Microsoft Excel for Data Management

Dr. K. A. Korb

University of Jos

Outline

- Introduction to Microsoft Excel
- Data Processing
- Data Checking

Definitions

- **Microsoft Excel:** Spreadsheet with rows and columns of a grid to store, organize, and analyse information
- **Workbooks:** Excel file; Also called *worksheet* or *spreadsheet*
  - Consists of separate **sheets**
- **Ribbon:** Tools and features at the top of the screen
  - Divided by tabs that have specific commands
- **Cells:** Noted by letter (column) and number (row)

Simple Commands

- New workbook: Control+N
- Save: Control+S
- Undo: Control+Z
- Scrolling
- Highlight one cell: Click
  - Highlight multiple cells: Click and drag OR Shift+Arrows
  - Highlight a row or column: Click on the letter/number

Cell Content

- **Text:** Characters only; String
- **Numbers**
  – Different number formats
- **Formulas:** Begin with =
- **Formatting: Bold**, *italics*, <u>underline</u>, etc.
  – Number of decimal points
  – Percent

Simple Commands

- Tab: Across
- Enter: Down
- Change column/row width
- Insert rows/columns
- Delete cells vs. Delete content in cells
- Freeze panes
- Hide columns/rows
- Wrap text
- Merge cells
- Add and name a new sheet

Simple Commands

- Copy and Paste
  – Paste will paste formulas, NOT values
- Paste Special:
  – **Formulas**: Excel will automatically update the cell referents
  – **Values**: Only numbers or text
  – **Transpose**: A row will be pasted as a column or vis versa
- **Fill handle:** Automatically copies and pastes down a row

Formulas

Maths Operators	In Excel
Addition	+
Subtraction	-
Multiplication	*
Division	/
Exponents	^

Creating a Formula

1. Click in the cell that will contain the formula
2. Type equals sign (=)
3. Click on the first cell in the equation OR type the cell address
4. Type the mathematical operator
5. Click on the second cell in the equation OR type the cell address
6. Press enter

Editing a Formula

- Double click within the cell
  - Excel will highlight all cells that are referenced
- Highlight the cell
  - Click in the formula bar

Errors

- *****Excel will do what you tell it to do. If you make an error, Excel will make an error.**
  - **Solution: Check, double check, and triple check your formulas**
- ####: There is not enough space for the number.
  - Solution: Increase the cell width
- #DIV/0!: The formula requires dividing by zero.
  - Solution: Fix the formula
- #NAME?: Invalid cell name
  - Solution: Check the cell referents in the formula

For More Tutorials in Microsoft Office…

- https://www.gcflearnfree.org/

Data Processing:
Introduction

- *It is more from carelessness about truth than from intentionally lying that there is so much falsehood in the world.* Samuel Johnson
- *Garbage in, Garbage out.* Proverb
- One thoughtless mistake in data processing and analysis can result in an entirely incorrect conclusion for the study
- Check, double-check, triple-check, and quadruple-check all steps in the data analysis process

Data Processing

- Step 1: Number the completed questionnaires
  - Rationale: Link data into the computer to the actual questionnaire
  - Generally use a letter related to the data, then 3 numbers, e.g., D001, D002, D003…D010, D011…
- Step 2: Code the data in Excel
  - Create a matrix with each questionnaire item across the top (column), and each completed questionnaire going down (row).

Tips for Coding Data

- *Coding the Data:* Each response is given a number or letter code and entered into the computer
- Make a reminder sheet of the codes given to each response for future reference
  - Some responses are already given a code (**a** for Male and **b** for Female)
  - If a code is not already given, use the first letter of the word (**M** for Males and **F** for Females)
    - Be careful that each response gets a unique code
  - Yes is coded as 1; No is coded as 0
  - For Likert-Scale items, enter the number of the response

Tips for Coding Data

- Be *very* careful when doing data entry to prevent computer typos
  - Incorrectly entering data is careless research and is therefore unethical
  - Double-check every questionnaire to ensure the correct responses are indicated in the correct columns

Missing Data

- Missing an item on the questionnaire
  - Discard the item for that particular participant
  - Enter **Z** for missing items
    - Do not leave empty cells
- Items where the response is unclear (e.g., circled two responses)
  - It is typically best to enter that item as missing, Z.
  - Possible Exception: If they circled 3 and 4, then you can enter 3.5

Missing Data

- Participants do not complete an entire portion of the questionnaire
  - Before the study, set criteria about when to discard data
    - Sample criteria: If all items for one variable are skipped, then discard
  - Keep track of how many are discarded and why
- Participants only complete one part of the study (in experiments)
  - Consider how to get their data with alternate procedures that do not confound the research design
  - If the data cannot be recovered or if they missed a substantial portion of the treatment, discard from the entire study
  - Report how missing a phase of the study was handled and how many were excluded because they did not complete the experiment

Reverse-Coded Items

- **Reverse-Coded Items:** Items that say the *opposite* of other items
  - Assign the opposite code to the response.
    - Agreeing to "I dislike school" is actually Disagreeing to the overall Enjoyment of School
    - Strongly Disagreeing to "I dislike school" is actually Strongly Agreeing to Enjoyment of School
  - Reverse coding is typically used to control for acquiescence bias
- It is easiest and more accurate to enter items exactly as they are completed on the questionnaire and then recode later
  - *If* function in Microsoft Excel
  - SPSS has a recoding function

Reverse-Coded Items

	Typical Code I like Maltina	Reverse Code I do NOT like Maltina
Strongly Agree	4	1
Agree	3	2
Disagree	2	3
Strongly Disagree	1	4

Typical Item:	Strongly Disagree	Disagree	Agree	Strongly Agree
Reverse Coded Item:	Strongly Agree	Agree	Disagree	Strongly Disagree

Step 3: Calculating Scores for Each Variable

- Total scores are needed for each variable
- Items that measure each variable are likely distributed across the questionnaire.
  - Identify which items measure each variable
  - Reverse-code items as necessary
- To avoid problems associated with missing items, calculating the mean (average) of the items that measure each variable is generally best

Data Checking

- After all data has been entered, run a few tests to ensure that there are no typographical errors
  - Check the maximum and minimum values for each numerical item
    - If the maximum/minimum is outside of the possible range of values, there is an error
  - Calculate the frequency of categorical values
    - Ensure the sum adds up to the total number of participants