

8 settembre 2022



Analisi dei dati di AlmaLaurea

Asaf Cohen, matr. 975599

Progetto per il corso di Visualizzazione Scientifica A.A. 2021-22

AlmaLaurea?

AlmaLaurea è un consorzio fondato nel 1994 da un gruppo di ricercatori universitari a cui aderiscono 75 atenei italiani e il Ministero dell'Istruzione, con l'obiettivo di realizzare **indagini statistiche** inerenti al mondo universitario italiano.

Il dataset

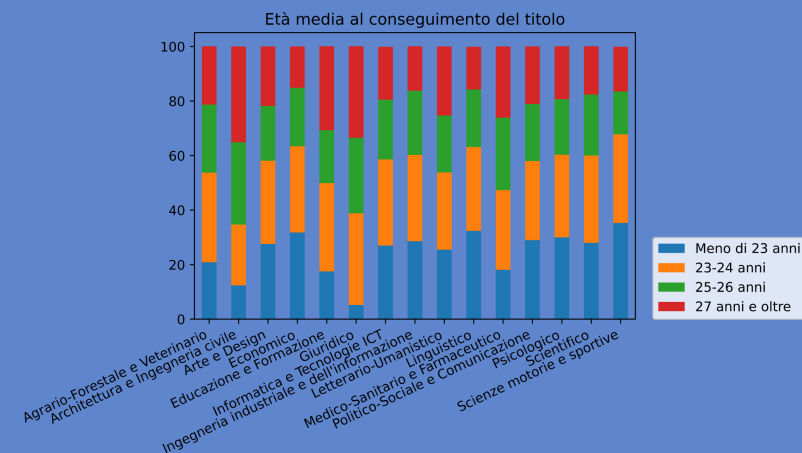
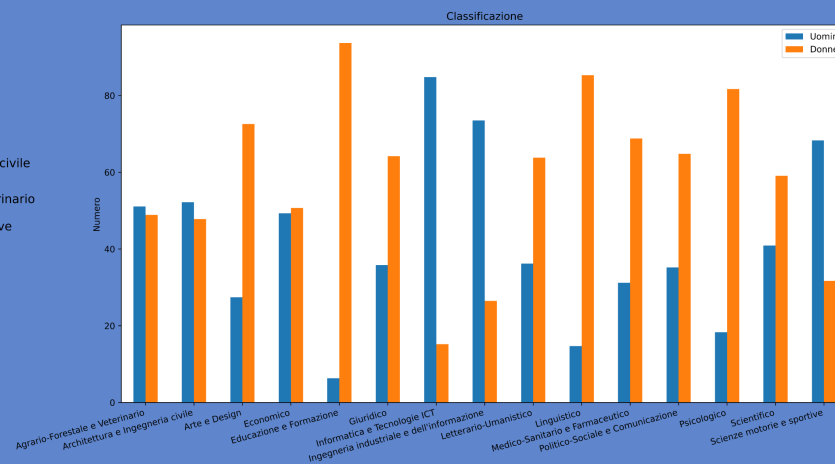
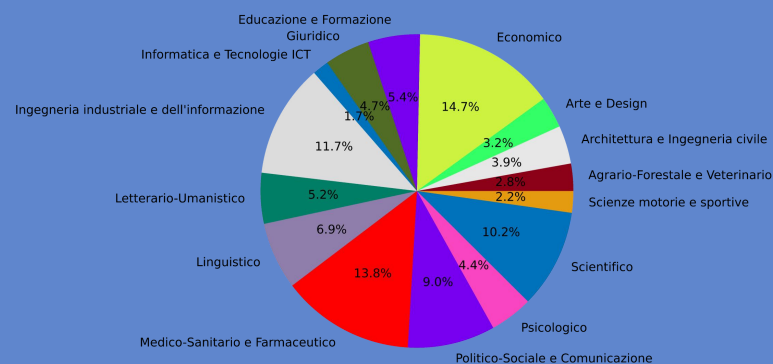
I dati sono stati prelevati dai seguenti link: [link1](#), [link2](#) di AlmaLaurea (relativi alla prima e alla seconda parte della presentazione), sono stati analizzati un totale di 30 CSV (due per gruppo disciplinare) in modo da poter avere termini di confronto

Tutto il [codice](#) utilizzato per generare i grafici di questa presentazione è liberamente accessibile nel mio account GitHub

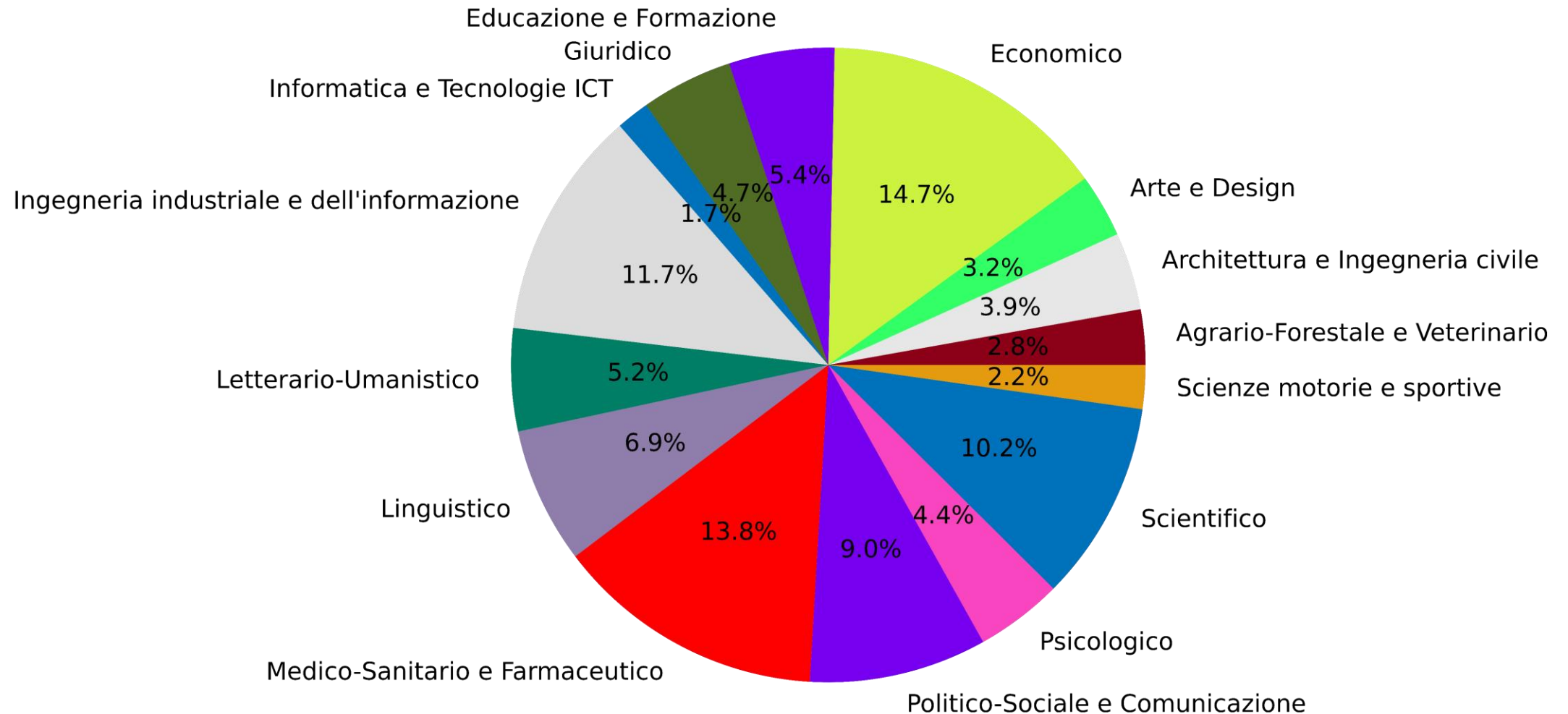
```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

# leggo i CSV di ogni gruppo disciplinare
agrario = pd.read_csv('./dati/Agrario-Forestale e Veterinario.csv', sep=";", decimal=".", na_values=' ')
architettura = pd.read_csv('./dati/Architettura e Ingegneria civile.csv', sep=";", decimal=".", na_values=' ')
arte = pd.read_csv('./dati/Arte e Design.csv', sep=";", decimal=".", na_values=' ')
economia = pd.read_csv('./dati/Economico.csv', sep=";", decimal=".", na_values=' ')
formazione = pd.read_csv('./dati/Educazione e Formazione.csv', sep=";", decimal=".", na_values=' ')
giurisprudenza = pd.read_csv('./dati/Giuridico.csv', sep=";", decimal=".", na_values=' ')
informatica = pd.read_csv('./dati/Informatica e Tecnologie ICT.csv', sep=";", decimal=".", na_values=' ')
ingegneria = pd.read_csv('./dati/Ingegneria industriale e dell'informazione.csv', sep=";", decimal=".", na_values=' ')
lettere = pd.read_csv('./dati/Letterario-Umanistico.csv', sep=";", decimal=".", na_values=' ')
lingue = pd.read_csv('./dati/Linguistico.csv', sep=";", decimal=".", na_values=' ')
medicina = pd.read_csv('./dati/Medico-Sanitario e Farmaceutico.csv', sep=";", decimal=".", na_values=' ')
scienze_politiche = pd.read_csv('./dati/Politico-Sociale e Comunicazione.csv', sep=";", decimal=".", na_values=' ')
psicologia = pd.read_csv('./dati/Psicologico.csv', sep=";", decimal=".", na_values=' ')
scientifico = pd.read_csv('./dati/Scientifico.csv', sep=";", decimal=".", na_values=' ')
scienze_motorie = pd.read_csv('./dati/Scienze motorie e sportive.csv', sep=";", decimal=".", na_values=' ')
```

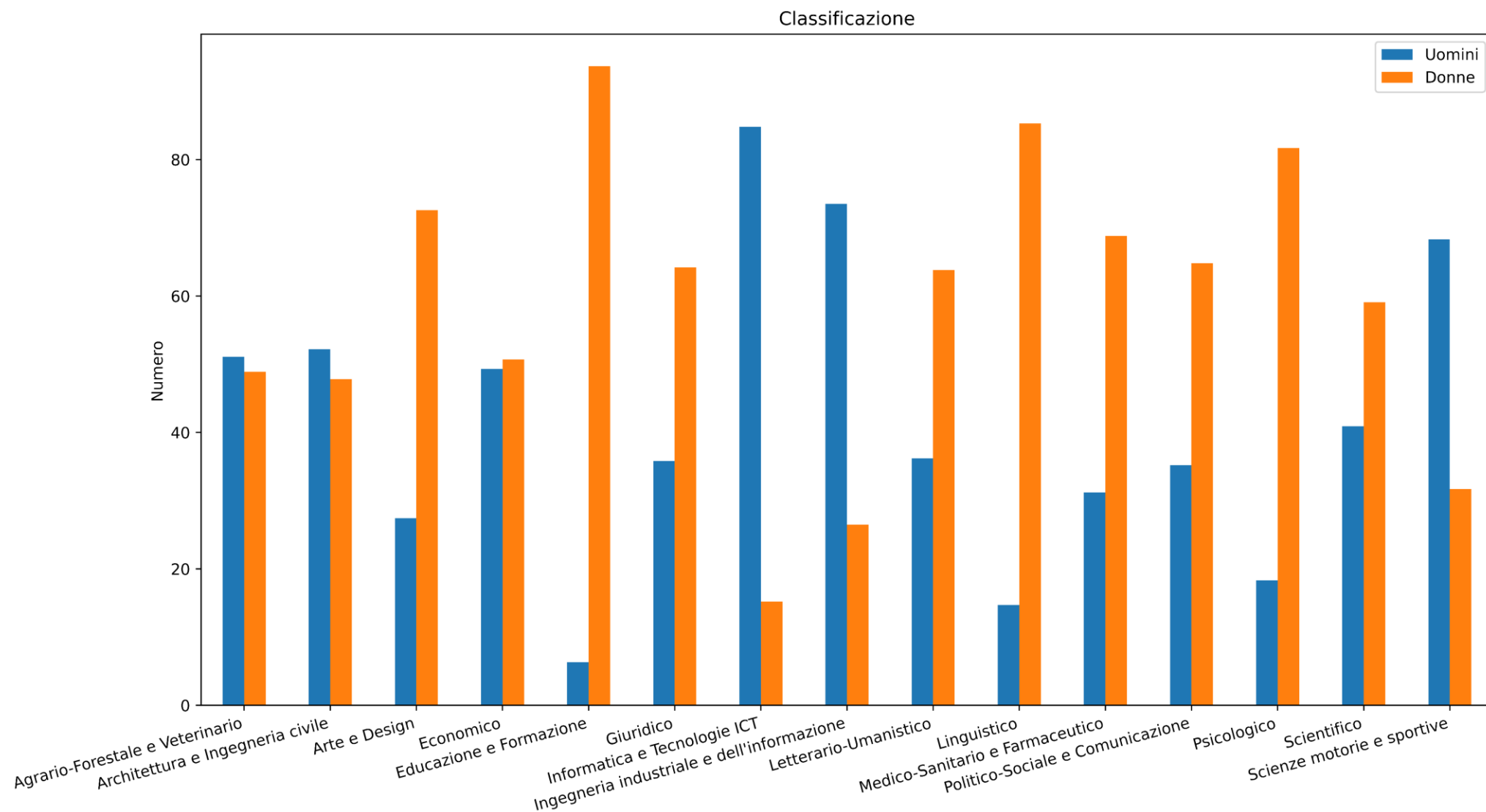
I dati durante gli studi



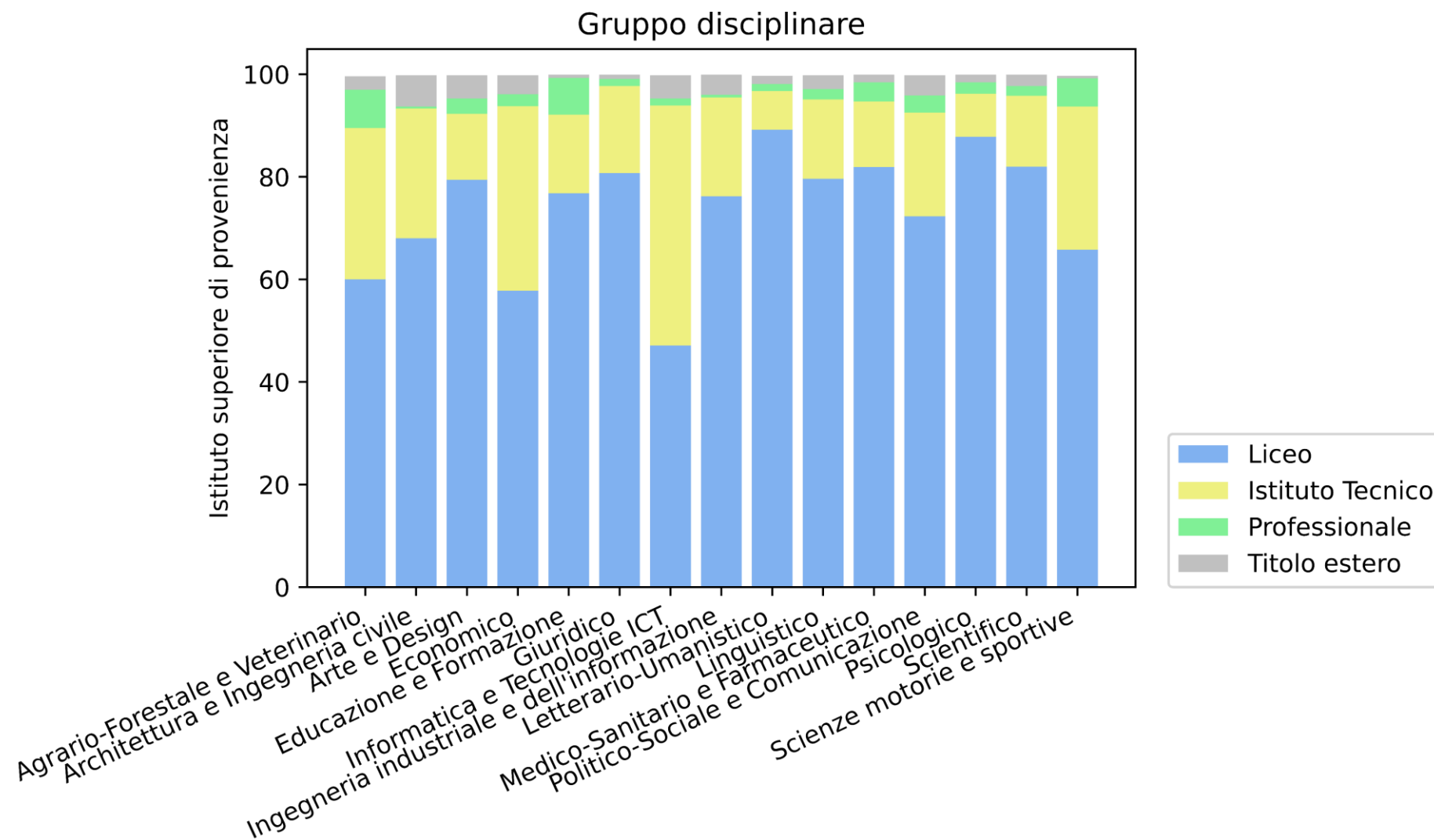
Quali sono le facoltà più frequentate?



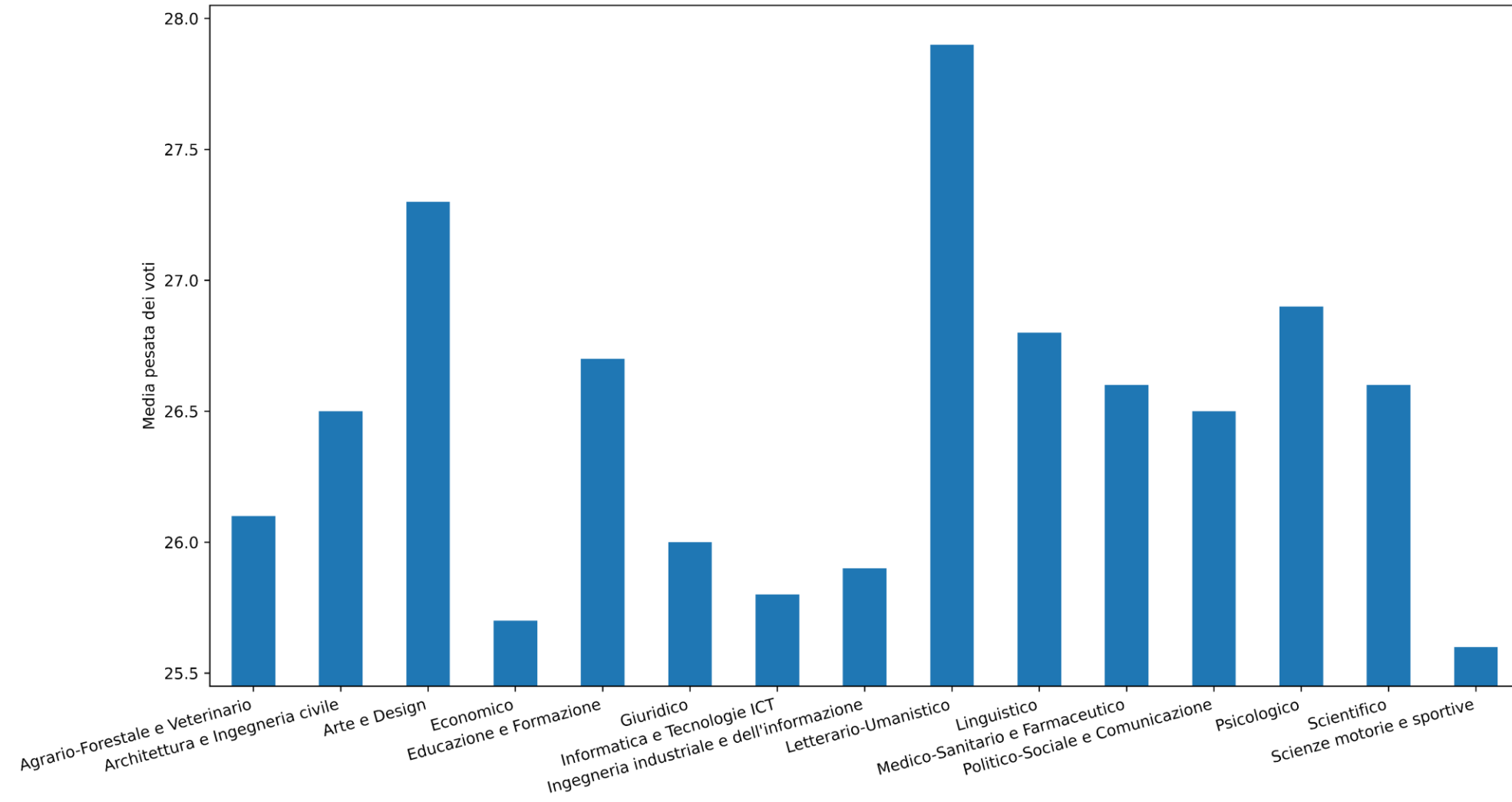
Disparità di genere nelle facoltà



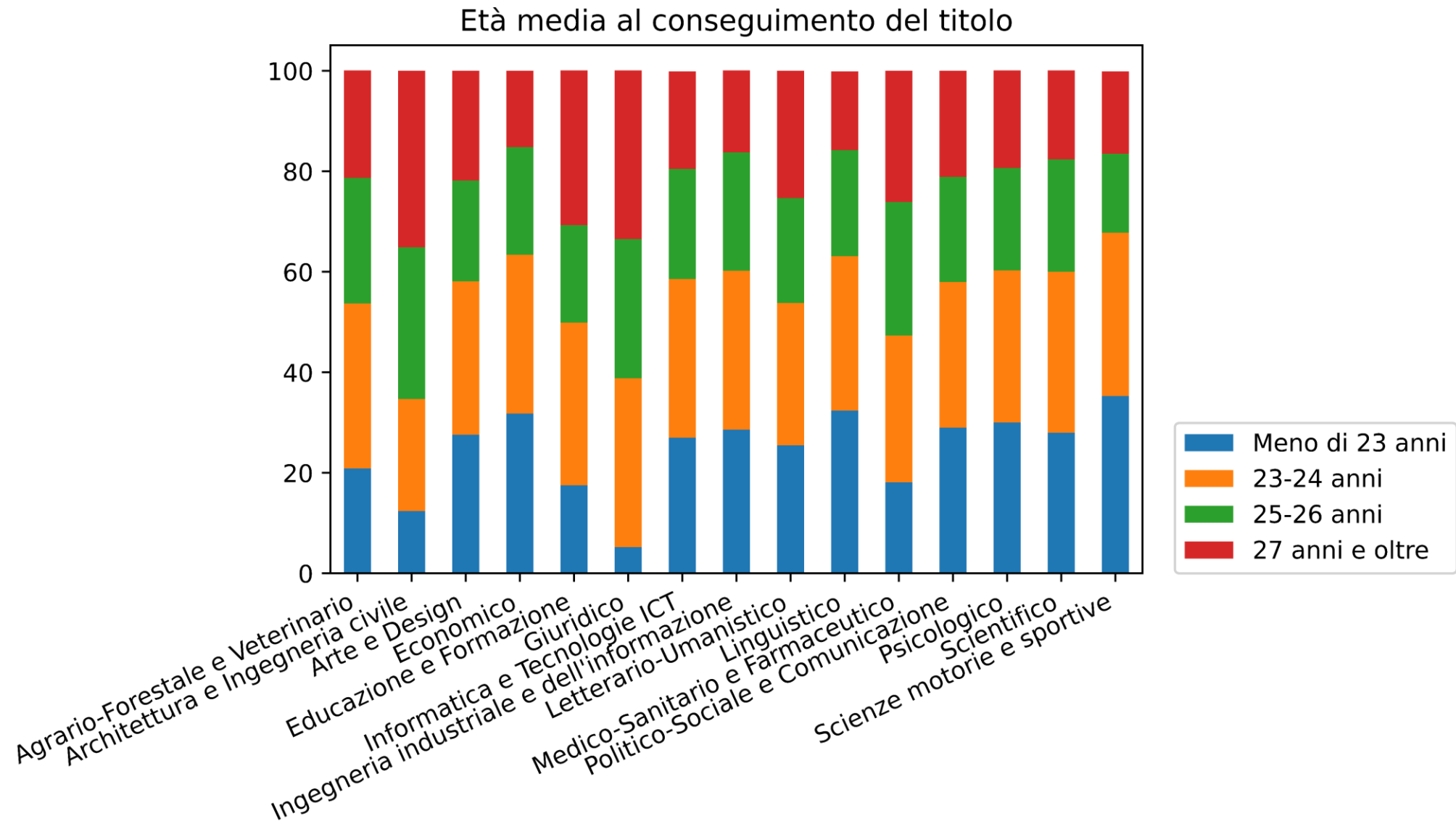
Che scuola superiore hai frequentato prima di iscriverti in università?



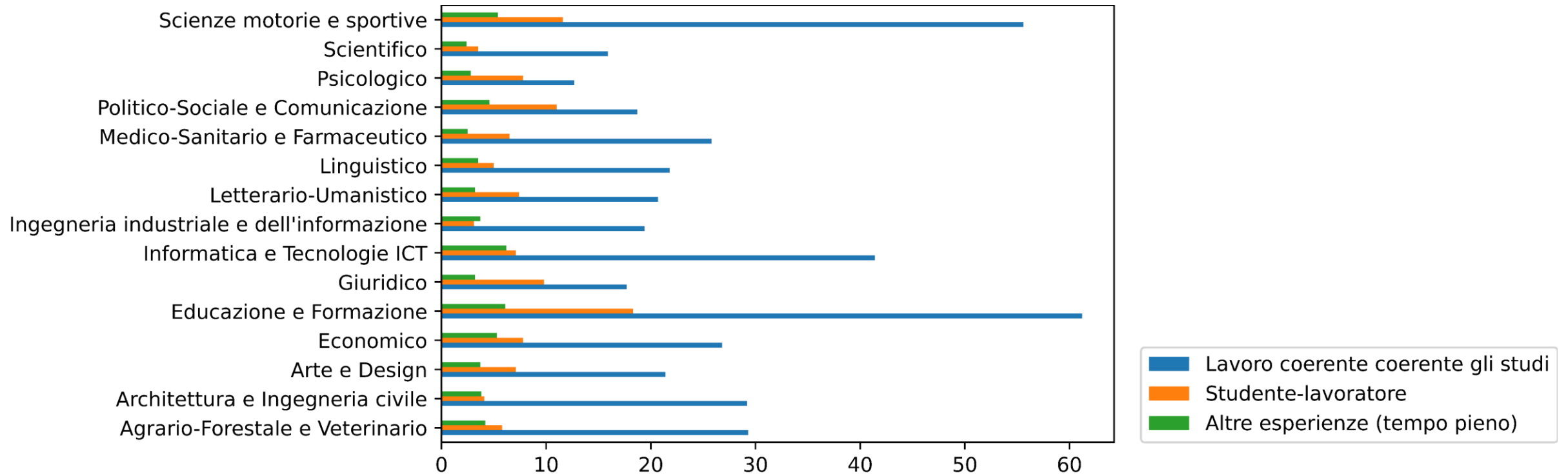
Ma... gli esami?



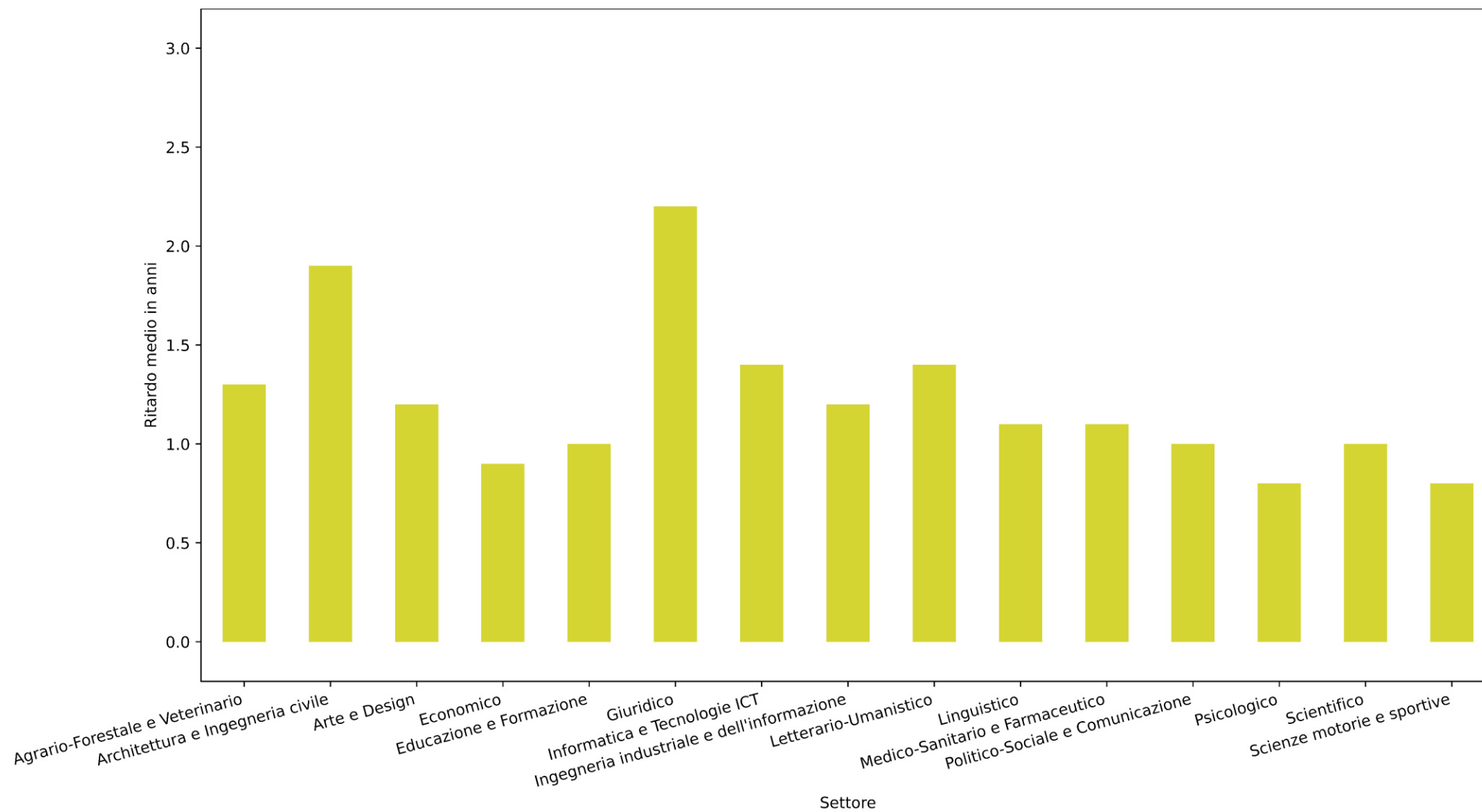
A che età si consegue, in media, la laurea?



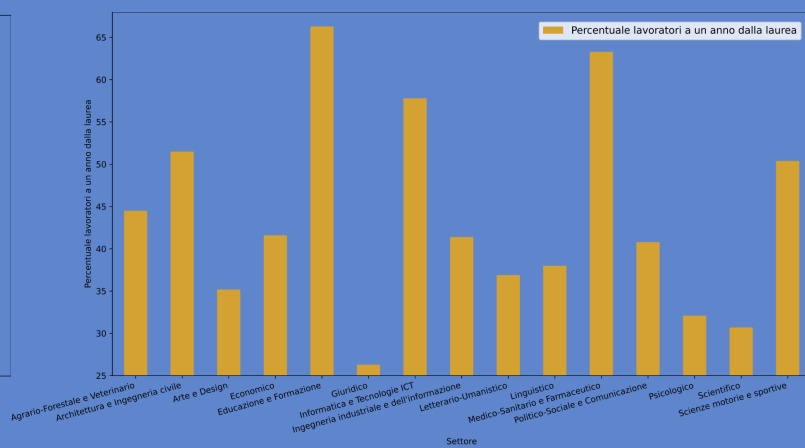
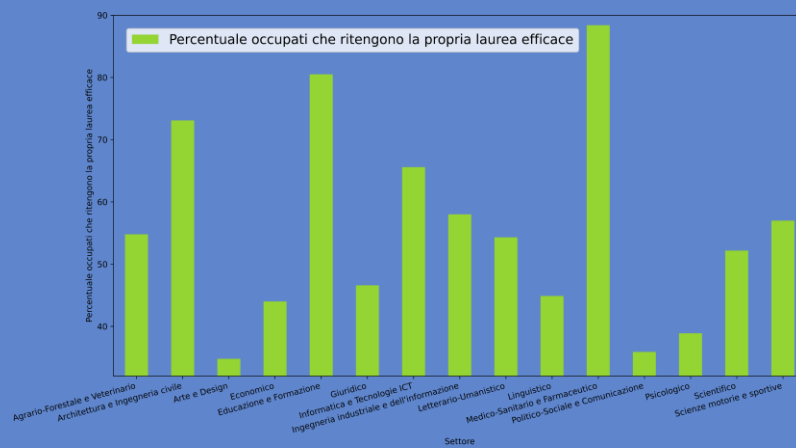
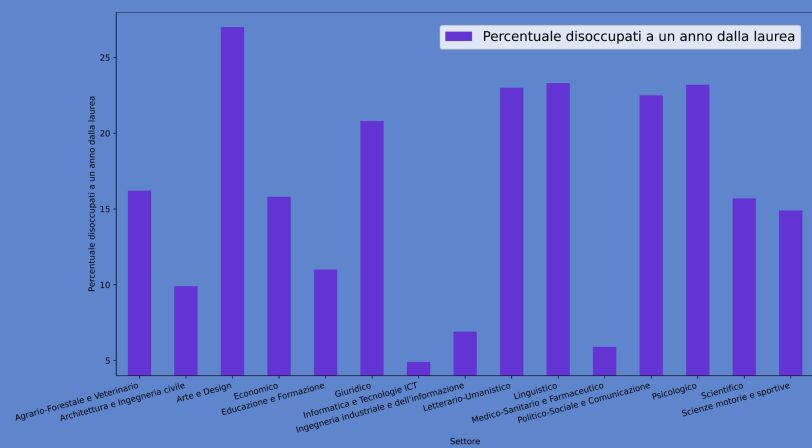
Lavoro durante gli studi



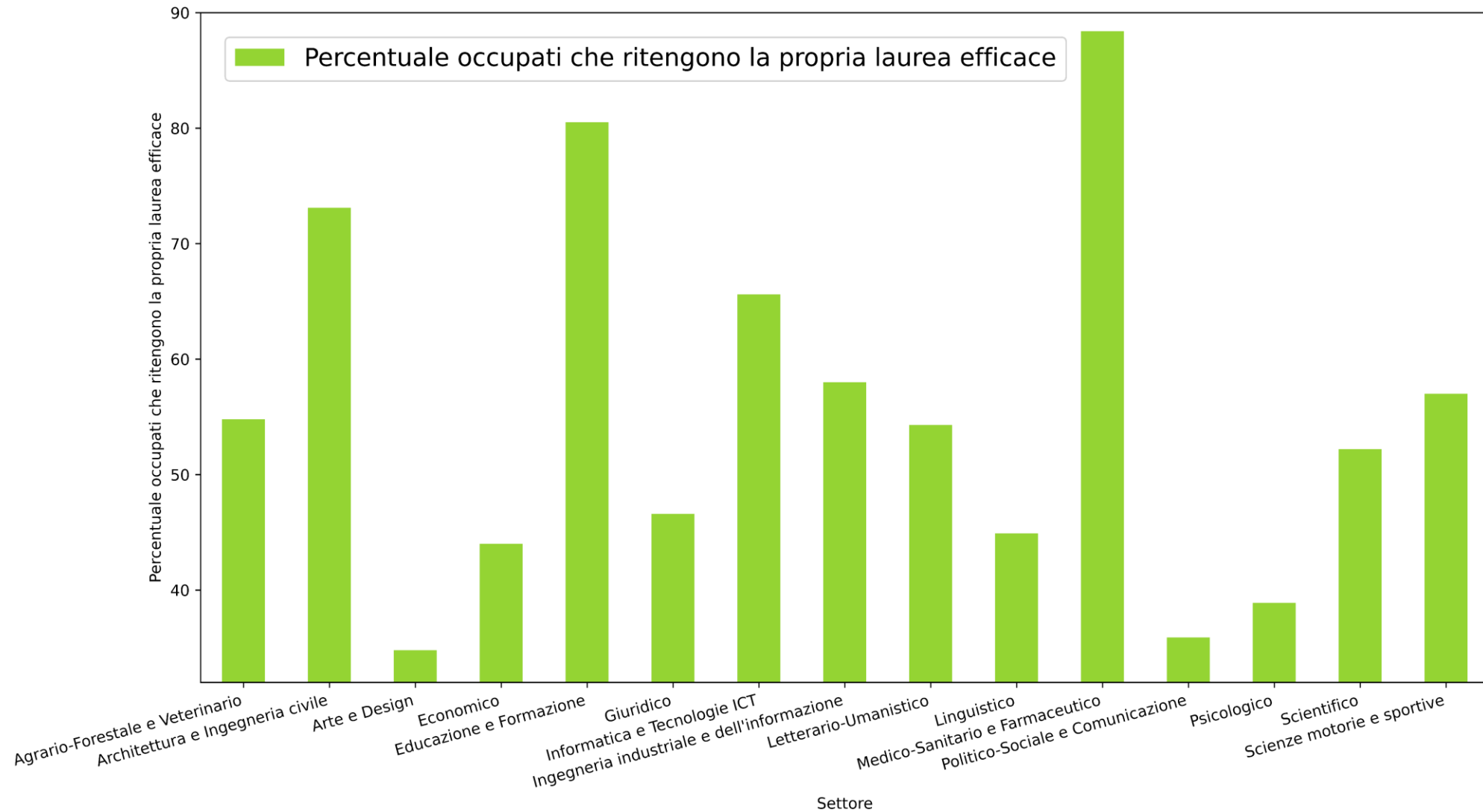
Anni fuori corso: ritardo medio



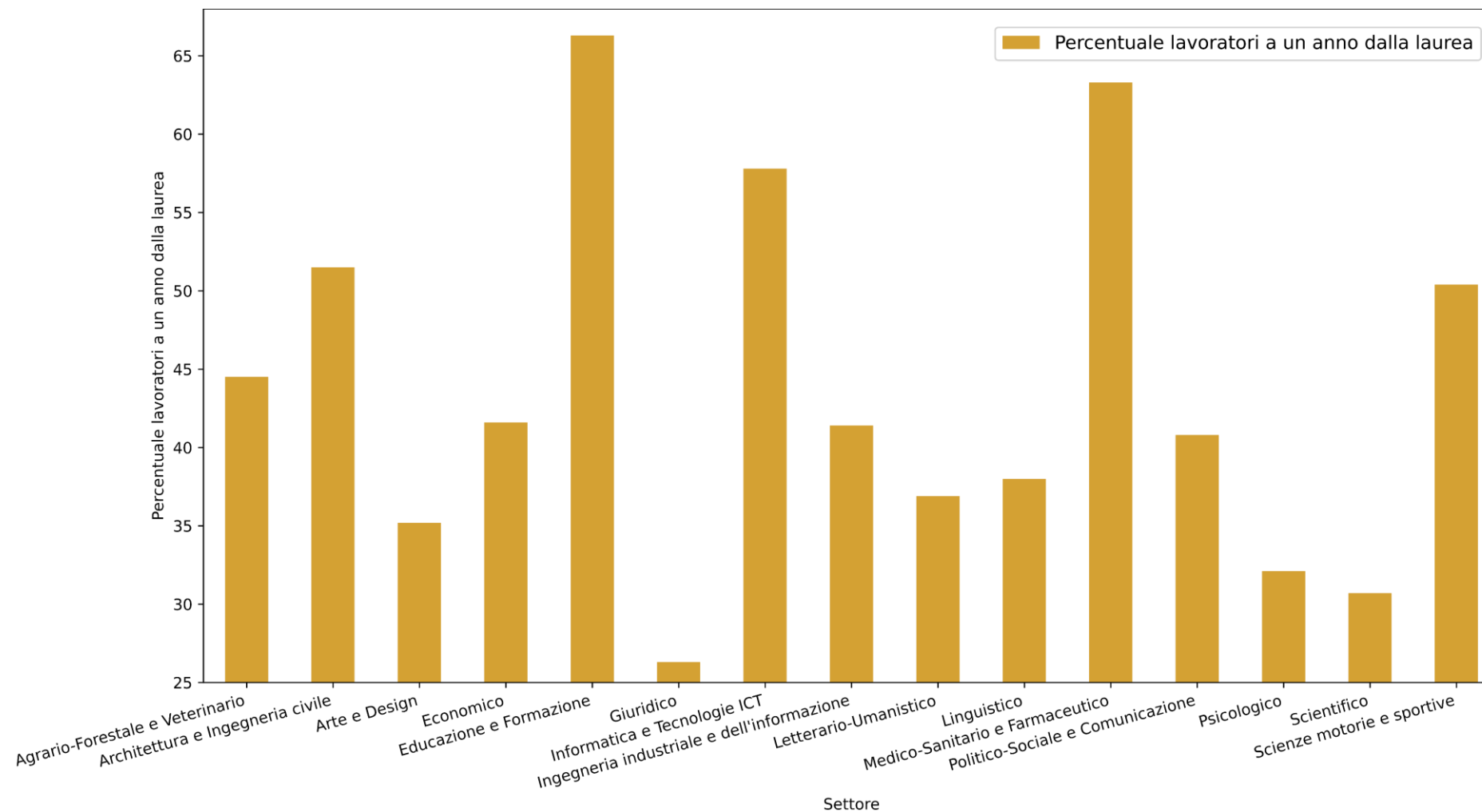
I dati occupazionali dopo gli studi



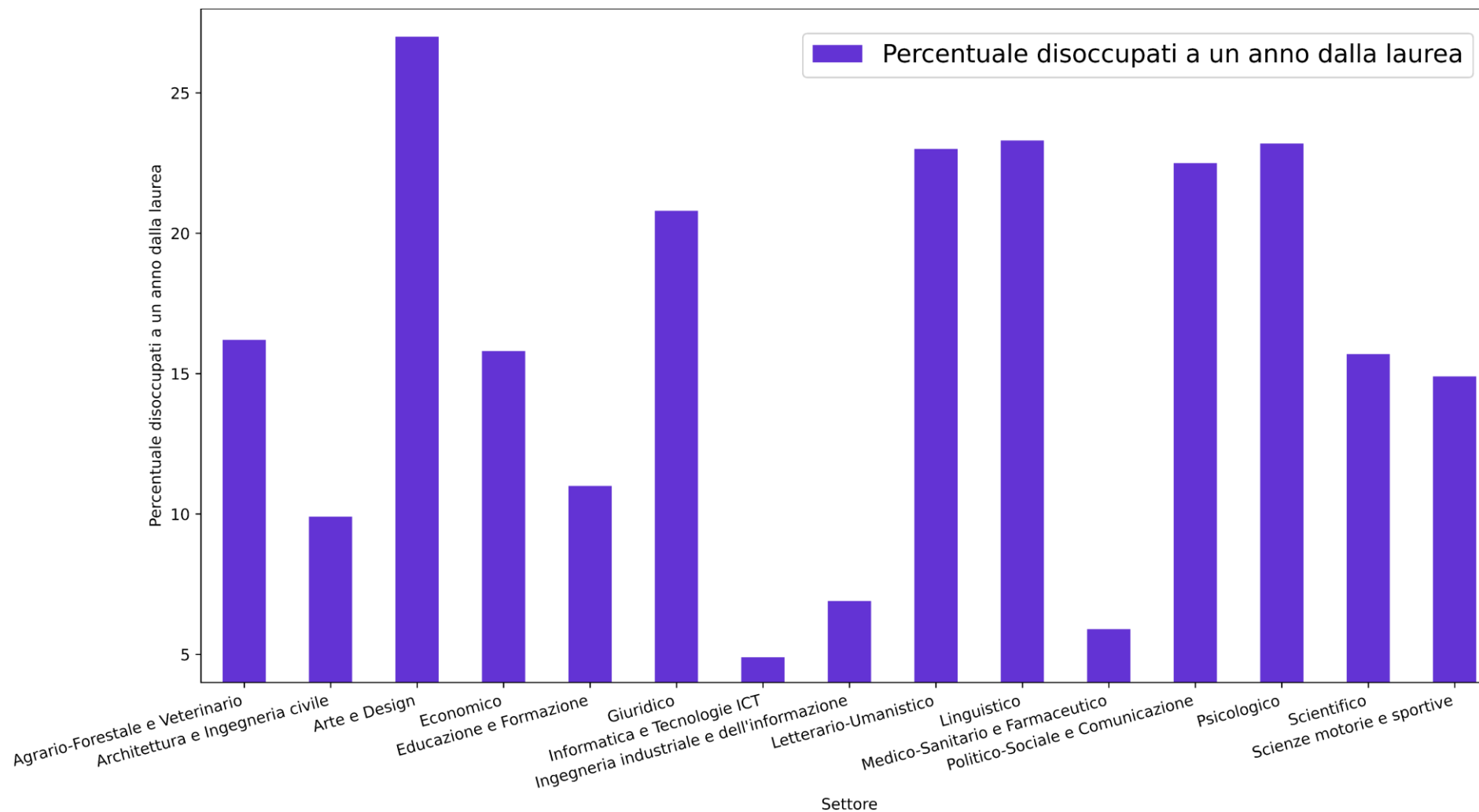
Percentuale occupati che ritengono la propria laurea efficace



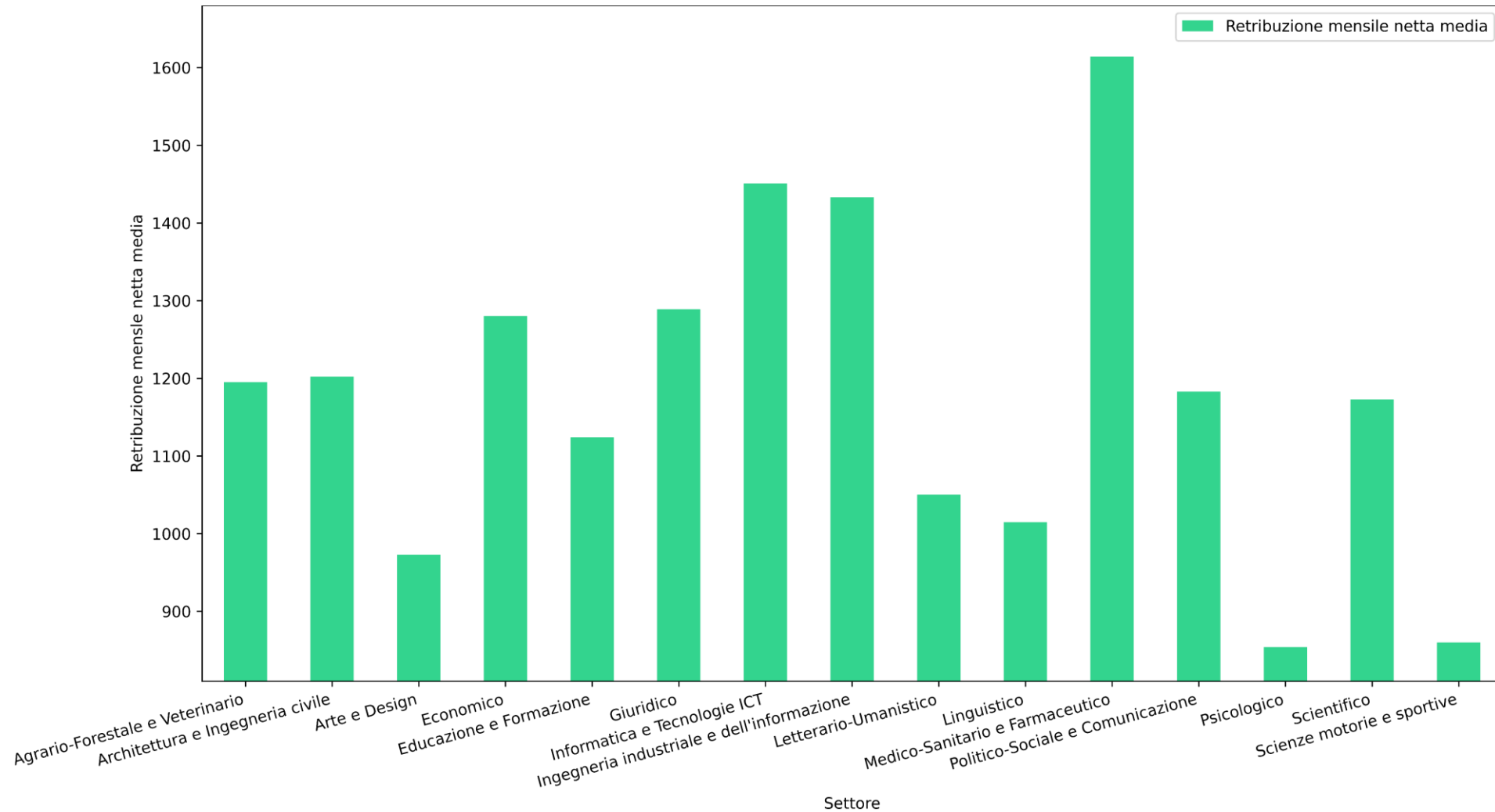
Lavoro a un anno dalla laurea



Disoccupazione a un anno dalla laurea



Retribuzione mensile netta media a un anno dalla laurea



Grazie per l'attenzione!