

Progetto finale – analisi dati e stima modello logistico

Indicazioni per report progetto finale (in gruppo o individuale)

1. Report: max 15 pag. inclusi grafici, tabelle e codice R
2. Inviare il report entro (max) la data ultima di iscrizione all'esame
3. Presentazione e discussione orale con slides (max 12)

Basi dati (file Descrizione basi dati 25_26 in moodle)

In gruppo (micro-dati indagini Banca d'Italia e Istat ad uso pubblico)

1. **Banca d'Italia:** Alfabetizzazione finanziaria dei giovani in Italia (2023)
2. **Banca d'Italia:** Alfabetizzazione finanziaria degli adulti in Italia (2023)
3. **Istat AVQ_2023_IT** Aspetti della vita quotidiana 2023
4. **Istat INTEG2GEN_2023_IT** Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri
5. **INCSTRAD_2023_IT** Rilevazione degli incidenti stradali con lesioni a persone 2023

Individuale: contattare la docente

Modello logistico

Data la variabile risposta Y indicata (file Descrizione basi dati 25_26):

- se **rilevata in modo diretto**: verificarne la natura e le modalità (se non dicotomica renderla tale)
- se **NON rilevata in modo diretto**: costruirla mediante opportune variabili rilevate
- individuare insieme di predittori (anche da costruire) da inserire nella costruzione di un modello logistico per Y

N.B.: Motivando la scelta, l'analisi può essere condotta (a seconda dell'indagine) su:

- dati non pesati o dati pesati o, anche, valutare i risultati in entrambi i casi
- intera base dati o sottogruppi di unità selezionate rispetto a qualche caratteristica (socio-demografica e/o territoriale, ...)

Report

4. Stesura report:

- a. descrizione indagine (disegno, modalità rilevazione, strumento di misura/questionario, campione)
- b. obiettivo vostra analisi e scelta variabili (descrizione dei dati)
- c. descrizione dettagliata della fase di *data preparation* (controllo e pulizia dei dati, ricodifica/trasformazione/creazione variabili, ...)

Esempi:

<https://www.kaggle.com/code/chebotinaa/cleaning-data-with-r>

<https://www.kaggle.com/datasets/vikramamin/students-performance>

- a. costruzione e stima modello logit
- b. interpretazione risultati

Gruppi TIS 2025/26

Gruppo 1

Nikolaj CERNIC
Francesco DE MARCO
Filippo GIACOMELLO
Simone SCUSSOLIN

Gruppo 2

Greta FAVARO
Giulia GOTTARDI
Cesare OLIVATI
Simone RONCATO
Francesca ZANE

Gruppo 3

Filippo CANDUSSI
Alessandro ERMACORA
Lorenzo PASINETTI
Tommaso VOLTAN

Gruppo 4

Simone BAVDAZ
Simone DUIZ
Cabrel KANA KENGNE
Giulio PETRONIO

Gruppo 5

Cesare CAIANI
Giada FUMOLO
Francesca PERIN
Alessia SILVESTRIN

> sample(5)
[1] 3 4 5 2 1

Assegnazione base dati

Gruppo 3 = AVQ_2023_IT
Gruppo 4 = INTEG2GEN_2023_IT
Gruppo 5 = INCSTRAD_2023_IT
Gruppo 2 = BI ADULTI
Gruppo 1 = BI GIOVANI