Comparison of 2018 & 2021 Survey Data from Bailey Library Clients

Elijah McClymonds, Jana Asher, & Amanda Goodrick
Department of Mathematics & Statistics, Slippery Rock University

Introduction:

Every three years, the SRU Bailey Library participates in a national survey that measures library user satisfaction, library quality, and library functionality. This survey was made available by the Bailey Library staff for the 2021 winter session DataCamp run through the SRU Mathematics/Statistics Department. This constructed study uses the survey data with over 1,000 respondents from the years 2018 and 2021. Each yearly dataset measures the minimum level(mn), desired level(de), and perceived level(pr) of a corresponding question. This data is further made measurable by the two included variables Adequacy Gap(ad) and Superiority Gap(su). These two variables are calculated using the corresponding user's perceived level minus the minimum level and the corresponding user's perceived level minus the desired level, respectively. The data is also divided into three distinct themes which are Affect of Service(AS), Information Control(IC), and Library as a Place(LP). Affect of Service questions deal with staff performance, Information Control questions deal with access to library resources, and Library as a Place questions deal with the environment and atmosphere the library creates for a user. Note that there are separate questions that are unique to a year's survey (i.e. L17359 is unique to the 2021 survey). There are many descriptive variables that reflect the user such as major, age, and position as well.

Objectives:

For this study, we want to:

i. compare 2018 and 2021 survey results
ii. determine the most meaningful questions in the most recent (2021) survey
iii. Investigate survey responses based on major (2021)

Methodology:

For i:

We will use SAS software to perform two sample t-tests. A two-sample t-test is a statistical hypothesis test that compares a mean of a sample across two classes to see if they yield different results. For every variable, we will compare the measurable variables' means across two different years (2018 vs 2021).

For ii:

We will continue using SAS software to perform paired t-tests. A paired t-test is another statistical hypothesis test that generally compares a mean of one variable to a second variable. In this circumstance, we will compute the mean of each measurable variable and compare the value to the mean of the corresponding hierarchal variable. In other words, ad will match up with AdqGap and su will match up with SupGap.

For iii:

We will use the statistical program R to perform a k-means data cluster. This test indicates if students answer similarly based on major. We plot the data clusters to visually illustrate the classification of students' responses by major.

Results:

2018 Vs 2021 Analysis

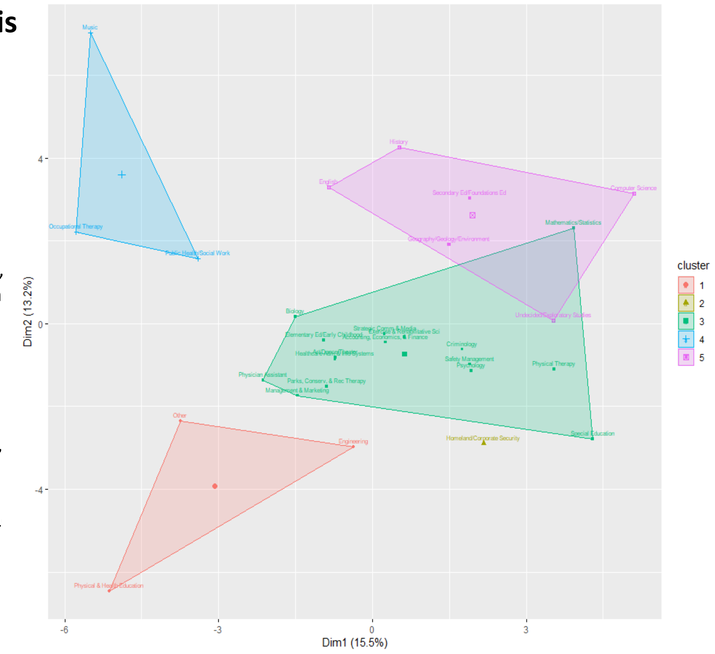
Variable	Label	Mean (2018)	Mean (2021)	P-Value
IC07_su	Question 8 - The printed library materials I need for my work (superiority gap)	0.0486	-0.3415	0.0024
IC10_ad	Question 11 - The electronic information resources I need (adequacy gap)	1.0672	0.8479	0.0017
IC10_su	Question 11 - The electronic information resources I need (superiority gap)	-0.1491	-0.3614	0.0007
IC14_ad	Question 17 - Modern equipment that lets me easily access needed information (adequacy gap)	0.7177	0.231	0.0001
IC14_su	Question 17 - Modern equipment that lets me easily access needed information (superiority gap)	-0.2744	-0.941	0.0001
AS18_ad	Question 22 - Willingness to help users (adequacy gap)	0.9382	0.6307	0.0083
AvgDes	Average desired expectation for first 27 questions	7.6079	7.4434	0.0019
AvgPer	Average perceived experience for first 27 questions	7.349	7.1539	0.0003
D2AvgPer	Average perceived experience regarding the library's information control	7.3342	7.0959	0.0001
D2AdqGap	Average adequacy gap regarding the library's information control	0.8472	0.6859	0.0031
D2SupGap	Average superiority gap regarding the library's information control	-0.2722	-0.4338	0.0007

From the table constructed from the two-sample t-test, we can note:

- ✓ Multiple variable means dropped significantly in value from 2018 to 2021.
- ✓ The survey data indicates overall perception and expectation levels went down from 2018 to 2021 and the amount of negligence people are willing to put up with regarding the library's information fell.
- ✓ Specific questions indicate the areas the Bailey Library should improve upon.

2021 Cluster Analysis

- ✓ K-means clustering is a technique in which we place each observation in a dataset into one of K clusters (in this case k=5).
- ✓ This K-means cluster is run comparing the means across different majors at SRU.
- ✓ Cluster 1 contains Engineering, Other, and Physical and Health majors
- ✓ Cluster 2 contains the Homeland/Corporate Security major
- ✓ Cluster 3 contains Accounting/Economy/Finance, Art/Dance/Theater, Biology, Criminology, Elementary Ed/Early Childhood, Exercise & Rehabilitative Sci, Healthcare Adm & Info Systems, Management & Marketing, Mathematics/Statistics, Parks, Conserve, & Rec Therapy, Physical Therapy, Physician Assistant, Psychology, Safety Management, Special Education, and Strategic Comm & Media majors.
- ✓ Cluster 4 contains the Music, Occupational Therapy, and Public Health/Social Work majors
- ✓ Cluster 5 contains the Computer Science, English, Geography/Geology/Environment, History, Secondary Ed/Foundations Ed and Undecided/Explored Studies majors



2021 Question Analysis

Survey Data Deficiencies

Variable	Label	Mean Difference	P-value
IC02_ad	Question 2 - Making electronic resources accessible from my home or office (Adequacy)	0.1389	0.011
IC07_ad	Question 8 - The printed library materials I need for my work (Adequacy)	0.1602	0.0113
IC14_ad	Question 17 - Modern equipment that lets me easily access needed information (Adequacy)	0.5727	0.0001
AS22_ad	Question 27 - Dependability in handling users' service problems (Adequacy)	0.2258	0.0001
L17363_ad	Question 26 - Access to innovative technology in Bailey Library (Adequacy)	0.3181	0.0001
IC05_su	Question 5 - A library Website enabling me to locate information on my own (Superiority)	0.1353	0.0055
IC10_su	Question 11 - The electronic information resources I need (Superiority)	0.074	0.0176*
IC14_su	Question 17 - Modern equipment that lets me easily access needed information (Superiority)	0.5937	0.0001
L17363_su	Question 26 - Access to innovative technology in Bailey Library (Superiority)	0.3652	0.0001

Survey Data Proficiencies

Variable	Label	Mean Difference	P-Value
LP03_ad	Question 3 - Library space that inspires study and learning (Adequacy)	0.1718	0.0001
AS04_ad	Question 4 - Giving users individual attention (Adequacy)	0.1177	0.0184*
L17361_ad	Question 14 - Instruction sessions conducted by librarians helping me to better understand the research process (Adequacy)	0.2401	0.0043
LP12_ad	Question 15 - A comfortable and inviting location (Adequacy)	0.14	0.005
AS13_ad	Question 16 - Employees who deal with users in a caring fashion (Adequacy)	0.0861	0.0032
LP21_ad	Question 25 - Community space for group learning and group study (Adequacy)	0.1474	0.0081
AS11_su	Question 13 - Employees who have the knowledge to answer user questions (Superiority)	0.1111	0.0093
AS13_su	Question 16 - Employees who deal with users in a caring fashion (Superiority)	0.117	0.0001

- ✓ The table constructed on the left indicates the questions that weigh the overall average down and by extension need improvement. The table on the right indicates the questions that weigh the overall average up and by extension deserve recognition.
- ✓ Information Control yields the most concern for the survey respondents, while employee service and the library environment yield the best results for the survey respondents.

Conclusion/Outcomes:

- ✓ The average respondent survey data yielded statistically significant results only for 2021 mean data less than 2018 mean data. Results suggest that desired expectations along with minimum expectations have decreased over time.
- ✓ The access to library resources and information is a growing problem as documented in the Data Deficiencies table.
- ✓ Employee interaction and the environment at the Bailey library are both strong areas for the library.
- ✓ The cluster analysis indicated that student responses across several majors were similar (See Cluster 3).

Future Work:

- ✓ Library user survey data could alternatively be analyzed using a time series method. If we have access to more of the data from the past, we could try and analyze the data to determine whether there is a trend.
- ✓ The results of the paired t-test can be extended further to see the driving forces regarding the thematic labels of each question.
- ✓ The data clustering methods used could be implemented and optimized greater than the generic k-means cluster.