

Analisis Korespondensi

Pengantar:

Analisis korespondensi adalah prosedur grafis yang mewakili asosiasi dalam tabel frekuensi atau jumlah. Keluaran dari analisis korespondensi mencakup representasi dua dimensi terbaik dari data, berikut dengan koordinat titik yang diplot, dan ukuran jumlah informasi yang disimpan di setiap dimensi (disebut inersia).

Contoh:

Suatu data frekuensi tentang hubungan antara kenyamanan kerja dengan pendapatan yang disajikan pada Tabel berikut :

Pendapatan	Kenyamanan Pekerjaan			
	Sangat Tidak Nyaman	Kurang Nyaman	Biasa Saja	Nyaman
< 25 juta per th	15	98	21	14
25-50 juta/th	4	49	125	250
> 50 juta/th	1	13	81	300

Konstruksikan Korepondensi antara Kenyamanan dengan Pendapatan!

Penyelesaian Menggunakan Program R:

Tahap 1. Input data

```
> n=read.table(header=T,text =
  "Pendapatan
  <25juta/th
  25-50juta/th
  >50juta/th
  ")
> data =
  read.table(header=T,text =
    "Sangat.Tidak.Nyaman  Kurang.Nyaman  Biasa.Saja  Nyaman
    15 98 21 14
    4 49 125 250
    1 13 81 300
    ")
> rownames(data)=n$Pendapatan
> data
      Sangat.Tidak.Nyaman Kurang.Nyaman Biasa.Saja Nyaman
<25juta/th             15           98         21      14
25-50juta/th             4           49        125     250
>50juta/th              1           13         81     300
```

Keterangan:

- input nama baris (**n**) dan kolom (**data**) menggunakan fungsi **read.table**
- menggabungkan nama baris ke dalam data menggunakan **rownames(data)=n\$Pendapatan**
- Munculkan data

Tahap 2. Analisis korespondensi

```
> library(ca)
> fit=ca(data)
> summary(fit)
```

Principal inertias (eigenvalues):

dim	value	%	cum%	scree plot
1	0.420978	97.1	97.1	*****
2	0.012713	2.9	100.0	*

Total: 0.433691 100.0

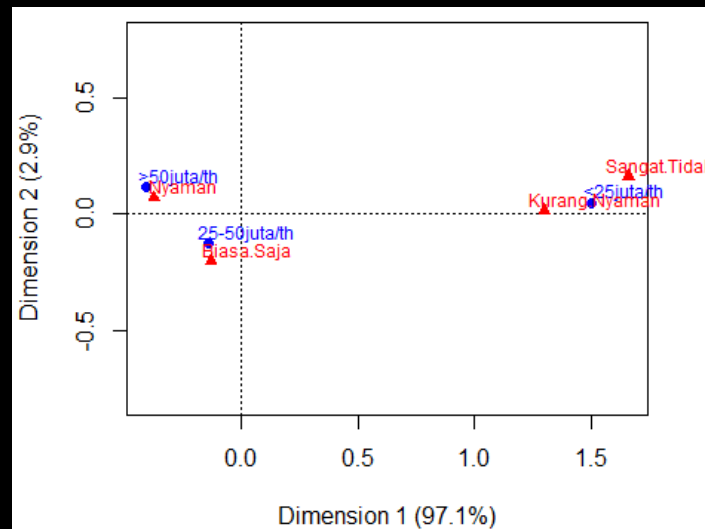
Rows:

	name	mass	qlt	inr	k=1	cor	ctr	k=2	cor	ctr
1	25jt	152	1000	795	1503	999	818	50	1	30
2	2550	441	1000	37	-143	570	21	-125	430	538
3	50jt	407	1000	169	-408	925	161	116	75	432

Columns:

	name	mass	qlt	inr	k=1	cor	ctr	k=2	cor	ctr
1	SnTN	21	1000	132	1662	991	135	162	9	43
2	KrnN	165	1000	642	1300	1000	662	16	0	3
3	BsSj	234	1000	31	-132	304	10	-199	696	731
4	Nymn	581	1000	195	-375	966	194	70	34	223

> plot(fit)



> chisq <- chisq.test(data)
> chisq

Pearson's Chi-squared test

data: data
X-squared = 421.11, df = 6, p-value < 2.2e-16

```
> #profil baris
> jml.baris<- apply(data, 1, sum)
> profil.baris<- data/jml.baris
> profil.baris
```

	Sangat.Tidak.Nyaman	Kurang.Nyaman	Biasa.Saja	Nyaman
<25juta/th	0.101351351	0.66216216	0.1418919	0.09459459
25-50juta/th	0.009345794	0.11448598	0.2920561	0.58411215
>50juta/th	0.002531646	0.03291139	0.2050633	0.75949367

```
> #profil kolom
> jml.kolom<- apply(data, 2, sum)
> profil.kolom<- data/jml.kolom
```

```

> profil.kolom
      Sangat.Tidak.Nyaman Kurang.Nyaman Biasa.Saja    Nyaman
<25juta/th      0.750000000      0.1737589 0.09251101 0.0875000
25-50juta/th      0.025000000      2.4500000 0.22163121 1.1013216
>50juta/th      0.004405286      0.0812500 4.05000000 0.5319149

> #rata2 profil kolom dan baris (= Mass pada fit)
> n <- sum(data)
> average.pb <- jml.baris/n
> average.pb
<25juta/th 25-50juta/th >50juta/th
  0.1524202   0.4407827   0.4067971

> average.pk <- jml.kolom/n
> average.pk
Sangat.Tidak.Nyaman  Kurang.Nyaman      Biasa.Saja      Nyaman
      0.02059732      0.16477858      0.23377961      0.58084449

> fit

Principal inertias (eigenvalues):
      1      2
value  0.420978 0.012713
Percentage 97.07% 2.93%

ROWS:
      <25juta/th 25-50juta/th >50juta/th
Mass      0.152420   0.440783   0.406797
ChiDist   1.503706   0.189847   0.424086
Inertia    0.344642   0.015887   0.073162
Dim. 1     2.316299  -0.220849  -0.628579
Dim. 2     0.442234  -1.104499   1.031075

Columns:
      Sangat.Tidak.Nyaman Kurang.Nyaman Biasa.Saja    Nyaman
Mass      0.020597      0.164779   0.233780 0.580844
ChiDist   1.669536      1.300129   0.238905 0.381200
Inertia    0.057412      0.278531   0.013343 0.084405
Dim. 1     2.560964      2.003652  -0.202865 -0.577577
Dim. 2     1.439706      0.145368  -1.768270 0.619405

```

Keterangan:

- `library(ca)` package R untuk mengaktifkan fungsi analisis korespondensi, jika belum tersedia `install.packages(ca)`
- `fit=ca(data)` melakukan analisis korespondensi. [Diperoleh total keragaman yang dapat dijelaskan adalah sebesar 97.1%]
- `plot(fit)` grafik analisis korespondensi.[Interpretasi:
 1. Pekerja yang merasa Sangat tidak nyaman dan kurang nyaman dengan pekerjaannya cenderung yang memiliki pendapatan < 25juta.
 2. Pendapatan diantara 25 juta hingga 50 juta cenderung merasa biasa saja terhadap pekerjaannya.
 3. Pendapatan yang lebih dari 50 juta merasa Nyaman dengan pekerjaannya.]
- `chisq <- chisq.test(data)` melakukan inferensia terhadap adanya hubungan antara baris dan kolom. [Uji Chi-square diperlukan untuk memeriksa apakah baris dan kolom pada tabel kontingensi secara statistik memiliki hubungan yang erat.

H0 : variabel baris dan kolom dari tabel kontingensi adalah independen.

H1 : variabel baris dan kolom dependen.

karena diperoleh $p\text{-value} < 2.2e-16$ maka H0 ditolak artinya terdapat hubungan yang erat antara pendapatan dan kenyamanan pekerjaan]

- Menghitung profil baris dan kolom secara manual karena pada fungsi tidak tersedia untuk matriks profil baris dan kolom. Pada output fungsi yang tersedia hanya rata-rata matriks profil baris dan kolom (Mass).

Latihan:

Lakukan analisis korespondensi dari data `housetasks` dalam Package `factoextra` !

Responsi Analisis Korespondensi