2.1 Different types of data

Luis Talavera

2022-12-08

Exercises

2.1 Categorical/quantitative difference
a. Explain the difference between categorical and quantitative variables.
b. Give an example of each.

Ans.

a. The possible values of a categorical variable are limited to an specific set of values. Meanwhile the possible values of a quantitative variable could be discrete quantities or continuous quantities.

b. A categorical variable could be the weather of an specific day (cloudy, sunny, rainy). A quantitative variable could be the temperature of an specific day (57, 75, 80, etc).

2.2 Common types of cancer in 2012 Of all cancer cases around the world in 2012, 13% had lung cancer, 11.9% had breast cancer, 9.7% had colorectal cancer, 7.9% had prostate cancer, 6.8% had stomach cancer and 50.7% had other types of cancer www.wcrf.org/int/cancer-facts-figures/ worldwide-data. Is the variable "cancer type" categorical or quantitative? Explain

Ans. The cancer type variable is categorical due to it can only take a value from an specific set of values (lung, breast, colorectal, prostate, stomach and other). The numbers are only the percentages of occurrence of each type of cancer.

2.3 Classify the variable type Classify each of the following variables as categorical or quantitative.
a. The number of social media accounts you have (Facebook, Twitter, LinkedIn, Instagram, etc.)
b. Preferred soccer team
c. Choice of smartphone model to buy
d. Distance (in kilometers) of commute to work

Ans.

a. Quantitative
b. Categorical
c. Categorical
d. Quantitative

2.4 Categorical or quantitative? Identify each of the following variables as either categorical or quantitative.
a. Choice of diet (vegan, vegetarian, neither)
b. Time spent shopping online per week
c. Ownership of a tablet (yes, no)
d. Number of siblings

Ans.

a. Categorical
b. Quantitative
c. Categorical
d. Quantitative

2.5 Discrete/continuous
a. Explain the difference between a discrete variable and a continuous variable.
b. Give an example of each type.

Ans.

a. A discrete variable refers to something countable. At the other side a continuous variable has a continuum of infinitely many possible values
b. Discrete variable: Number of siblings. Continuous variable: 200 m running time

2.6 Discrete or continuous? Identify each of the following variables as continuous or discrete.
a. The upload speed of an Internet connection
b. The number of apps installed on a tablet
c. The height of a tree
d. The number of emails you send in a day

Ans.

a. Continuous
b. Discrete
c. Continuous
d. Discrete

2.7 Discrete or continuous 2 Repeat the previous exercise for the following:
a. The total playing time of a CD
b. The number of courses for which a student has received credit
c. The amount of money in your pocket (Hint: You could regard a number such as $12.75 as 1275 in terms of "the number of cents.")
d. The distance between where you live and your statistics classroom, when you measure it precisely with values such as 0.5 miles, 2.4 miles, 5.38 miles

Ans.

a. Continuous
b. Discrete
c. Continuous
d. Continuous

2.8 Number of children In the 2008 General Social Survey (GSS), 2020 respondents answered the question, "How many children have you ever had?" The results were

No. Children	0	1	2	3	4	5	6	7	8+	Total
Count	521	323	524	344	160	77	30	19	22	2020

a. Is the variable, number of children, categorical or quantitative?

b. Is the variable, number of children, discrete or continuous?

c. Add proportions and percentages to this frequency table.

Ans.

a. Quantitative
b. Discrete
c.

No. Children	Count	Proportion	Percentage
0	521	0.258	25.8
1	323	0.16	16
2	524	0.259	25.9
3	344	0.17	17
4	160	0.079	7.9
5	77	0.038	3.8
6	30	0.015	1.5
7	19	0.009	0.9
8+	22	0.011	1.1
Total	2020	0.999	99.9

2.9 Fatal Shark Attacks Few of the shark attacks listed in Table 2.1 are fatal. Overall, 63 fatal shark attacks were recorded in the ISAF from 2004 to 2013, with 2 reported in Florida, 2 in Hawaii, 4 in California, 15 in Australia, 13 in South Africa, 6 in Réunion Island, 4 in Brazil, and 6 in the Bahamas. The rest occurred in other regions.
a. Construct the frequency table for the regions of the reported fatal shark attacks.
b. Identify the modal category.
c. Describe the distribution of fatal shark attacks across the regions.

Ans. The modal category is Australia and the places where most fatal attacks occur are Australia and South Africa

Region	Frequency	Proportion	Percentage
Florida	2	0.03	3
Hawaii	2	0.03	3
California	4	0.06	6
Australia	15	0.24	24
South Africa	13	0.21	21
Réunon Island	6	0.1	10
Brazil	4	0.06	6
Bahamas	6	0.1	10
Other	11	0.17	17
Total	63	1	100