edX   **UTAustinX:** UT.7.20x Foundations of Data Analysis - Part 2

Week 4: Hypothesis Testing (Categorical Data) > Lab > Reflect on the Question

🔖 Bookmarks

🔖 Bookmark

Reflect on the Question   Analyze the Data   Draw Conclusions

- ▸ 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)
- ▾ **Week 4: Hypothesis Testing (Categorical Data)**
  - **Readings**

Lab 4: Austin City Limits

Known as the "Live Music Capital of the World," Austin, Texas is also home to the longest-running music series in American television history, *Austin City Limits*. This dataset includes data on a sample of musicians that performed live on the PBS television series *Austin City Limits* over the last 10 years. Data on each artist include measures of commercial popularity, such as the number of social media followers on Twitter or Facebook, and their success in winning a Grammy Music Award.

(1/2 points)
Review of Chi Square Tests

In this lab, you will use **Chi Square Tests** to answer a question of interest. Let's start by remembering why we use each Chi Square test.

1a. In a Chi Square Goodness of Fit test, a proposed distribution model is compared to an observed

- ⚪ marginal distribution. ✔
- ⚪ mean.
- ⦿ conditional distribution. ✖
- ⚪ set of variables.

1b. Two categorical variables are said to be **independent** if their conditional distribution matches

- ⚪ the marginal distribution of one of the variables.
- ⦿ the distribution of expected counts, when the variables are assumed not to be related. ✔
- ⚪ the distribution of expected counts, where the counts are the same in every cell.

You have used 1 of 1 submissions

(2/2 points)
Lab Preparation

In this lab you will be working with data from the Austin City Limits.

1. Open RStudio. Make sure you've installed the SDSFoundations package.
2. Type `library(SDSFoundations)`. This will automatically load the data for the labs.
3. Type `acl <- AustinCityLimits`. This will assign the data to your Workspace.

Alternatively, you can use follow the steps in the "Importing a Data Frame" R tutorial video, and use the AustinCityLimits.csv file. (Right-click and "Save As.") Make sure to **name** the dataframe "acl" when importing.

1. Open RStudio.
2. Click on the "Import Dataset" button at the top of the workspace window. Choose *"from text file."*
3. Click on the location of the AustinCityLimits.csv file you just downloaded.
4. Click on the AustinCityLimits.csv file. Then, click Upload.

Feel free to use the script from the week's PreLab, which you can modify for use in this Lab.

The following **two** questions will be answered in lab. Match each question to its appropriate **Chi Square** test.

2a. Are each of the four musical genres equally represented on Austin City Limits?

[goodness of fit test ▼]   ✔   **Answer:** goodness of fit test

2b. Are some genres more likely to draw a large (100K+) Twitter following than others?

[test of independence ▼]   ✔   **Answer:** test of independence

You have used 1 of 1 submissions

POWERED BY
OPENedX