

# Indian Institute of Technology, Madras - BS in Data Science and Applications

## Notations :

- 1.Options shown in green color and with ✓ icon are correct.
- 2.Options shown in red color and with ✗ icon are incorrect.

Question Paper Name :	IIT M DEGREE AN2 EXAM QPE2 16 JULY 2023
Subject Name :	2023 July: IIT M DEGREE AN2 EXAM QPE2
Creation Date :	2023-07-10 17:54:46
Duration :	120
Total Marks :	575
Display Marks:	Yes
Share Answer Key With Delivery Engine :	Yes
Actual Answer Key :	Yes
Calculator :	Scientific
Magnifying Glass Required? :	No
Ruler Required? :	No
Eraser Required? :	No
Scratch Pad Required? :	No
Rough Sketch/Notepad Required? :	No
Protractor Required? :	No
Show Watermark on Console? :	Yes
Highlighter :	No
Auto Save on Console?	Yes
Change Font Color :	No
Change Background Color :	No

**Options :**

6406531933180. ✔ Both **x** and **y** are always reachable from **S** in the residual graph.

6406531933181. ✖ It is possible that both (S, x) and (S, y) are saturated with respect to **f**.

6406531933182. ✔ Both (x, T) and (y, T) are saturated with respect to **f**.

6406531933183. ✖ At most one of (x, T) and (y, T) can be saturated with respect to **f**.

## Data Viz

Section Id :	64065339134
Section Number :	10
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	5
Number of Questions to be attempted :	5
Section Marks :	100
Display Number Panel :	Yes
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065382982
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 186 Question Id : 640653579027 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0  
Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "DEGREE LEVEL : DATA VISUALIZATION DESIGN (PEN AND PAPER EXAM)"

ARE YOU SURE YOU HAVE TO WRITE EXAM FOR THIS SUBJECT?

CROSS CHECK YOUR HALL TICKET TO CONFIRM THE SUBJECTS TO BE WRITTEN.

(IF IT IS NOT THE CORRECT SUBJECT, PLS CHECK THE SECTION AT THE TOP FOR THE SUBJECTS REGISTERED BY YOU)

Options :

6406531933184. ✓ YES

6406531933185. ✗ NO

Sub-Section Number :	2
Sub-Section Id :	64065382983
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 187 Question Id : 640653579028 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 18

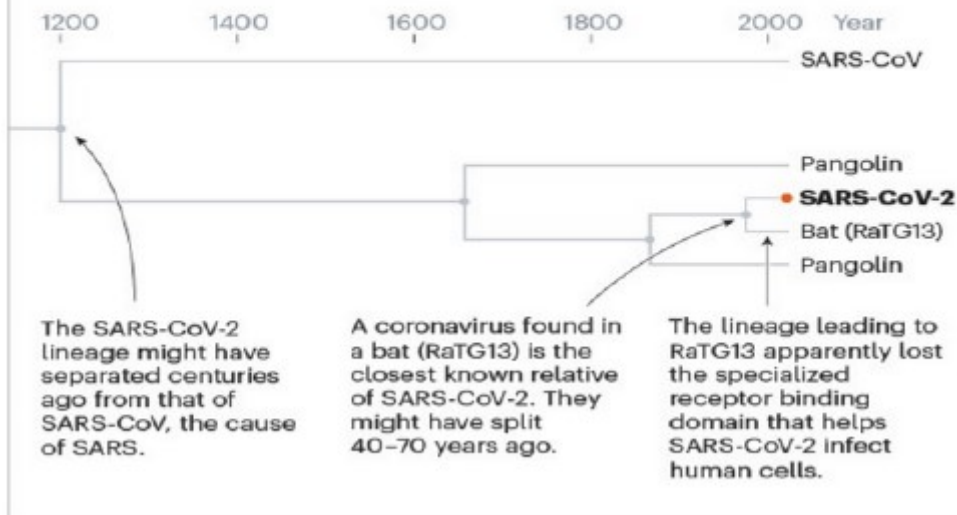
Question Label : Multiple Choice Question

Part A (Multiple Choice Questions): Each question has four choices, out of which only one is the correct answer.(12×3=36 marks)

1.What type of data is represented by the visualization shown in the figure given below?[MCQ: 3 marks]

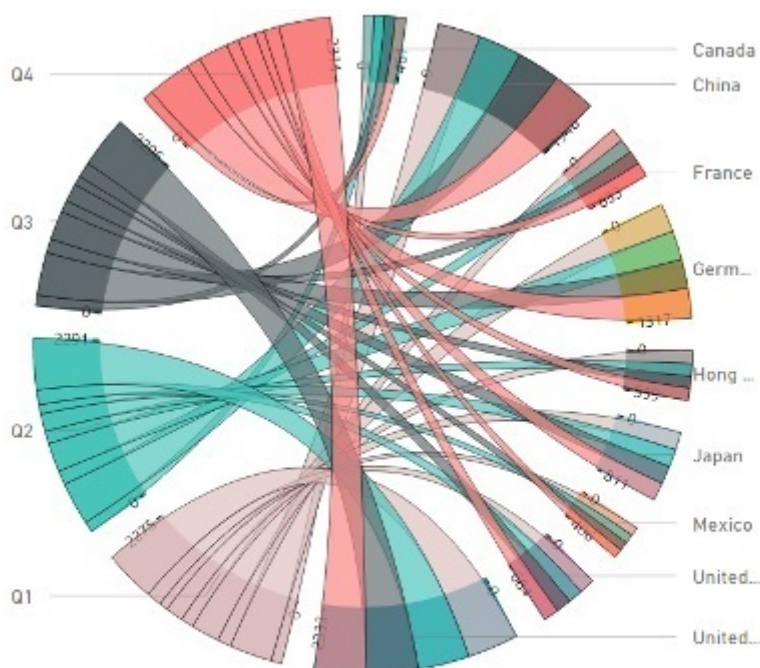
## A FAMILY OF KILLERS

Researchers are trying to work out the evolutionary relationships between the new coronavirus (SARS-CoV-2) and others in a large family of related viruses (a simplified version is shown below). A study, which has not yet been peer reviewed, suggests a possible scenario for how SARS-CoV-2 evolved.



- A. Temporal
- B. Hierarchical
- C. Network-based
- D. Geospatial

2.What does the diagram in the visualization shown below analytical represents?[MCQ: 3 marks]



Source: GitHub

- A. Correlation
- B. Distribution
- C. Flow
- D. Ranking

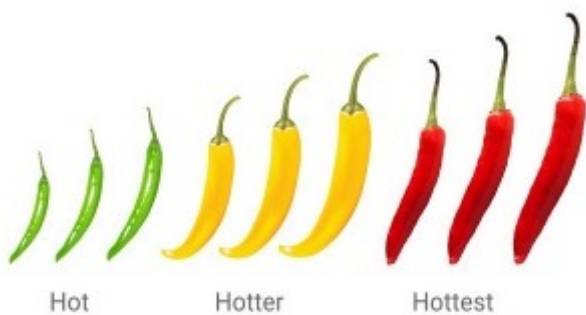
3. Which of the following statements is true for categorical data? [MCQ: 3 marks]

- A. The measurement scale 'Very satisfied, Satisfied, Neutral, Unsatisfied, Very unsatisfied' in a customer survey is an example of categorical data.
- B. Categorical data represents all its data variables quantitatively.
- C. Categorical data can be analyzed using descriptive statistics.
- D. Categorical data are non-numeric by nature.

4. Classification of blood into various blood groups represents which type of data? [MCQ: 3 marks]

- A. Ratio data
- B. Ordinal data
- C. Categorical data
- D. Interval data

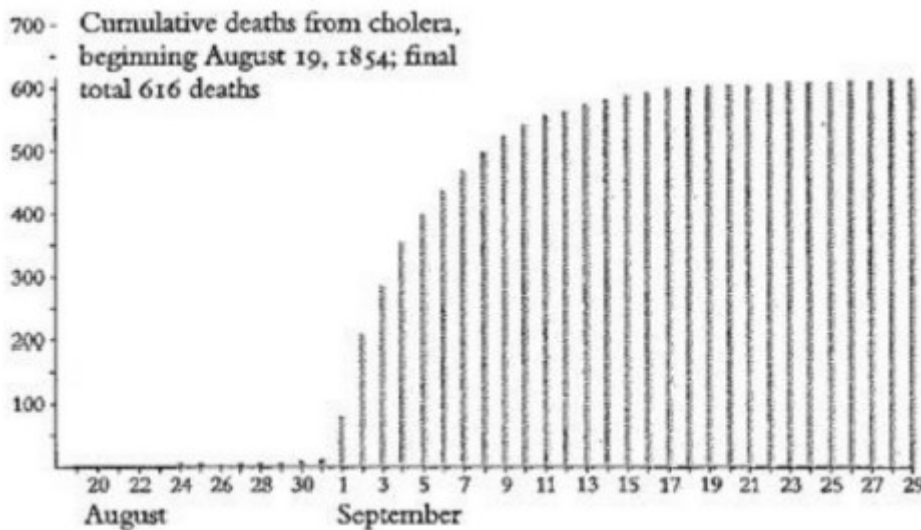
5. What type of data is represented by the visualization given below? [MCQ: 3 marks]



Source: <https://www.questionpro.com/blog>

- A. Ratio data
- B. Interval data
- C. Categorical data
- D. Ordinal data

6. With the idea of creating a casual theory on how the disease Cholera spread in the city, what was the major conclusion drawn by Dr. Snow based on the visualization given below? [MCQ: 3 marks]



- A. The deaths due to Cholera are coming to a saturation
- B. A sudden spike in deaths during the 1st week of September 1854
- C. The number of cumulative deaths in the months of August and September.
- D. The time series data is a poor match for the task.

Options :

6406531933186. ✓ I have written answers on the answer sheets

6406531933187. ✗ Not applicable

Sub-Section Number : 3

Sub-Section Id : 64065382984

Question Shuffling Allowed : No

Is Section Default? : null

Question Number : 188 Question Id : 640653579029 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 18

Question Label : Multiple Choice Question

7.Which of the following statements is correct for data visualization given in the figure below?[MCQ: 3 marks]

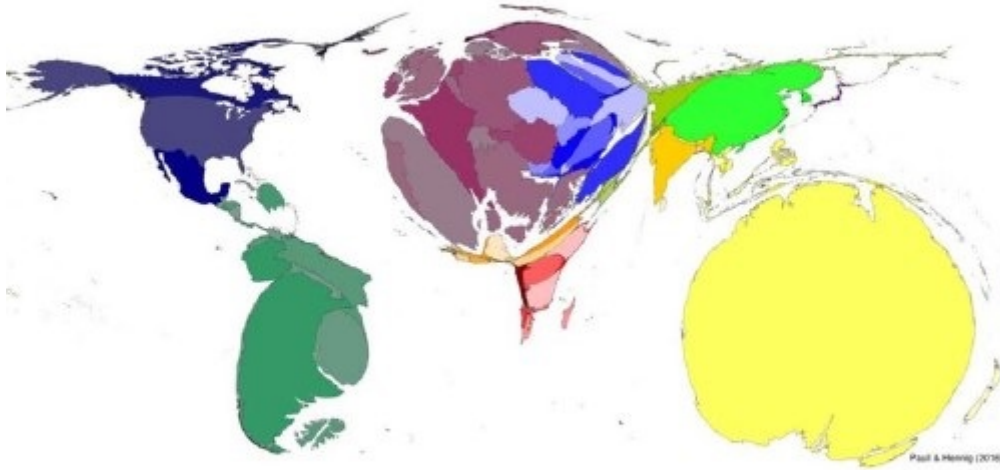


Source: <https://www.jotform.com>

- A. Correct assignment of color to both the data type and semantics.
- B. Incorrect assignment of color to both the data type and semantics.
- C. Incorrect data type assignment but correct semantic assignment of color.
- D. Correct data type assignment but incorrect semantic assignment of color.

8.This following map of the world is a \_\_\_\_\_.[MCQ: 3 marks]

## The World of Organic Agriculture (hectares)



Density equalising map

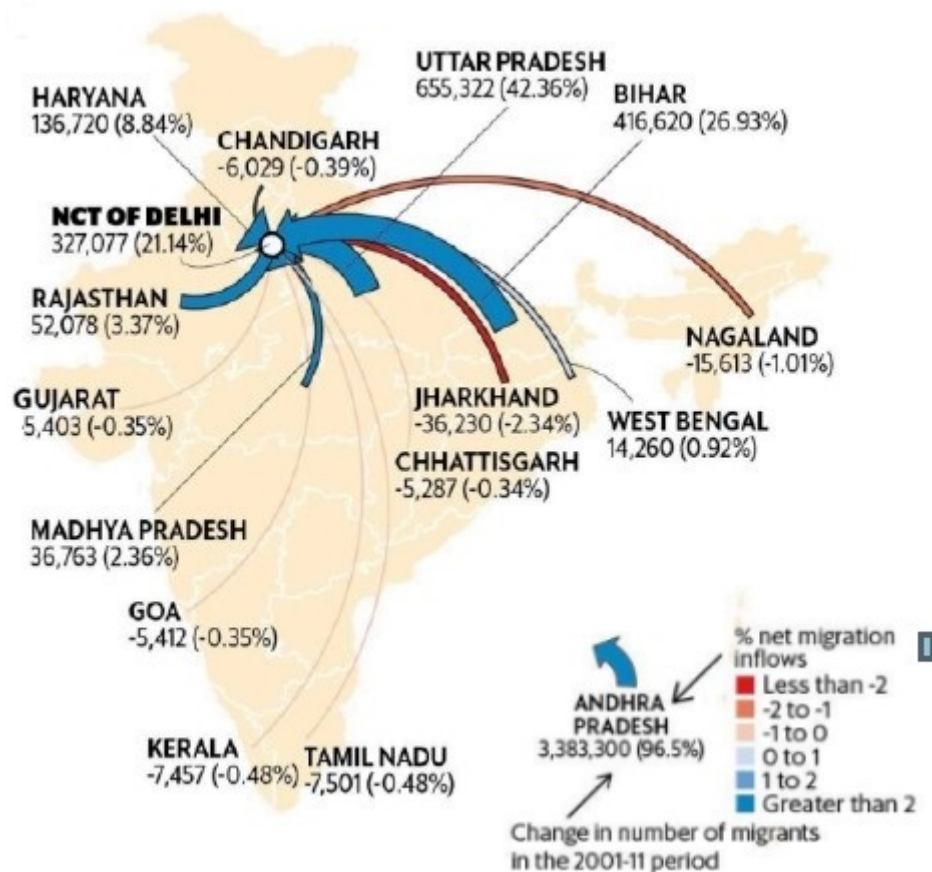
Paul & Hennig, 2016

Source: Map-of-World-Organic-Agriculture

- A. Isopleth
- B. Choropleth
- C. Cartogram
- D. Isochrone

9.The following map shows the migration of people from various cities to Delhi. It is an example of \_\_\_\_\_.[MCQ: 3 marks]





- A. Choropleth
- B. Cartogram
- C. Isopleth
- D. Flow maps

(Common Data for Q10, Q11) You are working as a data analyst for a social media platform. Your team is analyzing user engagement data for different types of content. The key insights are as follows:

- Video posts receive the highest average engagement rate of 12%.
- Infographic posts have an average engagement rate of 8%.
- Text-based posts have the lowest average engagement rate of 5%.

10. Based on the key insights, which visualization technique would be most effective in comparing the average engagement rates of different types of content?[MCQ: 3 marks]

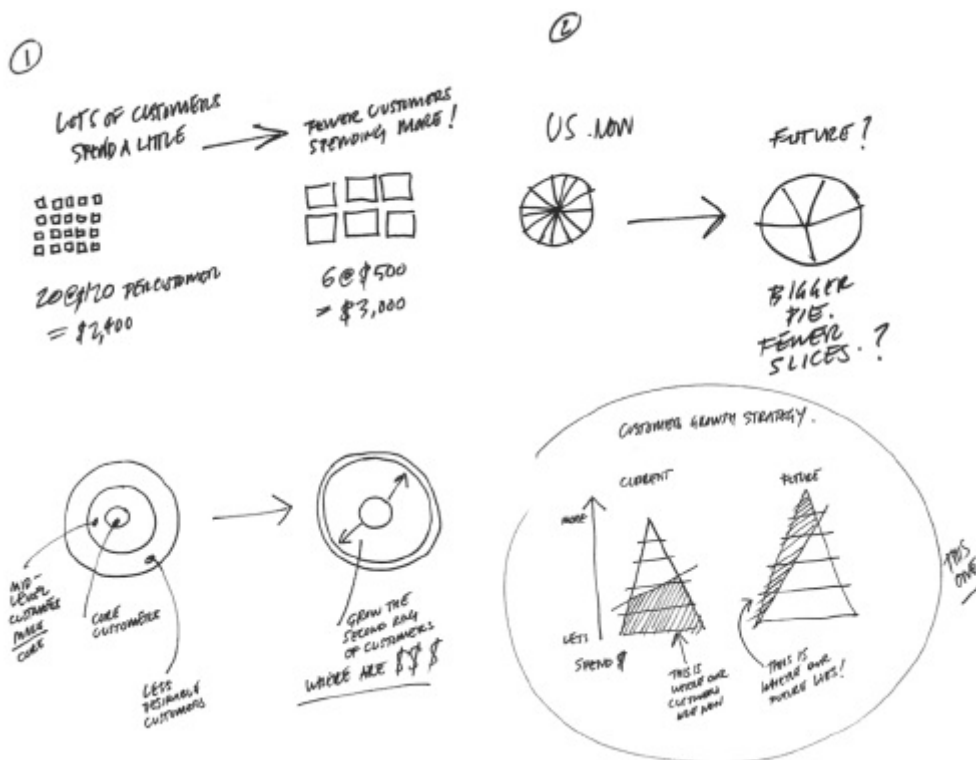
- A. Stacked bar chart
- B. Line graph

- C. Scatter plot
- D. Bubble chart
- E. None of these

11. To visualize the distribution of user engagement rates across different content types, which visualization technique would be most suitable for displaying both the average engagement rates and the relative proportions of each content type?[MCQ: 3 marks]

- A. A Sankey diagram representing the flow of engagement rates between different content types.
- B. A 3D pie chart displaying the average engagement rates as different slice depths.
- C. A scatter plot with engagement rates on the x-axis, content types on the y-axis, and markers representing the relative proportions.
- D. A heat map showing engagement rates as color gradients for each content type.

12. Which of the following best describes the given picture scenario?[MCQ: 3 marks]



- A. Visual confirmation
- B. Visual discovery
- C. Idea generation
- D. Idea illustration

**Options :**

6406531933188. ✔ I have written answers on the answer sheets

6406531933189. ✖ Not applicable

<b>Sub-Section Number :</b>	4
<b>Sub-Section Id :</b>	64065382985
<b>Question Shuffling Allowed :</b>	No
<b>Is Section Default? :</b>	null

**Question Number : 189 Question Id : 640653579030 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 32**

Question Label : Multiple Choice Question

Part B – **Multiple Select Question:** Each question has four choices, out of which one or more than one choice(s) is the correct answer. **(8×4=32 marks)**

1. Classification of the data based on eye color as one of its attributes is an example of which type of data?[MSQ: 4 marks]

- A. Ordinal Data
- B. Interval Data
- C. Categorical Data
- D. Nominal Data

2. Which of the following statements regarding the given visualizations is/are correct?[MSQ: 4 marks]

- A. Funnel charts can be used to represent the correlation between two variables.
- B. Sankey diagrams are generally used to represent the flow of information.
- C. Ridgelines are used to visualize the distribution of data in various categories.
- D. Choropleths display data in the chronological order of their occurrence.

3. Which of the following statements regarding the types of data visualization is/are correct?[MSQ: 4 marks]

- A. Idea illustrations are the abstraction of complex ideas into simple design conventions
- B. Decision trees are an example of Idea illustration
- C. Visual exploration confirms if something that is suspected to happen is true.
- D. Visual explorations are open-ended data-driven visualizations that are intended to produce insights that otherwise can not be inferred

4. Which of the following statements is/are true for data visualization?[MSQ: 4 marks]

- A. Data visualizations may involve a large amount of data.
- B. They are aesthetically rich.
- C. Data visualizations can display results based on real-time data.
- D. Data visualizations can be recreated with a different set of data.

5. Which of the following statements is/ are true about data type in the context of measurement scales?[MSQ: 4 marks]

- A. The data type determines the level of detail or precision of information captured by a variable.
- B. Data type describes the interpretation and meaning of the numbers assigned to variables.
- C. The data type is primarily concerned with the physical representation of data in a computer system.
- D. Data type influences the mathematical operations and statistical analyses that can be applied to a variable.

6. When considering the measurement of weight as a data type, Which of the following characteristics accurately describes weight as a data type? Select all that apply.[MSQ: 4 marks]

- A. Weight can be represented as interval data.
- B. Weight can be characterized as categorical data.
- C. Weight can be categorized as ordinal data.
- D. Weight can be categorized as ratio data.

7. Why is it necessary to understand data characteristics?[MSQ: 4 marks]

- A. Data characteristics provide insights into the underlying data distribution.
- B. Data characteristics help in predicting future outcomes based on historical data.
- C. Data characteristics help in selecting the appropriate data visualization techniques.
- D. Knowledge of data characteristics enhances creativity in data analysis.

8. Which of the following is/are true about Isopleths and Choropleths? (Select all that apply.)(MSQ: 4 marks]

- A. Isopleths are commonly used to depict population density or the distribution of continuous data.
- B. Isopleths are best suited for visualizing discrete and categorical data.
- C. A choropleth map visualizes a measure across a geographic area by displaying the range of variability through the use of color.
- D. Both refer to the same map technique.

**Options :**

6406531933190. ✓ I have written answers on the answer sheets

6406531933191. ✖ Not applicable

Sub-Section Number :	5
Sub-Section Id :	64065382986
Question Shuffling Allowed :	No
Is Section Default? :	null

**Question Number : 190 Question Id : 640653579031 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 32**

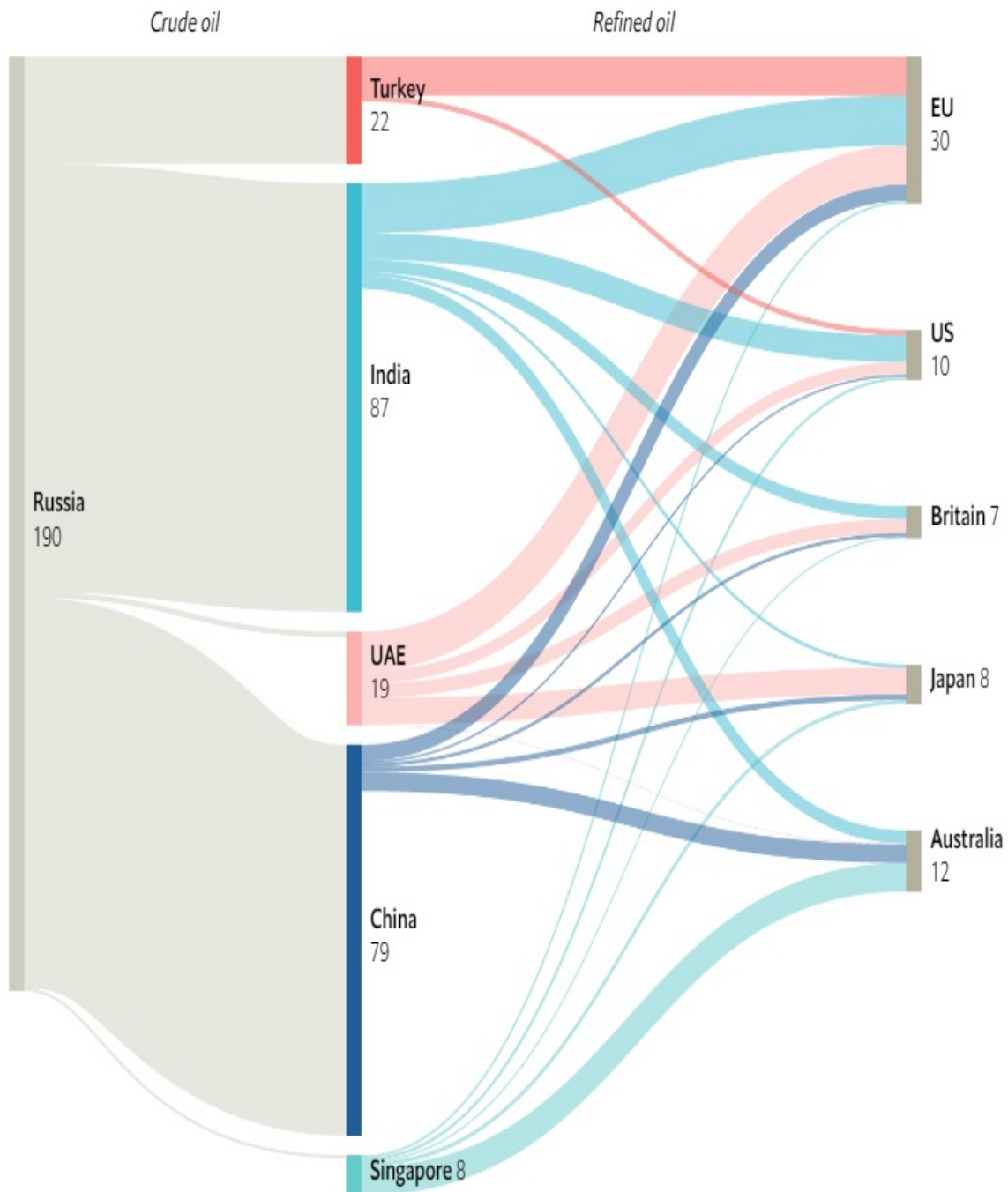
Question Label : Multiple Choice Question

Part C – **Short Answer Questions:** Write 100-150 words answers to each question. Answer using pen & paper using the paper provided to you. **(4×8=32 marks)**

1. Critically analyze this visualization. Critique its effectiveness, and suggest at least one improvement.

### Seaborne oil trade, million tonnes

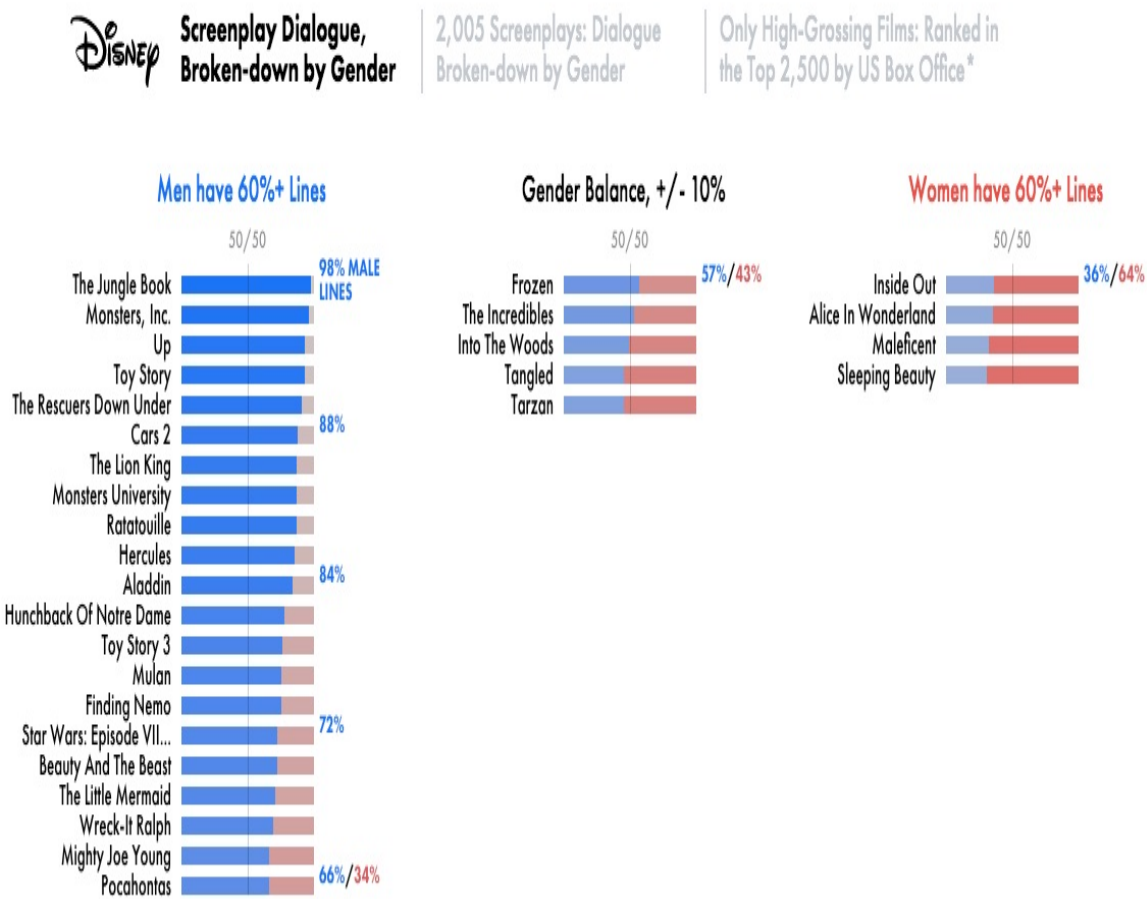
February 24th 2022-June 18th 2023



2. Consider the provided dataset containing information about housing properties in various cities across India. Each property is described by variables such as price, location, area, number of bedrooms, number of bathrooms, and the presence of a garden. Analyze the dataset and answer the following questions related to data types:
- Identify the categorical variables in the dataset.
  - Determine the continuous variables present in the dataset.
  - Point out any ordinal variables that can be found in the dataset.
  - Give 2 additional data that you can derive from it and how you will derive them.

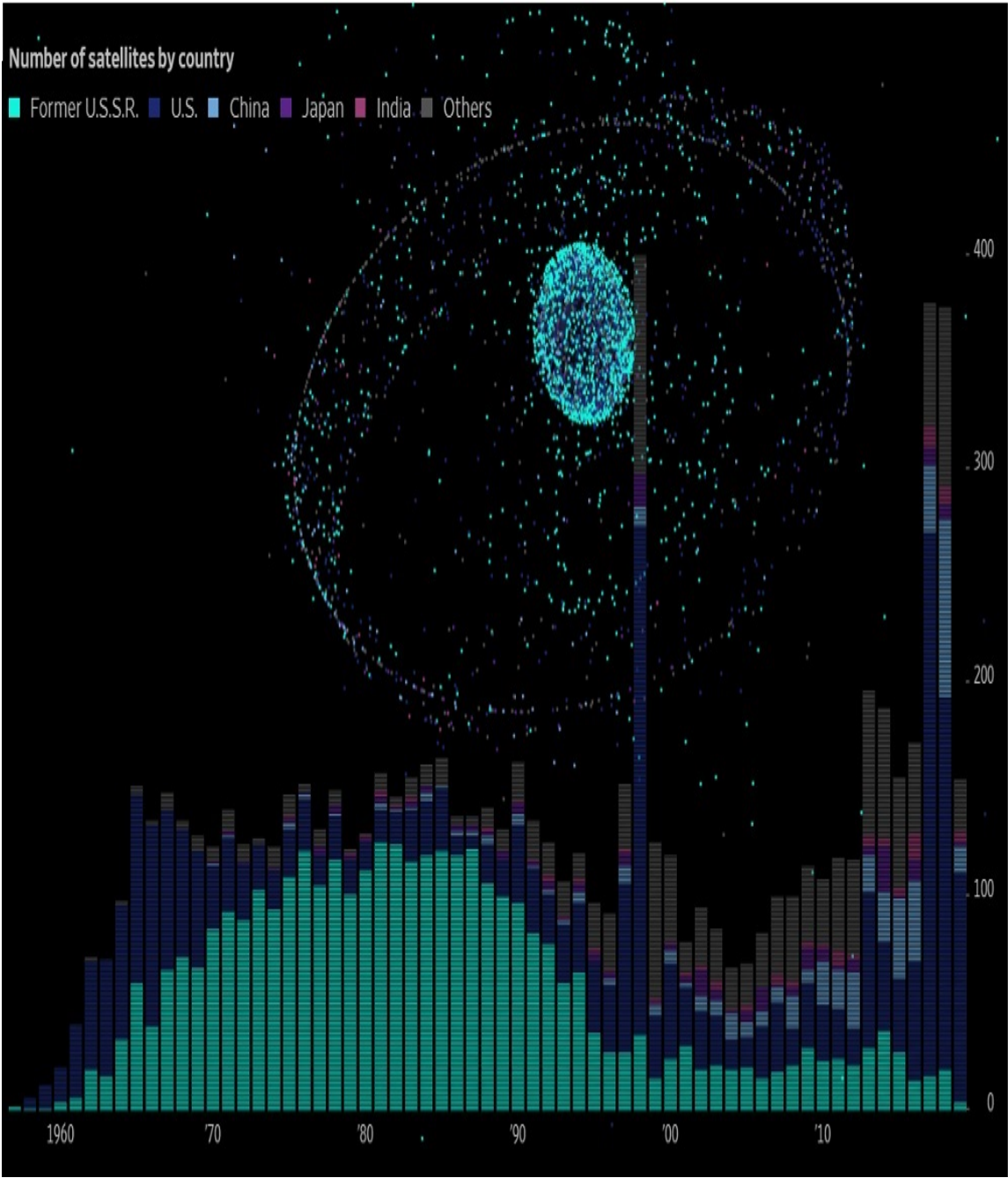
Property_ID	Price	Location	Area	Bedrooms	Bathrooms	Garden
1	250000	Saket, New Delhi	1500	3	2	1
2	350000	Koregaon Park, Pune	2000	4	3	0
3	180000	Marine Drive, Mumbai	1200	2	1	1
4	400000	Indiranagar, Bangalore	2500	4	3	0
5	280000	Salt Lake City, Kolkata	1600	3	2	1
6	320000	Kondapur, Hyderabad	1800	3	2	1
7	210000	Anna Nagar, Chennai	1300	2	1	0
8	390000	Koramangala, Bangalore	2400	4	3	1
9	260000	Malad, Mumbai	1400	3	2	0
10	370000	Gachibowli, Hyderabad	2200	4	3	1

3. Critically analyze this visualization. Critique its effectiveness, and suggest at least one improvement.





4. Critically analyze this visualization. Critique its effectiveness, and suggest at least one improvement.



Options :

6406531933192. ✓ I have written answers on the answer sheets

6406531933193. ✗ Not applicable

## Sem2 Statistics2

Section Id :

64065339135

Section Number :

11