

Indicazioni generali per i progetti di
Statistica Applicata (canali A-H e I-Z)

- Il progetto di Statistica Applicata verte sull'Analisi di Regressione di un semplice dataset per ogni gruppo di studenti (max quattro studenti).
- Gli studenti dovranno applicare le conoscenze acquisite durante il corso per l'analisi dei dati e presentare i risultati, quali, ad esempio, la statistica descrittiva per presentare i dati e prendere contezza di come sono distribuiti, l'analisi di correlazione delle variabili, i test di ipotesi e la stima degli intervalli di confidenza per valutare la significatività statistica dei parametri di un modello, la diagnostica dei modelli in esame, ecc...
- È suggerito l'utilizzo di grafici per la presentazione dei dati o per la valutazione della bontà di adattamento dei modelli ai dati in esame, quando la dimensione dei dati lo permette.
- Uno schema di riferimento per il progetto di Statistica Applicata è il seguente:
 1. Analisi preliminare dei dati, in forma grafica o tabellare, tramite gli strumenti di statistica descrittiva, quali scatter plot, istogrammi, box plot, ecc., e analisi di correlazione delle variabili presenti nel dataset, utile alla definizione dei modelli di regressione.
 2. Valutazione della relazione tra la variabile dipendente e le singole variabili indipendenti tramite scatter plot o regressione polinomiale.
 3. Definizione del modello statistico dei dati tramite regressione lineare multipla o di diversi modelli da confrontare.
 4. Stima ai minimi quadrati dei parametri del modello o dei modelli. Determinazione degli intervalli di confidenza dei parametri.
 5. Calcolo del coefficiente di determinazione e grafici diagnostici dei modelli statistici in esame.
 6. Dopo aver stabilito il criterio di scelta del modello (ad es. basato sui p-value dei test t , su AIC, su BIC, ...), si scelga il modello che meglio si adatta ai dati forniti, ad esempio tramite regressione *stepwise*. La scelta deve essere motivata da argomenti teorici e dalle evidenze empiriche emerse dall'analisi dei dati stessi.
 7. Presentazione ed analisi dei risultati.

- I dataset, uno per ogni gruppo, sono forniti dai docenti del corso. La preparazione delle routine in R per l'analisi dei dati verrà avviata durante le ore di Laboratorio in seno al corso di Statistica Applicata.
- *La consegna del progetto (PDF e codice) deve avvenire, via e-mail agli indirizzi dei docenti (fpostiglione@unisa.it, mdimauro@unisa.it), entro 3 giorni prima della data di esame, riportata nel calendario ufficiale, a cui il gruppo intende partecipare. Qualora gli studenti lo ritenessero utile, è possibile presentare il progetto in sede di esame utilizzando una presentazione (ad es., PowerPoint).*
- Per le comunicazioni:
 1. Ricordarsi di scrivere sempre a entrambi i docenti.
 2. È sconsigliabile cambiare la composizione dei gruppi in corso d'opera.