Here is a data matrix displaying some of the data collected in a survey designed to study smoking habits of Colorado residents.

	sex	age	weight	monthlyIncome	smokeAmount
1	Female	42	120	under $2000	12 cig/day
2	Male	33	150	$2000 to $3000	0 cig/day
3	Female	27	110	$5000 to $6000	0 cig/day
⋮	⋮	⋮	⋮	⋮	⋮

1. In this data set, what type of variable is `monthlyIncome`?

(A) Numerical, continuous

(B) Numerical, discrete

(C) Categorical, ordinal

(D) Categorical, nominal

Here is a data matrix displaying some of the data collected in a survey designed to study smoking habits of Colorado residents.

	sex	age	weight	monthlyIncome	smokeAmount
1	Female	42	120	under $2000	12 cig/day
2	Male	33	150	$2000 to $3000	0 cig/day
3	Female	27	110	$5000 to $6000	0 cig/day
⋮	⋮	⋮	⋮	⋮	⋮

1. In this data set, what type of variable is `monthlyIncome`?

(A) Numerical, continuous

(B) Numerical, discrete

(C) Categorical, ordinal

(D) Categorical, nominal

Follow-up. Classify all of the variables.

A political scientist is interested in the effect of government type on economic development. She wants to use a sample of 30 countries evenly represented among the Americas, Europe, Asia, and Africa to conduct her analysis.

2. What type of sampling strategy should she use to ensure that countries are selected from each region of the world?

(A) Simple random sampling

(B) Cluster sampling

(C) Stratified sampling

(D) None of the above

Participants in a study report whether or not they used a smart phone before bed and then rate their quality of sleep as either "very poor," "poor," "average," "good," or "very good." A statistician analyzes the data and determines that the two variables are associated.

3. Is it reasonable to conclude that smart phone usage before bed causes a change in quality of sleep?

(A) Yes

(B) No

Participants in a study report whether or not they used a smart phone before bed and then rate their quality of sleep as either "very poor," "poor," "average," "good," or "very good." A statistician analyzes the data and determines that the two variables are associated.

3. Is it reasonable to conclude that smart phone usage before bed causes a change in quality of sleep?

(A) Yes

(B) No

Follow-up. How might you design a study that can be used as evidence for a causal relationship?

Students at an elementary school are given a questionnaire that they are asked to return after their parents have completed it. One of the questions asked is, "Do you find that your work schedule makes it difficult for you to spend time with your kids after school?" Of the parents who replied, 85% said "no."

4. Is it reasonable to conclude that a great majority of parents have no difficulty spending time with their kids after school?

(A) Yes

(B) No

A study compared percent of internet users and life expectancy at birth in 208 countries and found a positive association.

5. What is a possible confounding variable that might explain this relationship?

Here is some data about the success/failure of space launches by different types of launching agencies.

	1957–1999		2000–2018	
	Failure	Success	Failure	Success
Private	13	295	10	562
State	281	3751	33	711
Startup	-	-	5	65

6. Describe a data matrix that one might use to create this summary table. How many variables are involved? What type is each variable? How might you use this table to figure out the total number of observations in the data matrix?

7. You would like to determine if your classmates prefer the taste of regular Coke or Diet Coke. Outline a design for this study.