Feedback — Week 3 Quiz

Help

You submitted this quiz on **Sat 6 Sep 2014 1:30 PM PDT**. You got a score of **20.00** out of **20.00**.

Question 1

We take a random sample of individuals in a population and identify whether they smoke and if they have cancer. We observe that there is a strong relationship between whether a person in the sample smoked or not and whether they have lung cancer. We claim that the smoking is related to lung cancer in the larger population. We explain we think that the reason for this relationship is because cigarette smoke contains known carcinogens such as arsenic and benzene, which make cells in the lungs become cancerous.

Your Answer		Score	Explanation
○ This is an example of an descriptive data analysis.			
○ This is an example of a predictive data analysis.			
○ This is an example of a mechanistic data analysis.			
◉ This is an example of an inferential data analysis.	✔	4.00	
Total		4.00 / 4.00	

Question 2

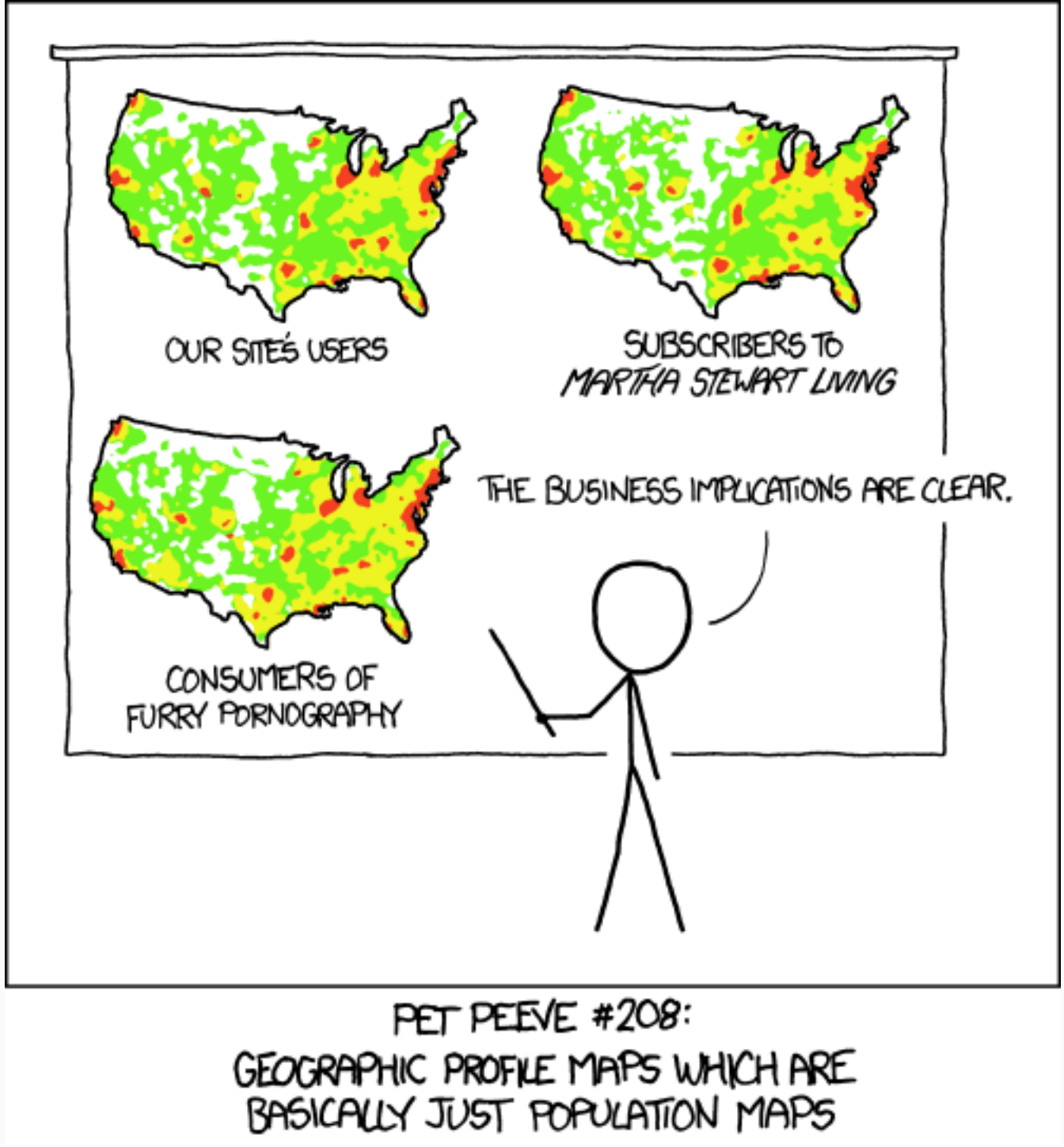
What is the most important thing in Data Science?

Your Answer	Score	Explanation

The question you are trying to answer.	✔	4.00
Hacking skills.		
Working with large data sets.		
Statistical inference.		
Total		4.00 / 4.00

Question 3

If the goal of a study was to relate Martha Stewart Living Subscribers to Our Site's Users based on the number of people that lived in each region of the US, what would be the potential problem?



Link to image

Source: http://xkcd.com/1138/

Your Answer		Score	Explanation
(•) There would be confounding because the number of people that live in an area is related to both Martha Stewart Living Subscribers and Our Site's Users.	✔	4.00	
( ) We would be performing inference on the relationship between Martha Stewart Living Subscribers and Our Site's Users.			
( ) We wouldn't be able to estimate the variability in Martha Stewart Living Subscribers.			
( ) We wouldn't be able to calculate descriptive statistics on Our			

Site's Users.

Total	4.00 / 4.00

Question 4

What is an experimental design tool that can be used to address variables that may be confounders at the design phase of an experiment?

Your Answer		Score	Explanation
○ Taking a convenience sample.			
○ Only using non-confounding variables.			
◉ Randomization	✔	4.00	
○ Using regression models.			
Total		4.00 / 4.00	

Question 5

What is the reason behind the explosion of interest in big data?

Your Answer	Score	Explanation
○ We recently discovered ways to use data to make predictions.		
○ There have been massive improvements in statistical analysis techniques.		
○ There have been massive improvements in machine learning algorithms.		

The price and difficulty of collecting and storing data has dramatically dropped.	✔	4.00
Total		4.00 / 4.00