MONASH University

FIT5147 Data exploration and visualisation TP2 2020

	, 5 April 2020, 4:28 PM
State	Finished
Completed on	Sunday, 5 April 2020, 5:58 PM
Time taken	1 hour 30 mins
Marks	33.75/65.00
Grade	**7.79** out of 15.00 (**52**%)

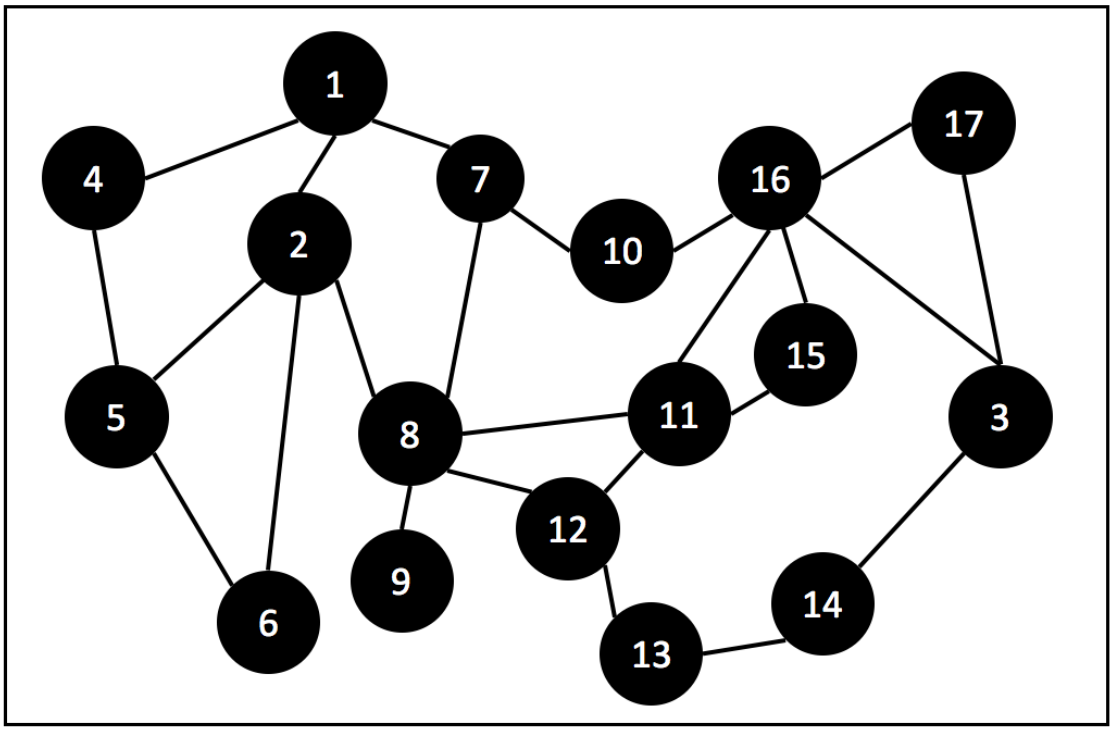
Print friendly format

Question **1**

Incorrect

Mark 0.00 out of 1.50

What is the graph theoretic distance between nodes **2** and **3** in the following network?



Answer: 5 ✖

Review **Network analysis and visualisation** in Module 4

The correct answer is: 4

Question **2**

What is a bowtie plot used on a data map for?

Select one:

a. Show uncertainty in observed magnitude ✖ Incorrect. Review Activity: How to design a data map topic of Module 3

b. Show compass bearing

c. Socioeconomic data

d. Show magnetic declination

e. Show uncertainty in observed direction

Your answer is incorrect.

The correct answer is: Show uncertainty in observed direction

Question 3

Incorrect

Mark 0.00 out of 1.00

What is iconic memory?

Select one:

a. Long term memory of spatial location ✖ Incorrect. Review The human visual system topic of Module 5

b. Short term auditory memory

c. Short term memory of the image on the retina

d. Long term auditory memory

e. Long term memory of geometric shapes

Your answer is incorrect.

The correct answer is: Short term memory of the image on the retina

Question 4

Correct

Mark 1.00 out of 1.00

Which of the following is an R package for visualising network data?

Select one:

a. tm

b. visGraph

c. graphDraw

d. igraph ✔ Correct.

e. NPL

Your answer is correct.

The correct answer is: igraph

Question 5

Incorrect

Mark 0.00 out of 1.00

What kind of data map would be best used to show average rainfall across Australia?

Select one:

a. Choropleth map ✖ Incorrect. Review Activity: How to design a data map topic of Module 3

b. Proportional symbol map

c. Cartogram

d. Flow map

e. Isarithmic map

Your answer is incorrect.

The correct answer is: Isarithmic map

Question 6

Complete

Mark 0.50 out of 4.50

The human visual system has 3 main levels or stages. List them down and briefly explain what each stage is about.

Shapes mainly to make distinction of two dimensional shapes
Colours to make distinction between colours
Long term memory to rely on historical cues helping with the dimensions, depth, etc.

1. Parallel processing to extract low level properties: colour, texture, lines and movement
2. Rapid serial processing divides the visual field into regions of similar colour or texture and achieves "proto-object" recognition of surfaces, boundaries and relative depth. This is driven both top-down by visual attention and bottom-up by low level properties.
3. Visual working memory: object recognition & attention, this is under conscious control

Comment:
Partially correct.

Question 7

Correct

Mark 1.00 out of 1.00

There are two parts of a CSS rule: selector and properties. The selector selects the HTML elements to operate. Which of the following allows the selector to select all level 1 headings?

Select one:

- a. h1 ✔ Correct. Selects all level 1 headings
- b. #H1
- c. #l1
- d. *
- e. p em

Your answer is correct.

The correct answer is: h1

Question 8

Incorrect

Mark 0.00 out of 1.00

Scalable Vector Graphics (SVG) is the web vector graphics format. The coordinate system of SVG canvas has the origin (0,0) at:

Select one:

- a. top left corner
- b. bottom right corner
- c. top right corner
- d. bottom left corner ✖ Incorrect. Review Activity 4 & 5 in Module 5.

Your answer is incorrect.

The correct answer is: top left corner

Question 9

Complete

Mark 3.00 out of 5.00

One of the major focuses in network analytics is network cohesion.

1) Explain what network cohension is about?

2) What are the possible ways to measure network cohension? List at least 4.

Network cohesion regers to how tightly the nodes are positioned in respect to other nodes. The desnsity being M /N*(N-1)/2 measures the cohesion.

Other measures such as degree showing the number of neighbours, betweenness showing the strength of how much a node is between other nodes and closeness how close a node is to other nodes are other aspects of cohesion.

1) Network cohension investigates how tightly connected the graph is.

2) There are few ways to measure network cohension, e.g.,

- Measure the number of highly connected sub-graphs, e.g., number of cliques and bi-cliques in the network
- Density, i.e., the frequency of actual edges in the network compared with the total number of possible edges (N(N-1)/2 where N is the number of nodes).
- The frequency of actual edges compared with the total number of edges.
- The clustering coefficient measures the frequency in which sub-graphs with 3 connected nodes are actually cliques, i.e., there is an edge between all 3 nodes.
- Vertex and edge connectivity measure how many nodes or edges need to be removed from a connected network before it is disconnected.

Comment:

Question **10**

Incorrect

Mark 0.00 out of 1.50

Which of the following is/are best to shown with a choropleth map? Select all that apply.

Select one or more:

- a. Flight routes across the globe
- b. Location of milkbars in Melbourne ✖
  Incorrect. Use a proportional symbol map

  Review **How to design a data map**.
- c. Population density of each Melbourne suburb. ✔ Correct
- d. Average rainfall across Victoria ✖
  Incorrect. Use a isarithmic map.

  Review **How to design a data map**.
- e. Migration between Europe and Australia

Your answer is incorrect.

The correct answer is: Population density of each Melbourne suburb.

Question **11**

Correct

Mark 1.00 out of 1.00

What is the most commonly used text analysis package in R called?

Select one:

- a. **itext**
- b. **igraph**
- c. **NLP**
- d. **tm** ✔ Correct
- e. **textm**

Your answer is correct.

The correct answer is: **tm**

Question 12

Correct

Mark 1.00 out of 1.00

Which of the following events fires after the whole web page is loaded:

Select one:

- a. window.mousedown
- b. window.onpageshow
- c. window.onload ✔ Correct. The onload event occurs when an object (i.e., the whole web page) has been loaded.
- d. window.ononline

Your answer is correct.

The correct answer is: window.onload

Question 13

Incorrect

Mark 0.00 out of 1.00

Which of the following is **not** a true statement about the human eye?

Select one:

- a. If you are focusing on a display at arms length, you see an area roughly 10cm by 10cm in great detail.
- b. When looking at a graphic, the eye frequently changes focus.
- c. Detailed vision relies more on cones than rods
- d. It contains about 20 times more rods than cones.
- e. The fovea is responsible for detailed vision. ✖
  Incorrect. This part of the eye is where most of the cones are.

  Review **The Human Visual System**.

Your answer is incorrect.

The correct answer is: If you are focusing on a display at arms length, you see an area roughly 10cm by 10cm in great detail.

Question 14

Correct

Mark 1.00 out of 1.00

Which of the following best describes tf-idf?

Select one:

- a. Measures the importance of a word in a document corpus, taking into account the word's distribution in the corpus.
- b. Measures the frequency of a word in a document, taking into account the word length.
- c. Measures the importance of a document in a document corpus, taking into account the number of words it shares with other documents.
- d. Measures the frequency of a word in a document, taking into account the word's frequency in the rest of the corpus. ✔ Correct

Your answer is correct.

The correct answer is: Measures the frequency of a word in a document, taking into account the word's frequency in the rest of the corpus.

Question 15

Correct

Mark 1.50 out of 1.50

The words "SUV car ute BMW Kia" occur in three documents with the following frequencies.

Which word will have the highest weight when computing tf-idf?

Document	SUV	car	ute	BMW	Kia
A	3	4	0	0	1

B	2	3	3	1	1
C	0	1	0	1	1

Answer: ute ✔

Review **Exploring text and document collections** topic of Module 4

The correct answer is: ute

Question **16**

Correct

Mark 1.00 out of 1.00

CSS is designed to describe how the contents of a web page should be presented. CSS has to be stored in a file separated from all the HTML files. This statement is:

Select one:

- True
- False ✔

Correct.

The correct answer is 'False'.

Question **17**

Which of the following is/are correct about Mercator projection? Select all that apply.

Select one or more:

- a. Regions drawn on the map locally retain their shape. ✔
  Correct
- b. It becomes the standard projection and it is used in most maps.
- c. It allows easily navigate between very distant locations ✔
  Correct.
- d. The vertical scale always remains the same as the horizontal scale, as it smoothly increases the vertical scale as the latitude increases. ✔
  Correct
- e. The scale is consistent across the map.

Your answer is partially correct.

You have correctly selected 3.
The correct answers are: It allows easily navigate between very distant locations, The vertical scale always remains the same as the horizontal scale, as it smoothly increases the vertical scale as the latitude increases., Regions drawn on the map locally retain their shape., It becomes the standard projection and it is used in most maps.

Question **18**

Correct

Mark 1.00 out of 1.00

As a data scientist, when you are designing a graphic to communicate what you have found to other people, which of the following is **least** important?

Select one:

- a. The design encourages the perceptual system to group data in meaningful ways
- b. Ensuring the design is visually novel ✔ Correct
- c. Ensuring the reader can easily discriminate and compare the data
- d. Taking into account limitations in working memory
- e. Directing attention to the most relevant elements in the graphic

Your answer is correct.

The correct answer is: Ensuring the design is visually novel

Question 19

Correct

Mark 1.00 out of 1.00

What kind of data map would be best used to show population density in different regions?

Select one:

- a. Proportional symbol map
- b. Choropleth map ✔ Correct
- c. Isarithmic map
- d. Nautical chart
- e. Flow map

Your answer is correct.

The correct answer is: Choropleth map

Question 20

What are the three opponent colour channels in the human visual system? Select all that apply.

Select one or more:

- a. red to green
- b. green to blue ✖ Incorrect. Review The human visual system topic of Module 5
- c. red to yellow ✖ Incorrect. Review The human visual system topic of Module 5
- d. purple to brown
- e. black to white ✔ Correct
- f. blue to yellow
- g. red to blue

Your answer is partially correct.

You have correctly selected 1.
The correct answers are: red to green, blue to yellow, black to white

Question 21

Correct

Mark 1.00 out of 1.00

What does ggmap's function **geocode** do?

Select one:

- a. Return the latitude and longitude of a geo-tagged image
- b. Map latitude and longitude to the visual or geometric attribute of a **facet**
- c. Create a dot density data map
- d. Map latitude and longitude to the visual or geometric attribute of a **geom**
- e. Return the latitude and longitude of a location or address ✔ Correct

Your answer is correct.

The correct answer is: Return the latitude and longitude of a location or address

Question 22

Incorrect

Mark 0.00 out of

Which of the following is/are a visualisation that one would use for textual documents? Select all that apply.

Select one or more:

1.50

- [x] a. word cloud ✔ Correct
- [x] b. word tree ✔ Correct
- [ ] c. word umbrella
- [ ] d. SOM
- [x] e. word graph ✖ Incorrect. This term is not a known visualisation.
  Review Exploring text and document collections topic of Module 4
- [ ] f. stream graph

Your answer is incorrect.

The correct answers are: word tree, word cloud, stream graph, SOM

Question **23**

Correct

Mark 1.50 out of 1.50

Which of the following is/are **not** usually an assumption made when two humans communicate? Select all that apply.

Select one or more:

- [ ] a. That the sentences will have a logical flow and connection.
- [ ] b. The information takes into account the listener's expectations and background
- [ ] c. That important information will be highlighted
- [x] d. The information will be entertaining ✔ Correct. Unfortunately we cannot assume this.
- [x] e. That the listener will enjoy the conversation ✔ Correct. Unfortunately we cannot assume this.
- [ ] f. The information provided is relevant

Your answer is correct.

The correct answers are: That the listener will enjoy the conversation, The information will be entertaining

Question **24**

Correct

Mark 1.50 out of 1.50

Which of the following is/are GIS vector data exchange format/s? Select all that apply.

Select one or more:

- [x] a. Keyhole Markup Language ✔
  Correct.
- [ ] b. GeoTIFF
- [x] c. Geography Markup Language ✔
  Correct.
- [x] d. Esri shapefile ✔
  Correct.
- [x] e. GeoJSON ✔
  Correct.

Your answer is correct.

The correct answers are: Keyhole Markup Language, Esri shapefile, Geography Markup Language, GeoJSON

Question **25**

Correct

Which of the following is/are a grouping principle of the visual system? Select all that apply.

Correct

Mark 1.50 out of 1.50

Select one or more:

- [x] a. Closure ✔
  Correct.
- [ ] b. Disjoint region
- [x] c. Proximity ✔
  Correct.
- [x] d. Shared fate ✔
  Correct
- [x] e. Similarity ✔
  Correct

Your answer is correct.

The correct answers are: Proximity, Similarity, Shared fate, Closure

Question 26

Correct

Mark 1.00 out of 1.00

What does **betweenness centrality** measure for a node in a network?

Select one:

- a. The average closeness to all nodes in the graph
- b. The number of important neighbours it has
- c. The number of neighbours it has
- d. The weighted average degree of the neighbour
- e. The number of shortest paths that pass through the node ✔ Correct

Your answer is correct.

The correct answer is: The number of shortest paths that pass through the node

Question 27

Incorrect

Mark 0.00 out of 1.00

Which of these is the most effective visual channel for showing values of an ordered attribute?

Select one:

- a. Length
- b. Colour (luminance)
- c. Tilt/angle
- d. Area ✖
  Incorrect - OK, but not the best

  Review **The Human Visual System** in Module 5.
- e. Colour (saturation)

Your answer is incorrect.

The correct answer is: Length

Question 28

Once you have decided the number of classes for drawing a choropleth map, you will have to decide on how to

Complete

Mark 0.00 out of 4.50

Once you have decided the number of classes for drawing a choropleth map, you will have to decide on how to partition the observations into these classes, i.e., compute the class intervals (the minimum and maximum value for each class). The class intervals should not overlap and should contain all observations.

What are the common techniques to do the partition? List at least 3 and explain briefly how each of them works.

Use measures of qaulity, quantity and relevance. Quality to showthe strength. Quantity to show the volume and relevance hoe strong or effective indicators they are.

- Equal Interval: Simply divide the data range into equally sized intervals.
- Equal Frequency: Compute class intervals so that they have the same number of observations. Thus the intervals correspond to quartiles. If there are 4 classes then the classes correspond to the quartiles.
- Jenks Optimisation: This chooses the classes so as to minimise the variance within each class while maximising the variance between the classes. An iterative algorithm is used to do this.

Comment:
Incorrect answer.

Question **29**

Incorrect

Mark 0.00 out of 1.00

Out of the following, which is the **most** effective way to encode quantitative data?

Select one:

- a. Length
- b. Volume ✖ Incorrect. Review The human visual system topic of Module 5
- c. Tilt/angle
- d. Colour saturation
- e. Area

Your answer is incorrect.

The correct answer is: Length

Question **30**

Which of the following is/are a measure of a node's centrality in a network? Select all that apply.

Select one or more:

- a. Density
- b. Status
- c. Degree ✔ Correct
- d. Betweenness ✔ Correct
- e. Closeness ✔ Correct
- f. Longest distance from other nodes in the network
- g. Median popularity

Your answer is partially correct.

You have correctly selected 3.
The correct answers are: Degree, Closeness, Betweenness, Status

Question 31

Correct

Mark 1.00 out of 1.00

What kind of map is this?

Select one:

- a. Non-contiguous cartogram ✔ Correct
- b. Isarithmic map
- c. Prism map
- d. Dasymetric
- e. Contiguous cartogram

Your answer is correct.

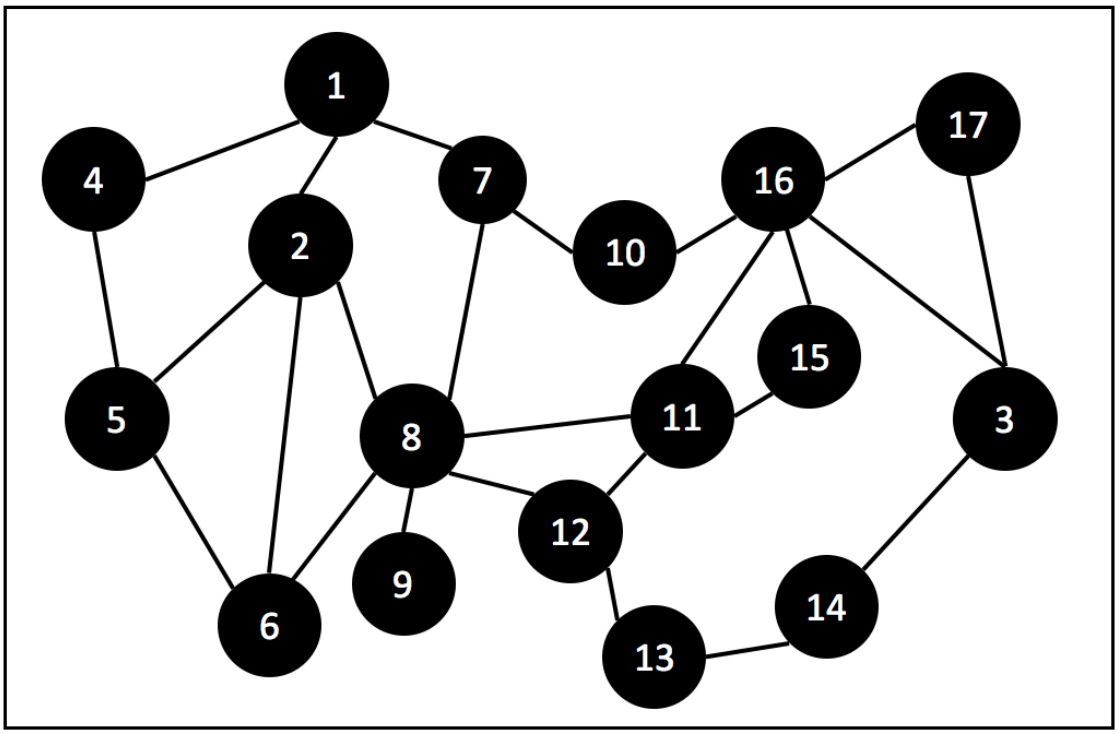
The correct answer is: Non-contiguous cartogram

Question 32

Correct

Mark 1.50 out of 1.50

In the following network, which node has the highest degree? (Write down the integer label of the node)



Answer: 8 ✔

Review **Network analysis and visualisation** in Module 4

The correct answer is: 8

Question 33
Correct
Mark 1.00 out of 1.00

Which of the following would be best suited for exploring how spatial distribution of wind farms is related to their generating capacity?

Select one:

- a. Isarithmic map
- b. Choropleth map.
- c. Flow map
- d. Contiguous cartogram.
- e. Proportional symbol map ✔ Correct. You could put a symbol for each wind farm and make the size proportional to the generating capacity.

Your answer is correct.

The correct answer is: Proportional symbol map

Question 34
Incorrect
Mark 0.00 out of 1.50

Which of the following is true about the R package ggmap? Select all that apply.

Select one or more:

- a. It is designed to work with ggplot
- b. It only supports vector GIS data ✖ Incorrect -- it mainly supports GIS raster data. Review **Data maps with R and ggmap.**
- c. You can set the level of zoom ✔ Correct.
- d. You can specify an address on the map ✔ Correct.
- e. It can download maps from Google
- f. The coordinate system is fixed to the Mercator projection.

Your answer is incorrect.

The correct answers are: It can download maps from Google, You can specify an address on the map, You can set the level of zoom, It is designed to work with ggplot, The coordinate system is fixed to the Mercator projection.

Question 35
Correct
Mark 1.00 out of 1.00

What does D3.js mean?

Select one:

- a. It is a JavaScript framework to display D3 models
- b. It is a JavaScript library for creating and manipulating documents based on data ✔ Correct.
- c. It is a JavaScript library for changing native objects to D3 objects
- d. It is Node.js to parse a server's data to objects with D3 features

Your answer is correct.

The correct answer is: It is a JavaScript library for creating and manipulating documents based on data

Question 36
Correct
Mark 1.00 out of 1.00

What is topic modelling based on?

Select one:

- a. Crowd-sourced annotations on text documents.
- b. Analysis of how news feed headlines change over time
- c. The frequency of unusual words in the document corpus.
- d. Finding clusters of words that often occur close to one another in a document corpus. ✔ Correct

Your answer is correct.

The correct answer is: Finding clusters of words that often occur close to one another in a document corpus.

Question 37
Correct
Mark 1.00 out of 1.00

You have data items with two categorical attributes and two quantitative attributes. You choose to map the quantitative attributes onto the x and y position of a symbol. What would be the best choice for encoding the two categorical attributes?

Select one:

- a. Colour hue and colour luminance of the symbol
- b. Colour and shape of the symbol ✔ Correct
- c. Size and orientation of the symbol
- d. Width and height of the symbol
- e. Motion and size of the symbol

Your answer is correct.

The correct answer is: Colour and shape of the symbol

Question 38
Incorrect
Mark 0.00 out of 1.00

Which of the following depth clues is most important for judging depth?

Select one:

- a. Depth blurring
- b. Familiar size
- c. Occlusion
- d. Linear perspective
- e. Cast shadows ✖ Incorrect. Review The human visual system topic of Module 5

Your answer is incorrect.

The correct answer is: Occlusion

Question 39
Incorrect
Mark 0.00 out of 1.00

What is a Sankey diagram used to visualise?

Select one:

- a. Magnitude of flow between processes
- b. Clusters in a network ✖
  Incorrect.

  Review **Network analysis and visualisation**.
- c. Topic maps
- d. Spatial distribution of a continuous variable

e. Set containment

Your answer is incorrect.

The correct answer is: Magnitude of flow between processes

Question **40**

Incorrect

Mark 0.00 out of 1.00

Which statement is true about the first stage of human visual processing?

Select one:

a. It is subconscious

b. Recognition occurs sequentially ✖ Incorrect. Review The human visual system topic of Module 5

c. It divides the visual field into regions of similar colour or texture

d. It recognises objects

e. It performs proto-object recognition

Your answer is incorrect.

The correct answer is: It is subconscious

Question **41**

Correct

Mark 1.00 out of 1.00

What is the name of the following graphic?

Select one:

a. Choropleth map

b. Prism map ✔ Correct

c. Cartogram

d. Proportional symbol map

e. Isarithmic map

Your answer is correct.

The correct answer is: Prism map

Question **42**

Incorrect

Mark 0.00 out of 1.00

Which of the following Gestalt laws best explains how we see clusters in scatter plots?

Select one:

a. Closure ✖ Incorrect. Review The human visual system topic of Module 5

b. Continuity

c. Proximity

d. Shared fate

e. Similarity

Your answer is incorrect.

The correct answer is: Proximity

Question **43**

Complete

Mark 4.00 out of 5.00

Pattern perception is used to extract objects from low-level visual features. There are many ways in which humans automatically organise such elements.

List at least 5 ways and briefly explain each of them.

By proximity: how close certain objects are together
By similarity: how similar one object is to another
Connectedness: How related and connected one object is to others
Symetry: Whether some objeects appear in any symetrical form
Continuity: How continuously an object appears.

- *Proximity*: Elements that are close together form groups.
- *Similarity*: Elements that are similar in some way such as colour or shape form a group.
- *Connectedness*: Connection by lines is a powerful way of grouping elements.
- *Continuity*: We tend to group regions and lines to form smooth and continuous shapes.
- *Symmetry*: We are good at recognising bilateral symmetry, especially around a horizontal or vertical axis and group the symmetric lines together to form an object.
- *Closure and common region*: we like to see closed contours and will mentally extend lines to close them. Being “inside” a closed contour is a very powerful grouping principle.
- *Shared fate*: Elements that move together are grouped together.

Comment:
Incomplete

Question **44**

Incorrect

Mark 0.00 out of 1.00

Tree maps are mainly used to show hierarchically organised qualitative data. This is statement is:

Select one:

- True ✖
- False

Incorrect. Tree maps are widely used to show hierarchically organised quantitative data, with the size of the rectangle being proportional to the data value.

Tree maps are widely used to show hierarchically organised quantitative data, with the size of the rectangle being proportional to the data value.

The correct answer is 'False'.

◄ Quiz 2 - Information Sheet

Jump to...

Assignment 3 | Data exploration and visualisation project final report - Description ►