

Audible

Project Presentation

Presented By: Azam Mustufa

06/10/2024.

Introduction

Escape into a world of stories with Audible, the leading provider of audiobooks and spoken-word entertainment. Brought to you by Amazon, Audible gives you access to a massive library of titles, from bestsellers and classics to exclusive podcasts and Audible Originals.

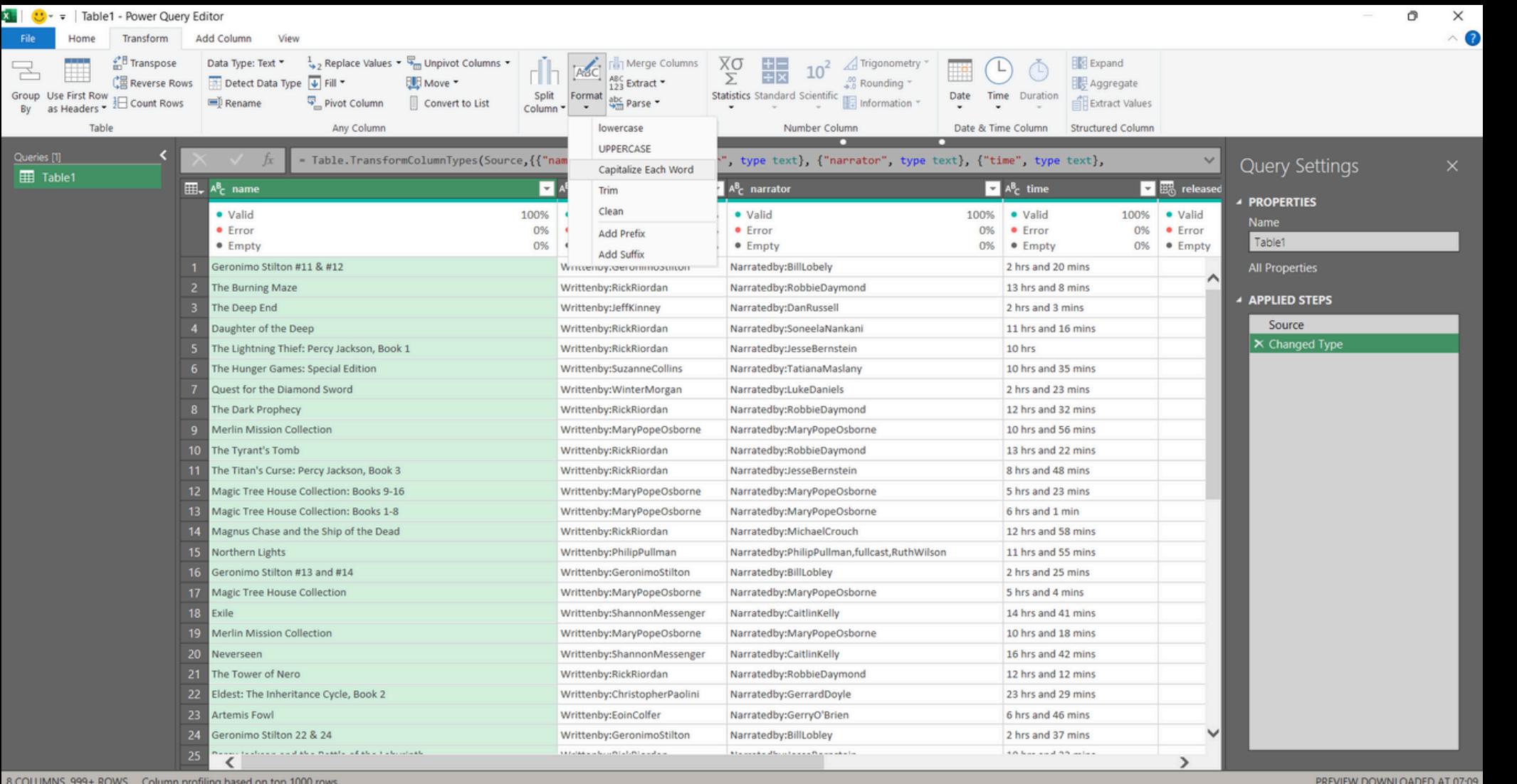
Perfect for your daily commute, workout routine, or just relaxing at home, Audible lets you enjoy captivating narratives and insightful content anytime, anywhere. With personalized recommendations and offline listening, discover new worlds and learn something new with Audible.



Link to Cleaned Dataset



Standardize the name column to ensure consistent title casing.



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

Screenshot of the Power Query Editor showing a query named "Table1". The query has 10 columns and 999+ rows. The "author_1_full_name" column contains combined first and last names, while other columns like "name" and "narrator" contain individual names. The "Applied Steps" pane shows the steps taken to process the data, including "Merged Columns2".

Query Settings:

- Name: Table1
- All Properties

Applied Steps:

- Source
- Changed Type
- Capitalized Each Word
- Trimmed Text
- Split Column by Delimiter
- Changed Type1
- Removed Columns
- Split Column by Delimiter1
- Changed Type2
- Removed Columns1
- Split Column by Character Tra...
- Trimmed Text1
- Merged Columns
- Split Column by Character Tra...
- Merged Columns1
- Split Column by Character Tra...
- Merged Columns2**

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

The screenshot shows the Microsoft Power Query Editor interface with two custom column dialogs open over a main query preview window.

Left Dialog (Custom Column):

- New column name:** hrs
- Custom column formula:**

```
= if Text.Contains([time], "hr") then  
    Number.FromText(Text.BeforeDelimiter([time], " hr"))  
else 0
```
- Properties:** Valid (100%), Error (0%), Empty (0%)
- Applied Steps:** Removed Columns3, Replaced Errors(free) with 0, Changed to dtype to cure..., Trimmed Text2, Split the ratings column, Replaced Value, Changed Type6, Renamed Columns, Split the other part of rating, Removed Columns4, Replaced Value1, Replaced Value2, Changed Type5, Renamed Columns1, Filtered Row, Replaced Value3, Added Custom, Reordered Columns, Renamed Columns2, **Added Custom1**, Reordered Columns1, Filtered Rows1

Right Dialog (Custom Column):

- New column name:** Custom
- Custom column formula:**

```
= if Text.Contains([time], "min") and not Text.Contains([time], "and") then  
    Number.FromText(Text.BeforeDelimiter([time], " min"))  
else if Text.Contains([time], "min") then  
    Number.FromText(Text.BeforeDelimiter(Text.AfterDelimiter([time], " and"), " min"))  
else 0
```
- Properties:** Valid (23%), Error (0%), Empty (77%)
- Applied Steps:** Merged Columns1, Removed Columns3, Replaced Errors(free) with 0, Changed to dtype to cure..., Trimmed Text2, Split the ratings column, Replaced Value, Changed Type6, Renamed Columns, Split the other part of rating, Removed Columns4, Replaced Value1, Replaced Value2, Changed Type5, Renamed Columns1, Filtered Rows, Replaced Value3, **Added Custom**, Reordered Columns, Renamed Columns2, Added Custom1, Reordered Columns1, Filtered Rows1

Main Query Editor:

- Table1** (15 columns, 999+ rows)
- Operations:** Conditional Column, Merge Columns, Trigonometry, Statistics Standard Scientific, Information, Date, Time, Duration, From Text, From Number, From Date & Time, From Date & Time
- Properties:** Name: Table1
- Applied Steps:** Merged Columns1, Removed Columns3, Replaced Errors(free) with 0, Changed to dtype to cure..., Trimmed Text2, Split the ratings column, Replaced Value, Changed Type6, Renamed Columns, Split the other part of rating, Removed Columns4, Replaced Value1, Replaced Value2, Changed Type5, Renamed Columns1, Filtered Rows, Replaced Value3, **Added Custom**, Reordered Columns, Renamed Columns2, Added Custom1, Reordered Columns1, Filtered Rows1

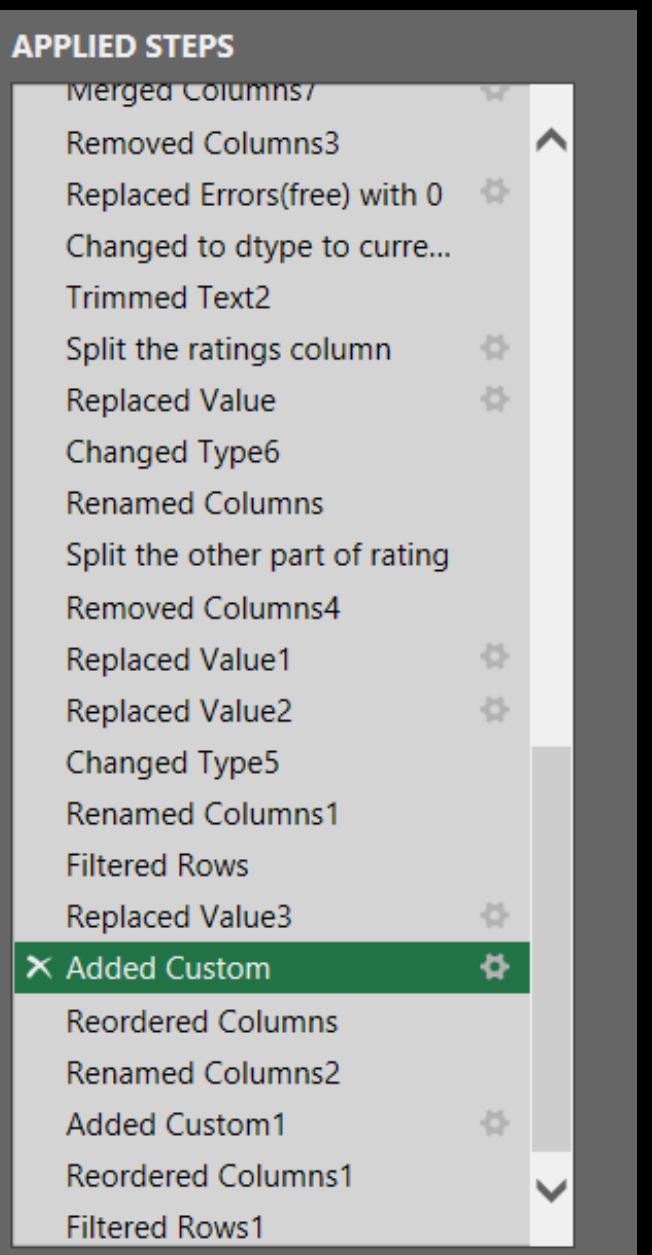
Ensure the price column is in a numeric format, and identify any non-numeric values.

The screenshot shows the Power Query Editor interface with the following details:

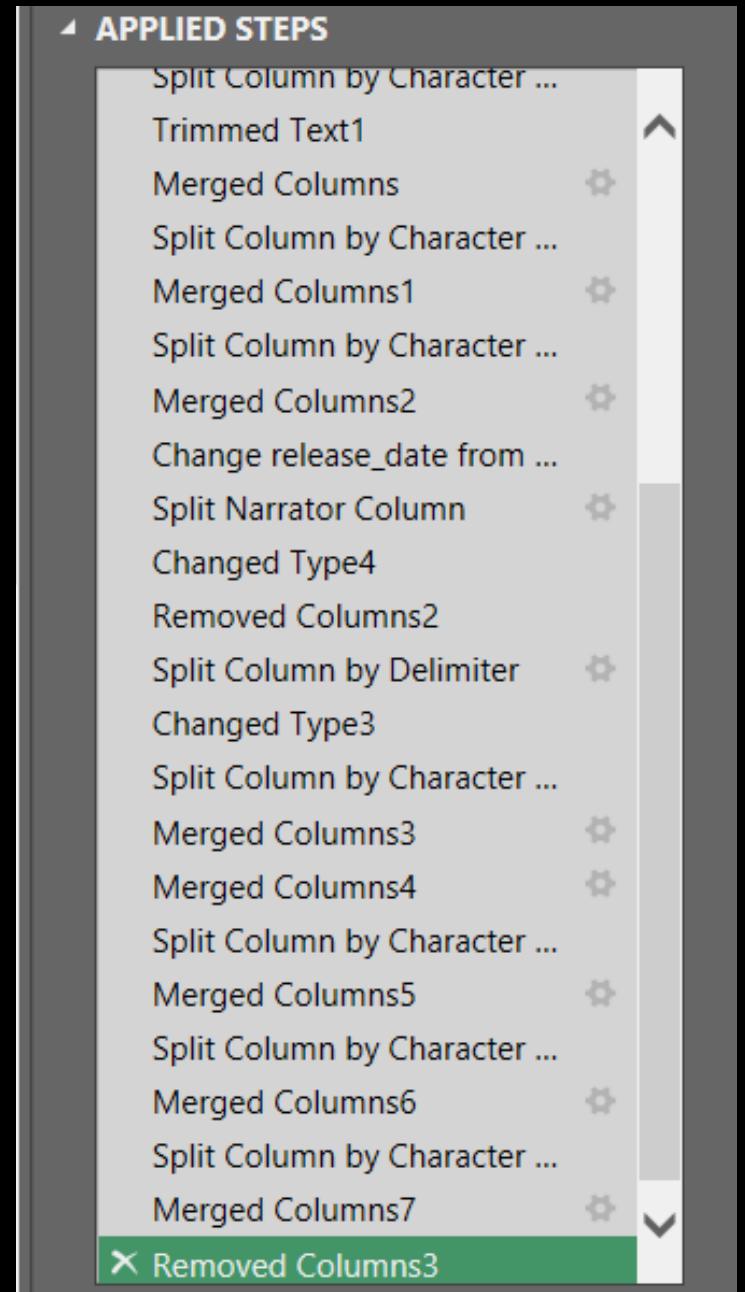
- File**, **Home**, **Transform**, **Add Column**, **View** tabs are visible at the top.
- Transform ribbon**: Data Type, Replace Values, Unpivot Columns, Detect Data Type, Fill, Move, Split Column, Format, Parse, Statistics, Standard, Scientific, Information, Date, Time, Duration, Expand, Aggregate, Extract Values.
- Queries [3]**: Table1, Errors in Table1, Table1_1 (selected).
- Table1_1 Data View**: A table with columns: narrator3_full_name, time, releasedate, language, stars, and price. The 'price' column is highlighted with a dropdown menu open, showing options like Decimal Number, Currency (highlighted), Whole Number, Percentage, Date/Time, Date, Time, Duration, Text, True/False, and Binary.
- Query Settings** pane:
 - Properties**: Name is set to Table1_1.
 - Applied Steps**: Shows the step "Changed Type".
- Bottom Status Bar**: 12 COLUMNS, 999+ ROWS, Column profiling based on top 1000 rows, PREVIEW DOWNLOADED AT 08:15.

Replaced Free with 0 and then converted whole column to currency

Convert text ratings in the stars column to numeric values.



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



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

The screenshot shows the Microsoft Power Query Editor interface. The main area displays a table with several columns: narrator2_full_name, narrator3_full_name, time, release_date_with_lang, stars, and price. A context menu is open over the 'release_date_with_lang' and 'stars' columns, with the 'Merge Columns' option selected. A 'Merge Columns' dialog box is centered, prompting the user to choose how to merge the selected columns. The 'Separator' dropdown is set to 'Comma', and the 'New column name (optional)' field contains the text 'release_date_with_lang'. The 'OK' button is highlighted with a green border. In the top right corner of the editor, the 'Query Settings' pane is visible, showing the 'Name' as 'Table1_1' and the 'APPLIED STEPS' list which includes 'Merged Columns'. The bottom status bar indicates '11 COLUMNS, 999+ ROWS' and 'Column profiling based on top 1000 rows'.

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

The screenshot shows the Microsoft Power Query Editor interface. The main area displays a table with several columns: narrator2_full_name, narrator3_full_name, time, release_date_with_lang, stars, and price. A context menu is open over the 'release_date_with_lang' and 'stars' columns, with the 'Merge Columns' option selected. A 'Merge Columns' dialog box is centered, prompting the user to choose how to merge the selected columns. The 'Separator' dropdown is set to 'Comma', and the 'New column name (optional)' field contains the text 'release_date_with_lang'. The 'OK' button is highlighted in green. In the top right corner of the editor, the 'Query Settings' pane is visible, showing the 'Name' as 'Table1_1' and the 'APPLIED STEPS' list which includes 'Merged Columns'.

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

Table1_1 - Power Query Editor

File Home Transform Add Column View

Data Type: Text Replace Values Unpivot Columns Merge Columns Statistics Standard Scientific Information

Transpose Reverse Rows Detect Data Type Fill Move Split Column Format ABC Extract 10² Trigonometry

Group Use First Row By as Headers Count Rows Rename Pivot Column Convert to List Text Column Number Column Date Time Duration

Table Any Column Date & Time Column Structured Column

Queries [3] Table1 Errors in Table1 Table1_1

AB_C narrator3_full_name AB_C time AB_C releasedate AB_C language AB_C stars AB_C price

	AB _C narrator3_full_name	AB _C time	AB _C releasedate	AB _C language	AB _C stars	AB _C price
1		2 hrs and 20 mins	04-08-2008 00:00:00	English	5 out of 5 stars	34 ratings
2		13 hrs and 8 mins	01-05-2018 00:00:00	English	4.5 out of 5 stars	41 ratings
3		2 hrs and 3 mins	06-11-2020 00:00:00	English	4.5 out of 5 stars	38 ratings
4		11 hrs and 16 mins	05-10-2021 00:00:00	English	4.5 out of 5 stars	12 ratings
5		10 hrs	13-01-2010 00:00:00	English	4.5 out of 5 stars	181 ratings
6		10 hrs and 35 mins	30-10-2018 00:00:00	English	5 out of 5 stars	72 ratings
7		2 hrs and 23 mins	25-11-2014 00:00:00	English	5 out of 5 stars	11 ratings
8		12 hrs and 32 mins	02-05-2017 00:00:00	English	5 out of 5 stars	50 ratings
9		10 hrs and 56 mins	02-05-2017 00:00:00	English	5 out of 5 stars	5 ratings
10		13 hrs and 22 mins	24-09-2019 00:00:00	English	5 out of 5 stars	58 ratings
11		8 hrs and 48 mins	14-01-2010 00:00:00	English	4.5 out of 5 stars	130 ratings
12		5 hrs and 23 mins	24-08-2011 00:00:00	English	5 out of 5 stars	6 ratings
13		6 hrs and 1 min	27-09-2011 00:00:00	English	5 out of 5 stars	7 ratings
14		12 hrs and 58 mins	03-10-2017 00:00:00	English	5 out of 5 stars	41 ratings
15	Ruth Wilson	11 hrs and 55 mins	24-06-2021 00:00:00	English	4 out of 5 stars	2 ratings
16		2 hrs and 25 mins	08-02-2008 00:00:00	English	4.5 out of 5 stars	33 ratings
17		5 hrs and 4 mins	26-12-2004 00:00:00	English	5 out of 5 stars	7 ratings
18		14 hrs and 41 mins	06-11-2018 00:00:00	English	5 out of 5 stars	20 ratings
19		10 hrs and 18 mins	02-05-2017 00:00:00	English	5 out of 5 stars	11 ratings
20		16 hrs and 42 mins	06-11-2018 00:00:00	English	5 out of 5 stars	13 ratings
21		12 hrs and 12 mins	06-10-2020 00:00:00	English	5 out of 5 stars	79 ratings
22		23 hrs and 29 mins	14-04-2009 00:00:00	English	4.5 out of 5 stars	47 ratings
23		6 hrs and 46 mins	11-04-2013 00:00:00	English	5 out of 5 stars	27 ratings
24		2 hrs and 37 mins	01-02-2011 00:00:00	English	4.5 out of 5 stars	25 ratings
25		40 hrs and 20 mins	21-01-2010 00:00:00	English	5 out of 5 stars	69 ratings

12 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

Query Settings

Properties Name: Table1_1

Applied Steps Source: Changed Type

PREVIEW DOWNLOADED AT 08:15

Thank You

