Homework #1

Instructor: Dr. Zafeirakis Zafeirakopoulos   *Name:* Yunus Emre Yumşak   *Student Id:* 1801042659

Assistant: Gizem Süngü   *1801042659*

Course Policy: Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- It is not a group homework. Do not share your answers to anyone in any circumstance. Any cheating means at least -100 for both sides.

- Do not take any information from the Internet.

- No late homework will be accepted.

- For any questions about the homework, come to my office hour.

- After the office hour, no questions about the homework by email will be responded.

- Submit your homework (both your latex and pdf files in a zip file) into the course page of Moodle.

- Save your latex, pdf and zip files as "Name_Surname_StudentId".{tex, pdf, zip}.

- The deadline of the homework is 22/04/21 23:55.

| **Problem 1** | (100 points) |

Homework 1 considers a Covid-19 dataset which is published on Github. Please download any document type that you prefer of the dataset from the links which are shown in Figure 1. The dataset is updated daily and
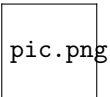
Figure 1: The complete dataset links

includes data on confirmed cases, deaths, hospitalizations, testing, and vaccinations as well as other variables of potential interest. The data set has the following basic columns:

- iso_code: Short name of the country

- continent: The continent where the country exists

- location: The country name

- date: The date when the data about various variables are taken.

You are responsible to implement a program which reads the given dataset from the file and computes the data for the following questions. Any programming language that you prefer will be accepted. Putting comments on your functions that you implement is must. Each question must be appended to a file which is called "output{.csv, .txt}". The file contains the first 18 questions listed below. The 18th question will be written in this document.

1. How many countries the dataset has?

2. When is the earliest date data are taken for a country? Which country is it?

3. How many cases are confirmed for each country so far? Print pairwise results of country and total cases.

Table 1: The format of the output for the questions 5, 6, 7, 8, 9, 10, 12, 13.

Country	minimum	maximum	average	variation
value	value	value	value	value

4. How many deaths are confirmed for each country so far? Print pairwise results of country and total deaths.

5. What are the average, minimum, maximum and variation values of the reproduction rates for each country?

6. What are the average, minimum, maximum and variation values of the icu_patients (intensive care unit patients) for each country?

7. What are the average, minimum, maximum and variation values of the hosp_patients (hospital patients) for each country?

8. What are the average, minimum, maximum and variation values of the weekly icu (intensive care unit) admissions for each country?

9. What are the average, minimum, maximum and variation values of the weekly hospital admissions for each country?

10. What are the average, minimum, maximum and variation values of new tests per day for each country?

11. How many tests are conducted in total for each country so far?

12. What are the average, minimum, maximum and variation values of the positive rates of the tests for each country?

13. What are the average, minimum, maximum and variation values of the tests per case for each country?

14. How many people are vaccinated by at least one dose in each country?

15. How many people are vaccinated fully in each country?

16. How many vaccinations are administered in each country so far?

17. List information about population, median age, # of people aged 65 older, # of people aged 70 older, economic performance, death rates due to heart disease, diabetes prevalence, # of female smokers, # of male smokers, handwashing facilities, hospital beds per thousand people, life expectancy and human development index.

Table 2: The format of the output for the question 17

Country	population	median age	# of people aged 65 older
value	value	value	value

18. Summarize all the results that you obtain by the first 17 questions (except question 2).

Table 3: The format of the output for the question 18

Country	q#3	q#4	q#5_min	q#5_max	q#5_avg	q#5_var
value	value	value	value	value	value	value

19. Comment the results based on your observations. Write your opinions about the reasons of increasing infection rates by giving examples from the results. Feel free to explain any situation that you observe. More observations more opportunities will bring you for the second homework.
(Solution)

- Hypothesis: Second dose vaccination prevents covid cases.

- Africa has a value of 11 million of vaccinated at least 1 dose
  and has a value of 4 million of vaccinated 2 dose. 0.36 of people those are vaccinated
  has vaccinated the second dose.
  Africa has 4.5 million cases total. Africa has 1.34 billion population. 3.4 out of 1000 people are covid
  positive.

- UK has a value of 33.5 million of vaccinated at least 1 dose
  and has a value of 1.2 million of vaccinated 2 dose. 0.036 of people those are vaccinated
  has vaccinated the second dose. UK has 4.5 million cases total. UK has 67 million population. 67 out of
  1000 people are covid positive.

  Africa has higher vaccination of 2. dose than UK. Africa has also less case by population rate than UK.

- Hypothesis: Intensive care unit usage does not decrease death rate.

- Cyprus has 60681 deaths and 875899 population. Ratio is 0.069.
  Cyprus has 8.6 average of weekly intensive care units patients. Cyprus has 49.5 average of weekly hospital
  patients.Ratio is 0.17.

- Czechia has 1615461 deaths and 10708982 population. Ratio is 0.15.
  Czechia has 973.3 average of weekly intensive care units patients. Czechia has 5144.3 average of weekly
  hospital patients. Ratio is 0.19.

- Usage of intensive care unit does not affect death ratio.