Statistical Theory – Final Course Project

- The project is to be done in **pairs**.
- **Choose a dataset** to analyze according to the instructions below
- **Final paper submission** in the format of a **scientific article (PDF, in English)**, including the following sections (maximum length: **8 pages**):

1. **Abstract** – Background, main results, and conclusions.
2. **Introduction** – Literature review on the topic.
3. **Results** – Presentation of the main hypotheses and their testing using course tools and new tools; support with tables, figures, visualizations.
4. **Methods** – Description of tools and calculations used in the project: algorithms, statistical tests, etc.
5. **Discussion** – Conclusions and limitations of your analysis.

- You must **attach a link to a GitHub repository** with:
    - A **well-organized Readme**.
    - **Precise instructions** for reproducing the full analysis.
    - The **dataset** should also be uploaded.
    - A python notebook containing all your code.

Make sure your work **tells a story**, rather than simply running a series of statistical tests without connection.

Clarification:

As part of the project analysis (if relevant to the research question), please make sure your research question is **comprehensive** and use **as many tools and concepts from the course** as possible to answer it.

Concepts you are expected to consider and use (where relevant):

- Hypothesis testing: null/alternative hypothesis, type I and II errors, p-value, confidence intervals, significance level.
- Tests: MP test, UMP test, GLRT, one-sided and two-sided tests, independent and paired samples t-tests, chi-square test, distribution fitting tests, F test, Wald test, Rao's score test, Wilks' test, CAN estimators, MSE, stopping times, parametric and nonparametric tests, correlation, multiple testing corrections, dependency, regression, classification, evaluation metrics, interaction terms, and any other concept that can give insights regarding your question.

Additional Guidelines:

- Graphs must be **clear**, with **labeled axes**, **captioned figures**, **legible fonts**, and **meaningful colors**.