**Piano di analisi**

# Step0: domande di ricerca e ipotesi da testare

**RQ: Quale presentazione dei cluster è migliore tra albero, parallel e dendrogramma?**

**HP1: Albero più usabile rispetto agli altri due metodi (SCS)**

**HP2: albero cognitivamente meno faticoso (NASA TLX)**

**HP3a: Albero minor numero di errori (performance)**

**HP3b: Albero decisioni più veloci (performance)**

*Su questa hp3 possiamo ragionare creando un indice di performance tra tempo e n di errori*

# Step1: Pulizia del dataset e preparazione dei dati

Abbiamo le seguenti informazioni che vanno codificate:

* NASATLX: Media ponderata per ogni metodo di visualizzazione. C’è una parte di pesi per la condizione sperimentale Easy e dei pesi per la condizione Complex. Quei pesi devono essere calcolati e ricodificati in nuove colonne per ogni variabile del NASA TLX
* SCS: media tra gli item per ogni metodo di visualizzazione

# Step2: Descrittive e outliers

1. Trasformare variabili categoriche in fattori (es. età, school major, ecc)
2. Creare colonne per la performance nei singoli task con misura binaria di successo/fallimento
3. Calcolare statistiche (media, mediana, sd)
4. Visualizzazione dei dati
   1. Visualizzazione della performance tra i le due difficolta e tra i tre task
   2. Visualizzazione dei risultati dei test (NASATLX e SCS) tra le due difficoltà e tra i tre task
   3. Capire la distribuzione dei dati
5. Pulizia degli outliers (es chi ci ha messo troppo poco o troppo tempo)

# Step3: Analisi

ANALISI PARAMETRICHE:

*Queste sono da provare per ogni ipotesi. Quindi:*

*HP1: ANOVA repeated measures ad una via su tre medie diverse tra le condizioni facile vs difficile(SCS score) (DA CONTROLLARE)*

*HP2: ANOVA repeated measures ad una via su tre medie diverse tra le condizioni facile vs difficile(NASATLX score) (DA CONTROLLARE)*

*HP3: ANOVA repeated measures ad una via su tre medie diverse tra le condizioni facile vs difficile (performance) (DA CONTROLLARE)*

Test delle assunzioni dell’ANOVA

1. Omogeneità delle varianze (boxplot, Bartlett’s test)
2. Normalità dei dati: Q-Q plot e Shapiro Wilk test per la normalità
3. Indipendenza delle osservazioni: garantita dal disegno sperimentale randomizzato

Se le nostre variabili dipendenti (SCS score, NASA TLX score, Performance,…) allora

ANALISI NON PARAMETRICHE

* Wilcoxon Test
* Kruscal Wallis

**Correzione di bonferroni!!! Da controllare**

# Step4: calcolare effect size per ogni ipotesi

* Cohen’s d
* Cohen’s f

# Step5: Tabelle e valori da riportare

Vedere APA style