**STUDIO DI FATTIBILITÀ**

**Apertura di un’attività di ristorazione nella provincia di Arezzo**

Nel file Excel *Studio fattibilità* è riportata una matrice di dati territoriali. Si tratta di un set di variabili rilevate sui 37 comuni della provincia di Arezzo. In particolare, sono rilevate le seguenti variabili: *abitanti*, *competitors*, *laureati*, *diplomati*, *totale laureati e diplomati, reddito comune, tasso di scolarizzazione, reddito pro-capite, arrivi italiani, arrivi stranieri, totale turisti, presenze italiani, presenze straniere, comuni limitrofi ad Arezzo*.

Al fine individuare tra i 37 comuni della provincia di Arezzo quello più adatto ad ospitare una nuova attività di ristorazione, occorre:

* **Analisi dell’offerta**: Individuare (anche da fonti statistiche secondarie oppure tramite web, pagine gialle, pagine bianche, etc.) i competitors (diretti e indiretti) di interesse. Tali informazioni sono state già rilevate e sintetizzate nella variabile **COMPETITORS.**
* **Analisi di domanda**: Individuare (tramite web, es.: siti della Regione, Provincia, Comune, Istat, …) dove si contestualizza il *tipo* e la *quantità* di clienti su cui si può contare.
* **Interazione tra le analisi di offerta e domanda** al fine di determinare il *luogo* dove vi sono *meno competitors* e *più clienti*(quindi, in teoria, il luogo più adatto per intraprendere l’attività). Se si usa la variabile abitanti come proxy dei clienti potenziali, tale interazione può essere *stimata*, ad esempio, tramite una *funzione* del tipo:

***n*° competitors = *f* (abitanti)** (1)

***n*° competitors = *f* (Tot\_turisti)** (2)

1. Sulla base delle informazioni riportate nel file Excel, fornire una descrizione della variabile ‘Competitors’ attraverso opportune misure di sintesi, di variabilità e di forma. Costruire e commentare il boxplot.
2. Rappresentare graficamente la distribuzione bivariata dei competitors (y) vs. abitanti (x)
3. Rappresentare graficamente la distribuzione bivariata dei competitors (y) vs. Tot\_turisti (x)
4. Calcolare la matrice di correlazione lineare tra ‘Competitors’ e le due variabili per capire se possano influenzare la variabile ‘Competitors’
5. Proporre i test di correlazione di cui al punto 4)
6. Stimare il modello di regressione lineare semplice di cui alla funzione (1) e interpretarne i risultati sia a livello descrittivo che inferenziale
7. Stimare il modello di regressione lineare semplice di cui alla funzione (2) e interpretarne i risultati sia a livello descrittivo che inferenziale
8. Rappresentare graficamente la retta di regressione (tracciare la retta di regressione nello scatter plot di cui al punto 2)
9. Cosa rappresenta la retta di regressione nel caso specifico? Dove si collocano i comuni in cui vi è un eccesso di offerta? E quelli in cui vi è un eccesso di domanda?
10. Stimare un modello di regressione lineare multiplo considerando come variabile dipendente la variabile ‘Competitors’ e come variabili esplicative “abitanti” e Tot\_turisti. Commentarne i principali risultati.
11. Stimare il modello di regressione lineare multiplo che **meglio** spieghi il comportamento della variabile ‘Competitors’, commentarne i principali risultati ed effettuare il confronto col modello precedente.
12. Verificare le ipotesi alla base del modello di regressione (normalità, omoschedasticità, indipendenza dei residui)
13. Calcolare e commentare *Tolerance* e VIF ed illustrare la relazione tra i due
14. Calcolare AIC e BIC del modello di regressione di cui al punto 11)