**DESCRIZIONE DELLA RICERCA**

La seguente ricerca è stata sviluppata per capire eventuali differenze di pressione sanguigna all’interno di diverse città nella regione Lombardia. All’interno del database sono presenti diverse informazioni di carattere demografico riguardanti i partecipanti della ricerca stessa. Il seguente file di dati è stato costruito appositamente per le esercitazioni, di conseguenza i dati non necessariamente corrispondono alla realtà dei fatti, essendo costruiti in maniera causale. Vi sono anche delle variabile che non utilizzerete in questa esercitazione.

**DESCRIZIONE DEI DATI:**

**id** = codice univoco per partecipante  
**genere** = genere del partecipante (1 maschio; 2 femmina)  
**raccolta** = anno di raccolta del questionario sottoposto ai partecipanti  
**nascita** = anno di nascita del partecipante  
**città** = città di provenienza del partecipante  
**fuma** = indica se il paziente è un fumatore o meno (1 fumatore; 2 non fumatore)  
**argom** = indice di argomentazione verbale  
**attenz** = indice di attenzione  
**memo** = indice di memoria

**DOMANDE:**

1. All’interno del file non esiste la variabile età, ma può essere creata a partire da altre informazioni, creala.
2. Crea un indice aggregato medio degli item che vanno da item1 a item4, in modo da creare l’indice per tutti coloro che hanno risposto ad almeno 3 dei 4 item, denominandolo “indice\_medio”.
3. Esegui le statistiche descrittive della variabile indice\_medio, ma delle sole donne.
4. Esiste una certa associazione statistica tra il genere dei partecipanti e il fatto di fumare o meno?
5. Controlla se possiamo affermare che vi sia una relazione tra l’indice di attenzione e quello di argomentazione verbale.