

SAYYED NIDA; COHORT-4



Standardize the name column to ensure consistent title casing.

Apply Title Case Transformation

- Click on "Transform" in the ribbon.
- Select "Format" > "Capitalize Each Word".



A ^B _C author.2.1	A ^B _C author.2.2	A ^B _C author.2.3
• Valid ● Error ● Empty	100% 0% 0%	91% 0% 9%
424 distinct, 266 unique	460 distinct, 325 unique	125 distinct, 84 unique
Geronimo	Stilton	null
Rick	Riordan	null
Jeff	Kinney	null
Rick	Riordan	null
Rick	Riordan	null
Suzanne	Collins	null
Winter	Morgan	null
Rick	Riordan	null
Mary	Pope	Osborne
Rick	Riordan	null
Rick	Riordan	null
Mary	Pope	Osborne
Mary	Pope	Osborne
Rick	Riordan	null
Philip	Pullman	null
Geronimo	Stilton	null
Mary	Pope	Osborne
Shannon	Messenger	null
Mary	Pope	Osborne
Shannon	Messenger	null
Emily Arnold	Mc	Cully
Saskia Maarleveld		Various
Kate Pankhurst		Adjoa Andoh
Bill Andrew	Quinn	Raphael Corkhill
Mark Bramhall		Sean Patrick

Separate combined first and last names in the author column if they are currently combined.

Split the Column

Go to the Transform tab.

Click "Split Column" > "By Delimiter".

Choose Space (" ") as the delimiter.

Click OK.

Rename Columns

Rename the new columns as "First Name" and "Last Name" (if they were properly split).

Handle Edge Cases

If some names have middle names, manually adjust them or use "Column from Examples" to extract first and last names correctly.



Ensure all entries in the releasedate column follow a consistent date format (DD-MM-YYYY).

Step 1: Change Data Type to Numeric

Open Power Query Editor (Data > Get Data > From Table/Range).

Select the Price column.

Go to Transform > Data Type > Select Decimal Number (or Whole Number if applicable).

Click Replace Current if prompted.

Step 2: Identify Non-Numeric Values

Check for Error values (this happens if Power Query fails to convert text to numbers).

Click the filter dropdown on the Price column.

Uncheck Numbers and only select Errors or Text values to review them.

Step 3: Fix or Remove Non-Numeric Values

Replace Non-Numeric Values Manually:

Select the Price column.

Go to Transform > Replace Values.

Replace symbols like \$, „ or unwanted text.

Extract Numeric Values from Mixed Data:

Go to Add Column > Column From Examples > From Selection.

Enter the correct numeric values manually for a few rows.

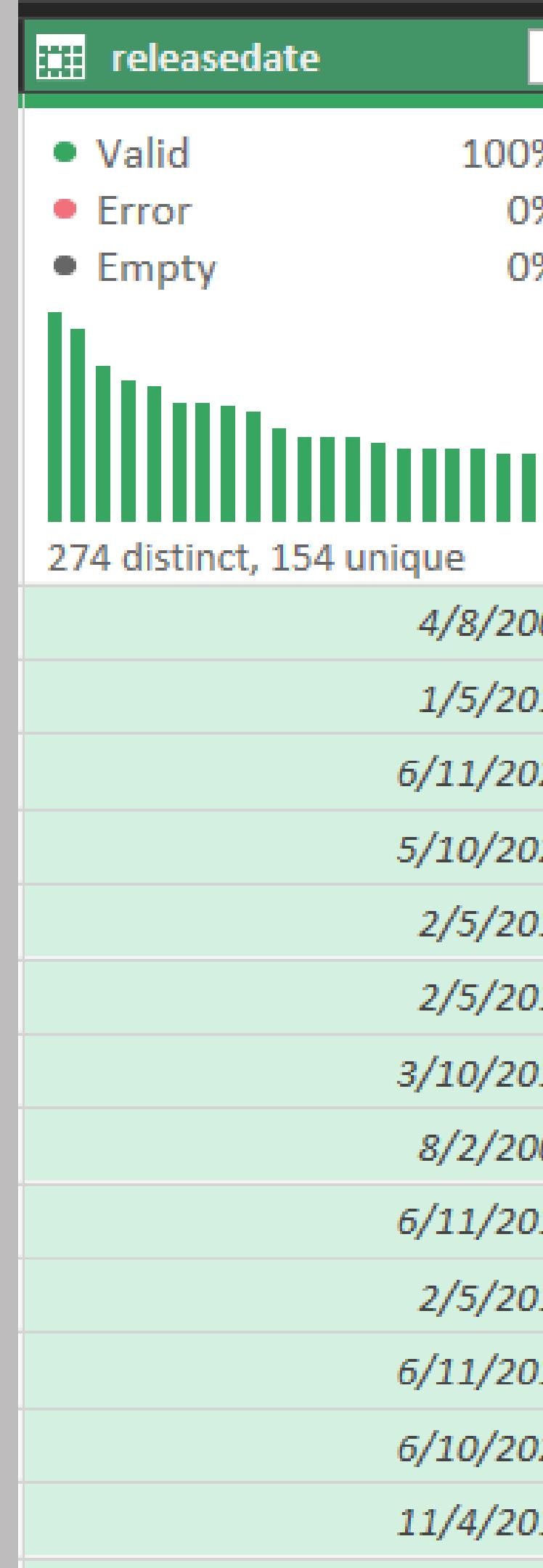
Power Query will detect the pattern and extract numbers from mixed values.

Remove or Replace Errors:

Select the Price column.

Go to Transform > Replace Errors.

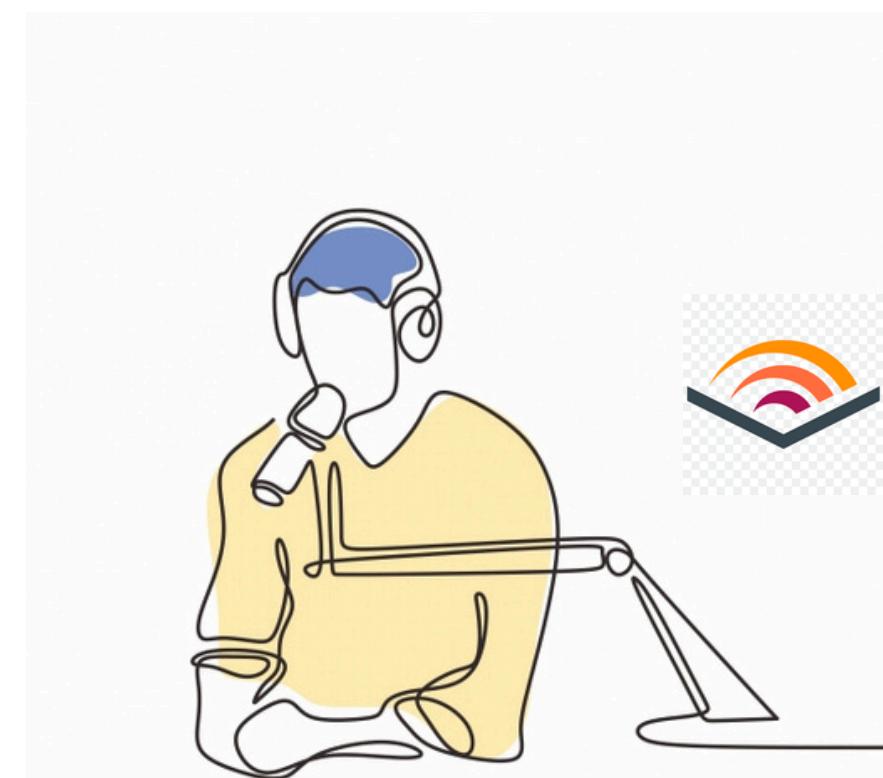
Enter 0 or another default value (optional).



Convert the time column from text format to a duration format that Excel recognizes.

A B C	time	Merged
● Valid	100%	● Valid 100%
● Error	0%	● Error 0%
● Empty	0%	● Empty 0%
455 distinct, 234 unique	455 distinct, 234 unique	
8 hr 32 min		0.08:32:00
7 hr 51 min		0.07:51:00
5 hr 10 min		0.05:10:00
5 hr 27 min		0.05:27:00
2 hr 54 min		0.02:54:00
1 hr 35 min		0.01:35:00
2 hr 23 min		0.02:23:00
33 min		0.00:33:00
3 hr 37 min		0.03:37:00
11 hr 19 min		0.11:19:00
2 hr 56 min		0.02:56:00
6 hr 58 min		0.06:58:00
5 hr 41 min		0.05:41:00

- Replaced Value
- Replaced Value1
- Replaced Value2
- Inserted Text Before Delimiter
- Inserted Text Between Delimiter...
- Replaced Errors
- Replaced Value3
- Changed Type2
- Replaced Value4
- Replaced Errors1
- Reordered Columns
- Merged Columns
- Changed Type3**



Ensure the price column is in a numeric format, and identify any non-numeric values.

123	price	▼
● Valid	100%	
● Error	0%	
● Empty	0%	
	468	
	820	
	410	
	615	
	820	
	1256	
	820	
	467	
	836	
	1256	
	1003	
	615	
	683	

Select the Price column.

Go to Transform > Data Type > Select Decimal Number (or Whole Number if applicable).

Click Replace Current if prompted.

Step 2: Identify Non-Numeric Values

Check for Error values (this happens if Power Query fails to convert text to numbers).

Click the filter dropdown on the Price column.

Uncheck Numbers and only select Errors or Text values to review them.

Step 3: Fix or Remove Non-Numeric Values

Replace Non-Numeric Values Manually:

Select the Price column.

Go to Transform > Replace Values.

Replace symbols like \$, ,, or unwanted text.

Extract Numeric Values from Mixed Data:

Go to Add Column > Column From Examples > From Selection.

Enter the correct numeric values manually for a few rows.

Power Query will detect the pattern and extract numbers from mixed values.

Remove or Replace Errors:

Select the Price column.

Go to Transform > Replace Errors.

Enter 0 or another default value (optional).



Convert text ratings in the stars column to numeric values.

1.2 stars	
● Valid	100%
● Error	0%
● Empty	0%
	5
	4.5
	4.5
	4.5
	5
	5
	5
	5
	5
	4.5
	5
	5

Step 1: Extract Numeric Ratings

Select the Stars column.

Go to the Transform tab and click Extract > Text Before Delimiter.

Enter " out of" as the delimiter.

Click OK.

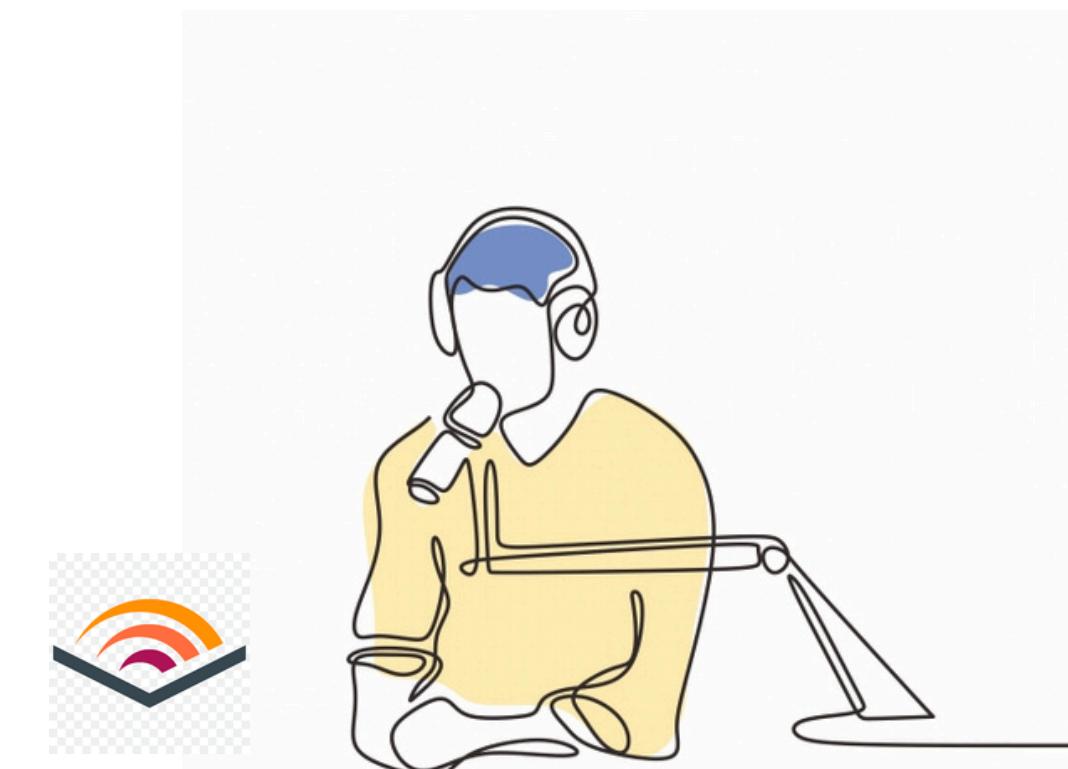
The column will now contain only the numeric ratings (e.g., 5, 4.5).

Step 2: Change Data Type to Numeric

Select the modified Stars column.

Go to Transform > Data Type > Select Decimal Number.

Click Replace Current if prompted.



Split the narratedby column into multiple columns if multiple narrators are listed.

Emily Arnold	Mc	Cully	
Saskia Maarleveld			Various
Kate Pankhurst			Adjoa Andoh
Bill Andrew	Quinn		Harriet Carmichael
Mark Bramhall		Raphael Corkhill	Sean Patrick

Select the narratedby column.

Go to Transform > Split Column > By Delimiter.

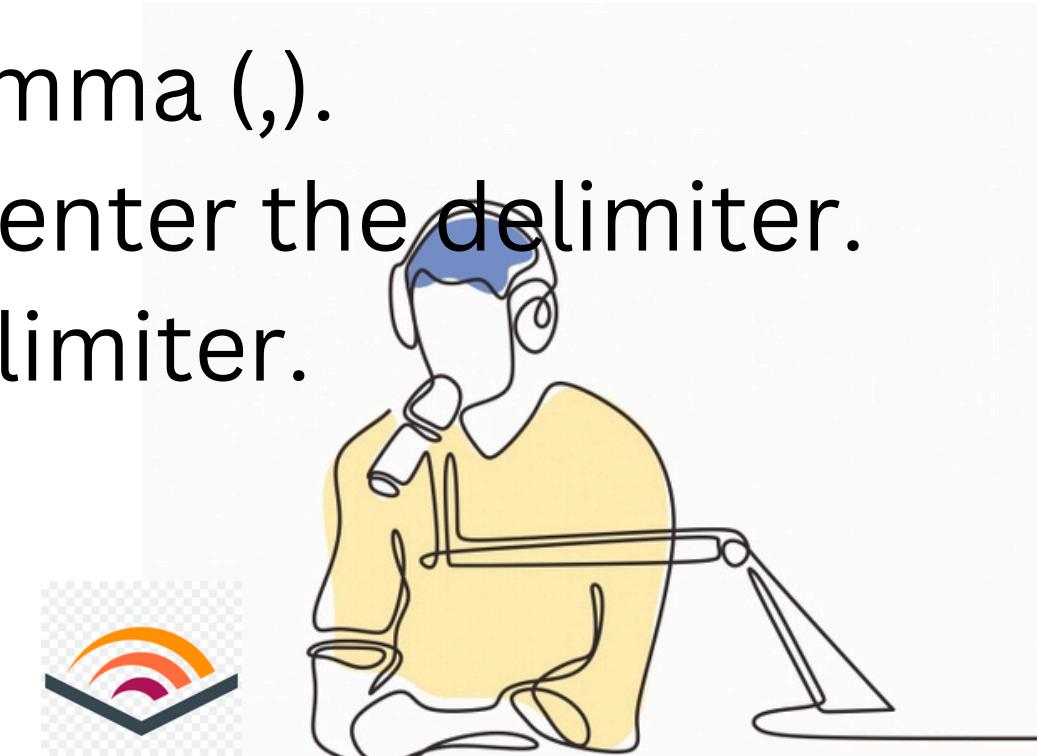
In the Choose Delimiter section:

If narrators are separated by commas, select Comma (,).

If separated by "and" or other symbols, choose Custom and enter the delimiter.

Under Split at, select Each occurrence of the delimiter.

Click OK.



Merge the releasedate and language columns into a single new column named releaseinfo with the format "DD-MM-YYYY, Language."

A	B	C	releaseinfo	D
•	Valid		100%	
•	Error		0%	
•	Empty		0%	
	4/8/2008,English			
	1/5/2018,English			
	6/11/2020,English			
	5/10/2021,English			
	2/5/2017,English			
	2/5/2017,English			
	3/10/2017,English			
	8/2/2008,English			
	6/11/2018,English			
	2/5/2017,English			
	6/11/2018,English			
	6/10/2020,English			
	11/4/2013,English			
	1/2/2011,English			

Select both the releasedate and language columns
(hold Ctrl while selecting).

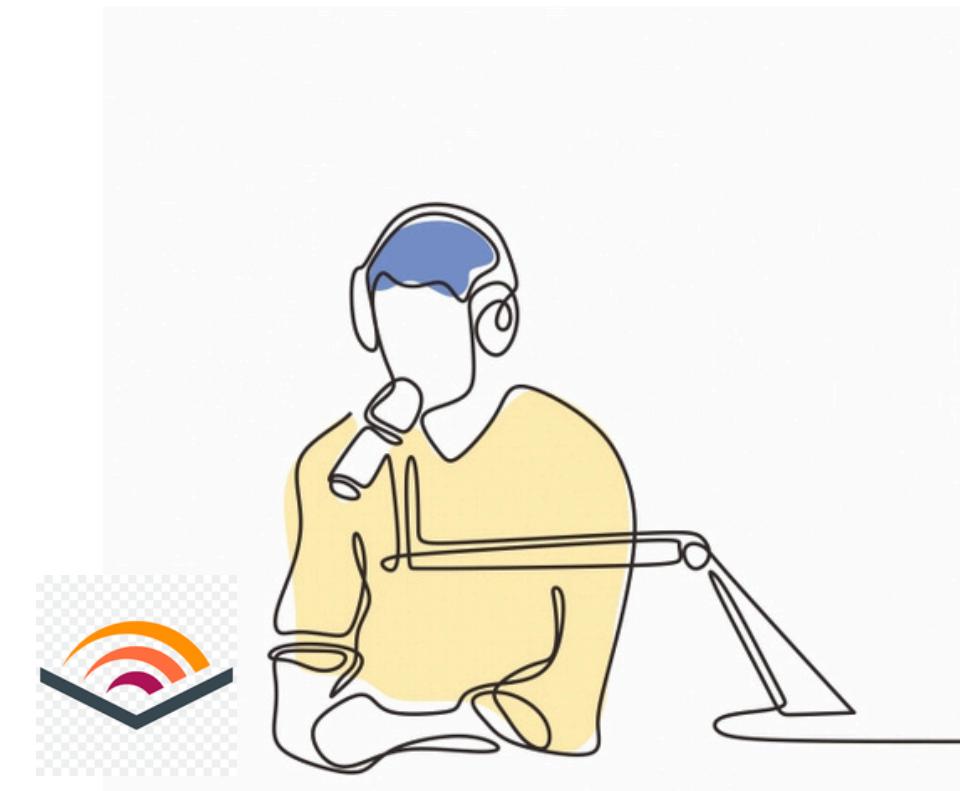
Go to Transform > Merge Columns.

In the Merge Columns window:

Select Custom Separator and enter ", " (comma and space).

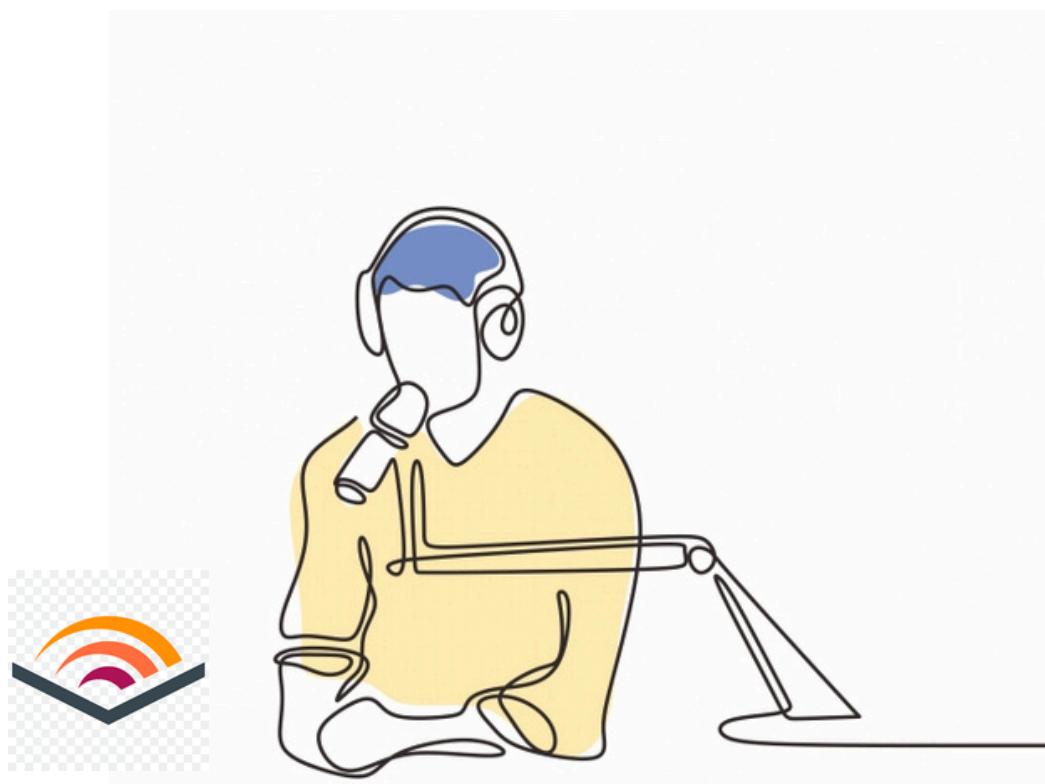
Name the new column releaseinfo.

Click OK.



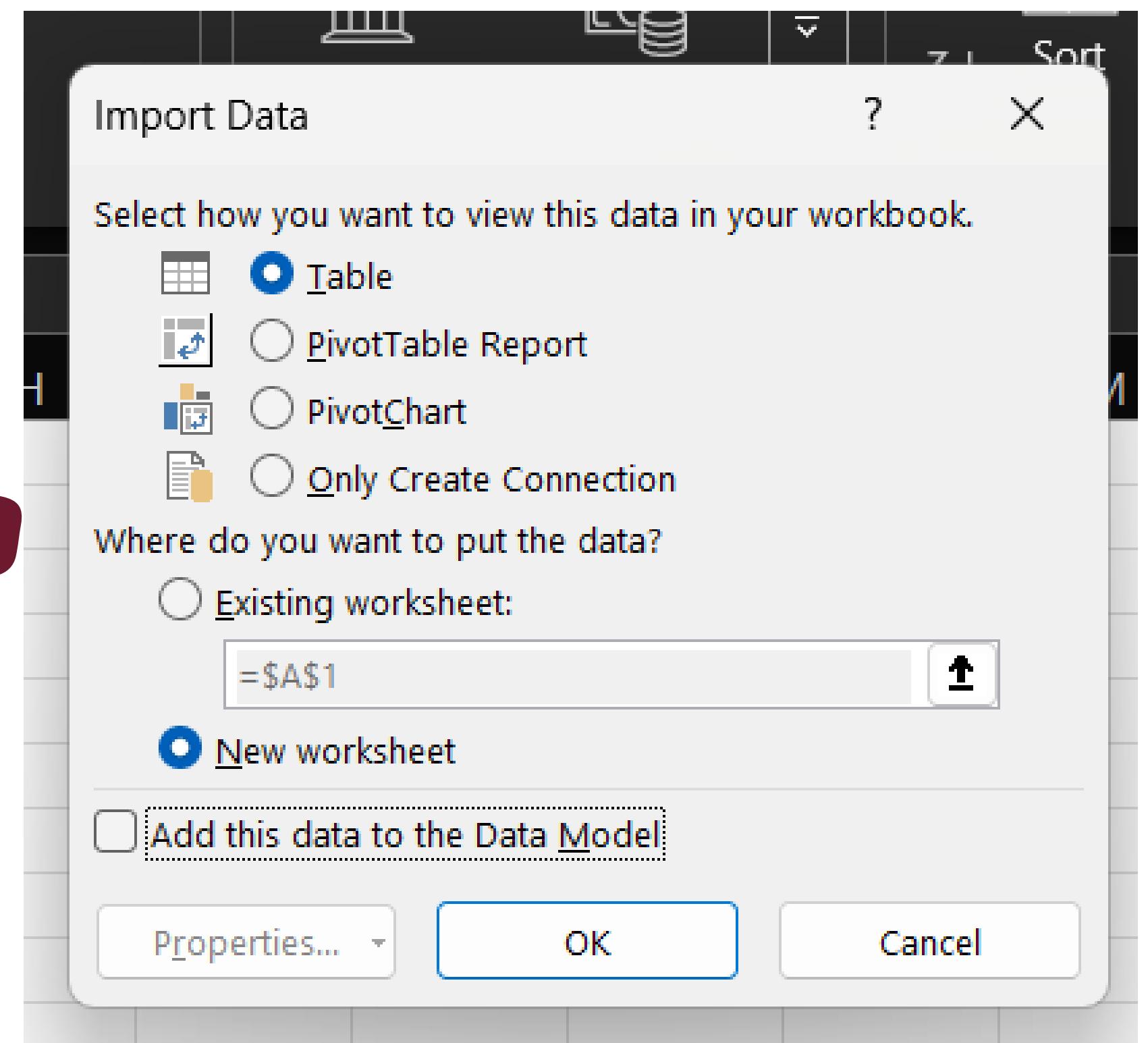
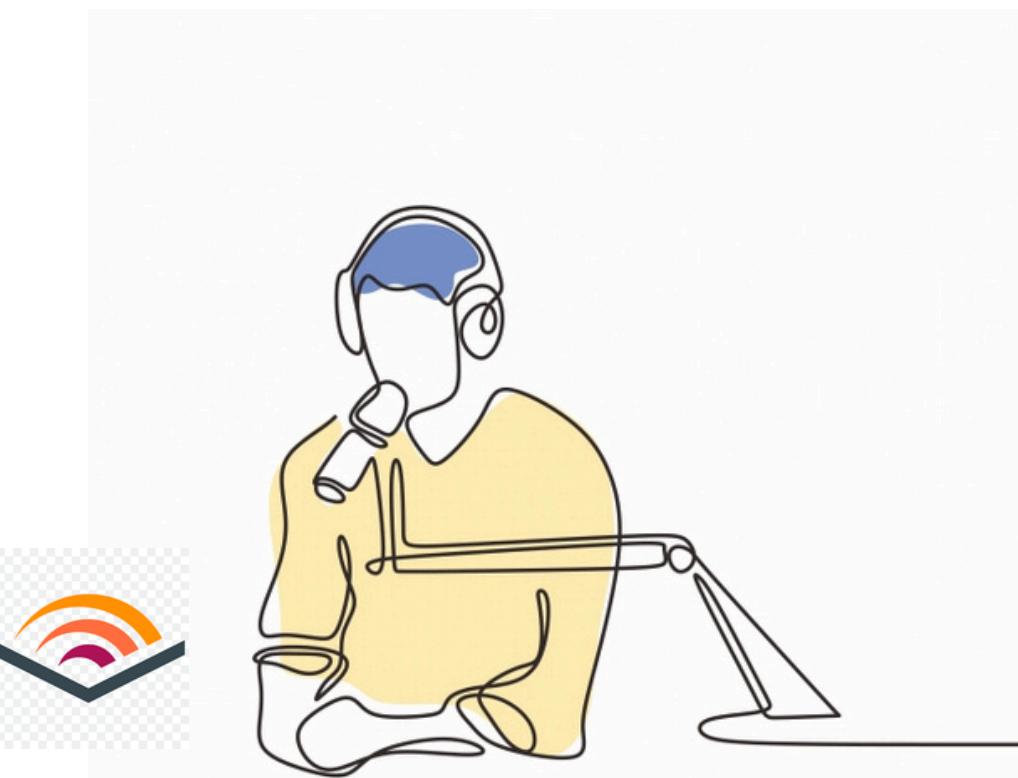
Ensure all currency values in the price column are formatted consistently with two decimal places.

Select the price column.
Go to Transform > Data
Type > Select CURRENCY.
Click Replace Current if
prompted.



\$	price	▼
●	Valid	100%
●	Error	0%
●	Empty	0%
	468.00	
	820.00	
	410.00	
	615.00	
	820.00	
	1,256.00	
	820.00	
	467.00	
	836.00	
	1,256.00	
	1,003.00	
	615.00	

**CLICK ON
close and load to....**



THANK YOU