MATH 118: Statistics and Probability **(Due: 26/04/21)**

Homework #1

Instructor: Dr. Zafeirakis Zafeirakopoulos *Name:* Muhammed Bedir ULUCAY *Student Id:* 1901042697
Assistant: Gizem Süngü *1901042697*

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 https://github.com/owid/covid-19-data/tree/master/publ
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 includes data on confirmed cases, deaths, hospitalizations, testing, and

[scale=0.5]pic.png

Figure 1: The complete dataset links

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

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

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)

death Rate Of Total Case = total death ÷ total case

Observation 1:
Countries with good economic performance kept the disease under better control.We can see the economic and death rate nearly reverse ordering this is proof our observation.As we seen table 4

Table 4: Observation 1

country	Death Rate Of Total Case	economic performance
Canada	0.021260	44017.591
Australia	0.030864	44648.710
Germany	0.025616	45229.245
Austria	0.016721	45436.686
Iceland	0.004619	46482.958
Denmark	0.010165	46682.515
Sweden	0.015419	46949.283
Netherlands	0.012167	48472.545
Saudi Arabia	0.016887	49045.411
Cayman Islands	NaN	49903.029
Bermuda	NaN	50669.315
United States	0.017948	54225.446
Hong Kong	NaN	56054.920
San Marino	0.017166	56861.470
Switzerland	0.016658	57410.166
Norway	0.006656	64800.057
Kuwait	0.005643	65530.537
United Arab Emirates	0.003144	67293.483

Observation 2:
If the number of elderly people is high, the death rate is higher than in other countries.We can see the age is when age is growing Death Rate also growing. As we seen table 5

Table 5: Observation 2

country	aged 65 older	Death Rate
Netherlands	18.779	0.012167
Lithuania	19.002	0.016185
Czechia	19.027	0.017711
Slovenia	19.062	0.018011
Austria	19.202	0.016721
Malta	19.426	0.013509
Spain	19.436	0.022634
Estonia	19.452	0.009157
Denmark	19.677	0.010165
France	19.718	0.019096
Croatia	19.724	0.021409
Latvia	19.754	0.018500
Sweden	19.985	0.015419
Greece	20.396	0.029996
Bulgaria	20.801	0.039134
Finland	21.228	0.010630
Germany	21.453	0.025616
Portugal	21.502	0.020417
Italy	23.021	0.030301
Japan	27.049	0.018218

Observation 3:
If hospital bed per thousand's rate is high patients could get better treatment this must be the decrase death rate.As we seen table 6

Table 6: Observation 3

country	hospital beds per thousand people	Death Rate
Sudan	0.80	0.067047
Bangladesh	0.80	0.014252
Myanmar	0.90	0.022481
Ghana	0.90	0.008367
Somalia	0.90	0.051102
Nicaragua	0.90	0.026557
Philippines	1.00	0.017245
Central African Republic	1.00	0.013200
Indonesia	1.04	0.027101
Gambia	1.10	0.029856
Bolivia	1.10	0.044049
Morocco	1.10	0.017703
Costa Rica	1.13	0.013471
Qatar	1.20	0.001841
United Arab Emirates	1.20	0.003144

Observation 4:
Smoking increases the mortality rate with the corona.Smoke rate and Death rate are right proportion. As we seen table 7.

Table 7: Observation 4

country	smokerRate(m+f smokers)	Death Rate
Ghana	8.0	0.008367
Ethiopia	8.9	0.013887
Nigeria	11.4	0.012561
Eritrea	11.6	0.002869
Panama	12.3	0.017183
Benin	12.9	0.012482
Ecuador	14.3	0.049205
Togo	15.1	0.009579
Niger	15.5	0.037160
Oman	16.1	0.010307
Barbados	16.4	0.011724
Senegal	17.0	0.027395
Eswatini	18.2	0.036389
Colombia	18.2	0.025819
Cape Verde	18.6	0.009505
Liberia	19.6	0.041626
Uganda	20.1	0.008191
El Salvador	21.3	0.030761
Kenya	21.6	0.016245
Iran	21.9	0.030283
India	22.5	0.012196
Bahamas	23.5	0.019989
Costa Rica	23.8	0.013471
Mali	24.6	0.033134

Observation 5:
Hitting a covid vacations is slows down the rate of coronary spreading.
We cannot see any lowing so we can say vaccation not effecting spreading.As we seen tables 8

Table 8: Observation 5

location	date	new$_cases$	people vaccinated
Slovenia	2021-01-17	569.0	48659.0
Slovenia	2021-01-18	299.0	48721.0
Slovenia	2021-01-19	1713.0	49276.0
Slovenia	2021-01-20	1714.0	50428.0
Slovenia	2021-01-21	1455.0	51750.0
...	...	...	...
Slovenia	2021-04-11	549.0	338961.0
Slovenia	2021-04-12	287.0	340604.0
Slovenia	2021-04-13	1050.0	343662.0
Slovenia	2021-04-14	1240.0	351326.0
Slovenia	2021-04-15	891.0	360451.0

Observation 6: Does hand washing rate reduce corono emission
We will compare two country Vietnam(85.847) and Zambia(13.938)
As we seen table 9 and 10 Vietnam population is so much bigger than Zambia but the different hand washing facilities is make huge different covid spreading.

Table 9: Observation 6

location	date	new cases	population	hand washing facilities
Zambia	2020-11-23	30.0	18383956.0	13.938
Zambia	2020-11-24	12.0	18383956.0	13.938
Zambia	2020-11-25	69.0	18383956.0	13.938
Zambia	2020-11-26	18.0	18383956.0	13.938
Zambia	2020-11-27	16.0	18383956.0	13.938
...	...	...	...	...
Zambia	2021-04-11	111.0	18383956.0	13.938
Zambia	2021-04-12	35.0	18383956.0	13.938
Zambia	2021-04-13	154.0	18383956.0	13.938
Zambia	2021-04-14	171.0	18383956.0	13.938
Zambia	2021-04-15	143.0	18383956.0	13.938

Table 10: Observation 6

location	date	new cases	population	hand washing facilities
Vietnam	2020-11-23	5.0	97338583.0	85.847
Vietnam	2020-11-24	4.0	97338583.0	85.847
Vietnam	2020-11-25	5.0	97338583.0	85.847
Vietnam	2020-11-26	10.0	97338583.0	85.847
Vietnam	2020-11-27	8.0	97338583.0	85.847
...	...	...	...	...
Vietnam	2021-04-11	1.0	97338583.0	85.847
Vietnam	2021-04-12	12.0	97338583.0	85.847
Vietnam	2021-04-13	9.0	97338583.0	85.847
Vietnam	2021-04-14	19.0	97338583.0	85.847
Vietnam	2021-04-15	25.0	97338583.0	85.847

Observation 7:
We are try to determine of heart problem is cause of death with covid.
As we seen in table 11 when heart problem death rate is increase also death rate is increase.
We can say that the relation betwenn right rate heart disase can cause of death.

Table 11: Observation 7

country	death rates due to heart disease	Death Rate
Colombia	124.239	0.025819
Ireland	126.459	0.019884
Portugal	127.842	0.020417
Chile	127.993	0.022480
Luxembourg	128.275	0.012161
Panama	128.346	0.017183
New Zealand	128.797	0.010035
Kuwait	132.235	0.005643
Sweden	133.982	0.015419
Nicaragua	137.016	0.026557
Costa Rica	137.973	0.013471
Ecuador	140.448	0.049205
Cyprus	141.171	0.005246
Austria	145.183	0.016721
United States	151.089	0.017948
Bahrain	151.689	0.003567
Mexico	152.783	0.092014
Slovenia	153.493	0.018011
Finland	153.507	0.010630
...	...	...
Somalia	365.769	0.051102
Philippines	370.437	0.017245
Romania	370.946	0.025287
Guyana	373.159	0.022816
Syria	376.264	0.068266
Guinea-Bissau	382.474	0.017809
Montenegro	387.305	0.014857
Lesotho	405.126	0.029415
Madagascar	405.994	0.017446
Moldova	408.502	0.022496
Fiji	412.820	0.029412
Morocco	419.146	0.017703
Pakistan	423.031	0.021447
Bulgaria	424.688	0.039134
Tajikistan	427.698	0.006763
Haiti	430.548	0.019600
Russia	431.297	0.022210
Sudan	431.388	0.067047
Central African Republic	435.727	0.013200