**Contact**

Teresa Neely, neely@unm.edu
Steve Koch, sjkoch@unm.edu

**Types of data files**

As of July 24, 2014, the data in this study are stored in .xlsx files, readable at a minimum by Microsoft Excel 2013. Alternatives for reading the data files are to import into an open source spreadsheet application, such as Open Office, or to read the files with open source libraries such as the `xlrd` module for python 2.7.

**Types of data spreadsheets**

There are three types of data spreadsheets in this study, identifiable by their column names or filenames:

1. **Bookshelves.** Columns include: Ranges, Sections per range, Shelves per section, Total number of shelves, empty shelves, linear feet, empty linear feet

2. **Cabinets.** Columns include: Number of cabinets per stack, cabinets per stack, number of cabinets, # drawers per cabinet, etc.

3. **Summary sheets.** These sheets total and sub-total data. The Excel files have links to cells in other named files. These files are identifiable by their name “Summary by...” and have columns such as “occupied shelves” that are not in the raw data files.

**Bookshelves spreadsheets**

The majority of the data in this study represent counts of bookshelves. The name of the Excel sheet describes the location, in combination with the filename. The location is also specified in the upper-left hand part of the sheet, usually cell A2 or A3. For example, in the file “Zim Government Information.xlsx”, there is a sheet named “ZimmB16.” The value of cell A3 is B16-Compact Shelving. A familiarity with UNM Libraries would enable one to determine that these counts are for Zimmerman Library, room B-16.

**Bookshelves column descriptions**

- Name of Collection – This column is not ideally named, and sometimes contains information about the collection described by that row of data, and sometimes contains notes to identify a change of aisle in the particular location, or beginning and ending call numbers for an aisle.
• Functional Responsibility – This column is often left blank and information is missing. It is not essential for the calculations. When present, it indicates who (as of Spring 2014) is responsible for the collection described by that row of data.

• Ranges – This column is a numeric value that indicates how many ranges are described by the data row. For this study, a “range” is one side of a bookshelf or compact shelving carriage. For example, a compact shelving carriage usually has two sides and thus two ranges. Shelving against a wall is single-sided and is one range. Most often, a range refers to one continuous shelving unit—for example, a compact shelving carriage or a single cantilever stack unit. For wall shelving, sometimes “range” refers to the entire face of the wall. Ambiguity can be reduced by reading the information in the “location” column, described below.

  – This column is summable. Row 2 of the spreadsheet has a value indicating the total number of ranges counted on the sheet.

• Sections per range – This column is a numeric value representing how many sections of shelves are in the range. In most cases, even for double-sided stacks or carriages, this is a valid number. A section has shelves stacked vertically, for example, six three foot shelves in a section. In rare cases, ranges have been grouped that do not have a common number of sections. In these cases, the value for this column is “n/a.” This column does not have a sensible sum.

• Shelves per section – This column is a numeric value representing how many shelves are in each section of the range(s) represented. In most cases, this is a valid number that is the same for all ranges described in the row. In some cases, range(s) have varying numbers of shelves per section, and the value is indicated as “n/a.” For example, many stacks have display shelves that are not used for book storage, so a range may have 6 sections with 6 shelves per section and 1 section with 5 shelves per section. Ambiguity can be reduced by reading the information in the “Location” column, described below. This column does not have a sensible sum.

• Total number of shelves – This column represents the total number of shelves for the range(s) described in the row. Most of the time, the preceding three columns have valid numerical values, and this column represents the product of those three values and thus the total number of shelves in the range(s). In cases where one of the three preceding columns has “n/a” as a value, the formula has been modified to account for this with hard-coded numerical values in the cell’s formula. Attempts have been made to color these cells blue (in the Excel format) and ambiguity can be reduced by reading the information in the “Location” column described below.

  – This column is summable. Row 2 of the sheet represents the total number of shelves accounted for on the sheet.

• Empty shelves – This column represents how many empty shelves were observed in the range(s) represented by the row of data. This is usually entered as a numerical value by hand, and often is noted in the “Location” column as well. In some cases, the value for “empty shelves” is calculated based on a sum noted in the “Location” column, or based on a percentage of the total shelves. For example, for a completely empty range, this value may be set to “=[total number of shelves]”. For this study,
to be counted as empty a shelf must have been completely empty of collections. We
did not count fractional emptiness of a given shelf.

- This column is summable. Row 2 of the sheet represenets the total number of
empty shelves accounted for on the sheet.

• Linear feet – This column represents the total linear feet of shelves for the range(s)
described by the data row. Almost always it is a simple calculation of [total number
of shelves] * [length of a single shelf]. Most often, the shelf length is 3 feet and is
hard-coded into the formula. In a few cases, a range or group of ranges had variable
shelf lengths. These situations were accounted for and attempts were made to color
the cells blue (in the Excel format) and to provide a description in the “Location”
column.

- This column is summable. Row 2 of the sheet represents the total linear feet of
bookshelves accounted for on the sheet.

• Empty linear feet – This column represents the total linear feet of empty bookshelves
for the range(s) described by the data row. It is calculated in the same manner as
“linear feet” above, with the same considerations for shelf length and odd shelf lengths.

- This column is summable. Row 2 of the sheet represents the total linear feet of
empty bookshelves accounted for on the sheet.

• Location – This column is not ideally named. It contains free-form text that usually
describes the range(s) represented by the data row. This information was typed while
data were being collected, so it is terse and varying in style. It will often note the
numerical label of the stack(s) or carriage(s). For double-sided stacks or carriages, we
used the convention “20.a” to represent the range closest to stack/carriage 19, and
“20.b” to represent the range closest to stack/carriage 21. The values for sections per
range and shelves per section are usually noted in this column, often using shorthand
such as “12x6” meaning 12 sections and 6 shelves per section. Other notes are also
sometimes included, such as notes about the collection, or the condition of the shelves.
For the CSEL collections, damaged shelves are colored yellow in this column (in the
Excel format).

- This column is not summable.

- This column is also used at the top or the bottom of the sheet for general infor-
mation about the room or the spreadsheet. For example, noting that the stacks
are numbered from east to west, starting in the southwest corner.

Cabinets spreadsheets

A minority of the data in this study represent counts of cabinets. The name of the Excel
sheet describes the location, in combination with the filename. The location is also usually
specified in the upper-lefthand part of the sheet, usually cell A2 or A3. For example, in the
file “Special Collections_all.xlsx”, there is a sheet named “B3 Cabinets” The value of cell
A2 is “B3 Cabinets.” A familiarity with UNM Libraries would enable one to determine that
these counts are for the cabinets in Zimmerman Library, room B-3. Compared with the
bookshelves data, these data are not as uniformly collected and thus not as well organized.
Contact the authors with questions that are not explained by the descriptions below.
Cabinets column descriptions

- Name of Collection – see same column in Bookshelves above.
- Functional Responsibility – see same column in Bookshelves above.
- Number of Stacks – This column represents the number of stacks of cabinets described by the row of data. A stack almost always contains cabinets of the same surface-area dimensions. For example, a stack of 3 map cabinets is common. Many “stacks” contain one cabinet, usually this represents an office-type file cabinet.
  - This column is summable. Row 2 of the sheet contains the total number of stacks accounted for on the sheet.
- Cabinets per stack – This column represents the number of cabinets in each stack represented by the row of data. Usually this is a numeric value that indicates the uniform number of cabinets in each stack. Occasionally, a row of data represents a collection of stacks that do not have uniform numbers of cabinets. In these cases, the value is “n/a.” This column does not have a sensible sum.
- Number of cabinets – This column represents the total number of cabinets represented by the row of data. Usually it is a calculation of the two preceding columns = [number of stacks] * [number of cabinets per stack]. In some cases where number of cabinets per stack is “n/a”, the formula is hard-coded.
  - This column is summable. Row 2 of the sheet for this column indicates the total number of cabinets accounted for on the sheet.
- # of drawers per cabinet – This column represents how many drawers each cabinet has. This column does not have a sensible sum.
- # drawers in stack – This column indicates the total number of drawers represented by the data row. Often it is a calculation of the two previous columns = [number of cabinets] * [# of drawers per cabinet]
  - This column is summable. Row 2 for this column represents the total number of drawers accounted for on this sheet.
- Width of cabinet in inches; Depth of cabinet in inches – These columns are numeric values representing the width and depth of the cabinet for the purpose of calculating the footprint of the stack(s). In many cases, these values were inferred from the textual data in the “location” column, and in some cases “depth” and “width” may be transposed, which did not have an impact on our calculation of surface area. These columns do not have a sensible sum.
- Footprint of stacks, sq. ft. – This column represents how much floor space (minimum) is occupied by the stack(s) represented by the data row. It is calculated as = [width of cabinet in inches] * [Depth of cabinet in inches] / 12 / 12.
  - This column is summable. Row 2 for this column represents the floor space (minimum) occupied by the cabinets accounted for on the sheet.
• Location – See the “Location” description for bookshelves above. Additionally, the “location” column for cabinets often contains free-form data describing the width, depth, number of drawers, possible contents, color, etc. This column does not have a sensible sum.

Summary spreadsheets

There are some spreadsheets that aggregate data from the individual sheets for shelves or cabinets. The two most important are “Summary by Location.xlsx” and “Summary by location_cabinets.xlsx”

“Summary by Location”

This sheet summarizes bookshelves by building and named room or location in the various UNM libraries. The columns are similar to Bookshelves, described above

• Column A (No Name) – This column is text that indicates the location described by the data row.
• Ranges – See same column name for bookshelves above
• Total number of shelves – See same column name for bookshelves above
• Occupied shelves – This column represents how many of the shelves are non-empty for the particular row of data. It is calculated as = [Total number of shelves] - [empty shelves]
• empty shelves – See same column name for bookshelves above
• Total linear Feet – See same column name for bookshelves above
• Occupied linear feet – This column represents the linear feet of totally empty shelves for the particular row of data. It is calculated as = [Total linear feet] - [Empty linear feet].
• Empty linear feet – See same column name for bookshelves above
• Percent empty – This column indicates what the percentage of empty linear feet relative to total linear feet for the data row. Note that this column is based on linear feet, not number of shelves.
• Row 43 represents a grand total for all bookshelves counted.
• Below row 43 are various other subtotals used for data analysis in the accompanying report.
“Summary by location_cabinets”

This sheet summarizes cabinets by building and room or location

- **Column A (No Name)** – This column contains text that describes the location of the cabinets. There are two buildings, Zimmerman and CSEL. The column indicates the building with centered and bold text (in Excel format).

- **Total number of cabinets** – This column represents the total number of cabinets for the location represented by the data row.

  - This column is summable. The column contains subtotals in row 9 for Zimmerman in row 14 for CSEL and a grand total in row 16.

- **Total number of drawers** – This column represents the total number of drawers for the location represented by the data row.

  - This column is summable. The column contains subtotals in row 9 for Zimmerman in row 14 for CSEL and a grand total in row 16.

- **Total footprint, sq. ft.** – See definition above for cabinets. This column represents the minimum floorspace occupied by the cabinets represented by the data row, in square feet.

  - This column is summable. The column contains subtotals in row 9 for Zimmerman in row 14 for CSEL and a grand total in row 16.