International University - VNUHCM
School of Computer Science and Engineering

Midterm Examination

Date: 18/11/2023; Duration: 90 min

Laptops are allowed, but no internet connection.

SUBJECT: Data Science and Data Visualization (IT138IU)	
Approval Signature Full name: Nguyen Thi Thuy Loan	Lecturer: Signature Full name: Trần Thanh Tùng
Proctor 1 Signature Full name:	Proctor 2 Signature Full name:

STUDENT INFO
Student name: **Student ID:**

INSTRUCTIONS: the total of point is 100 (equivalent to 30% of the course)

1. *Purpose:*
   - Test your knowledge on data science and data visualization in the following topics:
     - visualization: human perception, color, data type, data encoding and design (CLO1)
     - d3.js: basic syntax, handling data, scaling, drawing basic shapes and texts (CLO2)
   - Examine your skill in
     - draw charts with effect in D3.js (CLO2)
     - analyze and evaluate charts (CLO3)

Student Name:.....
Student ID:.....

1. **(30pts)** List 3 limitations of human visual system. For each of them, give an example and describe its implication for designing a visualization.
2. **(40pts)** Redesign. COVID-19 is known to be very infectious in closed area like in an aircraft cabin. The Figure 1 shows a case of COVID-19 spreading in an aircraft. However, the visualization is not compatible with gray-scale printing.

   Your task is to

   a. (10pts) What is its data (items/entities)? types of data?

   b. (15pts) Redesign the visualization in such a way that could show **different groups** in a gray-scale printing.

   c. (15pts) Justify your design: the reason for your selection of marks, channels, and others.

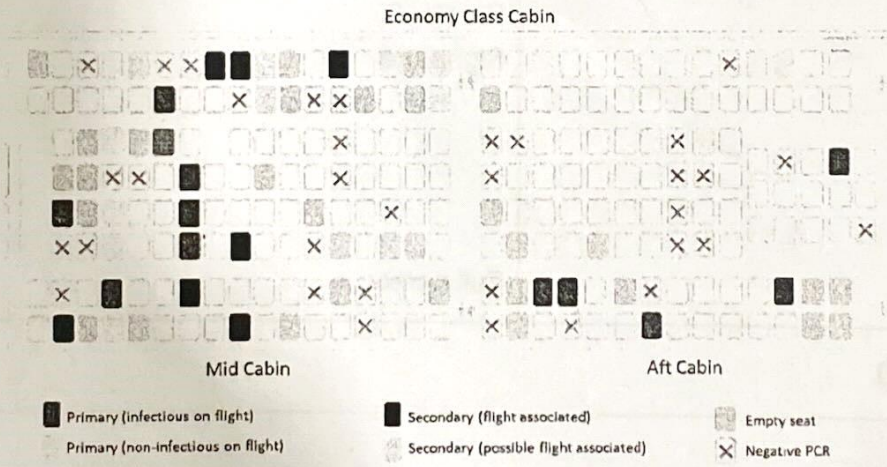

Figure 1 - COVID-19 patients in an aircraft cabin

3. (30pts) Create an HTML web page with the title "Midterm".

   a. Get data from https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv

   (in your code, you can refer to this source as "covid_global.csv")

   In the dataset,

   - "Province/State" and "Country/Region" are used as the key for each row.
   - The date is in US-format.

   Figure 2 is a sample from the dataset.

Student Name:.....
Student ID:.....

A	B	C	D	E	F	G	H	I	J
Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20
Northwest Territories	Canada	64.8255	-124.846	0	0	0	0	0	0
Nova Scotia	Canada	44.682	-63.7443	0	0	0	0	0	0
Nunavut	Canada	70.2998	-83.1076	0	0	0	0	0	0
Ontario	Canada	51.2538	-85.3232	0	0	0	0	1	1
Prince Edward Island	Canada	46.5107	-63.4168	0	0	0	0	0	0
Quebec	Canada	52.9399	-73.5491	0	0	0	0	0	0
Repatriated Travellers	Canada			0	0	0	0	0	0
Saskatchewan	Canada	52.9399	-106.451	0	0	0	0	0	0
Yukon	Canada	64.2823	-135	0	0	0	0	0	0
	Central African Republic	6.6111	20.9394	0	0	0	0	0	0
	Chad	15.4542	18.7322	0	0	0	0	0	0
	Chile	-35.675	-71.543	0	0	0	0	0	0
Anhui	China	31.8257	117.2264	1	9	15	39	60	70
Beijing	China	40.1824	116.4142	14	22	36	41	68	80
Chongqing	China	30.0572	107.874	5	9	27	57	75	110
Fujian	China	26.0789	117.9874	1	5	10	18	35	59
Gansu	China	35.7518	104.2861	0	2	2	4	7	14

Figure 2 - A sample from the dataset

b. (5pts) Write code to draw a horizontal bar chart to show COVID confirmed cases over the world on "04/04/2022". The chart must

R1. (5pts) have a fixed size (use scale to convert data),

R2. (5pts) have an axis with a title and ticks,

R3. (5pts) use Province/State and Country/Region as key/label for a row,

R4. (5pts) show only non-zero rows (Use filter function of arrays in JavaScript),

R5. (5pts) show value in the bar

Hint:

- Use rowConverter
- Use parseInt, parseFloat to convert strings to numbers.
- Filter function of arrays in JavaScript.

```
newDataSet = dataset.filter(d => d["04/04/2022"] > 0);
```

-- The end --