



[Course](#) > [Week...](#) > [Pre-L...](#) > [Prep...](#)

Prepare for the Analysis

Reflect on the Question

Analyze the Data

Draw Conclusions

Primary Research Question

For artists age 30 or older, do female artists play different kinds of music on Austin City Limits than male artists?

Breakdown Your Analysis

Let's break this analysis into its required steps:

1. Create a subset of the data for artists age 30 or older.
2. Create a table to show the marginal distribution for each variable.
3. Create a contingency table to show the conditional distribution for gender and genre.
4. Make a bar chart to better visualize how many male and female artists played in each genre.
5. Calculate $P(A)$: the probability of each type of music (genre) being played.
6. Calculate $P(A|B)$: the probability of each genre, given the artist's gender.
7. Interpret what these probabilities tell us about the relationship between genre and gender.

Here is the code you will use:

```
#Subset the data for artists age 30 or older
older <-acl[acl$Age>=30,]

# Create tables of marginal distributions
genre <- table(older$Genre)
genre
gender <- table(older$Gender)
gender

# Create contingency table
twoway <- table(older$Gender,older$Genre)
twoway

# Visualize the counts
barplot(twoway, legend=T, beside=T)

# Calculate P(A): the probability of each genre being played
prop.table(genre)

# Calculate P(A | B): the probability of each genre being played, given the artist's gender
prop.table(twoway,1)
```

problem

1/1 point (graded)

1) How many columns will be present in the table generated by the following line of code?

```
gender <- table(acl$Gender)
```

☐ 3

☒ 2 ✓

☐ 1☐ 4

You have used 1 of 1 attempt

problem

1/1 point (graded)

2) This code produces a bar chart with both a legend and side-by-side bars for each gender:

```
barplot(twoway, legend=TRUE, beside=TRUE)
```

What would the code look like if we wanted to keep the legend but stack the bars (instead of set them side-by-side)?

☐ `barplot(twoway, beside=TRUE)`☐ `barplot(twoway)`☒ `barplot(twoway, legend=TRUE)` ✓

You have used 1 of 1 attempt

problem

1/1 point (graded)

3) This line of code will produce four values, one for each genre of music:

```
prop.table(genre)
```

What value should you get if you sum the four values together?

☐ 0.25

☒ 1.00 ✓

☐ 0.50

☐ 4.00

Submit

You have used 1 of 1 attempt

problem

1/1 point (graded)

4) What does the value "1" refer to in this line of code:

```
prop.table(twoway, 1)
```

Hint: Go back and look at the comment line.

☒ The number 1 references the first variable (gender) listed in the contingency table code. ✓

- ☐ The number 1 refers to the first kind of question we typically ask of a contingency table.
- ☐ The number 1 refers to the fact that we typically code gender as (0,1) for (male,female).

Submit

You have used 1 of 1 attempt

problem

1/1 point (graded)

5) Suppose we ran the following code to find conditional probabilities. What error has caused the `prop.table` function to not work? (You may want to examine the dataset in R for help.)

```
acl <- AustinCityLimits
```

```
View(acl)
```

```
gender_grammy <- table(acl$Gender, acl$Grammy)
```

```
prop.table(acl$Gender, acl$Grammy, 1)
```

```
Error in prop.table(acl$Gender, acl$Grammy, 1) : unused  
argument (1)
```

- ☒ The `prop.table` function needs a table as an argument. ✓
- ☐ There is no "Gender" variable in our dataset.
- ☐ The "Grammy" variable in our dataset is not capitalized.

- We have used two quantitative variables instead of categorical variables.

Submit

You have used 1 of 1 attempt

© All Rights Reserved