Close

Week 1 Quiz

5 questions

*Passed*

*Passed, but unverified*

*Failed*

**/** points earned (%)

Quiz passed!

You haven't passed yet. You need at least  to pass.

Review the material and try again!  You have 3 attempts every 8 hours.

Identity-verification is pending: We will manually review your photo within a few days.

[Back to Week weekNumberReview Related Lesson](https://www.coursera.org/learn/data-cleaning/home/week/undefined)

**Be Recognized for Your Achievements.** "Course Certificates give you the recognition you need to get the job, the material gives you the skills to do the job. It makes you look more valuable because you are more valuable." - Peter B., USA, Software Developer

[**Showcase Your Accomplishment! Earn Your Course Certificate!**](https://www.coursera.org/certificate/data-cleaning)

loading   https://d3njjcbhbojbot.cloudfront.net/web/images/icons/loading.gif

Load Error!

1  
point

1.

The American Community Survey distributes downloadable data about United States communities. Download the 2006 microdata survey about housing for the state of Idaho using download.file() from here:

<https://d396qusza40orc.cloudfront.net/getdata%2Fdata%2Fss06hid.csv>

and load the data into R. The code book, describing the variable names is here:

<https://d396qusza40orc.cloudfront.net/getdata%2Fdata%2FPUMSDataDict06.pdf>

How many properties are worth $1,000,000 or more?



47



31



164



53

1  
point

2.

Use the data you loaded from Question 1. Consider the variable FES in the code book. Which of the "tidy data" principles does this variable violate?



Tidy data has one observation per row.



Each tidy data table contains information about only one type of observation.



Tidy data has one variable per column.



Each variable in a tidy data set has been transformed to be interpretable.

1  
point

2.

Use the data you loaded from Question 1. Consider the variable FES in the code book. Which of the "tidy data" principles does this variable violate?



Numeric values in tidy data can not represent categories.



Each variable in a tidy data set has been transformed to be interpretable.



Tidy data has one variable per column.



Tidy data has variable values that are internally consistent.

Use the data you loaded from Question 1. Consider the variable FES in the code book. Which of the "tidy data" principles does this variable violate?



Tidy data has variable values that are internally consistent.



Each variable in a tidy data set has been transformed to be interpretable.



Numeric values in tidy data can not represent categories.



Tidy data has one variable per column.

3.

Download the Excel spreadsheet on Natural Gas Aquisition Program here:

<https://d396qusza40orc.cloudfront.net/getdata%2Fdata%2FDATA.gov_NGAP.xlsx>

Read rows 18-23 and columns 7-15 into R and assign the result to a variable called:



1

dat

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

What is the value of:



1

sum(dat$Zip\*dat$Ext,na.rm=T)

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

(original data source: <http://catalog.data.gov/dataset/natural-gas-acquisition-program>)



36534720



154339



NA



33544718

1  
point

4.

Read the XML data on Baltimore restaurants from here:

<https://d396qusza40orc.cloudfront.net/getdata%2Fdata%2Frestaurants.xml>

How many restaurants have zipcode 21231?



127



100



156



17

1  
point

5.

The American Community Survey distributes downloadable data about United States communities. Download the 2006 microdata survey about housing for the state of Idaho using download.file() from here:

<https://d396qusza40orc.cloudfront.net/getdata%2Fdata%2Fss06pid.csv>

using the fread() command load the data into an R object



1

DT

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

The following are ways to calculate the average value of the variable



1

pwgtp15

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

broken down by sex. Using the data.table package, which will deliver the fastest user time?



DT[,mean(pwgtp15),by=SEX]



mean(DT$pwgtp15,by=DT$SEX)



tapply(DT$pwgtp15,DT$SEX,mean)



mean(DT[DT$SEX==1,]$pwgtp15); mean(DT[DT$SEX==2,]$pwgtp15)



rowMeans(DT)[DT$SEX==1]; rowMeans(DT)[DT$SEX==2]



sapply(split(DT$pwgtp15,DT$SEX),mean)



I, **Pradeep Sathyamurthy**, understand that submitting work that isn’t my own may result in permanent failure of this course or deactivation of my Coursera account. Learn more about Coursera’s Honor Code

4 questions unanswered

Submitting...Submit error! Please try again.Submit QuizSubmit Quiz

You pressed Close

Closing will discard your work so far.   
Are you sure you want to close this quiz?

Stay on this page Close Quiz

--> <img height="1" width="1" style="display:none" src="https://www.facebook.com/tr?id=946401778754875&amp;ev=PageView&amp;noscript=1">

Confirm Navigation

Are you sure you want to leave this page?

Stay on this Page  Leave this Page

Confirm Navigation

Are you sure you want to leave this page?

Stay on this Page  Leave this Page