

# ANALISI DELLE COMPONENTI PRINCIPALI sul dataset Open Manager

Gianluca Cerea mat.919779  
Federico Rausa mat. 919795  
Pablo Torres mat. \*\*\*\*\*

# Descrizione del dataset Open Manager

FONTE : *OPEN MANAGEMENT IN OPEN ORGANIZATION / Paolo Bruttini , 2022 Jun 20. 34. ciclo, Anno Accademico 2020/2021.*

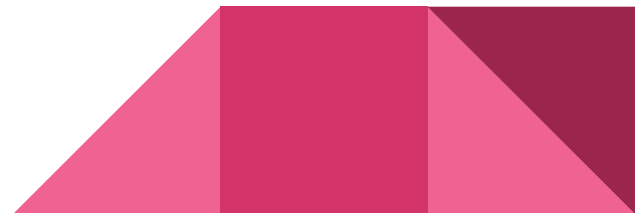
Dataset costituito da:

- 50 unità statistiche, persone che compilano il questionario
- ogni domanda presenta 4 possibili opzioni :
  - 1) fortemente in disaccordo
  - 2) parzialmente in disaccordo
  - 3) parzialmente in accordo
  - 4) fortemente in accordo
- 30 domande, di cui 4 “reverse”



# Fasi del lavoro

- 1) Pulizia del dataset
- 2) standardizzazione del dataset con Z-score
- 3) Verifica della presenza di correlazioni tra le variabili originarie
- 4) analisi delle componenti principali sul dataset standardizzato
- 5) selezione del numero di componenti con lo scree plot
- 6) emissione del nuovo dataset
- 7) interpretazione variabili del nuovo dataset, in base alle correlazioni con le originali
- 8) analisi distribuzioni delle prime componenti



# Workflow su KNIME, R e Excel

### Correlazioni fra le variabili originarie



R Plot

### Correlazioni tra componenti e variabili



R Snippet

top k correlate



R Snippet

Correlazioni Variabili-CP

### Rielaborazioni finali



Excel Writer

Esporta gruppi di variabili



Excel Writer

Esporta top k correlate



Excel Writer

Esporta Varianze



Excel Writer


Esporta dataset



Excel Writer


esporta statistiche

### Importa dati



Excel Reader

importa dati grezzi




Column Filter

Rimuovi reverse e codici



R Snippet


riduci nomi variabili



Normalizer


Z score

### screeplot e varianza% spiegata




R Plot

Screepilot




R Plot

Metodo del Gomito




Line Plot

Varianza-Cumulata




PCA Compute

Decomposizione PCA




Math Formula

prop var




Moving Average

cum var




Column Filter




Column Renamer

### Analisi distribuzione dei fattori




PCA

CP4 per individuo




Column Renamer


assegnazione nomi



Statistics




Column Filter



R Plot

istogrammi - distribuzioni



Statistics View

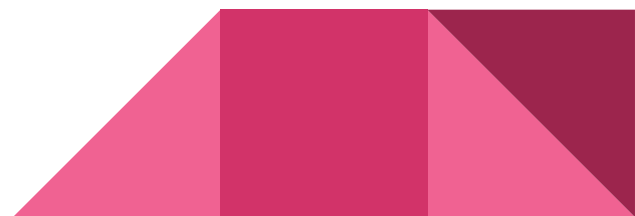
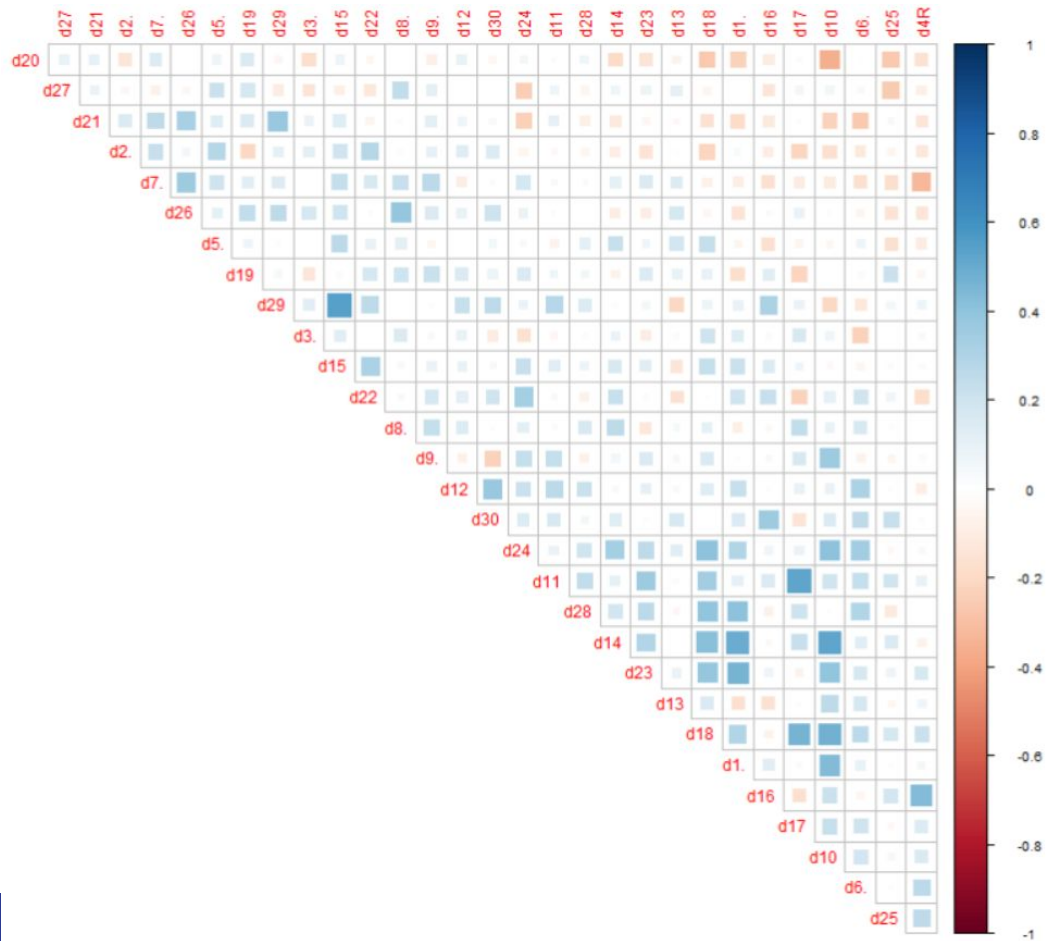
statistiche descrittive

# Pulizia del dataset

- rimuovere le 4 domande che non sono “reverse”
- rimuovere la colonna “codici”
- verificare la presenza di valori nulli all’interno del dataset, qualora siano presenti, vanno sostituiti con NA, ovvero chi compila il questionario non ha dato una risposta alla domanda

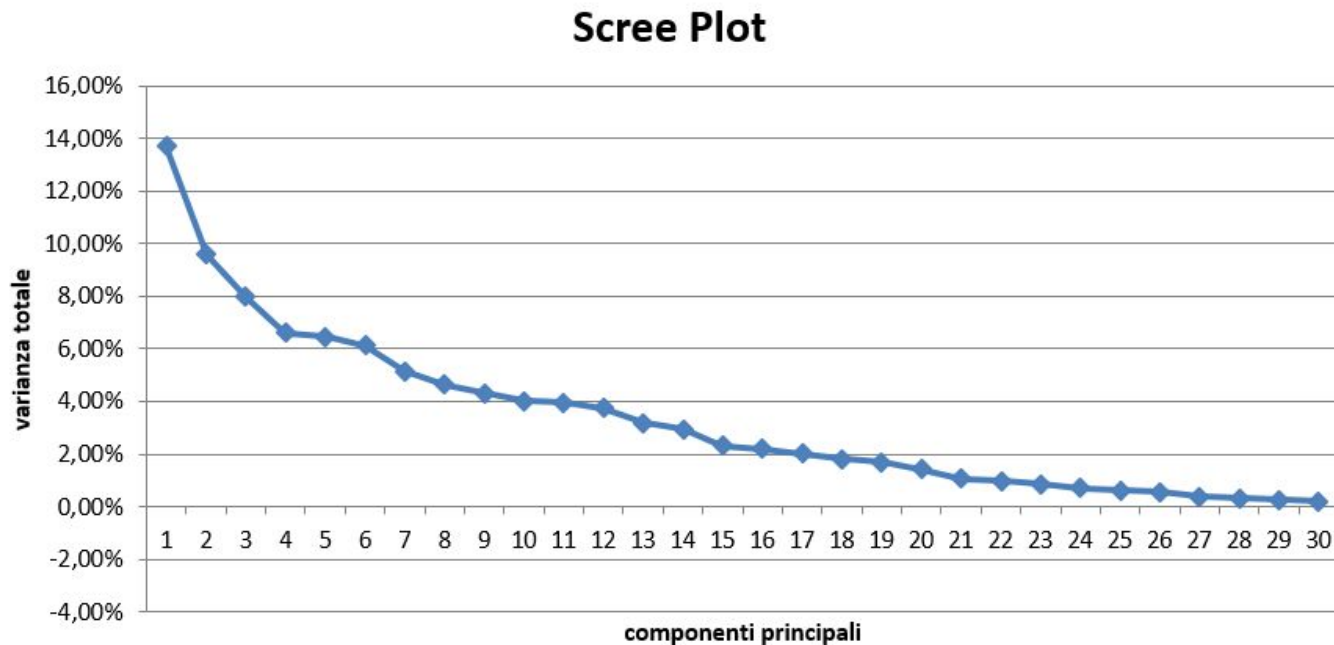


# Correlazioni fra le variabili originarie

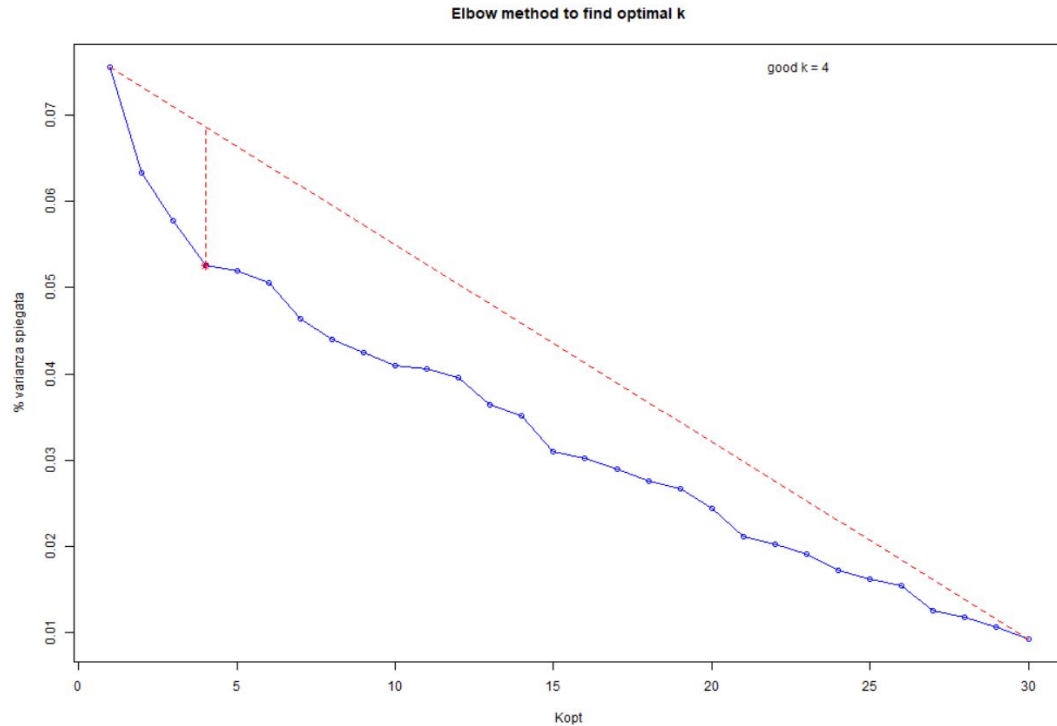


# Selezione n° componenti: scree plot

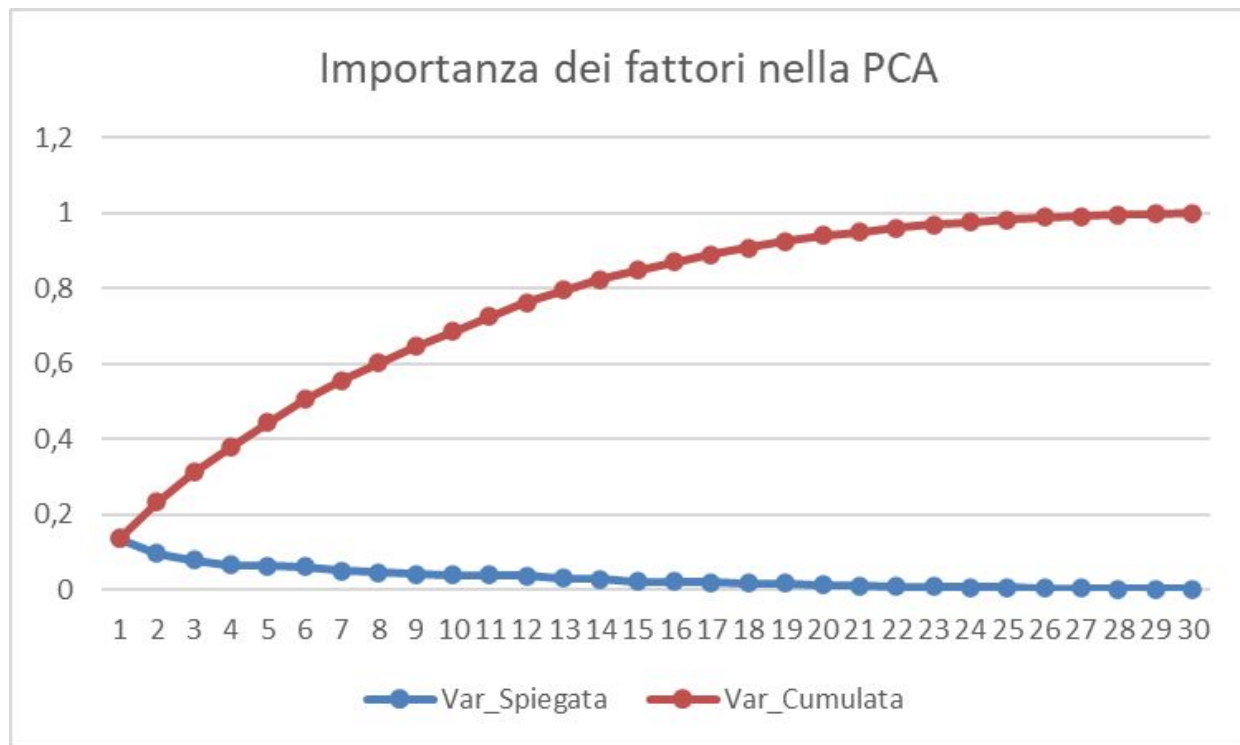
- eseguiamo lo scree plot per vedere quali componenti principali tenere



# Selezione n° componenti: metodo del gomito sullo scree plot



# Selezione n° componenti: criterio della varianza% minima spiegata



# Denominazione componenti principali

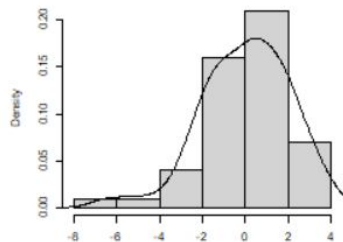
Utilizzando la matrice delle correlazioni di Pearson, interpretiamo le prime 8 domande che contribuiscono a spiegare ciascuna componente

Fiducia	Orgoglio	Autoritarietà	Pigrizia relazionale	Umiltà	Egoismo	Empatia	Sottomissione	Collaborazione	Spirito di squadra
PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
+d18	+d24	-d30	-d8.	-d11	+d29	+d15	+d30	+d27	+d27
+d25	-d25	-d29	+d20	-d17	-d19	-d8.	+d12	+d9.	+d5.
+d10	+d7.	+d18	-d26	-d20	-d25	+d29	-d9.	+d11	+d15
+d11	+d26	-d26	-d9.	+d22	-d24	-d4R	-d4R	+d23	+d14
+d24	+d8.	+d10	+d30	+d1.	+d16	-d16	-d16	+d10	+d8.
+d16	+d28	-d21	+d6.	+d2.	+d17	-d27	-d15	+d19	+d1.
+d30	-d4R	+d17	-d7.	+d10	+d4R	-d6.	-d24	-d4R	-d11
+d14	+d12	-d16	-d21	+d24	-d20	-d10	+d27	-d26	+d28

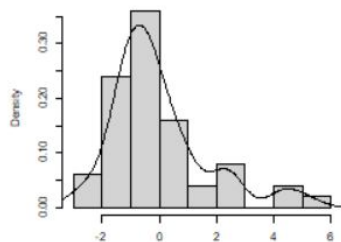


# Distribuzioni del dataset ridotto

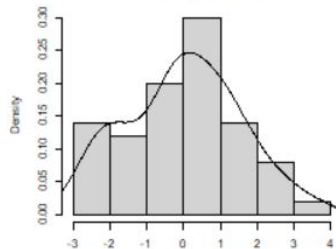
PC1 - Fiducia



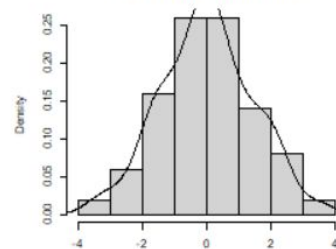
PC2 - Orgoglio



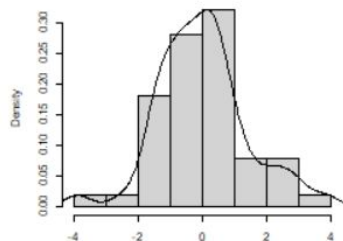
PC3 - Autoritaria



PC4 - Pigrizia relazionale



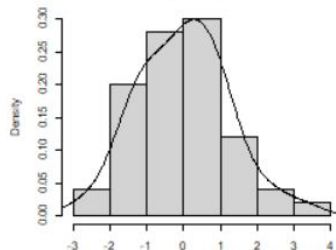
PC5 - Umilta



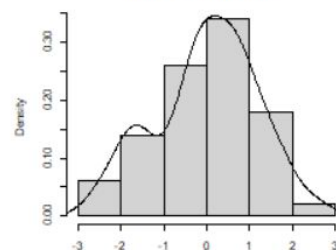
PC6 - Egoismo



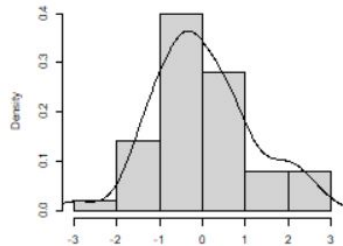
PC7 - Empatia



PC8 - Sottomissione



PC9 - Collaborazione



PC10 - Spirito di squadra

