

**ANALISI DEI DATI PER IL MARKETING – GENNAIO 2009 II****ESERCIZIO 1**

Si è effettuata un'analisi di 100 modelli di *smartphone* di marche differenti. Per i primi 3 modelli della matrice dei dati, si conoscono i seguenti valori:

MARCA E MODELLO	PREZZO IN €	PESO IN g
BLACKBERRY 8100	175	90
NOKIA 6210	300	120
TOSHIBA G910	360	180
. . .	. . .	

Sapendo che, **con riferimento all'intero data set**, il vettore delle medie è uguale a [390; 140] ed il vettore delle deviazioni standard è uguale a [150; 40],

- si calcolino le prime 3 righe della matrice delle distanze euclidee standardizzate e le prime tre righe della matrice delle distanze standardizzate della città a blocchi;
- si dica se tra queste parti delle matrici dei due tipi di distanza esiste una relazione monotona.

**ESERCIZIO 2**

Si è applicata l'analisi delle componenti principali ai 100 modelli di *smartphone* dell'esercizio 1, considerando ora 3 variabili ed estraendo una sola componente. Conoscendo la seguente tabella delle comunalità:

Comunalità

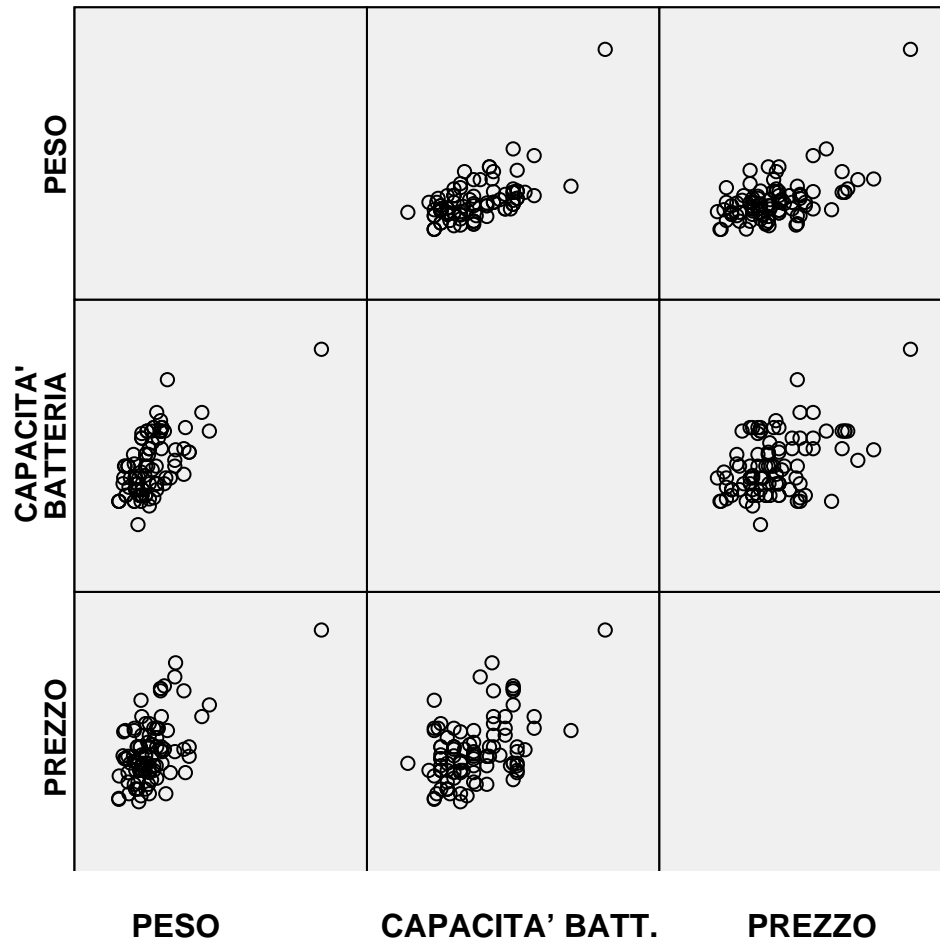
	Iniziale	Estrazione
PESO	1,000	,787
CAPACITA' BATTERIA	1,000	,771
PREZZO	1,000	,697

- si valuti la bontà del modello;
- si calcolino i coefficienti di correlazione (tutti positivi) tra la componente estratta e ciascuna delle tre variabili;
- si scriva l'**espressione** degli *scores* della prima componente, adottando il criterio di SPSS e calcolando i valori numerici dei relativi coefficienti.

### ESERCIZIO 3

Si è costruita la *scatterplot matrix* dei dati utilizzati nell'esercizio 2.

- Si descrivano tutte le informazioni che essa fornisce.
- Si illustri la strategia che deve adottare il ricercatore per migliorare la visualizzazione dei dati e si dica quali sono i risultati attesi.



## TRACCIA DI RISOLUZIONE

### ESERCIZIO I

Gli scostamenti standardizzati risultano i seguenti:

MARCA	Z(PREZZO)	Z(PESO)
Blackberry	-1.43	-1.25
Nokia	-0.60	-0.50
Toshiba	-0.20	1.00
. . .		

Si noti che tali scostamenti standardizzati non presentano media nulla e varianza unitaria, poiché la standardizzazione è stata effettuata con riguardo all'intero data set, e non alle prime tre righe della matrice dei dati.

Le matrici delle distanze sono:

	Distanza euclidea standardizzata		
	1:Blackberry	2:Nokia	3:Toshiba
1:Blackberry	,000	1,119	2,564
2:Nokia	1,119	,000	1,552
3:Toshiba	2,564	1,552	,000

	Distanza City Block standardizzate		
	1:Blackberry	2:Nokia	3:Toshiba
1:Blackberry	,000	1,580	3,480
2:Nokia	1,580	,000	1,900
3:Toshiba	3,480	1,900	,000

Considerando le distanze nel triangolo superiore di ciascuna matrice, si vede immediatamente che tra esse vi è una perfetta relazione monotona, poiché i posti d'ordine sono i medesimi.

### ESERCIZIO II

La quota di varianza totale spiegata è uguale al 75,17% ed è inferiore alla soglia di accettabilità, pari a  $0.95^3 \times 100 = 85,7\%$ , per cui l'analisi non può considerarsi soddisfacente.

#### Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% di varianza	% cumulata	Totale	% di varianza	% cumulata
1	2,255	75,170	75,170	2,255	75,170	75,170
2	,440	14,651	89,820			
3	,305	10,180	100,000			

### Matrice di componenti

	Componente
	1
PESO	,887
CAPACITA' BATTERIA	,878
PREZZO	,835

### Matrice dei coefficienti di punteggio dei componenti

	Componente
	1
PESO	,393
CAPACITA' BATTERIA	,389
PREZZO	,370

L'espressione degli *scores* è la seguente:

$$Y_{1i} = 0.393 z(\text{peso})_i + 0.389 z(\text{cap. batt.})_i + 0.370 z(\text{prezzo})_i \quad (i = 1, 2, \dots, n)$$

### ESERCIZIO III

In tutti i pannelli si nota chiaramente un valore anomalo, eccezionalmente grande.

La relazione tra capacità della batteria e peso e tra prezzo e peso è di tipo diretto e abbastanza forte; risulta invece più debole la relazione tra prezzo e capacità della batteria (nuvola dei punti più dispersa).

Il ricercatore può eliminare il punto anomalo dal grafico, cancellando la corrispondente unità nella matrice dei dati. Rifacendo il grafico, si ottiene quindi un effetto "zoom" sulla nuvola principale dei punti, che ne evidenzia meglio le caratteristiche.