Yuyi Zhang

Apr 21, 2019

Spring 2019 CPS Quarter Term A

Instructor: Steward Huang

Assignment 2

ALY 6015\_Intermediate Analytics

# Introduction

The Original report that being used is Adoptions by SFY, DCF Office, Gender and Length of Stay. All data retrieved from Department of Children and Families Office in Connecticut. The original report contains aggregate data concerning the number of children that exited DCF care to an adaption. In the original report, adoptions in 2000 to 2016 across 16 offices are being recorded in 3 different perspectives.

In this report, we are going to conduct a hypothesis test based on the original report. The key columns we are going to use are: SFY(State Fiscal Year), Office, LOS\_LessThan24Mnths, and LOS\_MoreThan24Mnths. In data preparation phase, I’ve also aggregated the demographic column because gender is not our concern at this time.

# Analysis

## Claim

One of the claims is that children stay in DCF care more than 24 months had better chance to be adopted.

So in this test we are going to concentrate on columns: Length of Stay\_Less Than 24 Months and LOS\_MoreThan24Mnths. We are going to collect data in these two columns aggregated by different office and compare the mean adoption numbers for two LOS by conducting a two-sample t-test.

## Data Preparation

### Read data

**Code:**

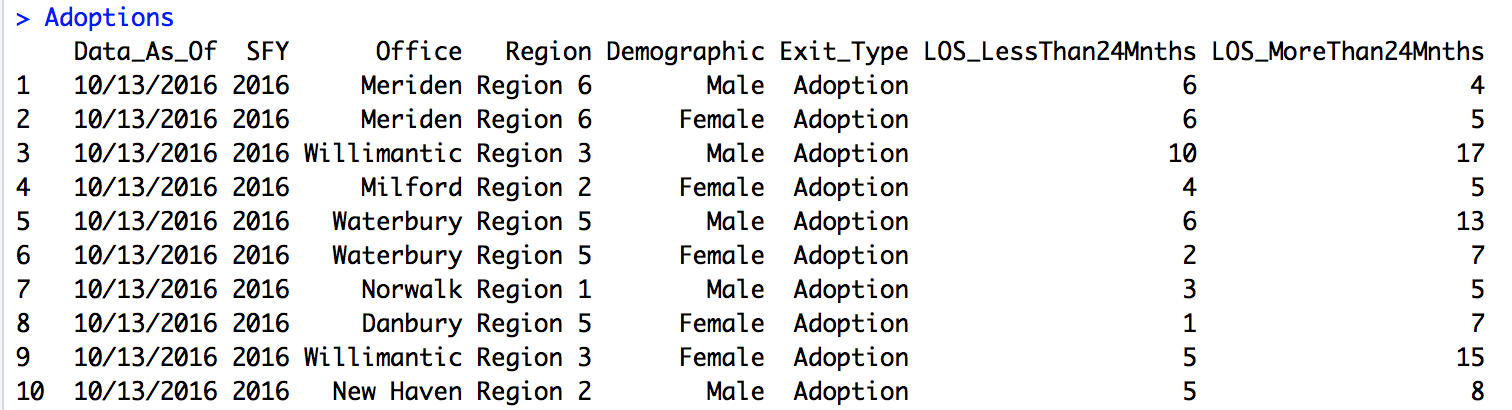
# Data Preparation

library(readr)

library(dplyr)

Adoptions <- read.csv("FCR/NEU/CPS/Analytics\_2018/ALY 6015\_Intermediate Analytics/Week 2/Adoptions\_by\_SFY\_\_DCF\_Office\_\_Gender\_and\_Length\_of\_Stay.csv", head(TRUE))

**Console:**



### Aggregate by office and sfy

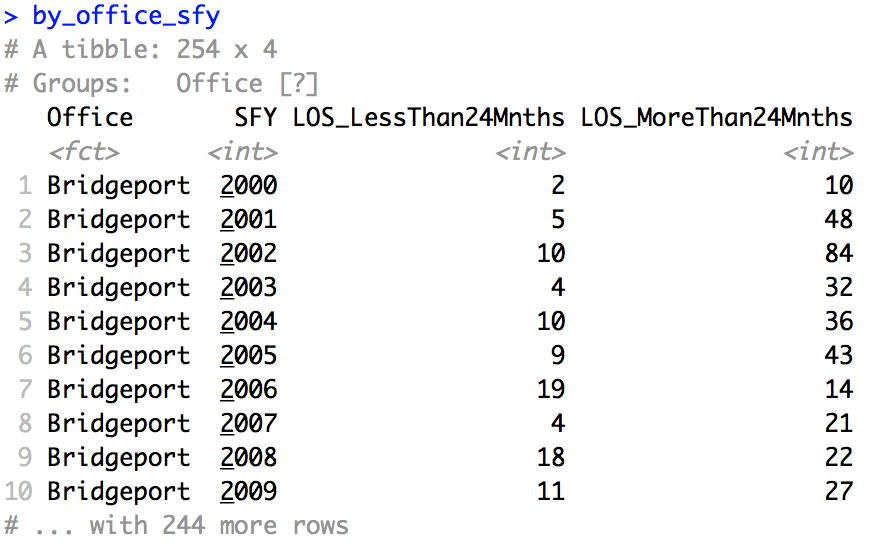
**Code:**

by\_office\_sfy <- Adoptions %>%

group\_by(Office, SFY) %>%

summarise(LOS\_LessThan24Mnths = sum(LOS\_LessThan24Mnths), LOS\_MoreThan24Mnths = sum(LOS\_MoreThan24Mnths))

**Console:**



## Conduct Hypothesis Test

### State null and alternative hypothesis.

We let μ1 and μ2 represent the population mean LOS less and more than 24 months. If the claim is true, then the mean LOS less than 24 months would be smaller than the mean in LOS more than 24 months.

Null Hypothesis H0: (μ1 - μ2) ≥ 0

Alternative Hypothesis Ha: (μ1 - μ2) < 0

### Select appropriate test statistic and probability distribution.

Two sample t-test.

### Select level of significance.

The significance level, α, would be setup as 0.05.

### Identify regions of rejection.

In this case, a left-tailed hypothesis test would be appropriate since we are interested in compare the two population means.

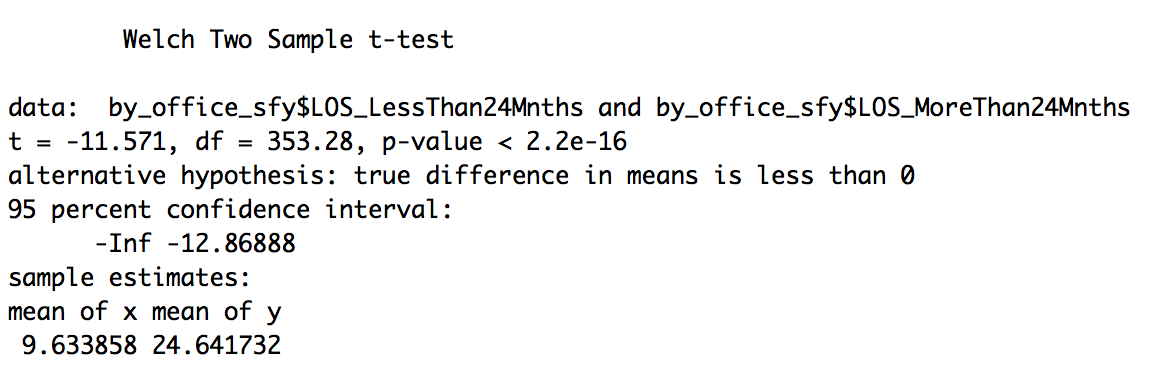
### Calculate test statistics in R.

**Code:**

#Two Sample t-test

t.test(by\_office\_sfy$LOS\_LessThan24Mnths, by\_office\_sfy$LOS\_MoreThan24Mnths, alternative = c("less"))

**Console:**



### Make a decision regarding null hypothesis.

There are two methods to decide whether reject H0. First, compare test statistic and critical value. In this case, test statistic (-11.571) is larger than critical value (-12.86888). So H0 is rejected. The second method is to compare P-value and significance value. In this case, P-value (2.2e-16) is smaller than significance level (0.05). So H0 is rejected. Both methods 1 and 2 of the previous step have efficient evidence to reject the Null hypothesis.

# Conclusion

In summary, we 95% confidence to assert the claim that children stay in DCF care for more than 24 months were more likely to be adopted.

Reference

1. Maindonald, J. H. (2008). *Using R for Data Analysis and Graphics.* Retrieved from <https://cran.r-project.org/doc/contrib/usingR.pdf>
2. Johnson,M. (2014). *Introduction to hypothesis testing with R*. Retrieved from <https://s3.us-east-1.amazonaws.com/blackboard.learn.xythos.prod/5a3148150d016/15190315?response-content-disposition=inline%3B%20filename%2A%3DUTF-8%27%27Introduction%2520to%2520Hypothesis%2520testing%2520using%2520R.pdf&response-content-type=application%2Fpdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Date=20190421T071013Z&X-Amz-SignedHeaders=host&X-Amz-Expires=21600&X-Amz-Credential=AKIAIL7WQYDOOHAZJGWQ%2F20190421%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Signature=19a9b9a10e201ae2c57ebbb76e58e66d22b203fb1856e9a61bd34b433afbe850>
3. Connecticut Open Data (Sep 23, 2017). *Adoptions by SFY, DCF Office, Gender and Length of Stay*. Retrieved from <https://data.ct.gov/Health-and-Human-Services/Adoptions-by-SFY-DCF-Office-Gender-and-Length-of-S/cyz8-6esi>