DILETTA ABBONATO

EDUCATION

Sapienza University of Rome, Rome, Italy *October 2019 - October 2020*
MSc in Data Science
Department of Information Engineering, Computer Science and Statistics

Sapienza University of Rome, Rome, Italy *October 2014 - December 2018*
BSc in Management Statistics
Department of Information Engineering, Computer Science and Statistics

WORK EXPERIENCE

Statistical Analysis at Sant'Eugenio Hospital, Rome, Italy
December 2018 - February 2019
Collaboration on the experimental thesis project 'The internship in CDL in nursing: needs perceived by the student and role of the clinical tutor', through the use of Python and R programming languages.

Statistical Analysis at the Regional Health Service of Emilia-Romagna, Ferrara, Italy
May 2018 - June 2018
Collaboration in the project 'Teenagers and the fear of others', through the use of R and Excel.

Scholarship at the Library of the Department of Computer Science Sapienza Unversity of Rome, Rome, Italy
January 2020 - Ongoing
Assistance to users; distribution of books; retrieval and relocation of books; room control.

Research fellowship at 'DIAG' Sapienza Unversity of Rome, Rome, Italy
May 2020 - August 2020
Winner of the research fellowship for the program 'Support to the preparation of a Data Quality Package for educational and research data' at the Department of Computer, Automatic and Management Engineering 'Antonio Ruberti'.

PROJECTS

Coca-Cola Vs Pepsi: A case study
Analysis of information on weekly sales of Coca Cola and Pepsi drinks in 86 stores, in the city of Chicago.
Programming language: R

Aptitude and reliability analysis: questionnaire on statistics
Analysis of a research conducted by an English teacher on 2751 students, carried out through a questionnaire, concerning their opinion on some applied sciences (statistics, mathematics and programming)..
Programming language: R

Item Response Theory: Didactic Evaluation
Analysis of a questionnaire of 10 questions submitted to 120 students concerning the evaluation of a university course they attended.
Programming language: R

Sampling analysis on A-grade players' scores
Analysis of the 2015/2016 football season, focusing on the average scores obtained by players who have made at least one presence in that season.
Programming language: R

Analysis of sales 'Mario Kart Wii'
Analysis of data regarding sales through the E-commerce of the Mario Kart video game for Nintendo Wii. The purpose of the analysis was to detect and compare the different pricing policies adopted by video game suppliers through linear regression.
Programming language: R

Analysis of salaries in the USA from 1976 to 1982
Analysis of a dataset containing information on the wages of 595 workers examined from 1976 to 1982 in the USA. The aim was to identify the factors that explain the differences in wages by focusing on ethnicity and different levels of education.
Programming language: R

Evaluation of the course and beauty of the teacher who delivers it: an analysis
Analysis of a research conducted on a sample of students at the University of Texas, with the aim of envy a correlation between the evaluation of the course by the students and the beauty of the teacher who provides the course.
Programming language: R

World00 dataset analysis
Analysis of a dataset of 103 countries with indicators on wealth, population, health and economy. The analyses correspond to the year 2000.
Software and Programming Language: Matlab, SPSS

Twitter Analysis: Models and Applications
Degree thesis on the diffusion of the hashatag 'StephenHillenburg' in the period between 27/11/2018 and 29/11/2018. After scraping through the Twitter API, it has been hypothesized that the network takes as model structure the one theorized by Albert-Làszló Barabsi.
Programming language: R

House Prices: Advanced Regression Techniques
Competition on Kaggle. The aim of the analysis was to predict the selling prices of houses based on features and to practice new feature engineering.
Programming language: Python, R

Soccer analytics
The goal was to perform an analysis of the largest open collection of football records ever released, containing all the space-time events (passes, shots, fouls, etc.) that occur during all the matches of the entire 2017-2018 season of seven competitions (La Liga, Serie A, Bundesliga, Premier League, Ligue 1, FIFA World Cup 2018, UEFA Euro 2016 Champions Cup).
Programming language: Python

What movie to watch tonight?
The objective of the analysis was to build a search engine on a list of movies that have a dedicated page on Wikipedia.
Programming language: Python

Explore California and Nevada with graphs
Building a system that provides users with information about roads in California and Nevada. In particular, the implementation of the system was composed in two parts: Backend, where algorithms implemented from scratch such as Dijkstra and Primm were used to define the functionality of the system; Frontend, where the visualization of the requests entered by the user was provided.
Programming language: Python

Randomize this...
Implementation of the 'one step randomized' algorithm and its n-dimensional version, the 'p-step randomize algorithm'.
Programming language: R

Stock, Dependency and Graphs
The dependency of a portfolio of stocks has been studied through a marginal correlation network.
Programming language: R

Deep-Sky-Bayes
Considering a log-normal model, the masses of the stars were analyzed from a Bayesian inference point of view and concluded with a psychic analysis.
Programming language and Software: R, OpenBUGS

How much does it cost me?
Starting from a survey carried out by the Ministry of Economy and Finance, concerning the variation of fuel prices in July 2019, an analysis was carried out to assess which factors were responsible for the variation of fuel prices.
Software and Programming language: R, OpenBUGS

Stats Database
Implementation of a database in MySQL and Neo4j, based on data shared by users on the 'Stats Stack Exchange' site.
Programming language and Software: MySQL, Neo4j

AWS
Building a cloud application on Amazon Web Service for a map site. It also evaluated the current status of the cloud by assessing whether the system complies with all the rules that make it safe and reliable.
Software: Docker

What ETER tells us about female and male university education
Using ETER and EUROSTAT data, gender differences within higher education were analysed.
Programming language: Python

Paper implementation of 'Deep Learning-Based Document Modeling for Personality Detection from Text
Implementation of the paper using the Python library 'Tensorflow' to identify, from a text, the 5 personality traits of an individual.
Programming language: Python

Breast Density Assessment and Breast Cancer Detection
The main goal of this project is to make a classification through breast clinical images. In this retrospective study , a Convolutional Neural Network finetuned models were trained to asses a diagnosis based on the original interpretation by experienced radiologist. The main tools used to carry on this study has been Google Colaboratory and open python libraries as Pytorch, PIL, OpenCV, NumPy, Pandas and Matplotlib.
Programming language: Python

IT SKILLS

Software and Programming Languages
R, SAS, SPSS, Java, Python, Matlab, MySQL, Neo4j, MongoDB, LaTeX, HTML, Microsoft Office Suite, OpenBUGS, Docker.

OTHER EXPERIENCE

SAS Training Camp June 2019
Rome,Italy

· Introduction to SAS Vija software.

Hash Code - Google's Coding Competitions February 2020
Rome,Italy

· Team competition using the Python programming language.

Internship at the University of Strasbourg March 2020 - July 2020
Strasbourg, France

· Internship for the Master's thesis in 'Economics of Science' at the department Bureau d'économie théorique et appliquée (BETA).

LANGUAGE SKILLS

Italian	*Native speaker*
English	*Fluent*
French	*School level*