

Scenario Demonstration Skill Builder 2 - Importing Messy Data

This guide is essential for anyone looking to efficiently import and clean messy data in Excel. It provides a step-by-step visual approach, making it easy to follow along and apply the techniques directly. By mastering these skills, users can enhance their data management capabilities, ensuring accuracy and organization in their datasets. Whether you're a beginner or looking to refine your skills, this guide offers valuable insights for effective data handling.

- 1 Let's start first of all by improving one of the tables which currently has poorly named column headers.

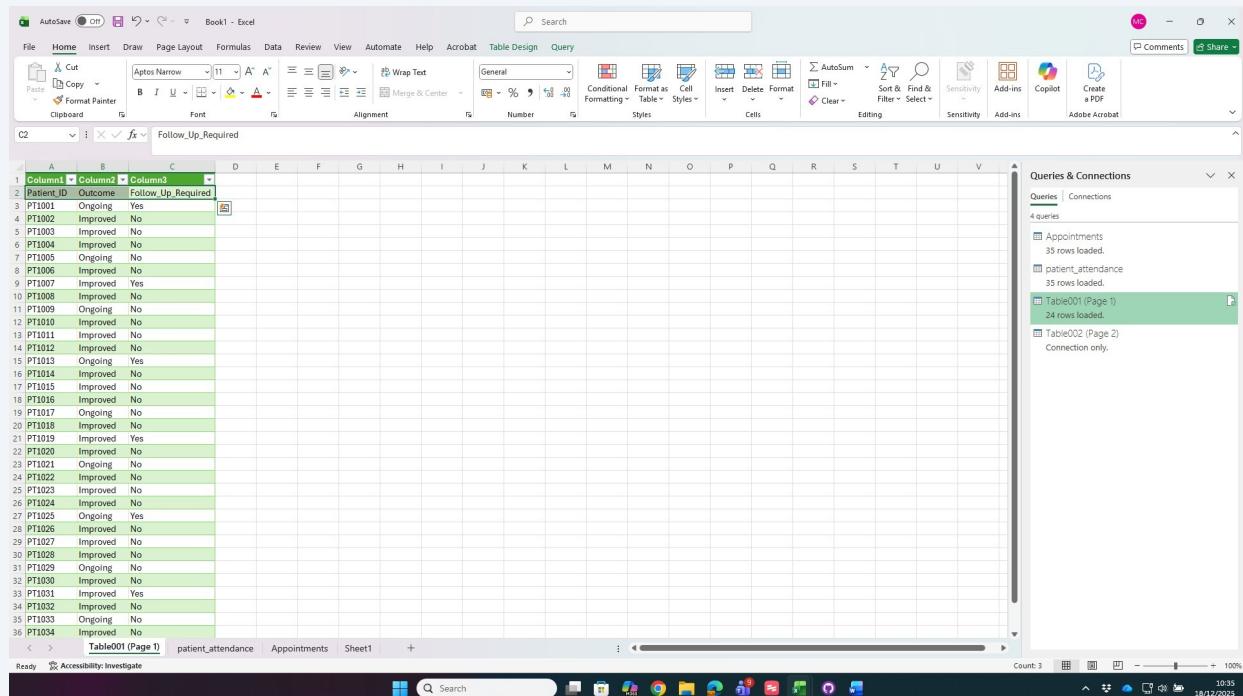
Select the **Table001** data import you did previously.

If you look carefully Row B contains information that should be the column headers in Row A.

The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel". The ribbon menu is visible at the top, with "Home" selected. The "Font" and "Alignment" toolbars are also visible. The worksheet contains a table with 11 rows and 3 columns. The first row (Row 2) contains the column headers: "Patient_ID", "Outcome", and "Follow_Up_Required". The second row (Row 3) contains the text "Patient ID", "Outcome", and "Follow Up Required". Rows 4 through 11 contain patient data. The entire range from A2 to C11 is highlighted with a green selection bar. An orange rounded rectangle highlights the header row (A2:C2). The data is as follows:

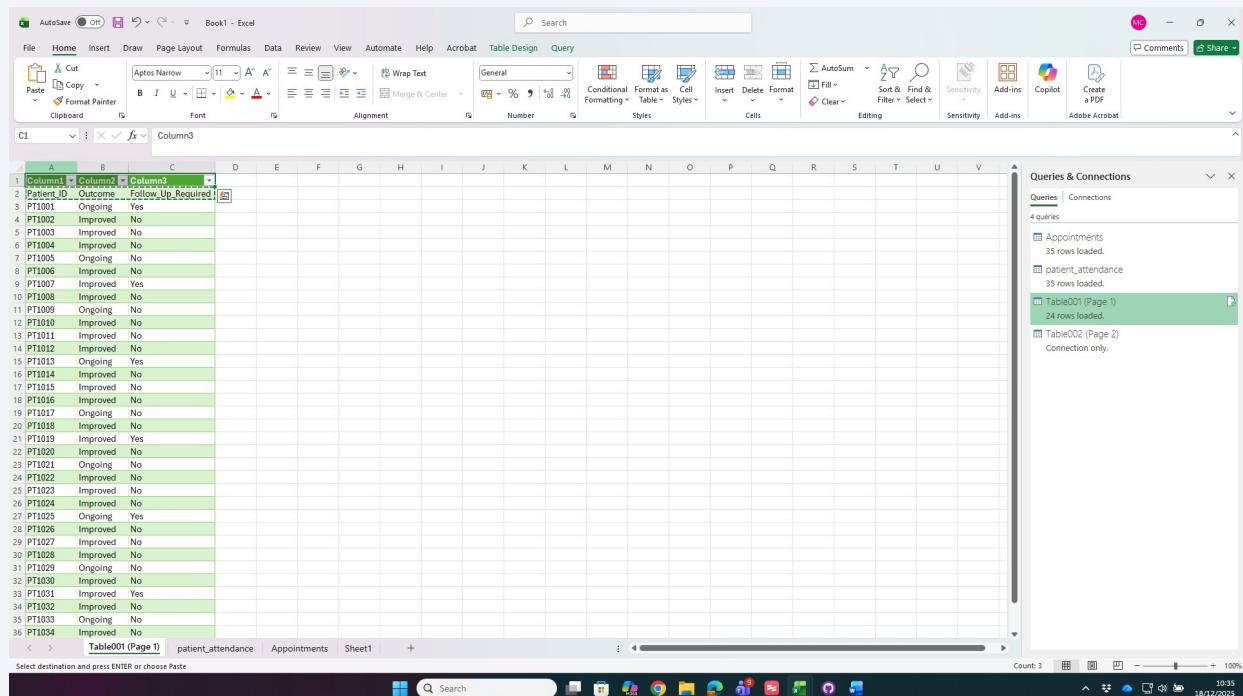
	A	B	C
2	Patient_ID	Outcome	Follow_Up_Required
3	Patient ID	Outcome	Follow Up Required
4	PT1002	Improved	No
5	PT1003	Improved	No
6	PT1004	Improved	No
7	PT1005	Ongoing	No
8	PT1006	Improved	No
9	PT1007	Improved	Yes
10	PT1008	Improved	No
11	PT1009	Ongoing	No

2 Highlight Row B and press **Ctrl + X**



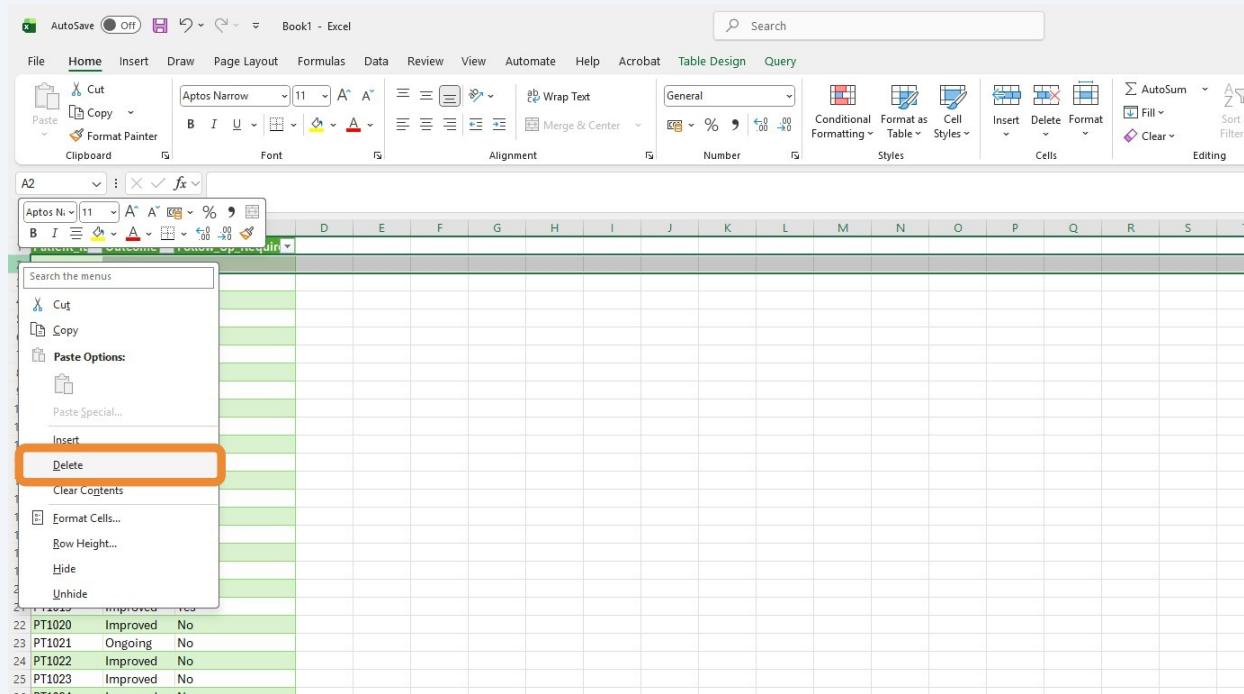
The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel". The data is organized into three columns: Patient_ID, Outcome, and Follow_Up_Required. Rows 1 through 36 contain patient information. Row B (row 2) is highlighted in green. The "Editing" tab is selected in the ribbon. The "Queries & Connections" pane on the right shows three loaded queries: Appointments, patient_attendance, and Table001 (Page 1). The status bar at the bottom indicates "Count: 3" and the date "18/12/2025".

3 Now highlight cells A1:A3 and press **Ctrl + V**



The screenshot shows the same Microsoft Excel spreadsheet as the previous one. In this step, cells A1, A2, and A3 have been selected and highlighted in green. The "Editing" tab is still selected in the ribbon. The "Queries & Connections" pane remains the same. The status bar at the bottom indicates "Count: 3" and the date "18/12/2025".

- 4 After you have pasted the correct value into the column headers, select row 2 and right-click then click **Delete**

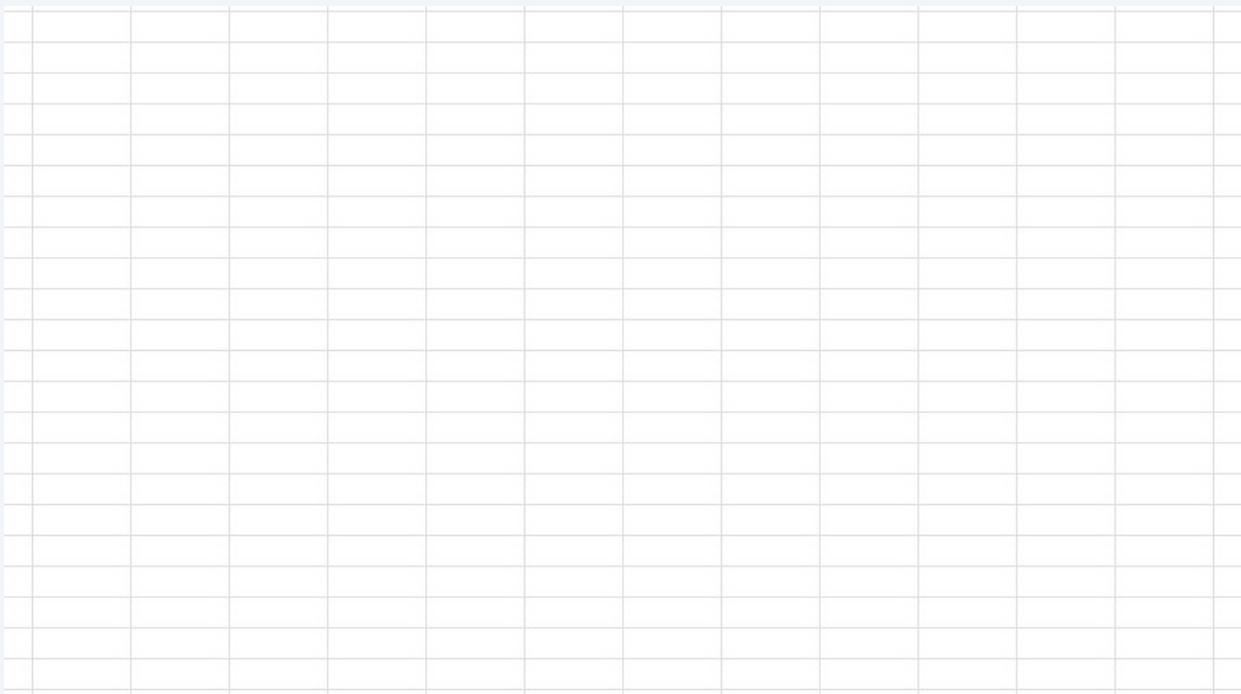


- 5 Let's name the sheet something more meaningful - double-click **Table001 (Page 1)** and rename it to **patient_outcome**

A screenshot of Microsoft Excel showing the sheet tabs at the bottom. The first tab is labeled 'Table001 (Page 1)', which has a double-headed orange arrow pointing to the second tab. The second tab is labeled 'patient_outcome'. Other visible tabs include 'patient_attendance', 'Appointments', and 'Sheet1'. The data in the sheet consists of 36 rows, each containing a patient ID (PT1015 to PT1035), their status ('Ongoing' or 'Improved'), and a response ('Yes', 'No', or ' ') in columns A, B, and C respectively.

	PT1015	Ongoing	res
15	PT1014	Improved	No
16	PT1015	Improved	No
17	PT1016	Improved	No
18	PT1017	Ongoing	No
19	PT1018	Improved	No
20	PT1019	Improved	Yes
21	PT1020	Improved	No
22	PT1021	Ongoing	No
23	PT1022	Improved	No
24	PT1023	Improved	No
25	PT1024	Improved	No
26	PT1025	Ongoing	Yes
27	PT1026	Improved	No
28	PT1027	Improved	No
29	PT1028	Improved	No
30	PT1029	Ongoing	No
31	PT1030	Improved	No
32	PT1031	Improved	Yes
33	PT1032	Improved	No
34	PT1033	Ongoing	No
35	PT1034	Improved	No
36	PT1035	Improved	No

- 6 Now let's bring in the last collection of data - click on **Sheet 1**

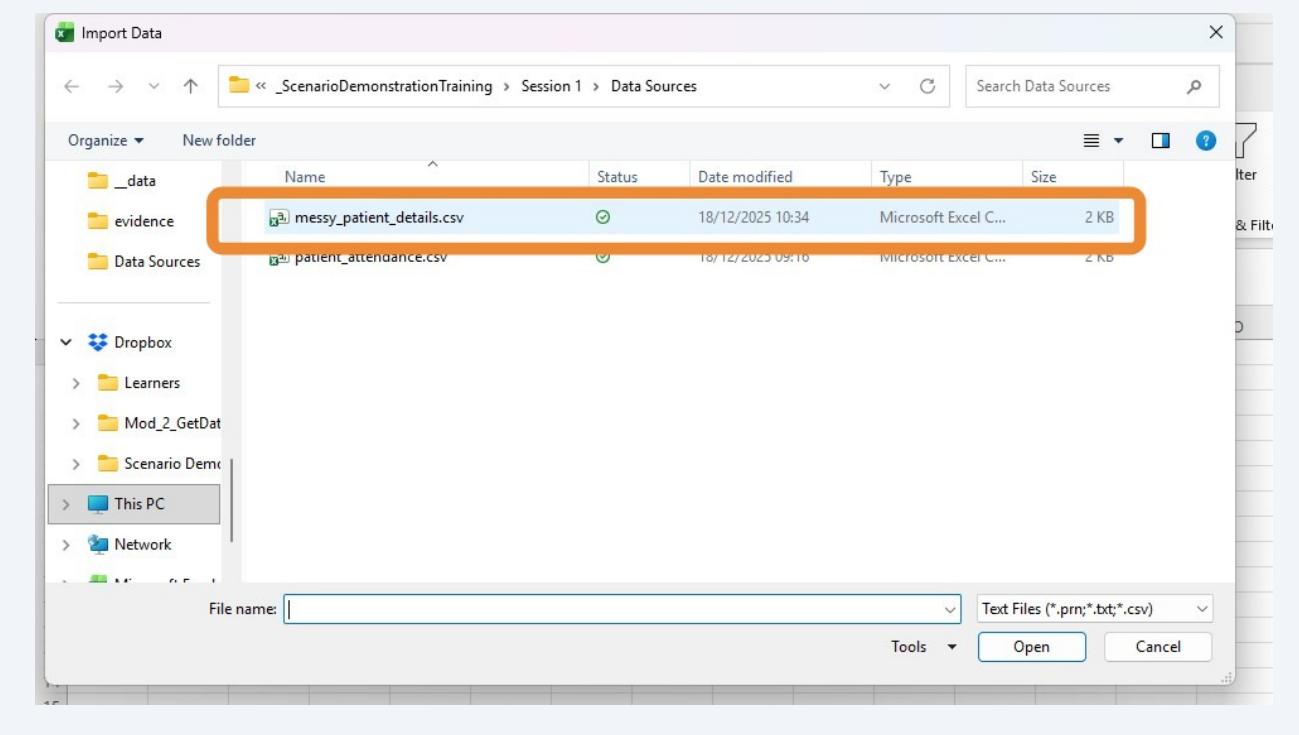


- 7 Click on **Get Data > From File > From Text/CSV.**

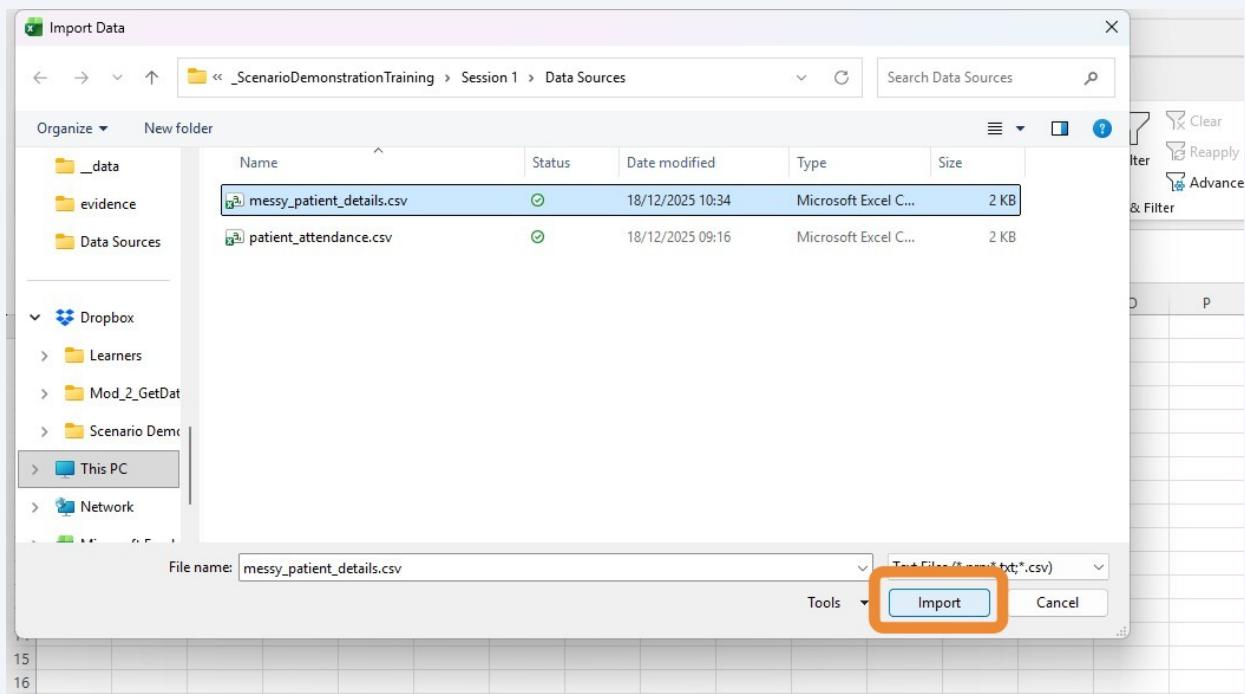
You shout be an expert at this by now.

The screenshot shows the Microsoft Excel ribbon with the 'Data' tab selected. On the far left, there is a dropdown menu labeled 'Get Data' with several options: 'From Text/CSV', 'From Picture', 'From Web', 'Recent Sources', 'From Table/Range', and 'Existing Connections'. Below this, a 'Get Data (Preview)' section is visible. Under 'From File', there is a list of options: 'From Excel Workbook', 'From Text/CSV', 'From XML', 'From JSON', 'From PDF', 'From Folder', and 'From SharePoint Folder'. An orange arrow points from the 'From Text/CSV' option towards the right side of the screen, where the main Excel workspace is visible with columns H through O.

8 Click messy_patient_details.csv



9 Click Import



10 In the Navigator window, click Load

The screenshot shows the Power BI Navigator window. A file named "messy_patient_details.csv" is selected. The "File Origin" dropdown shows "1252: Western European (Windows)". The "Delimiter" dropdown shows "Comma". The "Data Type Detection" dropdown shows "Based on first 200 rows". The preview pane displays the first 200 rows of the CSV file. At the bottom right of the preview pane, there is a "Load" button, which is highlighted with an orange box.

11 You should now find all your data has loaded, in all it's messy magnificence

The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel". The table is named "messy_patient_details". The table style is set to "Table Style Options" with a green header row and alternating green rows. The data consists of 25 rows, each containing a Patient ID, Full Name, Date of Birth, Gender, GP, and On Medication status. The columns are labeled A through Q.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Patient ID	Full Name	Date of Birth	Gender	GP	On Medication											
2	PT1001	john smith	12/05/1984	Male	HEALTH CENTRE NORTH	n											
3	PT1002	Jane Doe	01/06/1970	Male	Dr adams	N											
4	pt1003	Alice Brown	01/06/1970	f	Dr adams	n											
5	pt1004	Robert Jones	01/06/1970	F	health centre north	N											
6	PT1005	Sarah White	12/05/1984	male	HEALTH CENTRE NORTH	n											
7	PT1006	michael green	01/06/1970	F	health centre north	n											
8	PT1007	EMILY BLACK	01/06/1970	M	HEALTH CENTRE NORTH	Yes											
9	pt1008	THOMAS HARRIS	06/01/1970	Female	health centre north	No											
10	PT1009	Laura Wilson	12/05/1984	Male	Dr Adams	Yes											
11	PT1010	David Clark	23/11/1990	Female	dr adams	yes											
12	pt1011	EMMA LEWIS	23/11/1990	F	Health Centre North	No											
13	PT1012	daniel walker	12/05/1984	Male	Health Centre North	n											
14	pt1013	Olivia Hall	01/06/1970	male	health centre north	Yes											
15	PT1014	james young	01/06/1970	f	Health Centre North	Yes											
16	PT1015	SOPHIA KING	01/06/1970	male	dr adams	n											
17	PT1016	MATTHEW WRIGHT	01/06/1970	Male	DR ADAMS	No											
18	PT1017	isla lopez	01/06/1970	f	health centre north	yes											
19	PT1018	benjamin scott	23/11/1990	male	health centre north	Y											
20	PT1019	AMELIA ADAMS	01/06/1970	Male	health centre north	N											
21	pt1020	LUCAS BAKER	23/11/1990	M	HEALTH CENTRE NORTH	yes											
22	PT1021	grace nelson	12/05/1984	f	health centre north	No											
23	PT1022	henry carter	12/05/1984	Male	HEALTH CENTRE NORTH	Yes											
24	PT1023	Chloe Mitchell	01/06/1970	Male	DR ADAMS	N											
25	PT1024	Noah Perez	12/05/1984	F	Dr Adams	No											

12

Well done, you have completed **Scenario Demonstration Skill Builder 2**