Guide to recreate RNTV-EnumerateKey.xlsx

Steps

1 Load data from csv file

Import data from RNTV-TransformedData.csv into a new worksheet named 'Data' (step 1A of RNTV-PercentRankInc.pdf). The sum of E5:E31 should be 7957, the total number of NT verses in the RSV-2CE bible.

(Sheet Data)

Optional: A 'simpler' method of loading csv that should work, but runs into a known problem.

In a new worksheet named 'Error Data', and with cell C1 selected, import data from RNTV-TransformedData.csv (Data tab >> From Text/CSV >> Import >> Load To cell C1 in existing worksheet).

With the table selected (click on any cell in the table), remove its formatting (Table Design tab >> Table Style Options >> uncheck Header Row and Banded Rows), and convert to range (Table Design tab >> Tools group >> Convert to Range).
Remove row 1 (so that the ranges match the other workbooks). Select the entire worksheet, and format column width to autofit. Delete the query connection.

Note that the 'numbers' in columns D and E sum to 0; even when the cells are formatted as number, ISNUMBER() returns false. Using PQE to load csv helps to avoid these subtle problems, at least it did with this file. Hide this sheet.

(Sheet Error Data)

2 Enumerate key

Add a new sheet to the workbook, name it 'NTVerseKey'.

Open the VBA Editor (Developer tab >> View Code), insert a new module (Insert >> Module), and type the code in Figure 1 to declare a sub-procedure named 'Enumerate_NTVerseKey'.

Execute this subprocedure (In the VBA Editor, with the cursor in the sub-procedure code, hit F5), or run the newly created macro 'Enumerate_NTVerseKey' (Developer tab >> Macros >> select Enumerate_NTverseKey >> Run).

Go to cell A7958 (type cell address in Namebox), to check that the enumeration *ends* with value 7957 in cell A7958. Adjust column widths to accommodate contents.

With this NTVerseKey data, the BCV coordinates of a random verse can be looked up directly; more elaborate search methods such as the one described in RNTV_PercentRankInc.pdf can be avoided. However, enumeration is practical only for small search spaces.

Figure 1 VBA code for Enumerate_NTVerseKey subprocedure/macro

```
Sub Enumerate_NTVerseKey()

  'Ensure active sheet is 'NTVerseKey'
  Dim target_sheetname As String
  target_sheetname = "NTVerseKey"
  If (ActiveSheet.Name <> target_sheetname) Then
    MsgBox "Active sheet needs to be " + target_sheetname
    Exit Sub
  End If

  'Get data needed for enumeration
  Dim numBooks As Integer
  numBooks = 27
  Dim books, chapters, verses As Range
  Set books = Worksheets("Data").Range("C5:C31")
  Set chapters = Worksheets("Data").Range("D5:D31")
  Set verses = Worksheets("Data").Range("F5:AG31")

  'Prepare output sheet
  'Worksheets(target_sheetname).Activate
  Worksheets(target_sheetname).Cells.Clear
  Range("$A$1").Value = "CumVerse"
  Range("$B$1").Value = "BCV"
  Range("$C$1").Value = "Book"
  Range("$D$1").Value = "Chapter"
  Range("$E$1").Value = "Verse"

  Dim cumVerse, row As Integer
  cumVerse = 1
  row = 2
  For book = 1 To numBooks
    For chapter = 1 To chapters(book)
      For verse = 1 To verses(book, chapter)
        Range("$A$" & row).Value = cumVerse
        Range("$B$" & row).Value = books(book) & " " & chapter & ":" & verse
        Range("$C$" & row).Value = books(book)
        Range("$D$" & row).Value = chapter
        Range("$E$" & row).Value = verse
        cumVerse = cumVerse + 1
        row = row + 1
      Next verse
    Next chapter
  Next book
End Sub
```

Create a new name ‘BCV’ that refers to range NTVerseKey!$B$2:$B$7958 (Formulas tab >> Name Manager). Explore I5:K17, which demonstrate conversion between ‘CumVerse’ and ‘BCV’ coordinates with INDEX and MATCH functions.

Export the NTVerseKey worksheet as .csv filetype (NTVerseKey.csv).

Save the workbook as macro-enabled (RNTV-EnumerateKeyWithSamples.xlsm).

(Sheet NTVerseKey)

3 Generate sample data

The other sub-procedures in RNTV-EnumerateKeyWithSamples.xlsm, uses the 'Data' and 'NTVerseKey' sheets from the previous two steps, to generate samples of random NT verses, using the book-first and verse-first methods.

Run the 'Generate_Samples' macro to create three samples for each method. Each sample is created on a separate worksheet. Export these sheets individually as .csv files (e.g. book-first-Sample1.csv, verse-first-Sample1.csv); delete these sheets and the query connections from the workbook.

4 Preliminary analysis of the two methods

On a new worksheet in RNTV-EnumerateKey.xlsx, load a book-first sample and a verse-first sample.

Define names BF2_CumVerse and VF1_CumVerse for ranges $A$1:$A$10001 and $I$2:$I$10001 respectively (select a few top rows of each column, including the header; Formulas >> Create from Selection; use Name Manager to adjust the cell range).

Create two histograms: one for each of the two named ranges, to compare their distribution of random verses. Change bin width of both histograms to 400 (Double-click on the x-axis).

As expected, the verse-first sample exhibits a more uniform distribution than the book-first sample (Figure 2). A uniform distribution means that each NT verse is equally likely to be selected. According to the BF2-CumVerse histogram, the book-first method is most likely to pick verses in (7201, 7601].

Follow the procedure above to create another pair of histograms: one for BF2_BookIndex, and VF1_Book Index. Configure number of bins to 27, and format Number to 0 decimal places.

Now, the BF2-BookIndex histogram shows a more uniform distribution than the VF1-BookIndex histogram (Figure 2).

Use the book_first_Sample2 table to filter BF2_CumVerse to be in (7201, 7601]. The histogram charts respond to this change in range (Figure 3); the BF2-BookIndex histogram show the NT books with CumVerse within this range. The six stand-out data points in this range point to books 1 Peter (21), 2 Peter (22), 1 John (23), 2 John (24), 3 John (25), and Jude (26), which are NT books with fewer chapters and verses.

This preliminary analysis has revealed differences between the two methods for selecting an NT verse at random. Further questions about this dataset can be asked to determine which of the two methods are more suited to a purpose, with other tools such as Pivot Table and Chart (Figure 4).

(Sheets Prelim Analysis, BF2 Pivot, VF1 Pivot)

Figure 2 The two pairs of histograms with all data.

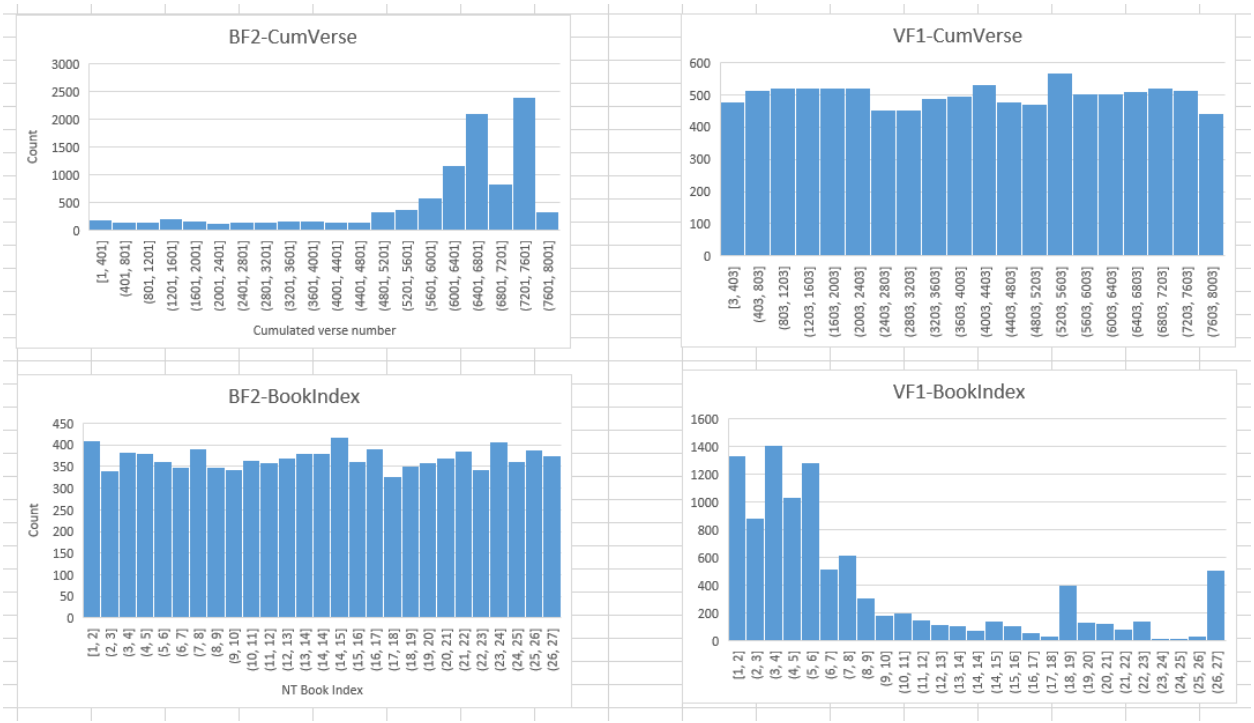

Figure 3 The histograms with filtered data.

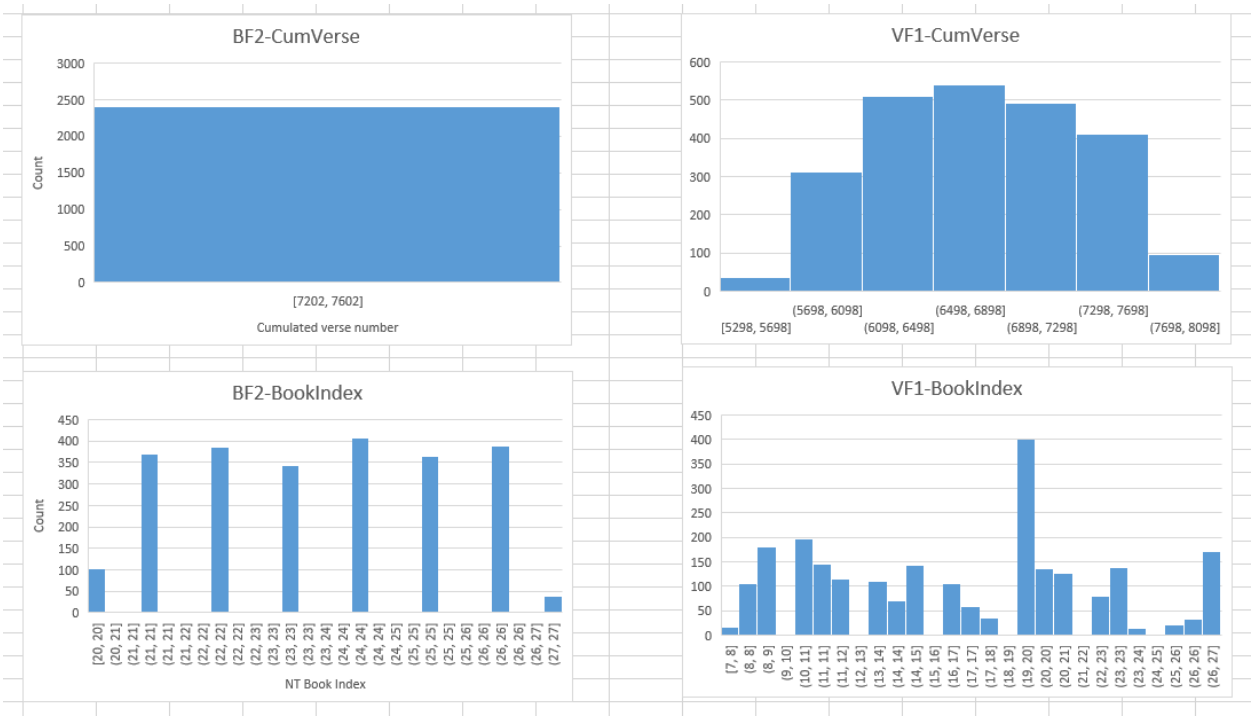

Figure 4 Example of two pivot charts created from pivot tables of the sample data.