk<-lda(Belanja~Bersih+Harga+Kasir+Lengkap+AC+Diskon+Parkir+Staf+Citra,data = dt.disk1)

> model.disk
```

```
Call:
lda(Belanja ~ Bersih + Harga + Kasir + Lengkap + AC + Diskon +
    Parkir + Staf + Citra, data = dt.disk1)

Prior probabilities of groups:
      0      1 
0.625 0.375 

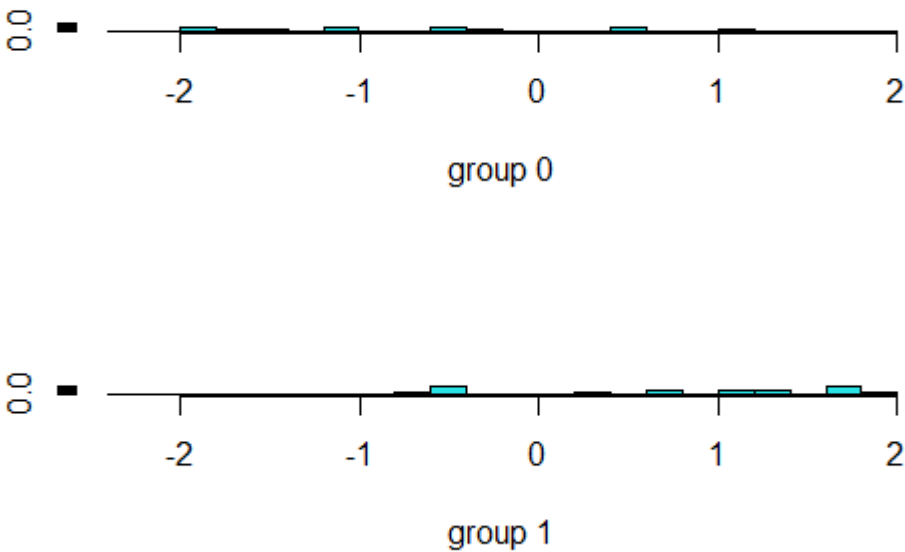
Group means:
      Bersih      Harga      Kasir  Lengkap      AC      Diskon      Parkir  Staf  Citra
0 3.160000 3.680000 3.880000      2.48 3.040000 3.440000 2.840000 3.44 3.36
1 2.933333 3.533333 2.866667      2.20 3.066667 3.066667 3.133333 3.20 3.60

Coefficients of linear discriminants:
              LD1
Bersih  -0.5615990
Harga   -0.3174686
Kasir   -0.6347040
Lengkap -0.7824461
AC       -1.1398210
Diskon  -0.6564489
Parkir   1.1488720
Staf     0.1672213
Citra   -0.1294279
```

Berdasarkan hasil pemodelan tersebut diperoleh fungsi diskriminan

$Y = -0.562 \text{ Bersih} - 0.317 \text{ Harga} - 0.635 \text{ Kasir} - 0.782 \text{ Lengkap} - 1.140 \text{ AC} - 0.656 \text{ Diskon} + 1.149 \text{ Parkir} + 0.167 \text{ Staf} - 0.129 \text{ Citra}$
Berdasarkan nilai Groups means bahwa variabel “Kasir,AC,Diskon,Parkir,Staf,Citra” untuk kelompok “Belanja sedikit” dan “Belanja banyak” mempunyai nilai yang hampir sama, sehingga dapat dikatakan bahwa dua variabel tersebut tidak mempengaruhi perilaku seseorang dalam Belanja. Sementara variabel “Bersih,Harga,Lengkap” mempunyai masing-masing nilai mean yang berbeda antara kelompok “Belanja sedikit” dan “Belanja banyak”, sehingga dapat dikatakan bahwa kedua variabel tersebut dapat mempengaruhi perilaku sesorang dalam membayr.

```
>plot(model.disk)
```



Bagan 2 Plot Hasil Fungsi Diskriminan

7. Mengecek Hasil Pengklasifikasian

```
> lda.klasik<-LdaClassic(dt.disk1[,-5],dt.disk1[,5])

> pred.klas<-predict(lda.klasik) > pred.klas
```

Apparent error rate 0.1

Classification table
Predicted
Actual 1 2 3 4
1 5 1 0 0
2 0 17 1 0
3 0 2 9 0
4 0 0 0 5

Confusion matrix
Predicted
Actual 1 2 3 4
1 0.833 0.167 0.000 0
2 0.000 0.944 0.056 0
3 0.000 0.182 0.818 0
4 0.000 0.000 0.000 1

Berdasarkan hasil pengklasifikasian diperoleh tingkat keakuratan $(0.833+0.944+0.818+1)/4 = 0.89$
(89%) di mana terdapat 3 pengamatan yang salah klasifikasi.

8. Memprediksi objek baru

ii..Layout Bersih Harga Kasir Lengkap AC Diskon Parkir Belanja Staf Citra pred.class

1	3	3	3	4	3	3	4	3	0	3	3	0
2	3	3	3	3	3	3	4	3	0	3	3	0
3	3	3	3	3	3	3	4	3	1	3	3	0
4	4	3	3	2	2	3	4	3	1	4	3	1
6	4	3	3	3	2	4	3	3	0	4	5	0
7	2	3	3	4	3	2	3	3	0	2	3	1
8	3	3	3	2	2	2	4	3	1	2	5	1
9	4	4	3	5	2	2	4	3	1	2	3	0
10	4	4	5	2	2	4	2	4	1	2	5	1
11	4	4	5	5	3	4	2	4	0	2	4	0
12	5	2	3	5	2	4	2	4	1	3	5	1
13	5	2	5	2	1	3	2	2	1	4	3	1
14	3	2	3	5	2	3	5	2	0	3	4	0
15	4	2	5	3	2	3	5	2	0	4	3	0
16	4	2	4	5	1	3	3	2	0	3	3	0
17	5	3	5	5	2	3	3	2	0	3	3	0
18	4	4	3	5	4	2	3	3	0	3	4	0
19	3	3	4	3	4	2	4	4	0	2	3	1
20	3	4	3	3	4	2	4	3	1	4	4	0
21	3	3	3	4	3	3	5	4	0	4	4	0
22	3	3	3	4	3	3	3	3	1	4	4	0
23	3	3	4	4	3	3	2	3	0	3	2	0
24	2	2	3	2	3	4	3	3	1	3	2	1
25	4	4	3	2	2	4	3	2	0	3	2	0
26	4	4	5	2	2	4	2	4	1	3	2	1
27	4	4	5	5	3	4	2	4	0	5	2	0
28	5	2	3	5	2	4	2	4	1	5	3	1
29	5	2	5	2	1	3	2	2	1	3	4	1
30	3	2	3	5	2	3	5	2	0	5	3	0
31	4	2	5	3	2	3	5	2	0	3	4	0
32	4	5	4	5	1	3	3	2	0	5	3	0
33	5	3	5	5	2	3	3	2	0	4	3	0
34	4	5	3	5	4	2	3	3	0	5	3	0
35	3	4	4	3	4	2	4	4	0	3	2	0
36	4	3	3	2	2	3	4	3	1	4	3	1
37	3	3	3	2	1	4	3	3	0	4	5	1
38	4	3	3	3	2	4	3	3	0	4	5	0
39	2	3	3	4	3	2	3	3	0	2	3	1
40	3	3	3	2	2	2	4	3	1	2	5	1

Berdasarkan hasil diperoleh bahwa analisis diskriminan dapat digunakan untuk memisahkan kelompok “Belanja sedikit” dan “Belanja banyak”, berdasarkan variabel ersih,Harga,Kasir,Lengkap,AC,Diskon,Parkir,Staf,Citra dengan tingkat akurasi 89%. Tiga variabel, yaitu Harga bersih lengkap yang paling berpengaruh terhadap pemisahan tersebut.