Working with Excel

Dr. Gauti Sigthorsson, Deputy Dean (Communications and Digital Media), School of Arts

# Intro

3-hour workshop (3 x 45 mins., with breaks)

Download and open the file titled Workbook (Working with Excel). It contains the worksheets we use for this session.

Basics

Spreadsheets are tools for creating, manipulating and analysing datasets. As such, it’s good practice to keep things consistent.

## Cell, row, column

“Tidy data” is a simple set of organizing principles for data. In tidy data every column is a variable, every row is an observation. Three rules make a dataset tidy:

* Each variable must have its own column.
* Each observation must have its own row.
* Each value must have its own cell.



Wickham and Grolemund, [R for Data Science (2017)](https://r4ds.had.co.nz/tidy-data.html)

## Data types

Ensure that all observations in a column are of the same data type. Use the “Format Cells” command (by column, for consistency).

In Excel, the default data type is “General”, which identifies your data type automatically (i.e., currency, number, text, etc.). If you find that Excel changes things automatically, it’s probably attempting to enforce a data type you don’t want.

## Tables

There are two ways to define a table in Excel: “Format as Table” and Filter (on/off).

**Task:** In the Suzy’s Shoes sheet, apply both table-methods to the data (from header row down) and note the difference in behaviour.

## Filters, formulas

**Worksheet:** Suzy’s Shoes

Turn on filters for the columns (Data menu).

Apply formulas to the sales figures for Suzy’s Shoes:

* Calculation: Add up the whole Sales column
* Function: Apply the SUM function to the Sales column

Add new calculated column showing change in sales from the previous month.

Create separate table: Summarise Suzy’s Shoes’ sales by quarter (i.e., 3-month blocks), using a formula.

Let’s make graphs: In the same sheet, create three different graphs. Visualise the sales by month (bar chart), by quarters (pie chart), and change month-to-month (line graph).

# Summarise and Pivot

**Worksheet:** Southwest Sandwiches

Summarise the values in the sheet: Minimum, maximum, median and mean (average) values; number of items. Please use the functions MIN, MAX, MEDIAN, AVERAGE, COUNT.

What are the best and worst days for Southwest Sandwiches?

## Filter

**Worksheet:** Southwest Sandwiches

Format as table. Turn on “Show Totals” for the table, make sure the bottom row shows the sum (other calculations are available).

## Pivot

Create a pivot table (put it in the same sheet).

Which of the Southwest Sandwiches shops has the best sales? Pivot by location and average of sales.

Which months are the best for the chain as a whole? Pivot by date and sum of sales.

## Distribution

Visualise the distribution. Select the data, create a histogram showing the distribution.

**Worksheet:** Goodness Gym

Format as table, add automatic totals at bottom.

Add a calculated column showing the change in membership for each week.

Visualise the change as a line graph.

# References

[Microsoft Excel Support Site (see “Getting Started” for beginner’s guide)](https://support.microsoft.com/en-us/excel)

Knaflic, C. N. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley.

Nelson, S. L., Nelson, E. C., & Nelson, E. C. (2015). *Excel Data Analysis for Dummies*. Wiley.