# Indicative exam for the Midterm ITC6001

*Declaimer: It is provided as an indicative form of the midterm. It is a previous year exam. Questions may-be different.*

# Instructions

# Time allocated: 2.5 hours

### Open notes exam

### Usage of Generative AI tools is prohibited

### Individual work exam

### Answer all questions

**Exam Instructions:**

You have 15 minutes to login in the Virtual Machines and start the relevant tools (e.g. Anaconda).

For each question: create a python file:

* Q1 🡪 Q1-Lastname-Firstname.py
* Q2 🡪 Q2-Lastname-Firstname.py
* Q3 🡪 Q3-Lastname-Firstname.py
* When finished: zip all the files in a file named: “**LastName-id-midTerm-Exam.zip**”
* Submit the file in the blackboard: **/Term Exams/Midterm**

The files should be able to be executed just running them: Any datafiles you have used have be in the same directory as the python files, no paths should be used.

**Coding:**

Use python, and related libraries, e.g., Json, csv, Pandas, numpy and a mysql (e.g. pymysql) connector. No other framework may be used.

**Grading scale: US-Scale:**

## Q1: Python: Pandas Data frames 30%

***Instructions for Q1***

* *Use Pandas Data frame to read the file “diamonds.csv”, and then write code in python for the following sub-questions. You will find diamonds online.*
* *All sub-questions (1-6) should be in the same python file*
* *For each question, before the code. paste the question itself in the python file as a comment*
* *You should define a variable for each of the questions 1,2,3,4,5,6 named as: result1, result2, result3, result4, result5, result6 which should contain the answer and should be printed. E.g.*

*# Q1-1.* Print the first 4 and last 4 rows of the Data frame

*code …*

*more code ….*

*result1= code …*

*print (result1)*

***Q1 work to be done***

1. Calculate the min, max and average value for the ‘dept’ column
2. List all the ‘cut’ values, each ‘cut’ value should appear once only
3. Count and print the number of ‘diamonds’ for each cut
4. Print the number of duplicated rows in the data frame
5. Remove the columns that are marked with x,y,z from the data frame
6. Print the rows for the diamonds where the cut is fair or very good

## ~~Q2: Regular expressions: not covered this year~~ 20%

## Q3: Database: 50%

Go to **/assignments/Midterm**. You will find files: step-1-Create-DB.py and step-2-Populate-DB.py. It contains code to create and populate the DB. The DB contains two tables: one is about employees, while the second is about mentorships, i.e. who supervises whom and in what project.

***Instructions for Q3*** *Create & Populate the database*

*(Note: the code has been tested with spyder, anaconda in Linux)*

1. The connector from Python to MySQL is: pymysql
2. Create the database and tables by running the: “step-1-create-DB.py”: you will need to set username and password of your own database
3. Populate the database by running “step-2-populate-DB.py”. Make sure that the directories are right, and you provide the correct password.
4. Inspect the database to make sure that all data have been inserted in the tables (you may need to run a select \* command)

**Q3: *work to be done***

Write queries on the database you have created. The queries should be in Python.

* *All sub-questions (1-10) should be in the same python file*
* *For each question, before the code, paste the question itself in the python file as comment*
* *You should define a variable for each of the questions 1,2,3,.. 10, named as: result1, result2, result3, … result10, which should contain the answer and should be printed. E.g.,*

*#* Display the names of the courses and the term they are offered

*code …*

*result1= more code*

*print (result1)*

1. Display the names of all suppliers, as well as their city
2. Display all the name of suppliers that start with ‘S’
3. Count the number of products and the number of unique products based on pname
4. Count the number of Red products
5. Which is the lightest product?
6. Sort products by product names, and then by product color
7. Sort product names by descending order
8. Display the number of products for each color;
9. What is the name of the supplier that supplied 200 units of product 1?