

Quesiti prova scritta

Selezione 2025S48



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

A handwritten signature in blue ink, likely belonging to a faculty member or administrator.

Indicazioni svolgimento:

I candidati sono invitati a sviluppare le seguenti domande:

- usando massimo 1 facciata per domanda foglio A4
- senza modificare le impostazioni (carattere, interlinea, margini, ecc)
- entro un tempo massimo di un'ora e mezza

Elenco n. 1

1. Describe the essential elements of an R package (e.g., source code, documentation, and tests) and explain how each contributes to making the package functional, easy to maintain, and suitable for sharing with others.
2. Explain what *software containers* (e.g., Docker, Singularity) are and how they contribute to code reproducibility and portability, particularly in the context of data analysis.
3. Explain why reproducibility is essential in bioinformatics research. In your response, identify common obstacles to achieving reproducible results and describe effective practices or tools that can help improve reproducibility.



A handwritten signature in blue ink, likely belonging to a faculty member or administrator.

Indicazioni svolgimento:

I candidati sono invitati a sviluppare le seguenti domande:

- usando massimo 1 facciata per domanda foglio A4
- senza modificare le impostazioni (carattere, interlinea, margini, ecc)
- entro un tempo massimo di un'ora e mezza

Elenco n. 2

1. Describe the benefits of organizing your code into an R package rather than relying on a set of separate R scripts for data analysis.
2. Describe the key difficulties involved in managing multiple software versions in a research setting, and explain which strategies or tools, such as *version control* systems, can help address these challenges.
3. Outline the main steps of a standard data analysis workflow (from raw data to scientific interpretation). For each step, specify the type of software or tools typically required, and explain the role of the software in the analysis.



Indicazioni svolgimento:

I candidati sono invitati a sviluppare le seguenti domane:

- usando massimo 1 facciata per domanda foglio A4
- senza modificare le impostazioni (carattere, interlinea, margini, ecc)
- entro un tempo massimo di un'ora e mezza

Elenco n. 3

1. Describe effective strategies for analyzing large-scale datasets in R, with a focus on the bioinformatics context. In your answer, discuss one or more approaches and provide examples where appropriate.
2. Explain how job schedulers (such as Slurm or PBS) are used to handle multiple jobs in a High-Performance Computing (HPC) cluster.
3. Describe the main challenges of developing and maintaining a data analysis platform that requires interoperability between programming languages.