

Letter to Director of Blue Springs Public Library

To: maria_potter@bluespringspubliclibrary.com

From: UNTGroup1@my.unt.edu

Subject: Library Science Masters students requesting permission to conduct a public use study.

Dear Maria Potter,

We are writing to you today to request permission to conduct a public library research study at your facility as part of an assignment for a Research Methods & Analysis course.

About the research project: Our research goal is to take a sample of the population of library users and apply statistical tools to evaluate if the gender of the population of library users is representative of the gender of the total population of Blue Springs. To accomplish this goal, we intend to collect observational data in two-hour segments on three separate days during times that are representative of typical library usage by counting visitors as they enter the library and recording the numbers of male and female visitors. We will do our best to be unobtrusive to both the staff and the library users during this time, collecting no personal or identifying data about users and not disrupting them in their activities in the library. Once we have collected our data, we will apply a Chi-Square Goodness of Fit statistical tool to determine if the proportion of nominal variables (male or female) found in the sample is likely to reflect the proportions of the population.

About the research team: Our research team consists of six graduate students from the University of North Texas: Madison Brents, Valerie Darling, Forrest Davidson, Megan Del Mar, Sarah Diaz, and Katie Fuentes. This project is part of a group activity in 5080 Research Methods & Analyses taught by Dr. Philip Turner and Dr. Agnes Percy. We are applying a hypothesis testing worksheet to the project to compare the research question with a null hypothesis (RQ: Is there a difference between the visitors observed using the Blue Springs Library in the duration of the study and the population of the servicing area based on gender? NH: There is not a difference between library users and the overall population of Blue Springs with regards to gender.). We will then examine the data, calculate the p-value, examine our research methods and data collection for possible threats to validity as well as for Type 1 and Type 2 errors, then write a report summarizing the research. We will send you a summary letter of our findings as well so you can see how the data we collected was handled and analyzed.

Thank you for your time, we look forward to hearing back and scheduling our observation periods.

Regards,

UNT Group 1

Study Regarding Proportion of Genders at Blue Springs Public Library

Introduction

The Mayor of Blue Springs wanted to know if the gender of visitors to the Blue Springs Public Library represented the proportion of gender of the population of Blue Springs. The Blue Springs Public Library is the sole public library in Blue Springs and is located at 2220 S. 7 Hwy, Blue Springs, Missouri. The library offers various library services for all ages and serves as the location for town halls and voting. With permission from the Library Director, Maria Potter, researchers observed and recorded the assumed gender and number of visitors entering the library on three separate days for two hours each day. Using the data, the researchers will aim to address whether the proportions of gender in Blue Springs Public Library are representative of the proportion of gender in Blue Springs' total population.

Methods

To gather the data, two researchers observed and recorded the assumed gender and number of visitors entering the library on three randomly selected days for two hours each day. The two-hour observational times varied to be representative of typical library usage. Efforts were made to ensure that observations were not conducted during special events. Researchers recorded only the total number of visitors entering the library and their assumed gender; no identifying information was gathered. As the study is non-experimental, no consent forms were needed and patrons were unaware of the researchers. Over the course of the observations, a total of 418 visitors were recorded entering the library. Of the 418 visitors, 265 (64%) were assumed female and 153 (36%) were assumed male.

Because the research question included one nominal variable (gender), the results were analyzed using Chi-Square Goodness of Fit to determine whether the proportion of gender in Blue Springs visitors is likely to reflect the gender proportions of the town's population.

Population Data

For this study, data was collected by observing the entryway of the library and recording the assumed gender of those who entered. This was done on three separate occasions, for two hours each time. The population data below was provided by the Mayor of Blue Springs, when he requested the study.

Population of Blue Springs:

- 50.5% females
- 49.5% males

Sample Data:

Observations were collected at the following times and dates:

- Saturday 4/3/2021 from 12:00PM - 2:00PM
 - Males: 58
 - Females: 88
- Tuesday 4/6/2021 from 2:00PM - 4:00PM
 - Males: 44
 - Females: 97
- Friday 4/9/2021 from 3:30PM - 5:30PM
 - Males: 51
 - Females: 80

Total: 418 visitors to the library

- 153 males (36%)
- 265 females (64%)

Instrumentation and Statistical Analysis of the Sample

There were more female visitors and fewer male visitors than would be expected if the null hypothesis was accepted. Based on the null hypothesis, the researchers expected a total of 211.09 female visitors and 206.91 male visitors. The actual results show that more females (265) and fewer males (153) than expected visited the library during the observed time periods. This indicates that the visitors to the Blue Springs Public Library do not reflect the population of the town of Blue Springs by gender, and as a result, we do not accept the null hypothesis.

To determine whether this difference is significant, the researchers selected the Chi-Square Goodness of Fit test to analyze this sample. This test is used to compare the observed quantities of a nominal variable, such as gender, with the expected results according to the Null Hypothesis. The formula for the Chi-Square Goodness of Fit test is as follows:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Further information about the results of the Chi-Square Goodness of Fit inferential statistical tool can be found in the Hypothesis Testing Worksheet and Results sections.

Hypothesis Testing Worksheet

1. Research Question:

Is there a difference between the visitors observed using the Blue Springs Library in the duration of the study and the population of the servicing area based on gender?

2. Null Hypothesis based on Research Question:

There is not a difference between library users and the overall population of Blue Springs with regards to gender.

3. Collect and Observe the data and all descriptive statistics

Females make up 50.5% of the Blue Springs population, while males make up 49.5%. We took a sample of 418 total visitors to the library within a 3 day non-consecutive time period. If the population proportion matched exactly, we would have expected to see rounded values around 211.09 women and 206.91 men enter the Blue Springs Library. What we observed were 265 females, 64%, and 153 males, 36%. In this study, the city of Blue Springs and the Library servicing area are one and the same, because the Blue Springs Library is the public library serving the city.

4. Based on observing the data and descriptive statistics, what do you think about the research question?

Observations made during the study show that there is a difference in the gender ratio of the visitors to the library and the library servicing area. The library users do not represent the entire population on the basis of gender, since there is a greater percentage of females observed visiting the library than expected.

5. Select the probability level (usually .05)

The probability level selected is .05.

6. Select the inferential statistical tool that you will use to infer from the sample to its population.

The researchers used the Chi-Squared Goodness of Fit Statistical Tool to make inferences between the sample and the Blue Springs population.

7. Calculate the statistic by using the appropriate online tool.

The calculation of Chi-Square shows the value of 27.3.

8. Find the obtained p (probability value) in the online output.

The probability value is $<.0001$.

9. Concluding statement. If the obtained p-value is less than .05, then reject the null hypothesis. If not less than .05, do not reject the null hypothesis.

The Null Hypothesis can be rejected, as there is at least 95% certainty that if another sample was taken of library users, there would be a difference of gender ratio in library users and in the population of Blue Springs.

Analysis for Type I & Type II Errors

The observers coordinated with the Library Director to ensure that the observations were scheduled for times when no group meetings, events, or special activities were occurring. Such gatherings may have inadvertently skewed the data that was collected and would therefore have prevented the researchers from attaining an accurate sample of the average proportion of library visitors by gender.

As the Null Hypothesis was rejected, Type 2 errors (when researchers fail to reject the null hypothesis when they instead should have rejected the null hypothesis) are not applicable to the research results.

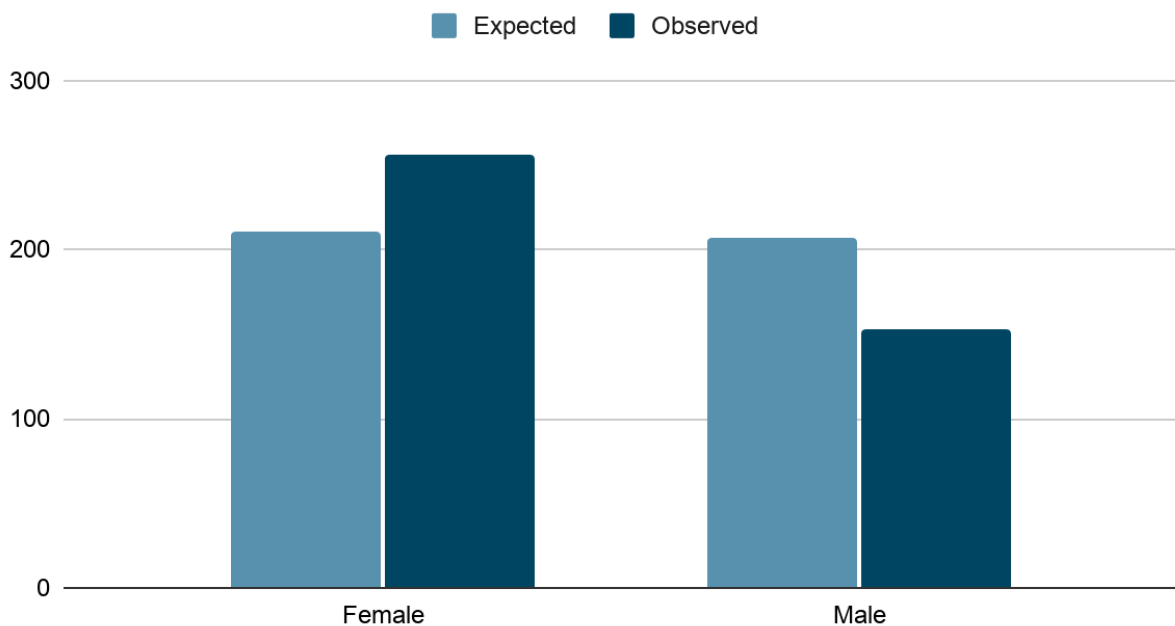
Internal Validity Analysis

One potential threat to the validity of this study is that we do not know the population size of Blue Springs, only the gender percentage. This means we cannot use the Table for Determining Sample Size to be sure that the 418 visitors we recorded is an accurate sample of the population. An additional threat to validity comes from the fact that there may be a difference between reporting and gathering gender information. The gender data was provided by the Mayor of Blue Springs who used self-reported data from the 2010 Census. In order to remain unobtrusive, the observers in the study used assumed gender to inform the data. This may prove an instrumentation threat to validity in that different observers may measure assumed gender differently. In a different sample, further differences in measurement might happen. Another threat could come from a potential difference in gender ratio between census data, and the population as not every resident participated, so the population gender ratio may be skewed from reality.

Results

The researchers found a difference between the observed percentages of male and female library users and the expected gender ratio of the general population in Blue Springs. As mentioned previously, 50.5% of city residents are female and 49.5% are male, whereas 64% of the library visitors in the sample were female and 36% were male. Using the Chi-Square Goodness of Fit inferential statistical tool, the Chi-Square value was 27.3. The p-value of less than 0.0001 falls well below the 0.05 probability level. This allows the researchers to conclude with at least 95% certainty that if further studies were launched, the results would continue to show that the gender ratio of library visitors is different from the gender ratio of Blue Springs.

Library Visitors by Gender



Conclusion

The analysis of the observed and expected data calculated using the Chi-Square Goodness of Fit inferential statistical tool shows that the researchers can reject the null hypothesis with 95% accuracy. It can be concluded that the percentage of library users by gender differs from the proportion of the population of Blue Springs. The results of the study show that there is a higher proportion of female visitors to the library when compared to the percentage of females living in the library serving area. In correlation, there is a lower ratio of male visitors to the library when compared to the male population of the city.

Communication of Project Results to Director

To: maria_potter@bluespringspubliclibrary.com

From: UNTGroup1@my.unt.edu

Subject: UNT MLS Public Library Use Study Summary

Dear Maria Potter,

Thank you so much for the opportunity to conduct our research at your facility. Below is a summary of our activities and findings.

Our research project was to determine if the proportion by gender of library users was representative of the proportion by gender of the population of Blue Springs. Data were collected in two-hour segments across three separate days by tallying the gender of visitors as they entered the library during data collection times. The resulting sample was 418 visitors to the library, 153 (36%) of which were identified as male, and 265 (64%) of which were identified as female. The Chi-Square Goodness of Fit inferential statistical tool was applied to this sample to determine if the results were statistically significant, and the p-value was calculated as $<.0001$. Based on this statistic, we can be at least 95% certain that the proportion of gender of the population of library users is different from the proportion by gender of the population of Blue Springs as a whole.

Thank you for allowing us to come and study your library.

Regards,

UNT Group 1

Madison Brents, Valerie Darling, Forrest Davidson, Megan Del Mar, Sarah Diaz, Katie Fuentes

Group Member Contributions

Madison Brents: Outline, Layout, Hypothesis testing, Internal Validity Analysis, Population Data, Editor

Valerie Darling: Introduction, Methods, Analysis for Type 1 and Type 2 Errors Copyediting, Editor

Forrest Davidson: Permission and Summary Letters, Editor, Hypothesis Testing Worksheet, Analysis for Type 1 and Type 2 Errors, Sample Data

Megan Del Mar: Instrumentation/Statistical Analysis, Results, Conclusion, Editor

Sarah Diaz: Instrumentation/Statistical Analysis, Results, Library Visitors By Gender Graph, Editor

Katie Fuentes: Hypothesis Testing Worksheet, Conclusion, Calculations, Editor