

# Manipulating dataset rows in Excel

This planning document is intended to support teachers who are delivering the NPA/PDA Data Science or for students who are learning independently. It also aligns with the Data Skills for Work framework.

**The lesson has been designed for learners using Microsoft Excel.** Most of the information in the lesson will work for other spreadsheets tools. However, if another tool is being used by the learners (such as Google Sheets) the step-by-step instructions will need to be checked.

## Contents

Lesson Description.....	1
Lesson Contents .....	2
Learning Intention .....	2
Success Criteria .....	2
Knowledge Prerequisites.....	2
Lesson Requirements.....	3
Task-types.....	4
Worksheet .....	5
How you can use this lesson.....	6

## Lesson Description

<b>Lesson Overview</b>	Subsetting Filtering Sorting Deduplicating
<b>Topic</b>	Data manipulation
<b>Book Chapter(s)</b>	“Data transformation and Manipulation”

<b>NPA level</b>	4, 5, 6
<b>PDA level</b>	7, 8
<b>Data skills for work level</b>	Core, Analysis

## Lesson Contents

This lesson consists of:

- A lesson plan (this document)
- A PowerPoint presentation, 'Manipulating datasets rows in Excel'
- Question worksheet on 'Manipulating datasets rows in Excel' in Excel
- Answers worksheet on 'Manipulating datasets rows in Excel' in Excel

## Learning Intention

We will be learning how to manipulate data in Excel, specifically,

- to be able to **filter and sort** rows
- to be able to **subset** data to select the parts of the data you are interested in
- to be able to **remove duplicates** from the data

## Success Criteria

I can *describe* what it means to sort, filter, subset and remove duplicates from a dataset.

I can *manipulate* data by sorting, filtering, subsetting and removing duplicates in Excel.

## Knowledge Prerequisites

Learners should know:

- How to open/save an Excel file
- That Excel documents have worksheets and use rows/columns
- To be able to enter data into cells
- To be able to select/highlight sections of data
- Headers on worksheet (e.g., A, B, C, D and rows 1,2,3,4)
- Excel has ribbons at the top for Home, Data, Formulas etc
- To be able to right-click on cells to see options

## Lesson Requirements

	<b>PDA</b>	<b>NPA</b>	<b>Data Skills for work</b>
<b>Qualification</b>	Yes	Yes	Yes
<b>Outcome ID(s)</b>	WD8.3b, WD8.3c, CD8.1g, WD7.2a, WD7.2b, CD7.3a	DS4.2c, DS4.3a, DS5.2c, DS5.3c, DS6.2b, DS6.3c	C2.1, A1.2, A2.3
<b>Outcome description(s)</b>	<p>WD8.3b Types of data transformation</p> <p>WD8.3c Transformations</p> <p>CD8.1g Preparing data for visualisation</p> <p>WD7.2a Types of data transformation</p> <p>WD7.2b Common transformations including filtering, sorting</p> <p>CD7.3a Preparing data for visualisation</p> <p><i>N.B. out of scope of this lesson,</i></p> <p><i>“WD8.3c ... including joins”</i></p> <p><i>“WD7.2b .... combining, separating and grouping”</i></p>	<p>DS4.2c Describe simple methods of cleaning and transforming data</p> <p>DS4.3a Perform simple data cleaning and structuring</p> <p>DS5.2c Describe methods of cleaning and transforming data</p> <p>DS5.3c Perform routine data cleaning and structuring.</p> <p>DS6.2b Explain techniques for data capture, cleaning and transformation including data modelling</p> <p>DS6.3c Perform data transformation to complete, correct and structure data</p>	<p>C2.1 Vocabulary used in data science and analytics</p> <p>A1.2 Data quality</p> <p>A2.3 Data calculation and manipulation</p> <p><i>N.B. out of scope of this lesson “A1.1.... quantitative and qualitative”</i></p>
<b>Level</b>	7, 8	4, 5, 6	Core, Analysis
<b>Software language</b>	Excel	Excel	Excel
<b>Required equipment /software for student</b>	Lesson: PowerPoint, Worksheet: Excel	Lesson: PowerPoint, Worksheet: Excel	Lesson: PowerPoint, Worksheet: Excel

## Task-types

In the worksheet for this lesson, there are up to 6 task-types to that become increasingly challenging to support the students learning. Based on the student's previous knowledge it is possible to select the task-types that are relevant to their stage.

Task-type	Description
<b>1. Recall</b>	To be able to recognise definitions or procedures.
<b>2. Define</b>	To be able to define definitions or procedures.
<b>3. Rephrase</b>	To be able to use their own words to describe definitions or procedures.
<b>4. Apply</b>	To be able to apply definitions or procedures to problem-solving activities.
<b>5. Create</b>	To be able to apply definitions or procedures and create their own solutions to a defined problem.
<b>6. Active</b>	Using knowledge from the lesson which they apply to scenarios they have researched/designed themselves.

## Worksheet

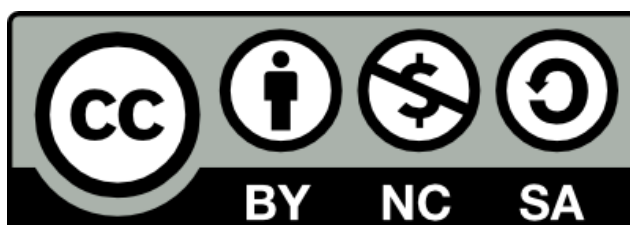
The worksheet associated with this lesson is available either in Excel or as a PDF that can be printed. The answer worksheet is available in both formats too.

Worksheet section ID	Description	Task-type	Number of questions
1.1	Sorting data	Recall	1
1.2	Sorting data	Define	2
1.3	Sorting data	Apply	5
1.4	Sorting data	Active	1
2.1	Filtering data	Recall	1
2.2	Filtering data	Rephrase	2
2.3	Filtering data	Apply	6
3.1	Subsetting data	Recall	1
3.2	Subsetting data	Rephrase	1
3.3	Subsetting data	Apply	6
4.1	Remove duplicates	Recall	1
4.2	Remove duplicates	Rephrase	1
4.3	Remove duplicates	Apply	1
4.4	Remove duplicates	Apply	1
<b>Total</b>			<b>30</b>

## How you can use this lesson

This lesson has been created by Effini in partnership with Data Education in Schools, The Data Lab and Data Skills for Work, with funding from the Scottish Government.

© 2021. This work is licensed under a [CC BY-NC-SA 4.0 license](#).



You are free to:

- **Share** – copy and redistribute the material in any medium or format
- **Adapt** – remix, transform and build upon the material

Under the following terms:

- **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **NonCommercial** — You may not use the material for [commercial purposes](#).
- **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.