# CONSTRUCTING RASCH SCALES

## PRE-REQUISITES

### THEORY

### SOFTWARE REQUIRED

The following software is required for the development of scales.

1) R/ R Studio\* or SPSS

2) Quest (ACER)

3) Microsoft Excel (2010 and above)

\* R and R-Studio scripts are optional as they help expedite the task of converting the raw data (files with responses) to suitable formats and automated build of script files for the Quest software. The format conversion of raw data for Quest compatible software can be done using SPSS also (Manual prepared by Kumaresh and Manjunath).

##### INSTALLATION

**For R installation**

https://cran.r-project.org/doc/manuals/r-release/R-admin.html#Installing-R-under-Windows

**R-Studio installation**

http://www.dummies.com/how-to/content/how-to-install-and-configure-rstudio.html

**R-scripts needed**

scale\_construction\_prep\_v31\_08\_2015.R

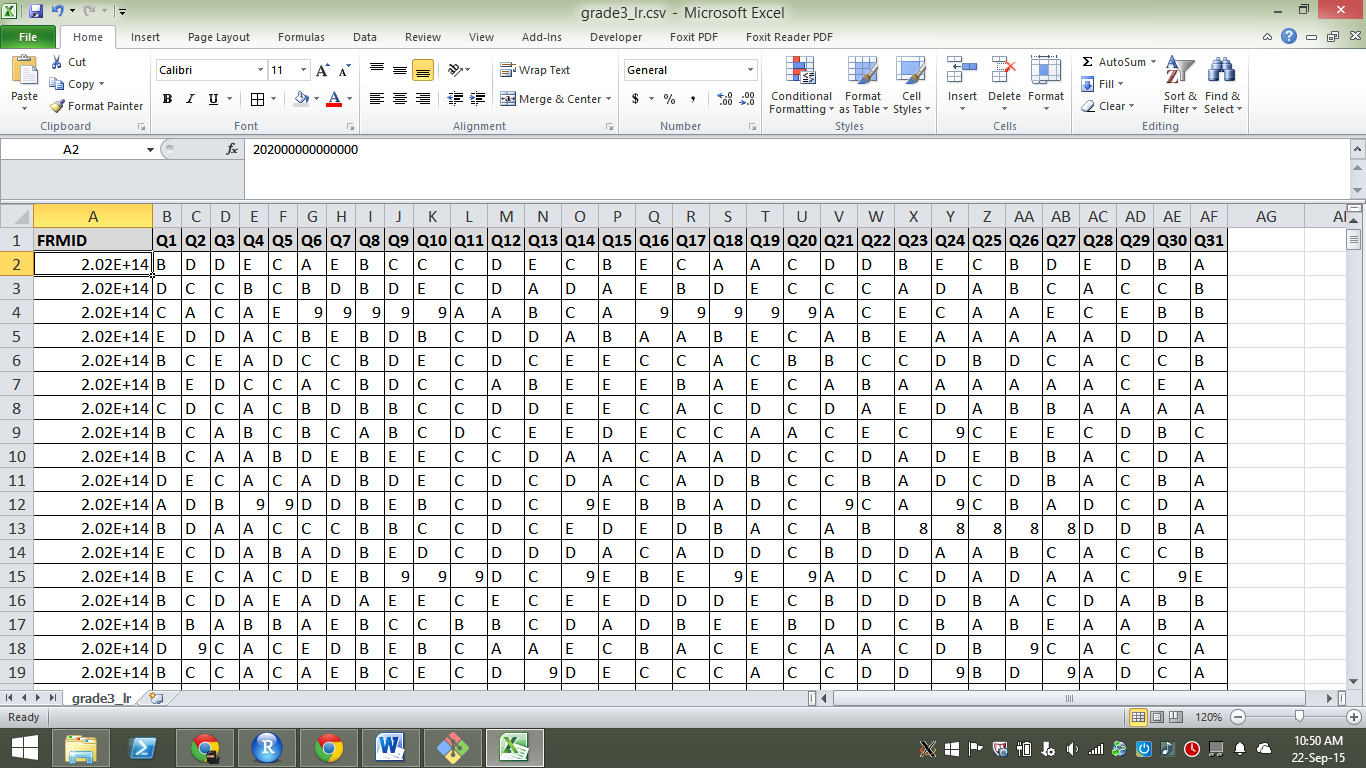
## DATA PREPARATION

The data for construction of scales should be prepared in the prescribed format.

1) Copy all the raw responses (either through OMR reading or manual tabulation) into a single sheet by grade and subject. Usually the scales are constructed for each subject by linking the grades from the lowest to the highest.

2) The spreadsheet should only have the formID(or any other identifier) and the responses to all questions. The responses on one student should be in one row.

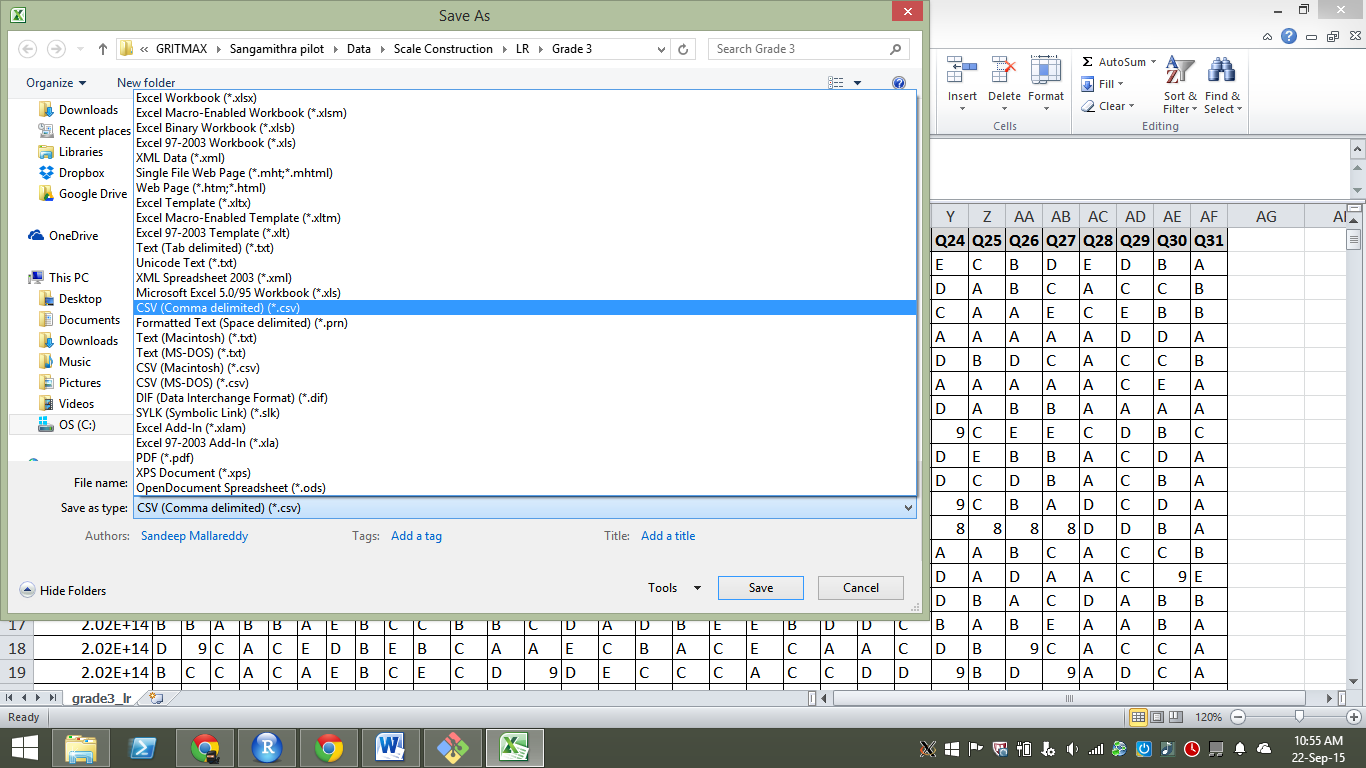
### Figure : Format of the data for scale construction



3) Save the file in the .**csv (comma separated values)**.

The file name should be ideally “grade {number}\_{subject}.csv”. *e.g.* grade3\_maths.csv

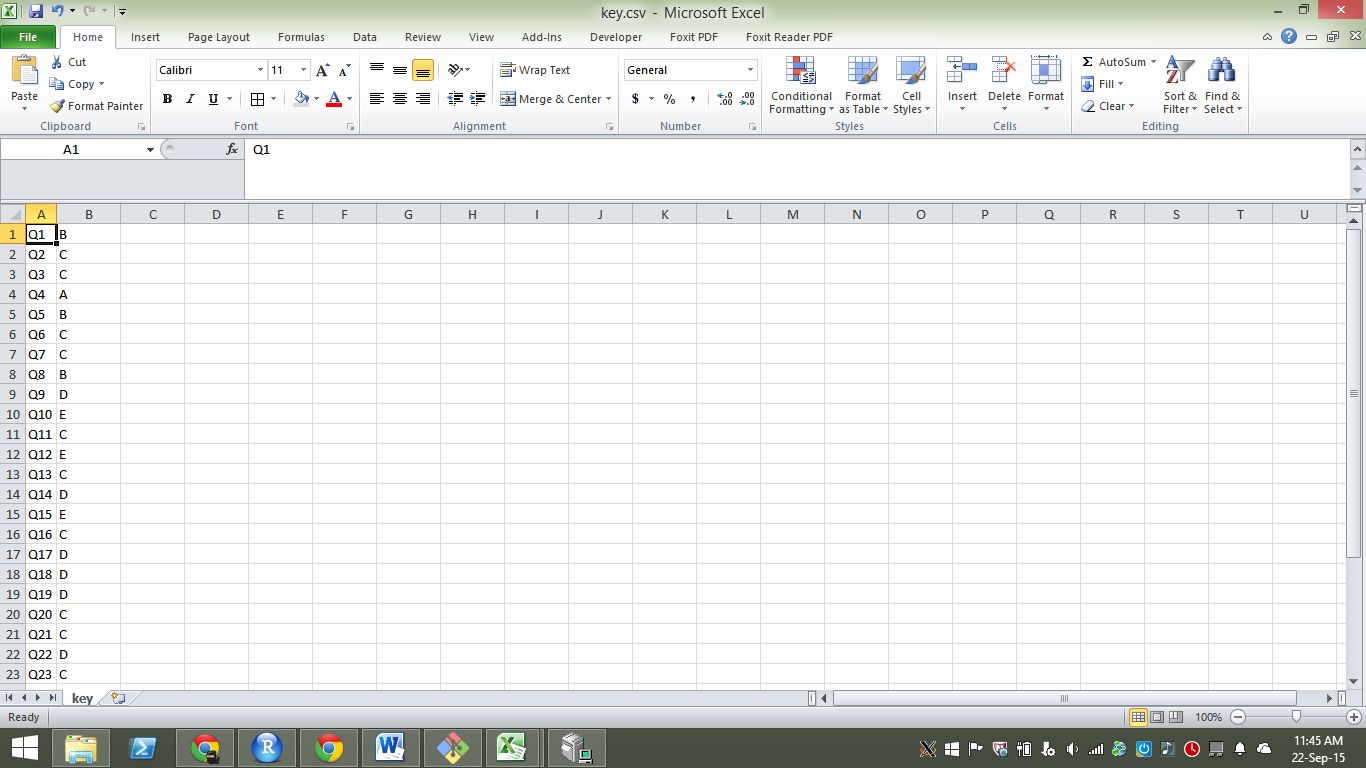
### Figure : Saving the file in csv format



**Note: The .csv format doesn’t allow more than one sheet to be stored.**

4) Create a **key.csv** which contains the questions and the correct solution key. There can be only one solution key and it is case sensitive. Use upper case alphabets for the solution key.

### Figure : Snapshot of key.csv



Note: The question names should be exactly the same as the column names in the data sheet containing the responses. The names are case sensitive, hence the best way to avoid any errors is to copy the row names and paste them as values and transpose.

## DATA CLEANING

The next step entails cleaning the data. The data cleaning step entails the following

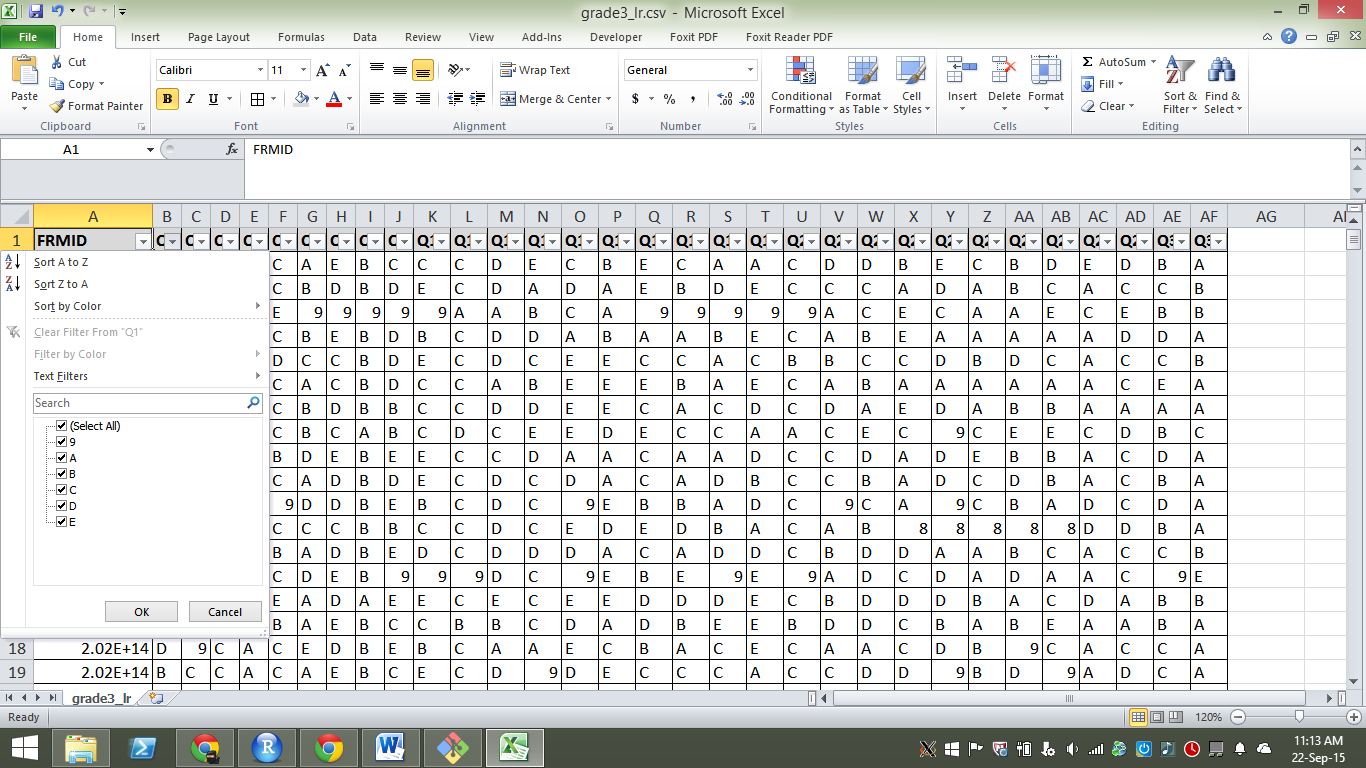
1) Convert all blank responses to number ‘9’

2) Replace multiple responses for a question (AB,ABC,ABCD, ABCDE etc) to’8’

3) For subjects which have only 4 options, if responses contain ‘E’, convert them to 8

4) Check all responses, question by question to ensure that only A,B,C,D,E,8,9 are the only responses in each column

### Figure : Checking for valid responses



## UNDERSTANDING THE CODEFRAME

The code frame is document listing the items used for testing. The code frame is a master list of the items used along with the Grade, Subject, Descriptor, Unique ID, question number in a particular question paper. This information helps us in identifying the links between various question papers and is very crucial for constructing RASCH scales spanning multiple grades.

**Figure 4** shows a part of the code frame. Question with S.no 2, is meant for Grade 4, 5 and belongs to Numbers/Algebra strand and the item tests the skill of ‘Adding a 5 digit number with a 4 digit number’.

This item also serves as a link item between Grade 4 & 5 question papers. This can be confirmed by the fact that this item record appears under Grade 4 and Grade 5 columns (extreme right). The number in each of the columns shows the question sequence number in the respective question papers. This question is Q.no 1 in both Grade 4 and Grade 5.

Similarly, S.no 6 serves as a link between 3,4. The item is Q.no12 in Grade 3 and Q.no18 in Grade 4.

Note: As a best practice, it is good to list down the link items and their sequence number before linking across grades in the following format

|  |  |
| --- | --- |
| Grade 3 | Grade 4 |
| 2 | 2 |
| 12 | 18 |
| 16 | 22 |

### Figure : Snapshot of a code frame

