

# The Data Scientists Toolbox - Week 2 Quiz

## 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.

- This is an example of a causal data analysis.
- This is an example of a predictive data analysis.
- This is an example of an inferential data analysis.
- This is an example of a mechanistic data analysis.

## Question 2

What is the most important thing in Data Science?

- Statistical inference.
- Machine learning and prediction.
- The data.
- The question you are trying to answer.

# The Data Scientists Toolbox - Week 2 Quiz

## Question 1

Which of the following commands will create a directory called data in your current working directory?

- `mkdir data`
- `pwd ../data`
- `pwd data`
- `mkdir /Users/data`

## Question 2

Which of the following will initiate a git repository locally?

- (i) `git merge origin master`
- (ii) `git push`
- (iii) `git boom`
- (iv) `git init`

## Question 3

Suppose you have forked a repository called datascientist on Github but it isn't on your local computer yet. Which of the following is the command to bring the directory to your local computer? (For this question assume that your user name is username)

- (i) `git pull datascientist master`
- (ii) `git init`
- (iii) `git push origin master`
- (iv) `git clone https://github.com/username/datascientist.git`

## Question 4

Which of the following will create a markdown document with a secondary heading saying "Data Science Specialization" and an unordered list with the following for bullet points: Uses R, Nine courses, Goes from raw data to data products

```
*h2 Data Science Specialization

* Uses R
* Nine courses
* Goes from raw data to data products
```

```
## Data Science Specialization

li Uses R
li Nine courses
li Goes from raw data to data products
```

```
*** Data Science Specialization

* Uses R
* Nine courses
* Goes from raw data to data products
```

```
### Data Science Specialization
```

```
* Uses R
* Nine courses
* Goes from raw data to data products
```

```
## Data Science Specialization

* Uses R
* Nine courses
* Goes from raw data to data products
```

## 0.1 Question 5

Install and load the KernSmooth R package. What does the copyright message say?

- (i) Copyright M. P. Wand 1990-2009
- (ii) Copyright M. P. Wand 1997-2009
- (iii) Copyright Coursera 2009-2013
- (iv) Copyright Matthew Wand 1997-2009