**Coursework (Econometrics and Statistics)**

1. The dataset *customers.red* contiene le seguenti variabili

* Hh\_key *→* identificativo della persona;
* monthnumb *→* tempo (1=Gennaio,…, 12 = Dicembre, 13 = Gennaio, etc.)
* pricepromo *→* numero di promozioni/sconti sul prezzo;
* retailpromo*→* promozioni varie del negozio;
* catalogpromo*→* numero di cataloghi inviati;
* item *→* numero di oggetti acquistati:
* gender *→* sesso
* married *→* 1 se sposato
* income *→* classe di reddito

1. Fornire le statistiche descrittive di tutte le variabili quantitative presenti nel dataset per genere e stato coniugale
2. Mostrare graficamente l’andamento temporale del numero degli oggetti acquistati
3. Identificare potenziali relazioni tra il numero degli oggetti acquistati e pricepromo, retailpromo, catalogpromo.
4. Dopo aver reso binaria la variabile item, creando una nuova variabile che assuma valore 1 se compro e 0 altrimenti, identificare potenziali relazioni tra il comprare/non comprare e pricepromo, retailpromo, catalogpromo.
5. Verificare attraverso simulazione che lo stimatore dei minimi quadrati dei parametri della retta di regressione è uno stimatore corretto. Sempre tramite simulazione, confrontare la varianza teorica dello stimatore con quella ottenuta tramite simulazione. Fornire anche una stima della varianza dei residui.
6. Scrivere una funzione per generare i dati dal seguente modello di regressione
   * yi = 0.5 + 2.5\*xi + errore, con X˜N(0;0.5) ed errore ˜N(0;0.25)
7. Scrivere una funzione che abbia come argomenti la numerosità campionaria, i parametri della retta di regressione e il numero di campioni da simulare (fissato a 1000) e che ritorni come output la distorsione e l’errore quadratico medio delle stime.