OBIETTIVI FORMATIVI

AIMS AND SCOPE

The aim of the course is to introduce the knowledge of the R language for statistical data analysis with special focus on descriptive statistics, probability distributions and statistical inference.

LEARNING OBJECTIVES

1. Knowledge and understanding (Conoscenza e capacità di comprensione). The students will learn the basic concepts behind the R language for statistical data analysis with special focus on descriptive statistics, probability distributions and statistical inference.

2. Applying knowledge and understanding (Capacità di applicare conoscenza e comprensione). The students will be able to utilize the R language for: i) providing basic statistical analyses of data; ii) simulating data according to given probability distributions; iii) applying main methods of statistical inference.

3. Making judgements (Autonomia di giudizio). The students will learn to independently analyse data to extract knowledge from it through statistical analyses in R.

4. Communication skills (Abilità comunicative). The students will acquire the necessary communication skills and the appropriate use of technical language to present the results from the statistical analyses, based on the use of the statistical software R.

5. Learning skills (Capacità di apprendimento). The students will learn to use the statistical software R for basic data analysis and modeling. They will also acquire the competences needed to learn new data analysis and presentation techniques through the statistical software R.

MODALITÀ DI SVOLGIMENTO DELL’INSEGNAMENTO

Lectures and practical activities and data analysis in R. Should teaching be carried out in mixed mode or remotely, it may be necessary to introduce changes with respect to previous statements, in line with the programme planned and outlined in the syllabus.
**PREREQUISITI RICHIESTI**
Basics of linear algebra and statistics.

**FREQUENZA LEZIONI**
Mandatory.

**CONTENUTI DEL CORSO**

**Use of the statistical software in R regarding:**


**Probability.** Random number generation and data modeling according to different probability distributions: uniform, binomial, Poisson, Gaussian.


**TESTI DI RIFERIMENTO**


**ALTRO MATERIALE DIDATTICO**
Slides nad notes shared by the teacher through Studium.
### PROGRAMMAZIONE DEL CORSO

<table>
<thead>
<tr>
<th>Argomenti</th>
<th>Riferimenti testi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Introduction to R, Basic Commands in R, Indexing Data, Matrices and</td>
<td>Sections 2.3 and 2.4 of [1]</td>
</tr>
<tr>
<td>Lists, Loading Data;</td>
<td></td>
</tr>
<tr>
<td>2  Charts and Data Visualization</td>
<td>Lecture notes</td>
</tr>
<tr>
<td>3  Mean, Median, Variance, standard deviation, quantiles, percentiles,</td>
<td>Lecture notes</td>
</tr>
<tr>
<td>interquartile distance, boxplot, outlier detection</td>
<td></td>
</tr>
<tr>
<td>4  Bivariate analysis, statistical inference, contingency table, joint</td>
<td>Section 3.6 of [1], lecture notes</td>
</tr>
<tr>
<td>probability, marginal probability, chi-squared test, t-test, linear</td>
<td></td>
</tr>
<tr>
<td>regression.</td>
<td></td>
</tr>
</tbody>
</table>

### VERIFICA DELL'APPRENDIMENTO

**MODALITÀ DI VERIFICA DELL'APPRENDIMENTO**

Practical activity and data analysis with R. Learning assessment may also be carried out on line, should the conditions require it.

**ESEMPI DI DOMANDE E/O ESERCIZI FREQUENTI**

- Perform a statistical analysis of a dataset using the statistical software R;
- Fit a linear regression and evaluate the significance of the regression coefficients;
- Compute descriptive statistics of a dataset and produce visualizations of the data.