

Homework 2: Pandas!

Introduction

The purpose of this exercise is to assess your understanding of **Pandas DataFrames** and your ability to work with them effectively. You will implement a series of functions that perform **data manipulation tasks**, including **cleaning, merging, and analyzing** DataFrames.

Submission Guidelines

1. This is an assignment to be done in pairs of exactly 2 students, single submission shall be penalized by 20%. Students who are caught cheating will obtain a score of 0 points. This homework is worth 20% of your final grade.
2. Comment your code properly, which includes naming your variables in a meaningful manner. Badly documented code will be penalized.
3. There is a 25% penalty for deliveries at most 2 days late. Homeworks submitted after this will be graded with 0 points.
4. Submit a single zipped folder following the naming convention **student1id_student2id.zip** containing the 20 individual .py functions, which are produced during the execution of the **assignment.ipynb** notebook when running it from start to finish. Make sure that all of them are inside the zipped folder. No additional files should be present, and the submission **must** be in **.zip** format — other compression formats such as **.tar** are not accepted.
5. The submission deadline is **Tuesday, 21st of October 2025, at 23:59**.
6. One group member shall submit the **.zip** file to the corresponding Moodle activity. It is not necessary for both group members to submit.

Setup

Before you begin, ensure that you answer every question in the **assignment.ipynb** notebook. You must keep all provided function names unchanged and make sure your outputs match the requested formats; otherwise, your answer to that question will not be considered.

Your task is to implement the required functions in the **assignment.ipynb** notebook. Read the instructions carefully and follow all specified requirements and restrictions.

Restrictions

You are not allowed to use third-party libraries. Only use libraries that are part of the Python standard library (i.e. that do not require a **pip install** or **conda install**).

Files Provided

You are given 5 files:

1. **assignment.ipynb** – Contains the assignment instructions and exercises to be completed.
2. **onlineretail.csv** – Dataset used for **Part I** and **Part II** of the assignment.
3. **customer_info.csv** – Dataset also used for **Part II** of the assignment.
4. **Managers.csv** – Dataset used for **Part III** of the assignment.
5. **Teams.csv** – Dataset also used for **Part III** of the assignment.