International University - VNUHCM
School of Computer Science and Engineering

Midterm Examination

Date: 30/03/2023; Duration: 90 min

SUBJECT: Data Science and Data Visualization (IT138IU)	
Approval Signature Full name: Nguyễn Thị Thúy 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)
2. *Requirement:*
   - Write the answers and draw models CLEAN and TIDY directly in the exam paper

HCMC National University
International University

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

1. **(35pts)** The chart in Figure 1 is to illustrate the relationship of human's organs, indicators and methylation.
   a. (10pts) What is its data? types of data?
   b. (5pts) How did the author use marks and channels to encode the data?
   c. (10pts) The visualization has many problems, describe them.
   d. (10pts) Redesign it.

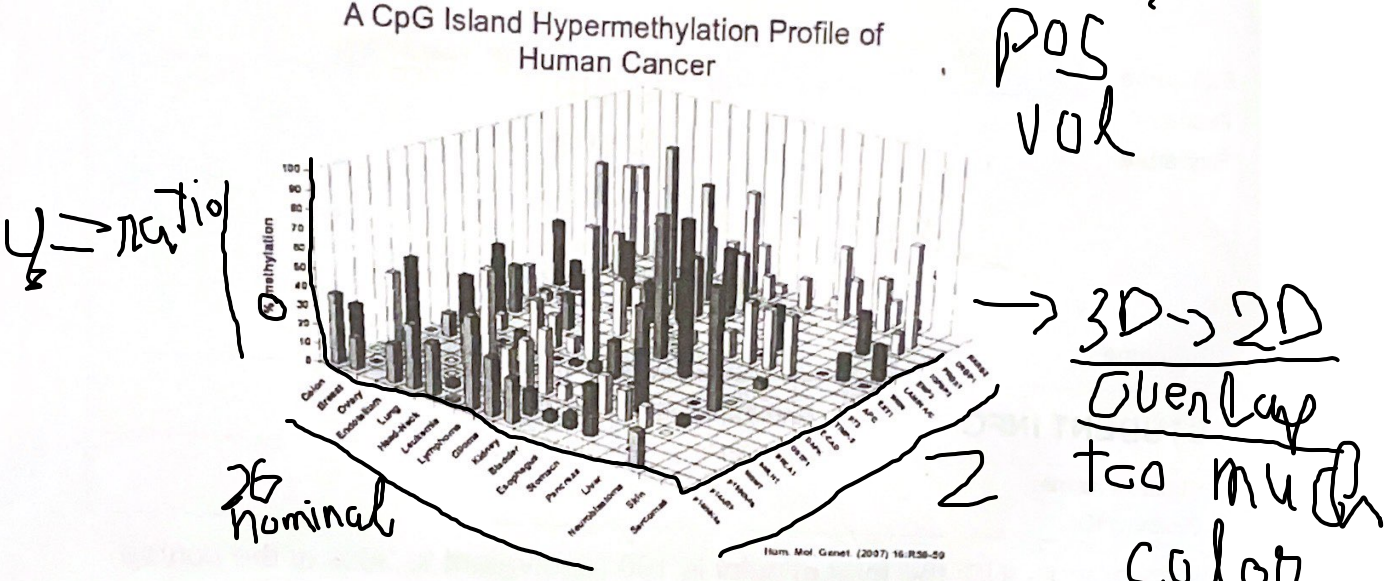

Figure 1

2. **(35pts)** Design a visualization to show students in a class.
   Each student has a GPA (10pts), a major (10pts), a group's ID (10pts), and a seat number.
   Show all attributes of each student and show links between all members of a group in the classroom arrangement given in Figure 2.

Index	GPA	Major	GroupID	Seat
1	80	IT	1	2
2	70	CS	1	3
3	90	CS	1	4
4	100	IT	1	12
5	60	DS	2	15
6	70	DS	2	16
7	90	DS	2	25
8	90	CS	2	26
9	50	CS	3	9
10	60	IT	3	10

Index	GPA	Major	GroupID	Seat
11	85	CS	3	11
12	90	DS	4	31
13	70	DS	4	32
14	70	DS	4	39
15	80	CS	5	5
16	90	IT	5	6
17	100	IT	5	7
18	70	DS	6	34
19	90	CS	6	35
20	60	CS	6	36

Classroom arrangement

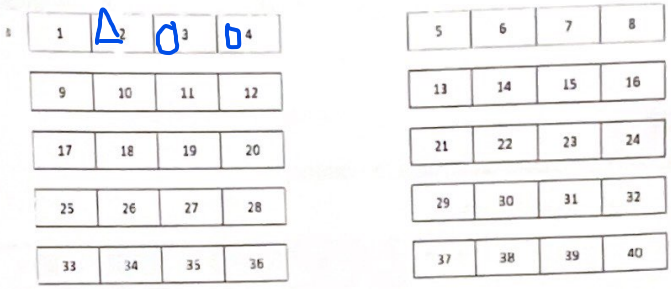

Figure 2

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 3 is a sample from the dataset.

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	6	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 3 - A sample from the dataset

   b. (5pts) Write code to draw a horizontal bar chart to show COVID confirmed cases over the world on "03/03/2023". The chart must
      R1. (5pts) have a fixed size (use scale to convert data)
      R2. (5pts) have an axis with a title and ticks

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

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["03/03/2022"] > 0);
```

-- The end --