

```
% ESAME DI SDE: novembre 2025
% Durata della prova: 75 minuti.
% Indicare nome GitHub ed eventuale presenza del nome GitHub nella pagina
% di errata corrige
% Indicare l'eventuale partecipazione al seminario
% Si assume che sia installata FSDA versione almeno 8.7.9.9
```

Caricare in memoria la table AgeCat tramite l'istruzione

```
load AgeCat
```

Questa table contiene la tabella di contingenza riferita al segmento di età ed alla categoria di acquisto prevalente.

```
'The contingency table summarises the joint distribution of customer groups
and product categories in a sample of 1,000 observations.
The counts indicate how frequently each age segment (Young (18-29),
Adult (30-64), Senior (65 or more))
purchased items from the five product categories considered
in a popular on line site'
```

Calcolare e mostrare nella Command Window la frequenza teorica in corrispondenza della modalità di riga Young a della modalità di colonna Electronics. (punti 4)

```
load AgeCat;
theoFreq = sum(AgeCat{:, 'Electronics'}) * sum(AgeCat{'Young',:}) /
sum(AgeCat{:,:},'all');
disp(['Frequenza teorica per Young e Electronics: ', num2str(theoFreq)]);
```

Frequenza teorica per Young e Electronics: 67.089

```
disp(['Frequenza effettiva per Young e Electronics: ',
num2str(AgeCat{'Young','Electronics'})]);
```

Frequenza effettiva per Young e Electronics: 76

```
% Commentare il valore ottenuto (punti 3)
%
% La frequenza empirica è maggiore della frequenza attesa. In altri termini,
% chi appartiene alla categoria Young tende ad acquistare prodotti di
% elettronica con frequenza leggermente maggiore di quella attesa sotto
% l'ipotesi di
% indipendenza tra i due fenomeni. Per capire se le differenze tra le
% frequenze teoriche e le frequenze effettive sono significative occorre
% fare riferimento al test Chi2
```

```
% Calcolare manualmente il p-value dell'indice Chi2 (punti 4)
out=corrNominal(AgeCat);
```

Chi2 index
5.1793

pvalue Chi2 index
0.7383

Phi index
0.0720

Cramer's V
0.0509

Test of H₀: independence between rows and columns

	Coeff	se	zscore	pval
CramerV	0.050888	0.025964	1.96	0.05
GKlambdayx	0.015171	0.014194	1.0688	0.28516
tauyx	0.001304	0.001126	1.1581	0.24684
Hyx	0.0016314	0.0014132	1.1543	0.24836

Indexes and 95% confidence limits

	Value	StandardError	ConflimL	ConflimU
CramerV	0.050888	0.025964	0	0.063412
GKlambdayx	0.015171	0.014194	0	0.042991
tauyx	0.001304	0.001126	0	0.0035109
Hyx	0.0016314	0.0014132	0	0.0044012

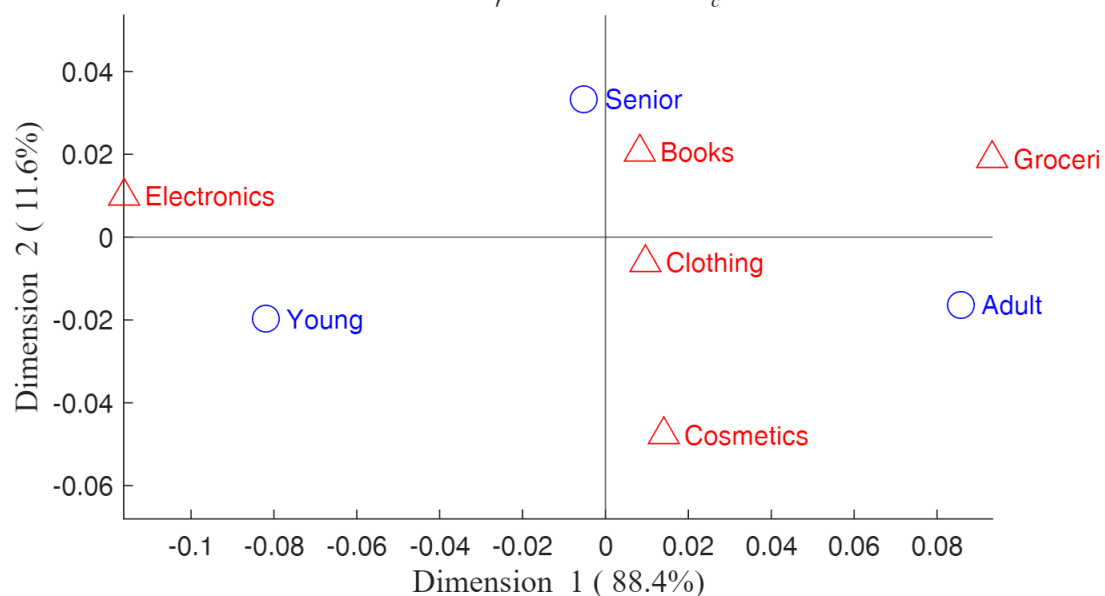
```
Chi2stat=out.Chi2;  
[I,J]=size(AgeCat);  
df = (I - 1) * (J - 1);  
pValue = 1 - chi2cdf(Chi2stat, df);  
disp(['p-value del test Chi2: ', num2str(pValue)]);
```

p-value del test Chi2: 0.73825

```
% Commentare il p-value ottenuto (punti 3)  
% L'associazione tra le due variabili (età e tipologia di acquisto  
prevalente) non è significativa  
% non è significativa  
%  
% Effettuare l'analisi delle corrispondenze e mostrare il relativo grafico  
% (punti 3)  
%  
outCorAna=CorAna(AgeCat);
```

French symmetrical model: rows and cols in principal coordinates.

Plot of $X = D_r^{-1/2}UT$ and $Y = D_c^{-1/2}VT$



Summary

	Singular_value	Inertia	Accounted_for	Cumulative
dim_1	0.067651	0.0045766	0.88364	0.88364
dim_2	0.024549	0.00060266	0.11636	1

ROW POINTS

Results for dimension: 1

	Scores	CntrbPnt2In	CntrbDim2In
Adult	0.085741	0.52687	0.965
Senior	-0.0051737	0.0020529	0.02353
Young	-0.081953	0.47108	0.94507

Results for dimension: 2

	Scores	CntrbPnt2In	CntrbDim2In
Adult	-0.01633	0.14513	0.035003
Senior	0.033329	0.64695	0.97647
Young	-0.019758	0.20792	0.054929

COLUMN POINTS

Results for dimension: 1

	Scores	CntrbPnt2In	CntrbDim2In
Electronics	-0.11624	0.61703	0.99276
Groceries	0.093393	0.36782	0.96054
Clothing	0.0096182	0.0041842	0.7119
Books	0.0081997	0.0030558	0.13724
Cosmetics	0.014059	0.0079032	0.079774

Results for dimension: 2

	Scores	CntrbPnt2In	CntrbDim2In
Electronics	0.009929	0.034189	0.0072435

Groceries	0.018928	0.11474	0.039456
Clothing	-0.0061187	0.012859	0.2881
Books	0.020559	0.14588	0.86276
Cosmetics	-0.047749	0.69233	0.92023

Overview ROW POINTS	Mass	Score_1	Score_2	Inertia	CntrbPnt2In_1	CntrbPnt2In_2	CntrbDim
Adult	0.328	0.085741	-0.01633	0.0024987	0.52687	0.14513	0.9
Senior	0.351	-0.0051737	0.033329	0.00039928	0.0020529	0.64695	0.023
Young	0.321	-0.081953	-0.019758	0.0022813	0.47108	0.20792	0.945

Overview COLUMN POINTS	Mass	Score_1	Score_2	Inertia	CntrbPnt2In_1	CntrbPnt2In_2	Cnt
Electronics	0.209	-0.11624	0.009929	0.0028445	0.61703	0.034189	
Groceries	0.193	0.093393	0.018928	0.0017525	0.36782	0.11474	
Clothing	0.207	0.0096182	-0.0061187	2.6899e-05	0.0041842	0.012859	
Books	0.208	0.0081997	0.020559	0.0001019	0.0030558	0.14588	
Cosmetics	0.183	0.014059	-0.047749	0.00045341	0.0079032	0.69233	0

Legend

Row scores in principal coordinates

Column scores in principal coordinates

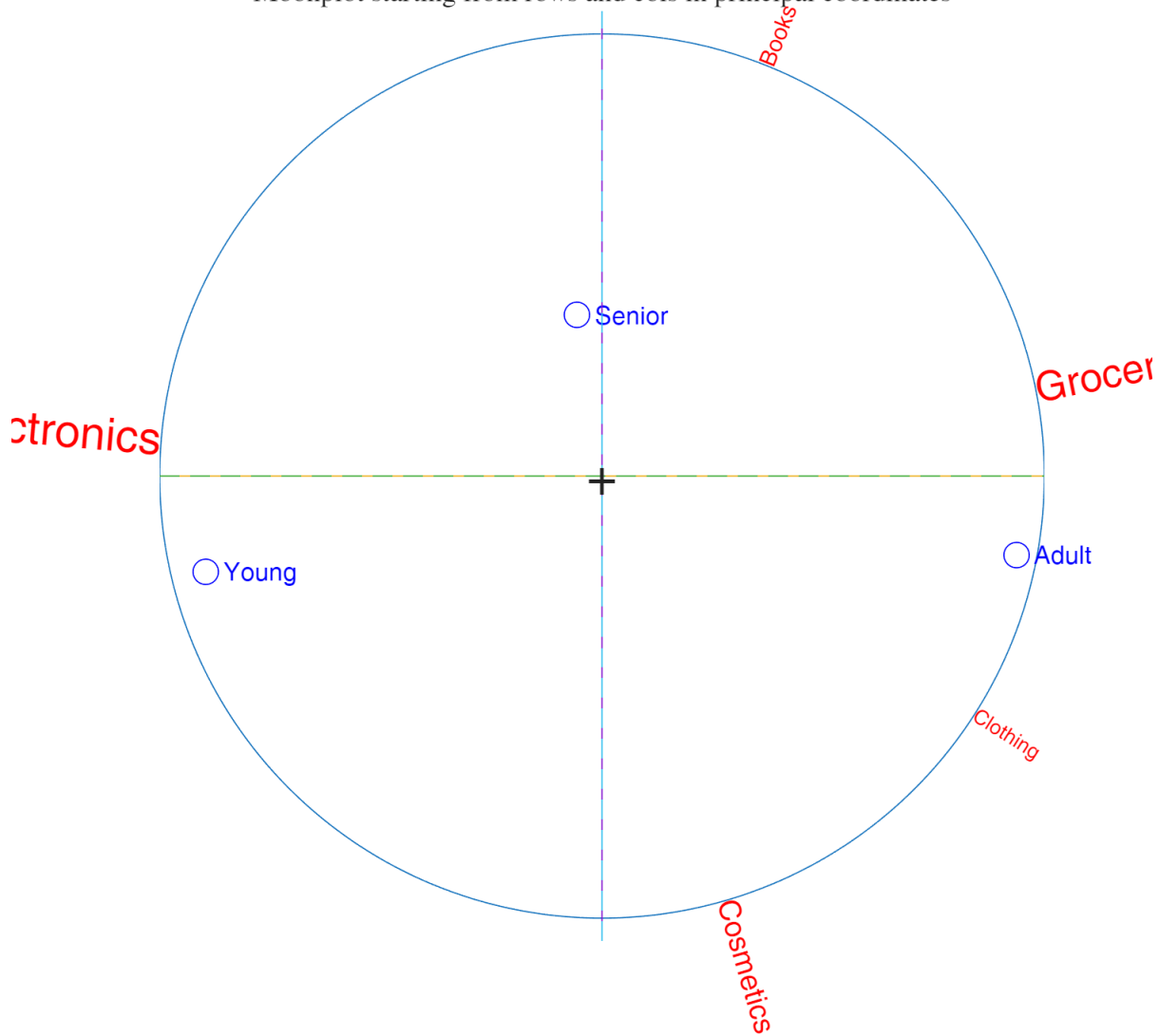
CntrbPnt2In = relative contribution of points to explain total Inertia of the latent dimension
The sum of the numbers in a column is equal to 1

CntrbDim2In = relative contribution of latent dimension to explain total Inertia of a point
 $CntrbDim2In_1 + CntrbDim2In_2 + \dots + CntrbDim2In_K = 1$

```
% Commentare il grafico (punti 5)
% I giovani acquistano prevalentemente prodotti di Electronics.
% I senior acquistano prevalentemente books.
% Il punto riga clothing (acquisto di generi di abbigliamento) è
% equidistante dalle 3 classi di età.
% Le prime due dimensioni latenti spiegano il 100 %
% dell'inerzia totale
```

```
% Creare il moonplot (punti 3)
moonplot(outCorAna)
```

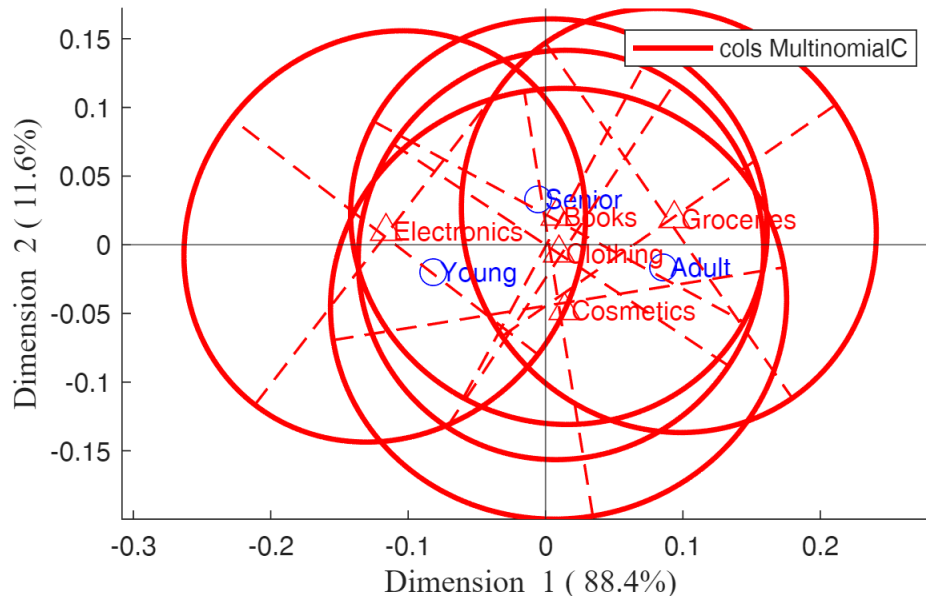
Moonplot starting from rows and cols in principal coordinates



```
% Ripetere il grafico aggiungendo l'ellisse di confidenza ai punti colonna.  
% (punti 5)  
% Vedi p.533  
confellipse=struct;  
confellipse.selCols=1:5;  
confellipse.selRows='';  
CorAnaplot(outCorAna,"confellipse",confellipse)
```

French symmetrical model: rows and cols in principal coordinates.

Plot of $X = D_r^{-1/2}U\Gamma$ and $Y = D_c^{-1/2}V\Gamma$



% Le ellissi di confidenza delle 5 tipologie di prodotti si sovrappongono.
La relazione non è
% significativa.