**Operations Research Fall 2022**

**Community-Engaged Learning Case Study Project**

**--------------------------**

**Allocating Space for Southwestern University (SU)**

**Library and Special Collections**

**--------------------------**

**Team: Clayton Nolen, Henry Fisher, Kyle Porter, Mark Mueller**

**Dec 2, 2022**

**1. Abstract**

Southwestern University’s Special Collections, located within the A. Frank Smith, Jr. Library Center, is undergoing foundational work, resulting in the required relocation of all of its contents. Thousands of books, documents, and artifacts must be removed from this part of the library to ensure their safety during the renovations. Our mission was to account for all of the contents in Special Collections, and be able to store them elsewhere in the library. We created two linear programs to aid us in our calculations. This first program, created in Glpsol, focuses on minimizing the amount of boxes needed to store all of the books. We found that an estimated 3,701 total boxes are required to store all of the books within this wing of the library. Our second program, created in Excel, takes the volumes of the contents into consideration and ensures that it will fit within the four outlined rooms that can be used to store Special Collections. Our evidence suggests that the contents will theoretically fit within three of the rooms, but four rooms is optimal to ensure everything is stored safely. This and many other considerations are outlined below.

**2. Executive Summary**

**Problem Statement**

Southwestern University is home to the A. Frank Smith, Jr. Library Center. This building is located on the southside of campus and has parts of its foundation sinking. The university has planned reconstruction to start in the year of 2023. These areas are made up of the Special Collections and neighboring parts of the library that are used to house books, shelves, periodicals, and other various artifacts that were donated to the library. These items are planned to be moved into on-campus storage, if possible. Various rooms are laid out in an attempt to store all of Special Collections so that the university does not have to pay the costs of off-campus storage. Our mission is to ensure that everything will fit, and provide analytical support to the library staff to make informed decisions.

**Proposed Solution**

To facilitate this, items will be put in boxes to and moved into storage locations throughout the library. We originally considered using PODS as a form of storage, but due to finding additional storage inside the library we do not think they will be required. Our objective is to minimize the amount of boxes used, which inherently maximizes the amount of items stored within the boxes. This will give us an estimate as to how many boxes the university should order, as well as how much space they may take up in on-campus storage.

**Model**

Our model is split into two linear programs: one that optimizes the books into boxes, and one that optimizes the boxes into storage. The first program focuses on minimizing the amount of boxes used, which also implicitly maximizes the amount of books that go into a single box. The second program takes into account these boxes, as well as the rest of the contents in Special Collections, and assigns them to one of the four to five possible rooms in which everything is to be stored.

**Results**

Considering each box is 10” x 12.5” x 15” we conclude that we will need an estimated amount of 3,701 boxes to adequately store all the books. This, coupled with the rest of the contents in Special Collections, will take up at least three of the allotted rooms in the library. This, however, is calculated by completely filling up the room to a height of 60”. So while it is doable in three rooms, our team recommends using the four possible rooms, or using the Prothro Room, to store everything as safely as possible.

**Value**

The items housed in Special Collections are made up of books and records about Southwestern University, Texas history, and education. If these items were not books, they were historical pieces. All items tell the story of the University’s history, the state’s history, or aspects of education gathered throughout the years. To conduct this project was of extreme importance so that the collections would be stored safely and not risk being ruined or damaged. In addition, there were opportunities to ensure the collection’s contents were stored safely and to save money.

**Conclusion**

Our solution currently finds the space within the library sufficient to hold the boxes of books during reconstruction. It should be made aware that this solution does not include all artifacts including, but not limited to, artifacts that are not box-like in shape, or items that are more fragile. We advise that an additional empty room, on top of the three aforementioned rooms, are all used to store Special Collections safely.

**3. Technical Report**

**3.1 Introduction**

The Frank A Smith Library of Southwestern University houses thousands of books and items including rare and old artifacts in Special Collections. The building foundation below Special Collections and surrounding rooms is sinking and therefore needs to be restored. This means thousands of rare, antique, and potentially fragile items must be moved into safe and secured storage. The university, its staff, students, faculty, and other stakeholders all have a vested interest in preserving Special Collections. Additionally, items in Special Collections are a valuable asset for learning, as well as exposing individuals to history and different cultures.

The methodology, at least for the first linear program, is structured around books, as they are the vast majority of items in Special Collections. We created five different categories based on the area of the spine of a book, because of its fairly strong correlation to the book’s volume. Each measured book is added to a category in order to simplify our data collection and improve efficiency. Using these categories, our team can extrapolate the sample results to the total number of books counted in any given room.

Our first linear model assigns a collection of books, based on their category, into storage boxes. Next, our second linear program attempts to fit the boxes into the rooms made available to us for storage. Our program is able to accurately place books within boxes, and those boxes into storage spaces. However, not having every data point collected, and focusing entirely on books is a limitation of our model.

**3.2 Methods**

Throughout formulating an effective strategy to tackle the problem at hand, our methodology undertook multiple stages. As data collection was a significant endeavor, and one of the biggest hurdles, it became our primary focus. Special Collections is primarily made up of books and smaller boxes of various sizes. Therefore, those items were our primary consideration in both of our linear models.

**3.2.1 Data Collection Structure**

Initially, we planned on measuring a shelf's dimensions, finding their volumes, and calculating the percentage of the shelving taken up by books or boxes. This was designed under the assumption that the volume of books should be reasonably correlated with their weight, which was a vital consideration due to the 40lbs weight constraint of the storage boxes. However, this idea was quickly scratched as it turned out to be too extensive and prone to inaccuracies. This is because we would have had to proportionalized the number of books by each shelf. This would have led to our data collection process, as well as a potential linear program, to become overly inaccurate due to the numerous different shelves within Special Collections and the amount of different types and sizes of books. Because of this and that this strategy was oversimplified and would not have transferred well into a linear program, the idea was scrapped.

Next, our group determined a different course of action was to collect as much accurate data as possible within our given time constraint. We decided to categorize each book within five different classifications: Smallest, Small-Mid, Medium, Mid-Large, LargestFigure 1. These categories split the books into volume categories starting at 100 inches cubed and increasing by 100 each time, and into weight categories, starting at half a pound and incrementing by another half a pound per category. Each book would be measured based on the area of the spine in inches, as we found a fairly strong correlation between that specific area of the book and its volume with a coefficient r = 0.9346, and a R2 value of 0.875Figure 2.

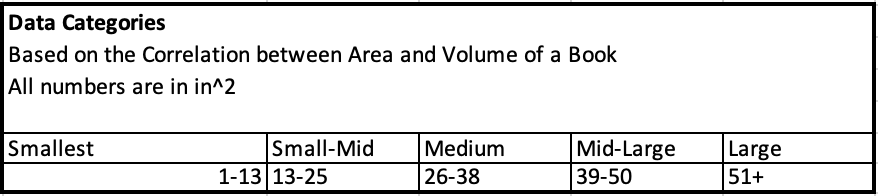


Figure 1-Data Categories

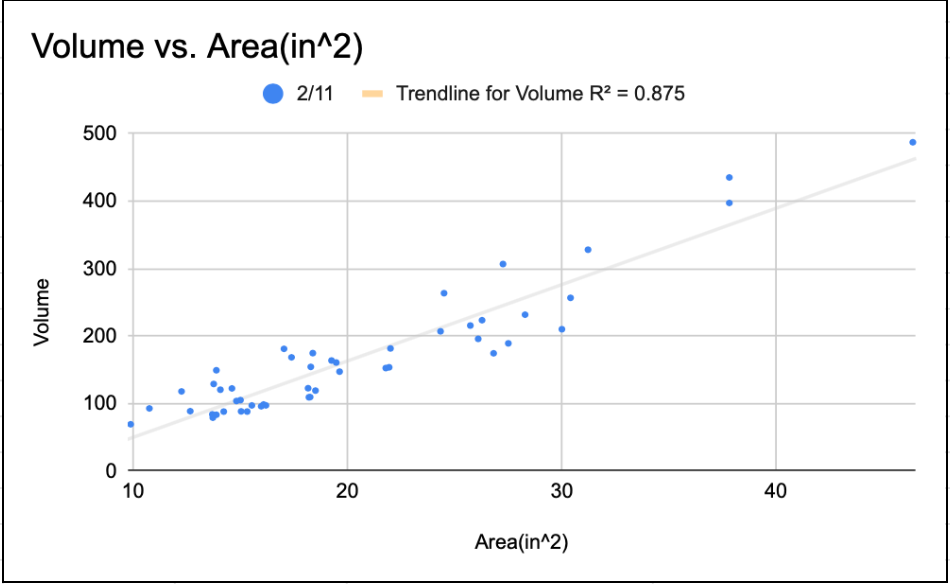


Figure 2-Correlation: Volume vs. Area of Spine

This process greatly furthered our ability to collect data efficiently, and led to increased confidence in our statistics. Creating categories also allowed us to measure a randomly selected amount of books per room, either 50 or 100, and to extrapolate those numbers out for the entire room. The books were selected through a dice roll to ensure controlling for biases in our collection process.

Another correlation our group had to consider was the relationship between the volume of the book and the weight of the book. The primary reason for this was the specific storage boxes, with a 40 lbs weight limit, that were provided to us by the library. This meant our data collection method had to ensure that this constraint was satisfied, and did not create future problems when the items were set to be moved. The sample size, which was randomly selected, showed a strong correlation between the volume of the books and the weight of the books. This was evident through its strong correlation coefficient r=0.9466, and its R2 value of 0.896Figure 3. This ensures that as a book’s size increases, its weight increases at a similar rate.

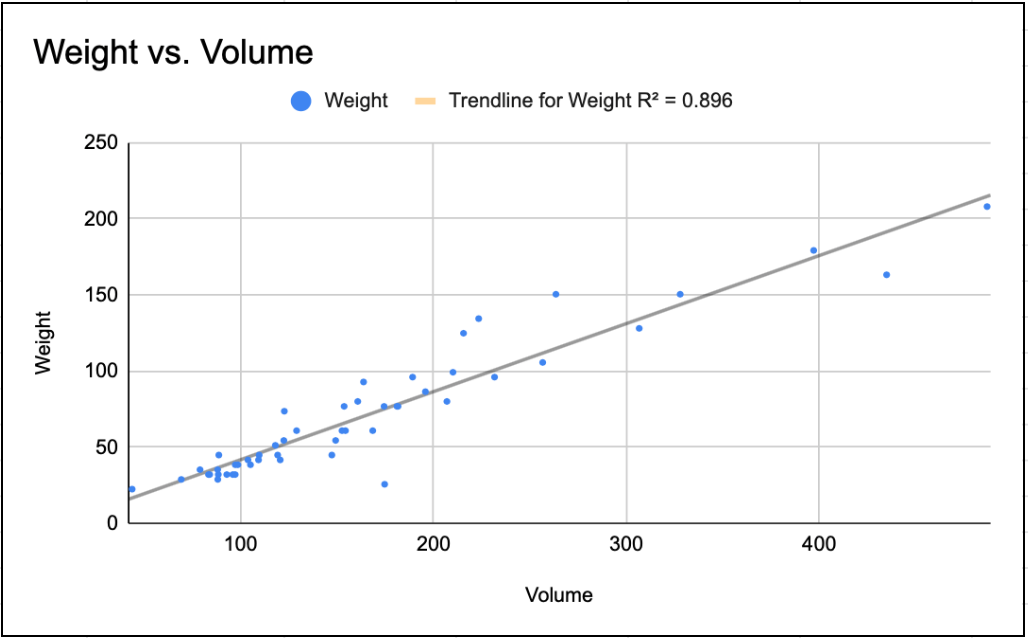


Figure 3 - Correlation: Weight vs. Volume

These results gave further validity to our data collection process as being feasible and accurate. Since the library was very clear about their requirement to store all possible books in the provided storage boxes, the close relationship between weight and volume of the books was an integral checkpoint in the process of the data collection. It also sped up the data collection process, as we were able to only measure the spines of the book, and were then able to get estimates that could be applied to an entire room.

**3.2.2 Individual Room Data**

**3.2.2.A Low-Ceiling Room**

The majority of contents in the Low-Ceiling room consist of books and bound periodicals. Both types of items fit well within the main structure of our data collection process. Therefore, our data collection method was applied. The sample size was limited to 50, while the total count of times was 3,971Figure 4. Figure 4 represents the data collected for the Low-Ceiling Room, including the amount of books per category and their respective proportions of our sample. Half of our sample (50%) was categorized in the Small-Mid group representing the largest valueFigure 4. Mid-Large and Large had the least amount of items categorized within them, both having 3 items representing 6% eachFigure 4.

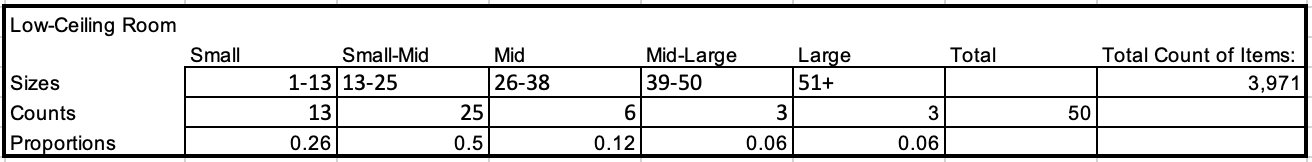


Figure 4 - Low Ceiling Room

**3.2.2.B Periodicals**

The size of the room known as periodicals was the largest collection of books. Nearly all of the area in the room was taken up by movable shelves jam packed with books. Our first step was to discover the total amount of books in the room by combing the rows with counters to keep track of the large number of items. We discovered a total of 17,363 (considering human error)Figure 5. After doing so, we went back through each row, using the dice method for randomization, to measure 100 books. During the planning for the process of data collection in this room we followed the same methods as others. However, we do worry that our sample size is too little to represent the total number of books in the room (~ 0.57%). We would like to suggest further research to consider this and evaluate a proportional sample size. The results gathered for this room are represented in Figure 5, and were in the Excel/GLPSOL linear programs.

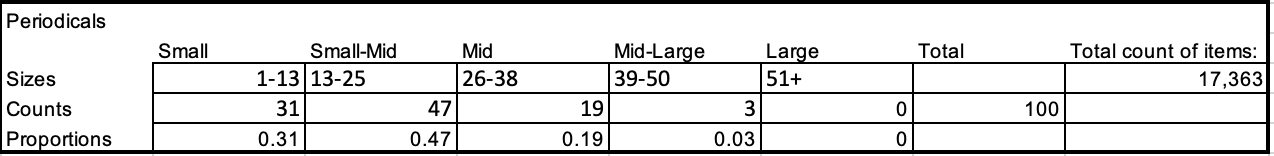
****

Figure 5 - Periodicals

**3.2.2.C Room 151**

According to the Library staff, Room 151 represents one the largest amounts of books per room. The majority of items within the room are books, many of them religious texts, and boxes from Senator John Tower. These books vary in size, shape and weight. This represented a challenge as we had to adjust our collection process slightly to incorporate them in our analysis. In total we counted 833 boxes of varying size and 5,178 books in Room 151Figure 6. Data on books was collected similarly to the other rooms. Our team sampled a total of 100 books, 50 (50%) falling into the Smallest category, 32 (32%) into Small-Mid. Mid-Large and Large represent the smallest category both only having 4 (4%) each of our sampleFigure 6.

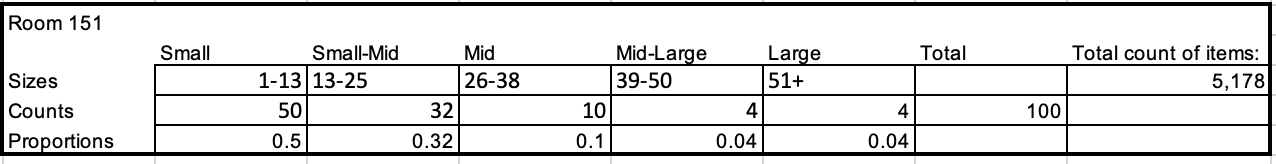


Figure 6 - Room 151

**3.2.2.D Clark Texana**

Although all rooms held subjects of books and items, this room was specific towards books related to Texas history. We recorded a total of 4688 books in this room which made it one of the smaller onesFigure 7. From the sample size collected only the height and width of the spine was recorded to find its area. What we found to be housed in the Texana room was 55% small, 40% mid-small, 4% mid, 1% mid-large, and 0% largeFigure 7.

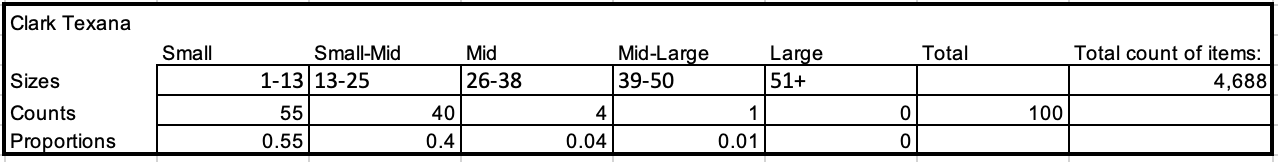


Figure 7 - Clark Texana

**3.2.2.E Reading Room**

The Reading Room is the most unique and the flagship room of Special Collections and the library. It displays collections from Senator John Tower’s journey, various books and other prime artifacts. As our focus is books, and specific boxes, much of the Reading Room’s unique items were left out of this analysis. Our team counted a total of 2,675 booksFigure 8. This is a low number compared to other rooms, and can be explained by the amount of other items that make up a large section of the Reading Room. Therefore, we sampled 50 books rather than 100. This is a limitation of our research, and could be improved on with more extensive data collection. The largest amount of data falls in the Small-Mid category with 24 (48%) booksFigure 8. Mid-Large and Largest are the smallest categories, representing 2% and 0% of our sample respectivelyFigure 8.

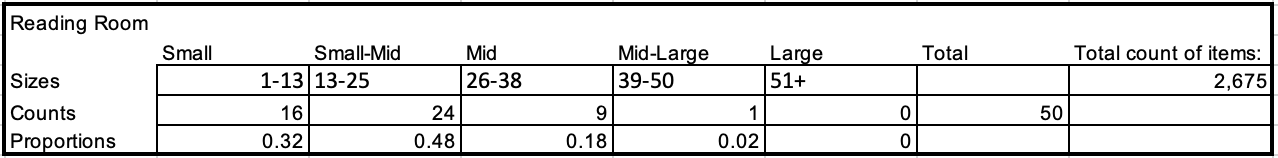


Figure 8 - Reading Room

**3.2.2.F 3rd Floor Room**

The 3rd Floor technically has two rooms, but we combined them for the sake of simplicity, as one was relatively small. A large share of the 3rd Floor Room consisted of books and it’s one of the more obscure rooms in Special Collections. Our team counted a total of 3,558 booksFigure 9, making it the second smallest room based on books counted. The individual books were ranging entirely from Smallest to Mid. The Mid-Large and Large categories did not have any data points. Small and Small-Mid had an equal amount of books at 36 (36%), while Mid had 28 (28%)Figure 9. This room is another unique collection, and contributes greatly to Special Collections’ variety.

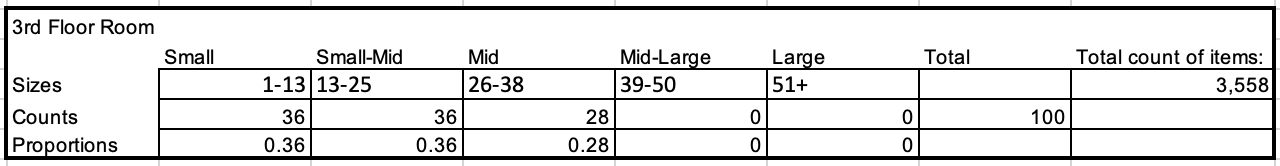
****

Figure 9 - 3rd Floor Room

**3.3 Model Formulation**

Formulating the model was broken up in two main parts: sorting books into boxes, and allocating the boxes into the rooms. There are two files that house the linear programs that calculate these respectively, and then several other helper files that we have used to aid us in calculations. These files are: DimensionsOfContents, RoomContents, BooksToBoxes20BoxModel.lp, BooksToBoxes40BoxModel.lp, and BoxesToStroage.

**3.3.1 Books into Boxes**

The first part of our plan was calculating how many boxes it would take to fit all of the books into the boxes provided by the university. These boxes are 10 inches x 12.5 inches x 15 inches, or 1,875 inches cubed in volume, and they can hold up to 40 pounds. These would both serve as constraints for our linear program, but the volume would almost always be filled up before the weight limit was reached. It’s also important to note that while there were other items in the rooms, this model was made specifically for putting books into boxes, as they took up by far the most amount of space in Special Collections. Document boxes and binders were also taken into consideration, but since they are already protected, and relatively large, they don’t need to be put into these boxes with the books.

We chose to use the program Glpsol to calculate the amount of boxes needed because we ran into various issues with Excel. Excel has a limit of 200 changing variable cells, however, because many cells called on one another, this bogged it down and would crash with programs larger than five boxes large. Glpsol is more lightweight and can handle much larger programs because it was specifically designed for linear programming. It is a bit harder to set up, but it gives us our desired results. Set up and usage of Glpsol can be found in Appendix A.

Our goal was to minimize the amount of boxes used to store all of the books, which also implicitly maximizes the amount of books per box. Because each room needs at least 250 boxes to store all of the books, we have upscaled the weights and volumes of each box by ten so that each box in the model actually represents 10 boxes. This saves us from having programs with thousands of boxes to fill. This means each “box” can hold 400 pounds and 18,750 square inches, these numbers can also be modified to increase or decrease the amount of boxes they represent.

Our first set of constraints are making sure that the weights of the books in any given box do not exceed 400 pounds. A similar set of constraints underneath it represents the volumes not exceeding 18,750 inches cubed. Each box in the program has one of each. After the weights and volumes are accounted for, A constraint for each size of books ensures all books of each size are accounted for. These are the size categories from above. A final constraint column documents whether or not a box is used by adding all of the books in a given box and subtracting 10,000 multiplied by a binary variable. These binary variables indicate whether or not a box is used, and there is also one for each box in the program. These binary variables are also minimized in the object function, meaning that an optimal solution is found when the lowest number of boxes are used. In addition, There are also a great deal of changing variables, five for each box to represent the count of the size of books that go into each individual box. These changing variables are not bounded to be integers due to other software limitations in Glpsol. On occasion it may give a fractional value, but we feel it is easy enough to round up and continue to use the data instead of waiting for it to find a genuine integer solution. Further information on usage, customization, and reading results can be found in Appendix A.

**3.3.2 Boxes to Storage**

Once we have calculated the amount of boxes of books we will have, we can apply that, as well as the counts of the other types of boxes, into the second linear program. We used a transportation model to take all of the items out from Special Collections and surroundings and put it back into the 4-5 rooms designated rooms in the library, the sizes of which are calculated int the Excel file. We used Excel for this problem because the problem was simple enough for Excel to handle it, and it is easier to visualize a transportation problem this way.

The top rows add the amount of boxes of books per room together to be used as one of the content types. The other types of boxes we took into consideration were document boxes, small document boxes, thin boxes, binders, and large boxes. Each was measured and counted per room. These sizes can be found in the DimensionsofContents file, and the counts for the books can be found in the BoxesToStorage file. The counts of each individual box type can be altered on the right side of the flows section, all except for the amount of books, as that is calculated above. The gray boxes represent the changing variables, which must equal the right side to ensure all of the contents are moved from Special Collection and its surroundings. All of the dimensions for each of the different contents are in the Volume Considerations matrix which is used for calculating the total volume needed for storage.

Above the flows section are the calculations of the room sizes. We split them up by their dimensions to allow for easy alterations in the event that boxes can’t be stacked higher than a certain level (We have it set to 60 inches at the moment), or if not all the space can be utilized solely on storing items. The Prothro Room is in the shape of an octagon, so it is not calculated the same as the other rooms. Vertically, the changing variables represent the counts of the contents going into that room, and it cannot exceed the lowest column, which represents the total working volume for the room. These are multiplied by the Volume Considerations matrix to get the volumes and ensure they don’t exceed the volume in the given room. Users can alter which rooms are and aren’t used by simply changing the volume available to 0. The objective adds up all the volume used and maximizes it, ensuring every item is accounted for. More information on usage can be found in Appendix B.

**3.3.3 Additional Files**

The other two files we provided have various reference values used throughout the rest of the paper. DimensionsOfContents has the dimensions of all of the different types of boxes that we have measured. RoomContents has the entire contents of each room, including the amount of boxes needed to store the books, as well as an estimate for the count of any other types of boxes.

**3.4 Results**

This section maps out the results we found purely from running our linear models. Interpretations will come in the next section.

**3.4.1 Books into Boxes Results**

After collecting data, plugging it into our program and running it, we were able to determine the estimated amount of boxes needed to store special collections and its surroundings. They are as follows: the low ceiling room requires 490 boxes, Clark Texana needs 376, periodicals needs 1798, room 151 needs 480, the reading room needs 271, and the third floor needs 367. This brings the total number of boxes needed to 3,701, at least by our estimates.

**3.4.2 Boxes to Storage Results**

Once we calculated the total number of boxes needed, we could plug that, as well as the counts of the other types of boxes into our other program. The counts are specified both in the BoxesToStorage file, as well as in the Room Contents file. When plugging in the contents, we are able to see that with Rooms 201, 207, 224, and 254 available, we are able to fit everything into rooms 201, 207, and 254. And this is without considering the Prothro room as a possibility. Despite the Prothro room being much bigger, it cannot store all of the items, but it could likely store about two room’s worth of content.

**3.5 Conclusions and Recommendations**

According to our calculations, all of the contents that we have measured in special collections will fit in the allotted spaces, oftentimes with at least an empty room to spare. While this is encouraging to hear, there is a slight caveat that there are additional items that still need to be accounted for, such as artifacts that are not box-like in shape, or items that are more fragile. This extra empty room should still be taken into account for all of the additional items that still need to be accounted for and stored.

With this margin of error taken into account, our recommendation for storing special collections is that at a bare minimum, three of the rooms will need to be cleared out to store everything. If possible, a fourth is highly recommended to be able to store everything safely. If the Prothro Room is available, it would be able to hold a lot more boxes than the other rooms, so it would be beneficial and recommended, but not necessary. In addition, as long as these rooms are available, off campus storage will not be needed, which will save the university thousands of dollars. As long as rooms 201, 207, and 254 are cleared and usable for storage, the university should have no problems fitting all of Special Collections into the main section of the library while Special Collections undergoes its renovations.

**Honor Code:**

We have acted with honesty and integrity in producing this work and are unaware of anyone who has not.

X. Henry Fisher X. Clayton Nolen

X. Mark Mueller X. Kyle Porter

**Appendix A**

Appendix A focuses on installing and using Glpsol for running our Books to Boxes program.

**A.1 Installing Glpsol**

Installing Glpsol on your machine is different depending on if you are running Windows or MacOs. In depth instructions on how to install Glpsol can be found here: <https://kam.mff.cuni.cz/~elias/glpk.pdf> in Appendix A. If you have any issues, feel free to consult any of the team members for additional assistance.

**A.2 Using Glpsol**

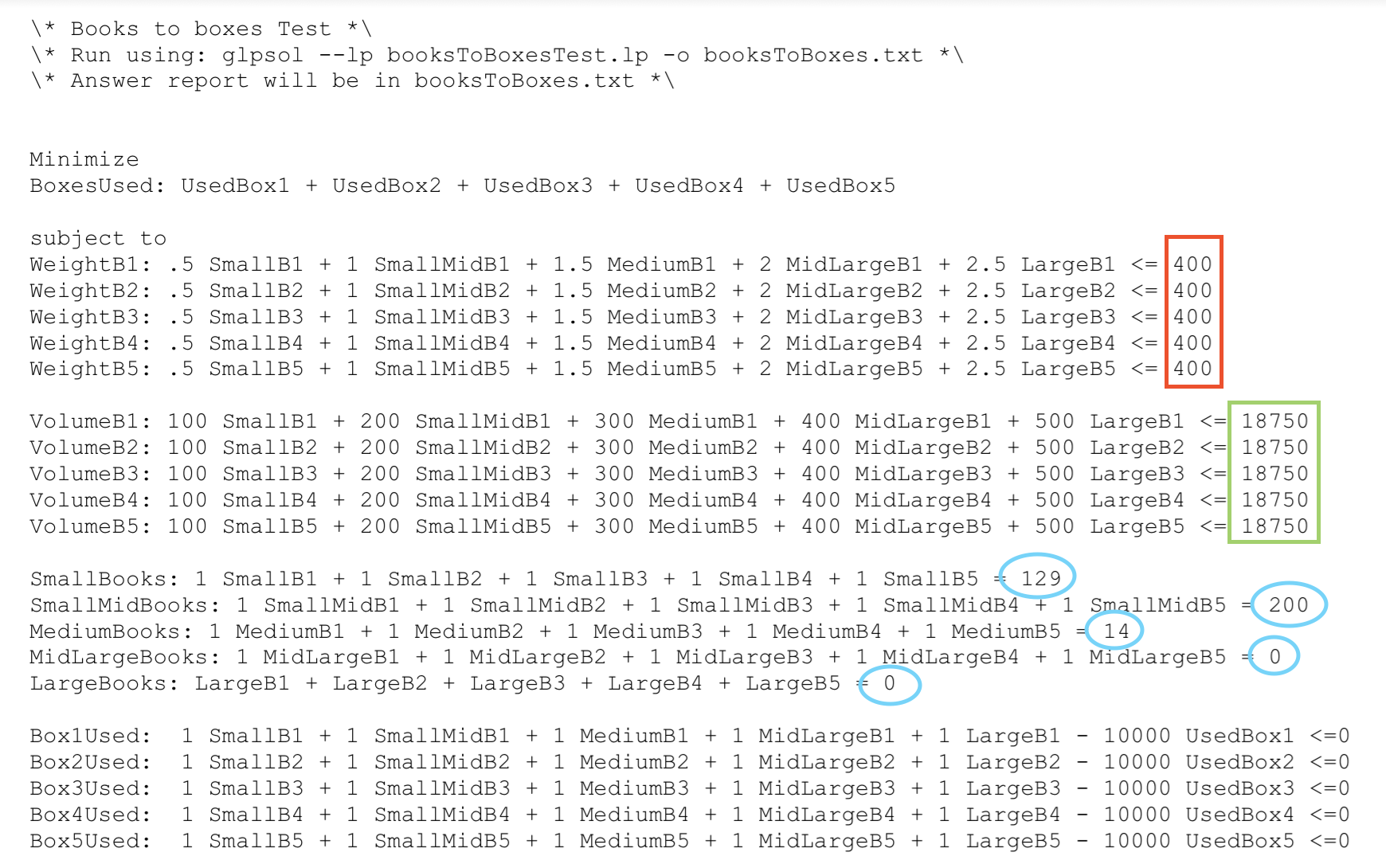
This guide will be tailored to MacOs users, as that is what was used for the project. The process for running on Windows shouldn’t be much different, but the guide above should also be able to aid in the event of confusion.

First, ensure that you are in the terminal and that you have a .lp file saved to your computer. An .lp file is just a .txt file that Glpsol traverses through to run its program. In the terminal, type the command: glpsol --lp /Path/To/file.lp.txt -o [fileToSaveTo].txt. –lp is used to indicate the file you want to run the linear program on, and -o is to specify the output file.

**A.3 Editing Books to Boxes**

We have provided you with two files for running calculations on books to boxes: BooksToBoxes20BoxModel.lp and BooksToBoxes40BoxModel.lp. Functionally, they are the same, but the 40 box model is twice the size of the 20 box model, and it is also much slower at calculating the linear program. Our group has found the most success in splitting up large linear programs into smaller ones, and then combining the box count to avoid long wait times on the large program. There are various aspects that can be altered in the program and they can all be changed in the .txt file. Simple programs such as TextEdit or Notepad are recommended, as more complex programs like Word can add special characters that will make Glpsol crash.

**A.3.1 Editing Constraints**

The most common and basic edit that will be made will be changing the constraints. Namely the volume, weight, and count constraints. Below is an image of a five box linear program that is being used to illustrate what each change will affect:

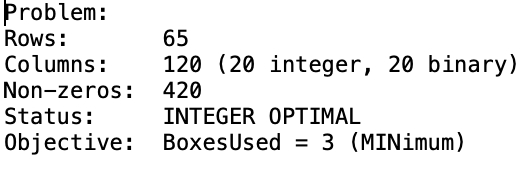
The red box represents the weight values, and the green box represents the volume. If one is changed, the other should be changed to the same magnitude as they are interconnected. Also, all values in the red box should be the same, as should all the values in the green box. All of these need to be changed manually, unfortunately. This is due to certain software limitations posed by Glpsol. Each of the values circled in blue represents a count for a respective book size. These can all be different, as seen here, as it represents the number of each size of book that needs to be put into boxes. The rows of Box#Used should not be altered, as it is a binary flag that registers whether or not a box is used. The only time it will be altered is in the event you want to add additional boxes.

**A.3.2 Adding Additional Boxes**

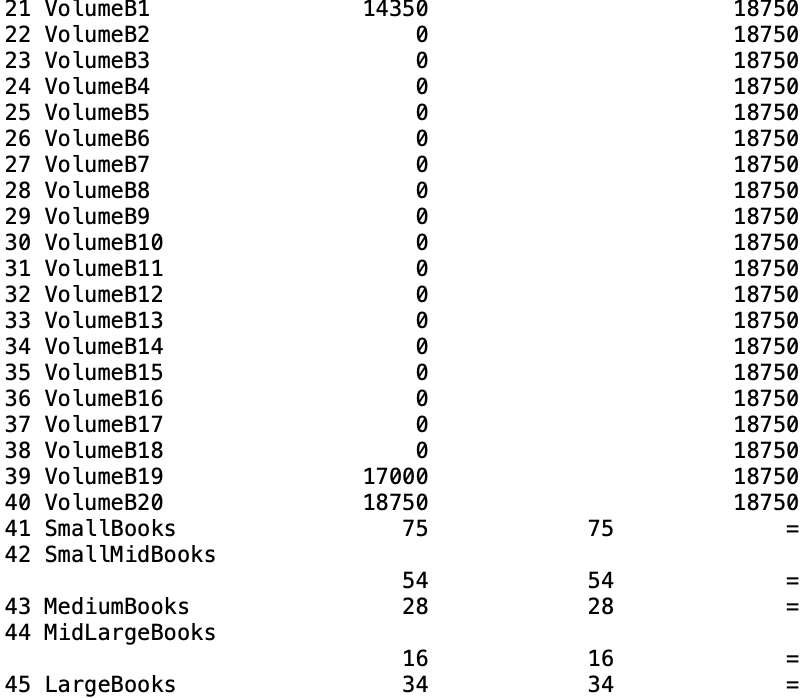
Adding additional boxes is a much more arduous task as it alters every section of the code in some way. We have found it easier to start from the bottom and work our way up through the code. For every box, you’ll need to add six variables, one binary, which is UsedBox#, and 5 under bounds: one for each size. The easiest thing to do is copy the 5 listed and increase the box number by one. Then we get into constraints. You’ll want to add a Box#Used constraint and make sure all the variables correspond to your newly added ones. The 5 rows above that all focus on accounting for all the books in each size, so for each of the five constraints, add a copy of the newly created variables for each size. A weight and volume row will also have to be added for each box, and the ordering can be followed exactly as seen in the examples above. The final thing that must be added in a copy of the binary variable in the objective. If it is not added it is not accounted for in the solution.

**A.4 Interpreting Results**

After the program is run, the results are outputted to the .txt file that was specified. Below are some images to aid in understanding how to interpret the results.



These are the first 5 rows, they showcase the amount of rows, columns, and variables used, as well as the solutions, which is listed under objective. We can see the boxes used in this example is 3, which actually means 30 boxes were used to store all of the books, due to the upscaling. Below this are the volume and weight . The first column represents how much is going into the boxes, and the second column represents the weight/volume limits. This example is using the 20 box model, thus there are 20 rows of weights and volumes.



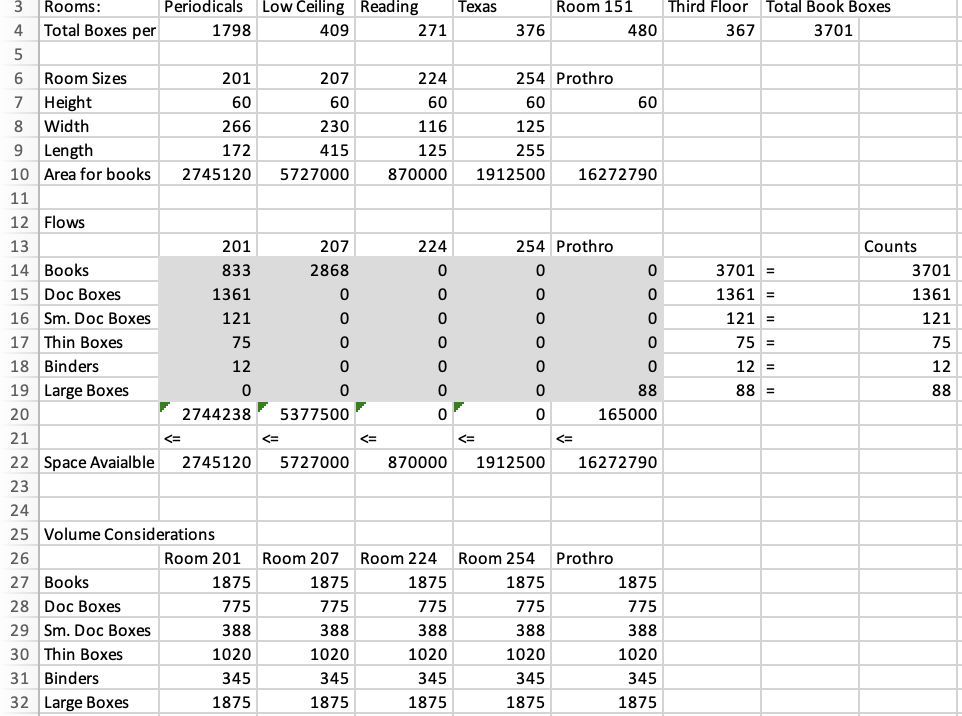
Below that we can see the counts of each of the books, and that the two columns are identical, which ensures all the books specified were accounted for and were put into boxes. The rest of the rows are usually not as informative. Below this are the binary value calculations, which if it is a large negative number, then it was one of the boxes used. Below that are each of the variables and their values, signifying how much of each book goes into any given box. One could follow this model exactly, but realistically, getting an estimated number of boxes was the main goal, and the packers can pack the books as they see fit.

**Appendix B**

Appendix B focuses on using the Boxes to Storage program. Make sure to have Excel installed on your machine, as well as having the Solver plugin installed.

**B.1 Editing Boxes to Storage**

There are lots of values that can be edited in this program to ensure the optimal solution is met for any needs. Below is an entire overview of the program for reference.



Row 4 represents the count of boxes of books, which can be changed if more exact, or additional boxes are added. This sum is put into the counts column at Row 14. The rest of the counts are editable by the user to account for increases or decreases in contents. Above the flows column, we can see the calculations for the total area of each room. They are easily editable, which are then reflected in Space Available in row 22. Note that the Prothro Room is octagonal, meaning it is not calculated the same way. In the event that one does not want to consider a room in the calculator, zero out the value in row 10. This signifies to Excel that there is no room for anything in the room, so nothing will be put in. Volume Considerations have all of the volumes for each different type of item and are able to be modified, but it is unlikely that it will be needed. Below this is the total volume, which is the objective function. It ensures all the volume is accounted for, and how much volume the contents make up.

**B.2 Adding Additional Rows or Columns**

Adding a row represents adding a new type of contents needed for storage, and adding a column represents using an additional room. Both are realistic to be done and need to be handled cautiously to ensure it is done correctly. Adding a row requires shifting rows 20-22 down by one value. A name will need to be added in the first row and the next 5 columns will be changing variable columns. The next column will be the sum of the previous columns which will equal whatever count is set. In volume considerations, you will need to add the volumes the item takes up and then you’ll also have to update what was in row 20 by adding the additional row to its sum product. In Solver, make sure that the changing variables are updated to include the new row, and that they are integral as well. And finally, ensure that the new row is added to the constraints column.

Adding a column follows the same steps, but it instead causes the columns to shift over, and the value in row 20 will be the sum product with the Volume Considerations matrix, rather than just the sum. All of the columns to the right need to account for the new column in their sums. And all of the considerations still need to be made in solver as well. All of these steps need to be done at once to ensure the program works. Feel free to contact any of the group members if additional assistance is needed.