





Bookmarks

- ▶ Important Pre-Course Survey
- ▶ Contact Us
- ▶ How To Navigate the Course
- ▶ Discussion Board
- ▶ Office Hours
- ▶ Week 0: Introduction to Data (Optional Review)
- ▶ Week 1: Sampling
- ▶ Week 2: Hypothesis Testing (One Group Means)
- ▼ **Week 3: Hypothesis Testing (Two Group Means)**

Readings

Reading Check due
May 03, 2016 at 17:00
UTC 

Lecture Videos

Comprehension Check
due May 03, 2016 at
17:00 UTC 

Week 3: Hypothesis Testing (Two Group Means) > Pre-Lab > Prepare for the Analysis



Bookmark

Reflect on the Question

Analyze the Data

Draw Conclusions

Primary Research Questions

1. Who is happier at the beginning of the semester: under-classmen or upper-classmen?
2. Does student happiness change from the beginning of the semester to the end?

Breakdown Your Analysis

Let's break this analysis into its required steps:

Question 1: Independent t-test

1. Make a vector of happiness scores for each sample (under- and upper-classmen).
2. Generate histograms to check the Normality assumption.
3. Run an independent t-test.
4. Interpret the results.


Question 2: Dependent t-test

1. Make a vector of difference scores for student happiness from the beginning to the end of semester.
2. Generate a histogram of the difference scores to check the Normality assumption.
3. Run a dependent t-test.
4. Interpret the results.

Here is the code you will use:

R Tutorial Videos


Pre-Lab

Pre-Lab due May 03, 2016 at 17:00 UTC 

Lab

Lab due May 03, 2016 at 17:00 UTC 

Problem Set

Problem Set due May 03, 2016 at 17:00 UTC 

- ▶ Week 4: Hypothesis Testing (Categorical Data)

Lab Question 1

Make a vector of happiness scores for each sample

```
underclass_happy <-
```

```
post$happy[post$classification=='Freshman' | post$classification=='Sophomore']
```

```
upperclass_happy <-
```

```
post$happy[post$classification=='Junior' | post$classification=='Senior']
```

Check the normality assumption

```
hist(underclass_happy, xlab='Underclassman Happiness', main='Percent of Time Happy')
```

```
hist(upperclass_happy, xlab='Upperclassman Happiness', main='Percent of Time Happy')
```

Run independent t-test

```
t.test(underclass_happy, upperclass_happy)
```

Lab Question 2

Make a vector of difference scores

```
post$diff_happy <- post$happy - post$post_happy
```

Check the normality assumption

```
hist(post$diff_happy, xlab= 'Difference in Happiness over the Semester', main = 'Happy-Post Happy')
```

Run dependent t-test

```
t.test(post$happy, post$post_happy, paired=T)
```

(1/1 point)

1. Which classifications of students are considered **upperclassmen**, according to the code above?

☐ seniors only

☒ juniors and seniors 

☐ sophomores, juniors and seniors

[Click here for a video explanation of how to answer this question.](#)

You have used 1 of 1 submissions

(1/1 point)

2. How many **sample means** are being compared in the t-test for Lab Question 1?

☐ three

☒ two ✓

☐ one

[Click here for a video explanation of how to answer this question.](#)

You have used 1 of 1 submissions

(1/1 point)

3. What does this line of code do?

```
post$diff_happy <- post$happy - post$post_happy
```

☐ Calculates how happy each student was at the end of the semester

☒ Creates a new variable for each student in the dataset ✓

☐ Finds the average difference in happiness for all students in the dataset

[Click here for a video explanation of how to answer this question.](#)

You have used 1 of 1 submissions

(1/1 point)

4. A student was happy 75% of the time at the beginning of the semester and 90% at the end of the semester. What will be the value of **post\$diff_happy** for this student?

☐ 0

☒ -15 ✓

☐ +15

[Click here for a video explanation of how to answer this question.](#)

You have used 1 of 1 submissions

(1/1 point)

Suppose we wanted to test the happiness scores of those who live on campus against those who live off campus. What has caused the error below?

```
post <- PostSurvey
on_campus <- post[post$live_campus == 'yes',]
off_campus <- post[post$live_campus == 'no',]
on_campus_happy <- on_campus$happy
off_campus_happy <- off_campus$happy
t.test(on_campus_happy, off_campus_happy, paired = T)
```

Error in complete.cases(x, y) : not all arguments have the same length

☒ We ran the wrong type of test. ✓

☐ We told R to look in the wrong dataset for the "happy" variable.

☐ We did not specify the value of mu.

☐ The responses for the "live_campus" variable have "Y" and "N" as responses, not "yes" and "no".

[Click here for a video explanation of how to answer this question.](#)

You have used 1 of 1 submissions

© All Rights Reserved



© edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY
OPENedX

