

## Data Analytics

Academic Year 2024-25

### Course Assignment N. 13: College network

Prof. Fabio Crestani & Dr Ana-Maria Bucur

TAs: Lili Lu & Enrico Verdolotti

For this assignment you will work in pairs to carry out simple tasks of data analysis given a specific dataset. The goal of this assignment is to use Python and complementary libraries on a given dataset to *explore* and *analyze* the given data and *draw conclusions*.

#### Description

This dataset is comprised of private messages sent on an online social network at the University of California, Irvine. Users could search the network for others and then initiate conversation based on profile information. An edge  $(u, v, t)$  means that user  $u$  sent a private message to user  $v$  at time  $t$ . There are 1899 nodes and 59835 edges. Your goal is to:

- Explore and describe the data (preprocess the data, visualize the variables with different graphs, distribution of the variables). While exploring the data, define research questions and answer them such as which is the most *active* user/s? Which are highly connected or isolated from others? Etc.
- Plot the graph that shows the links between the different users, *i.e.*, how the users are connected. Use graphics to enlarge the users that have most centrality, etc.

#### Submission procedure and evaluation

You should produce a report of your work and its evaluation along with the source code. It will be a concise explanation of how you tackled the different tasks, the reasons of your choices, successive conclusions, graphs you produced, results of the decisions and their accuracy, *etc.*

Use Jupyter Notebook to produce results of the commands in a single .ipynb file. For more information check: <https://jupyter.org/documentation>

The report (max 5 pages) and the code of the project need to be submitted via iCorsi.

Please, upload all the required items in a single file and name it following the structure: **no\_Project.[zip|tar.gz|7z]**. For instance, 05\_projectname.tar.gz

The dataset regarding this project can be downloaded from: [Dataset](#)