

For Lab 11 and 12, you will collect your own data. For Lab 11, we will analyze the colors of Starburst. For Lab 12, we will analyze the personality types of statistics students, but we will collect the data now.

To collect the personality type data, please open the following link in your browser window:

<http://www.16personalities.com/free-personality-test>, and complete the 10-15 minute personality test. Once you complete the test, your personality will be classified into one of 16 “types,” but we will consider only the 4 main personality categories: Analysts, Diplomats, Sentinels, and Explorers. The figure below shows someone classified as an “Architect,” which is one of the Analyst types as shown by the highlighting. Note that you may have to click on “Start Reading” to find this type. While you are waiting for the rest of the class to finish, feel free to peruse your “results,” or to click on the Personality Types tab to look at all of the possible types.



**Question #1** Once you obtain your individual results, complete the Questionnaire on Titanium for your personality type, major, and the number of Starbursts in each color you received.

In this lab, we will test the claim that all four main colors are equally likely in the population of all Starburst.

**Question #2** Write out the null hypothesis for this goodness-of-fit test.

There should be an equally likelihood of obtaining each color.

**Question #3** If the null hypothesis is true, how many of each Starbursts type would we expect to see in our sample?

There should be a 25% chance of being a certain color Starbursts.

**Question #4** If the sample size assumptions are met, what would be the sampling distribution of your test statistic (i.e., what is the type of distribution and the degrees of freedom)?

The degrees of freedom would be number of categories - 1 which is going to be  $4 - 1 = 3$ . There are 3 degrees of freedom.

Let's put in our data and running the goodness-of-fit test in R.

```
> Starburst <- c(11, 13, 17, 21) # fill this in with the actual numbers
> probs <- c() # fill this in with your probabilities under H0
```

**Question #5** Are the sample size assumptions are met for a chi-square goodness of fit test? Show how you checked the assumptions.

The sample size assumptions are met because simulation does not provide a degree of freedom and because the expected count of each color is more than 5.

If the sample size assumptions are met, run the following code to perform a chi-square goodness of fit test:

```
> chisq.test(Starburst, p = probs)
```

Otherwise, add the following argument to perform a goodness of fit test by simulation:

```
> chisq.test(Starburst, p = probs, simulate.p.value = TRUE)
```

**Question #6** Copy the RStudio output below.

Chi-squared test for given probabilities

data: star

X-squared = 7.4545, df = 3, p-value = 0.05874

**Question #7** What is the value of the chi-square test statistic as computed by R?

The value of chi-squared will be 7.4545

**Question #8** What is the p-value for this test?

The p-value is 0.05874.

**Question #9** Using a 5% significance level, what can you conclude about the distribution of Starburst?

We just barely crossed the threshold of it being reasonable so we can still keep the assumption that it is equally likely to pick one of four colors which is 25%.

**Question #10** Do you believe that your conclusion (from **Question #9**) applies to the population of all Starburst? (HINT: Think about the sample we used and the way we collected the data)

This will most likely not represent the entire population of all Starburst because there is not enough elements in the sample.