



Project presentation

*Presented by MOHAMMED
ASIF QURESHI*



INTRODUCTION



The primary goal of this project was to enhance the usability and value of an Audible dataset by cleaning and standardizing its data. By addressing common data quality issues, such as inconsistent formatting, missing values, and combined fields, we aimed to create a dataset that is suitable for efficient data analysis and exploration.



Project vision and mission

1. Standardize the name column to ensure consistent title casing.
2. Separate combined first and last names in the author column if they are currently combined.
3. Ensure all entries in the releasedate column follow a consistent date format (DD-MM-YYYY).
4. Convert the time column from text format to a duration format that Excel recognizes.
5. Ensure the price column is in a numeric format, and identify any non-numeric values.
6. Convert text ratings in the stars column to numeric values.
7. Split the narratedby column into multiple columns if multiple narrators are listed.
8. Merge the releasedate and language columns into a single new column named releaseinfo with the format "DD-MM-YYYY, Language."
9. Ensure all currency values in the price column are formatted consistently with two decimal places.



Standardize the name column to ensure consistent title casing.



Screenshot of Microsoft Power Query Editor showing a table named "audible_uncleaned". The table has three columns: "name", "Writtenby", and "Narratedby". The "name" column contains various book titles. The "Format" context menu is open over the first row of the "name" column, with the "UPPERCASE" option selected. A tooltip for "UPPERCASE" states: "Convert the first letter of each word in the selected columns into an uppercase letter."

name	Writtenby	Narratedby
Geronimo Stilton #11 & #12	Rick Riordan	Dan Russell
The Burning Maze	Rick Riordan	Soneela Nankani
The Deep End	Suzanne Collins	Tatiana Maslany
Daughter of the Deep	Rick Riordan	Jesse Bean
The Lightning Thief: Percy Jackson, Book 1	Rick Riordan	Robbie Daynes
The Hunger Games: Special Edition	Mary Pope Osborne	Mary Pope Osborne
Quest for the Diamond Sword	Winter Morgan	Luke Daniels
The Dark Prophecy	Rick Riordan	Robbie Daynes
Merlin Mission Collection	Mary Pope Osborne	Mary Pope Osborne
The Tyrantâ€™s Tomb	Rick Riordan	Robbie Daynes
The Titan's Curse: Percy Jackson, Book 3	Rick Riordan	Jesse Bean
Magic Tree House Collection: Books 9-16	Mary Pope Osborne	Mary Pope Osborne
Magic Tree House Collection: Books 1-8	Mary Pope Osborne	Mary Pope Osborne
Magnus Chase and the Ship of the Dead	Rick Riordan	Michael Gilmore
Northern Lights	Philip Pullman	Philip Pullman
Geronimo Stilton #13 and #14	Geronimo Stilton	Bill Lobley
Magic Tree House Collection	Mary Pope Osborne	Mary Pope Osborne
Exile	Shannon Messenger	Caitlyn Johnson
Merlin Mission Collection	Mary Pope Osborne	Mary Pope Osborne
Neverseen	Shannon Messenger	Caitlyn Johnson





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

Any Column

Text

fx = Table.PromoteHeaders(Source, [PromoteAllScalars])

A _B C author	A _B C narrator
Writtenby:GeronimoStilton	Narratedby:Billobley
Writtenby:RickRiordan	Narratedby:RobbieDaymond
Writtenby:JeffKinney	Narratedby:DanRussell
Writtenby:RickRiordan	Narratedby:SoneelaNankani
Writtenby:RickRiordan	Narratedby:JesseBernstein
Writtenby:SuzanneCollins	Narratedby:TatianaMaslany
Writtenby:WinterMorgan	Narratedby:LukeDaniels
Writtenby:RickRiordan	Narratedby:RobbieDaymond
Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne
Writtenby:RickRiordan	Narratedby:RobbieDaymond
Writtenby:RickRiordan	Narratedby:JesseBernstein
Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne

Text Column

Length

First Characters

Last Characters

Range

Text Before Delimiter

Text After Delimiter

Text Between Delimiters

Return the text that follows the delimiter.

Query Editor

Any Column

Text Column

Number Column

Date & Time Column

Structured Column

fx = Table.TransformColumns(#"Promoted Headers", {"author", each Text.AfterDelimiter(_, ":"), type text})

A _B C author	A _B C narrator	A _B C time	A _B C releasedate	A _B C language
GeronimoStilton	Narratedby:Billlobley	2 hrs and 20 mins	04-08-08	English
RickRiordan	Narratedby:RobbieDaymond	13 hrs and 8 mins	01-05-18	English
JeffKinney	Narratedby:DanRussell	2 hrs and 3 mins	06-11-20	English
RickRiordan	Narratedby:SoneelaNankani	11 hrs and 16 mins	05-10-21	English
RickRiordan	Narratedby:JesseBernstein	10 hrs	13-01-10	English
SuzanneCollins	Narratedby:TatianaMaslany	10 hrs and 35 mins	30-10-18	English
WinterMorgan	Narratedby:LukeDaniels	2 hrs and 23 mins	25-11-14	English
RickRiordan	Narratedby:RobbieDaymond	12 hrs and 32 mins	02-05-17	English
MaryPopeOsborne	Narratedby:MaryPopeOsborne	10 hrs and 56 mins	02-05-17	English
RickRiordan	Narratedby:RobbieDaymond	13 hrs and 22 mins	24-09-19	English
RickRiordan	Narratedby:JesseBernstein	8 hrs and 48 mins	14-01-10	English
MaryPopeOsborne	Narratedby:MaryPopeOsborne	5 hrs and 23 mins	24-08-11	English
MaryPopeOsborne	Narratedby:MaryPopeOsborne	6 hrs and 1 min	27-09-11	English
RickRiordan	Narratedby:MichaelCrouch	12 hrs and 58 mins	03-10-17	English
PhilipPullman	Narratedby:PhilipPullman,fullcast,RuthWilson	11 hrs and 55 mins	24-06-21	English
GeronimoStilton	Narratedby:Billlobley	2 hrs and 25 mins	08-02-08	English
MaryPopeOsborne	Narratedby:MaryPopeOsborne	5 hrs and 4 mins	26-12-04	English
ShannonMessenger	Narratedby:CaitlinKelly	14 hrs and 41 mins	06-11-18	English
MaryPopeOsborne	Narratedby:MaryPopeOsborne	10 hrs and 18 mins	02-05-17	English
ShannonMessenger	Narratedby:CaitlinKelly	16 hrs and 42 mins	06-11-18	English
RickRiordan	Narratedby:RobbieDaymond	12 hrs and 12 mins	06-10-20	English
ChristopherPaolini	Narratedby:GerrardDoyle	23 hrs and 29 mins	14-04-09	English
EoinColfer	Narratedby:GerryO'Brien	6 hrs and 46 mins	11-04-13	English
GeronimoStilton	Narratedby:Billlobley	2 hrs and 37 mins	01-02-11	English
RickRiordan	Narratedby:JesseBernstein	10 hrs and 32 mins	21-01-10	English
A.A.Milne	Narratedby:PeterDennis	2 hrs and 46 mins	28-12-03	English
MaryPopeOsborne	Narratedby:MaryPopeOsborne	7 hrs and 9 mins	01-09-20	English

Based on top 1000 rows

Search

Excel



WE HAVE TO SPLIT AUTHOR COLUMN BASED ON COMMAS.

Innocent Boy

Query Editor

Type: Text Replace Values Unpivot Columns

Detect Data Type Fill Move

Rename Pivot Column Convert to List

Any Column

Text Column Number Column Date & Time Column Structured Column

ABC 123 Extract abc Parse

XO Σ 10² Trigonometry Statistics Standard Scientific Information

Date Time Duration Expand Aggregate Extract Values

Table.TransformColumns#"Promoted Headers", {"author", each Text.AfterDelimiter(_, ":"), type text})

A ^B C author	A ^B C narrator	A ^B C time	A ^B C releasedate	A ^B C language
GeronimoStilton	Narratedby:BillLobely	2 hrs and 20 mins	04-08-08	English
RickRiordan	Narratedby:RobbieDaymond	13 hrs and 8 mins	01-05-18	English
JeffKinney	Narratedby:DanRussell	2 hrs and 3 mins	06-11-20	English
RickRiordan	Narratedby:SoneelaNankani	11 hrs and 16 mins	05-10-21	English
		13-01-10		English
		30-10-18		English
		25-11-14		English
		02-05-17		English
		02-05-17		English
		24-09-19		English
		14-01-10		English
		24-08-11		English
		27-09-11		English
		03-10-17		English
		24-06-21		English
		08-02-08		English
		26-12-04		English
		06-11-18		English
		02-05-17		English
		06-11-18		English
		06-10-20		English
		14-04-09		English
		11-04-13		English
		01-02-11		English
		21-01-10		English
		28-12-03		English
A.A.Milne	Narratedby:PeterDennis	2 hrs and 46 mins		
MaryPopeOsborne	Narratedby:MaryPopeOsborne	7 hrs and 9 mins	01-09-20	English

Split Column by Delimiter

Specify the delimiter used to split the text column.

Select or enter delimiter: Comma

Split at: Each occurrence of the delimiter

Advanced options

Quote Character: "

Split using special characters: Insert special character

OK Cancel

Based on top 1000 rows

Search Microsoft Edge File Explorer Task View Google Chrome Microsoft Excel

Query Editor

Type: Text Replace Values Unpivot Columns

Detect Data Type Fill Move

Pivot Column Convert to List

Any Column

Text Column Number Column

ABC 123 Extract abc Parse

XO Σ 10² Trigonometry Statistics Standard Scientific Information

Date Time Duration Expand Aggregate Extract Values

= Table.TransformColumnTypes#"Split Column by Delimiter",{{"name", type text}, {"author.1", type text}}

A ^B C author.1	A ^B C author.2	A ^B C author.3	A ^B C author.4
GeronimoStilton		null	null
RickRiordan		null	null
JeffKinney		null	null
RickRiordan		null	null
RickRiordan		null	null
SuzanneCollins		null	null
WinterMorgan		null	null
RickRiordan		null	null
MaryPopeOsborne		null	null
RickRiordan		null	null
RickRiordan		null	null
MaryPopeOsborne		null	null
MaryPopeOsborne		null	null
RickRiordan		null	null





IN AUTHOR 2 AND AUTHOR 3 COLUMN WE HAVE TO REPLACE NULL VALUES TO NOT APPLICABLE



Replace Values

Replace one value with another in the selected columns.

Value To Find: null

Replace With: NOT APPLICABLE

OK Cancel

author.1 author.2 author.3 narrator time

author.1	author.2	author.3	narrator	time
GeronimoStilton	null	null	Narratedby:BillLobley	2 hrs and 20 mins
RickRiordan	null	null	Narratedby:RobbieDaymond	13 hrs and 8 mins
JeffKinney	null	null	Narratedby:DanRussell	2 hrs and 3 mins
RickRiordan	null	null		5 mins
SuzanneCollins	null	null		5 and 35 mins
WinterMorgan	null	null		and 23 mins
MaryPopeOsborne	null	null		5 and 32 mins
RickRiordan	null	null		5 and 56 mins
RickRiordan	null	null		5 and 22 mins
RickRiordan	null	null		and 48 mins
RickRiordan	null	null		and 23 mins
RickRiordan	null	null		and 1 min
RickRiordan	null	null		5 and 58 mins
PhilipPullman	null	null		5 and 55 mins
GeronimoStilton	null	null		and 25 mins
MaryPopeOsborne	null	null		and 4 mins
CaitlinKelly	null	null		14 hrs and 41 mins
MaryPopeOsborne	null	null		10 hrs and 18 mins
CaitlinKelly	null	null		16 hrs and 42 mins
RobbieDaymond	null	null		12 hrs and 12 mins
GerrardDoyle	null	null		23 hrs and 29 mins
GerryO'Brien	null	null		6 hrs and 46 mins
BillLobley	null	null		2 hrs and 37 mins
JesseBernstein	null	null		10 hrs and 32 mins
PeterDennis	null	null		2 hrs and 46 mins
MaryPopeOsborne	null	null		7 hrs and 9 mins

uncleaned - Power Query Editor

Transform Add Column View

Replace Values

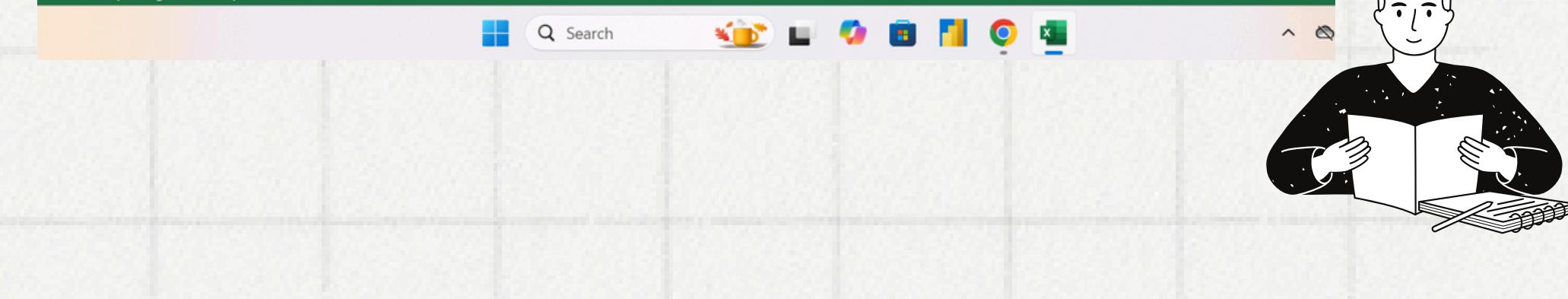
= Table.ReplaceValue(#"Removed Columns",null,"NOT APPLICABLE",Replacer.ReplaceValue,{"author.2", "author.3"})

Properties Name: audible_uncleaned

Applied Steps

Source Promoted Headers Extracted Text Split Column Changed Type Filtered Rows Removed Columns

author.1	author.2	author.3	narrator	time
GeronimoStilton	NOT APPLICABLE	NOT APPLICABLE	Narratedby:BillLobley	2 hrs and 20 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:RobbieDaymond	13 hrs and 8 mins
JeffKinney	NOT APPLICABLE	NOT APPLICABLE	Narratedby:DanRussell	2 hrs and 3 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:SoneelaNankani	11 hrs and 16 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:JesseBernstein	10 hrs
SuzanneCollins	NOT APPLICABLE	NOT APPLICABLE	Narratedby:TatianaMaslany	10 hrs and 35 mins
WinterMorgan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:LukeDaniels	2 hrs and 23 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:RobbieDaymond	12 hrs and 32 mins
MaryPopeOsborne	NOT APPLICABLE	NOT APPLICABLE	Narratedby:MaryPopeOsborne	10 hrs and 56 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:RobbieDaymond	13 hrs and 22 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:JesseBernstein	8 hrs and 48 mins
MaryPopeOsborne	NOT APPLICABLE	NOT APPLICABLE	Narratedby:MaryPopeOsborne	5 hrs and 23 mins
MaryPopeOsborne	NOT APPLICABLE	NOT APPLICABLE	Narratedby:MaryPopeOsborne	6 hrs and 1 min
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:MichaelCrouch	12 hrs and 58 mins
PhilipPullman	NOT APPLICABLE	NOT APPLICABLE	Narratedby:PhilipPullman,fullcast,RuthWilson	11 hrs and 55 mins
GeronimoStilton	NOT APPLICABLE	NOT APPLICABLE	Narratedby:BillLobley	2 hrs and 25 mins
MaryPopeOsborne	NOT APPLICABLE	NOT APPLICABLE	Narratedby:MaryPopeOsborne	5 hrs and 4 mins
ShannonMessenger	NOT APPLICABLE	NOT APPLICABLE	Narratedby:CaitlinKelly	14 hrs and 41 mins
MaryPopeOsborne	NOT APPLICABLE	NOT APPLICABLE	Narratedby:MaryPopeOsborne	10 hrs and 18 mins
ShannonMessenger	NOT APPLICABLE	NOT APPLICABLE	Narratedby:CaitlinKelly	16 hrs and 42 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:RobbieDaymond	12 hrs and 12 mins
ChristopherPaolini	NOT APPLICABLE	NOT APPLICABLE	Narratedby:GerrardDoyle	23 hrs and 29 mins
EoinColfer	NOT APPLICABLE	NOT APPLICABLE	Narratedby:GerryO'Brien	6 hrs and 46 mins
GeronimoStilton	NOT APPLICABLE	NOT APPLICABLE	Narratedby:BillLobley	2 hrs and 37 mins
RickRiordan	NOT APPLICABLE	NOT APPLICABLE	Narratedby:JesseBernstein	10 hrs and 32 mins
A.A.Milne	NOT APPLICABLE	NOT APPLICABLE	Narratedby:PeterDennis	2 hrs and 46 mins
MaryPopeOsborne	NOT APPLICABLE	NOT APPLICABLE	Narratedby:MaryPopeOsborne	7 hrs and 9 mins



Now again we have to split author 1 column (by lower case to uppercase) and merge them based on space and trim them.



Litator

View

Replace Values Unpivot Columns

Data Type Fill Move

Pivot Column Convert to List

Split Column

Merge Columns

Statistics Standard Scientific Trigonometry

10²

Date Time Duration

Expand Aggregate Extract Values

Number Column Date & Time Column Structured Column

By Delimiter By Number of Characters By Positions

By Lowercase to Uppercase

Split Column by Lowercase to Uppercase

Split values in the selected column based on transitions from a lowercase letter to an uppercase letter.

A^B author.2 A^B author.3 A^B narrator

	A ^B author.2	A ^B author.3	A ^B narrator
Geronimo	NOT APPLICABLE	Narratedby:BillLobley	
Rick	NOT APPLICABLE	Narratedby:RobbieDaymond	
Jeff	NOT APPLICABLE	Narratedby:DanRussell	
Dick	NOT APPLICABLE	Narratedby:SoneelaNankani	
RickRiordan	NOT APPLICABLE	Narratedby:JesseBernstein	
SuzanneCollins	NOT APPLICABLE	Narratedby:TatianaMaslany	
WinterMorgan	NOT APPLICABLE	Narratedby:LukeDaniels	
RickRiordan	NOT APPLICABLE	Narratedby:RobbieDaymond	
MaryPopeOsborne	NOT APPLICABLE	Narratedby:MaryPopeOsborne	
RickRiordan	NOT APPLICABLE	Narratedby:RobbieDaymond	
RickRiordan	NOT APPLICABLE	Narratedby:JesseBernstein	
MaryPopeOsborne	NOT APPLICABLE	Narratedby:MaryPopeOsborne	
RickRiordan	NOT APPLICABLE	Narratedby:RobbieDaymond	
RickRiordan	NOT APPLICABLE	Narratedby:JesseBernstein	
MaryPopeOsborne	NOT APPLICABLE	Narratedby:MaryPopeOsborne	
RickRiordan	NOT APPLICABLE	Narratedby:RobbieDaymond	
RickRiordan	NOT APPLICABLE	Narratedby:MichaelCrouch	
PhilipPullman	NOT APPLICABLE	Narratedby:PhilipPullman,fullcast,RuthV	
GeronimoStilton	NOT APPLICABLE	Narratedby:BillLobley	
MaryPopeOsborne	NOT APPLICABLE	Narratedby:MaryPopeOsborne	
ShannonMessenger	NOT APPLICABLE	Narratedby:CaitlinKelly	
MaryPopeOsborne	NOT APPLICABLE	Narratedby:MaryPopeOsborne	
ShannonMessenger	NOT APPLICABLE	Narratedby:CaitlinKelly	
RickRiordan	NOT APPLICABLE	Narratedby:RobbieDaymond	
ChristopherPaolini	NOT APPLICABLE	Narratedby:GerrardDoyle	
EoinColfer	NOT APPLICABLE	Narratedby:GerryO'Brien	
GeronimoStilton	NOT APPLICABLE	Narratedby:BillLobley	
RickRiordan	NOT APPLICABLE	Narratedby:JesseBernstein	
A.A.Milne	NOT APPLICABLE	Narratedby:PeterDennis	
MaryPopeOsborne	NOT APPLICABLE	Narratedby:MaryPopeOsborne	

in top 1000 rows

Search

Values Unpivot Columns

Move Convert to List

Split Column Format Merge Columns

Text Column Statistics Standard Scientific Trigonometry

10²

Date Time Duration

Expand Aggregate Extract Values

Number Column Date & Time Column Structured Column

Replacer.ReplaceValue,{"author.2", "author.3"} = Table.ReplaceValue(#"Removed Columns", "author.2", "author.3")

	A ^B author.1.1	A ^B author.1.2	A ^B author.1.3	A ^B author.2	A ^B author.3	A ^B narrator
Geronimo				null	NOT APPLICABLE	NOT APPLICABLE
Rick				null	NOT APPLICABLE	NOT APPLICABLE
Jeff				null	NOT APPLICABLE	NOT APPLICABLE
Dick				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
SuzanneCollins				null	NOT APPLICABLE	NOT APPLICABLE
WinterMorgan				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
MaryPopeOsborne				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
MaryPopeOsborne				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
MaryPopeOsborne				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
MaryPopeOsborne				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
PhilipPullman				null	NOT APPLICABLE	NOT APPLICABLE
GeronimoStilton				null	NOT APPLICABLE	NOT APPLICABLE
MaryPopeOsborne				null	NOT APPLICABLE	NOT APPLICABLE
ShannonMessenger				null	NOT APPLICABLE	NOT APPLICABLE
MaryPopeOsborne				null	NOT APPLICABLE	NOT APPLICABLE
ShannonMessenger				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
ChristopherPaolini				null	NOT APPLICABLE	NOT APPLICABLE
EoinColfer				null	NOT APPLICABLE	NOT APPLICABLE
GeronimoStilton				null	NOT APPLICABLE	NOT APPLICABLE
RickRiordan				null	NOT APPLICABLE	NOT APPLICABLE
A.A.Milne				null	NOT APPLICABLE	NOT APPLICABLE
Mary				null	NOT APPLICABLE	NOT APPLICABLE
Shannon				null	NOT APPLICABLE	NOT APPLICABLE
Mary				null	NOT APPLICABLE	NOT APPLICABLE
Shannon				null	NOT APPLICABLE	NOT APPLICABLE
Rick				null	NOT APPLICABLE	NOT APPLICABLE
Christopher				null	NOT APPLICABLE	NOT APPLICABLE
Eoin				null	NOT APPLICABLE	NOT APPLICABLE
Geronimo				null	NOT APPLICABLE	NOT APPLICABLE
Rick				null	NOT APPLICABLE	NOT APPLICABLE
A.A.Milne				null	NOT APPLICABLE	NOT APPLICABLE
Mary				null	NOT APPLICABLE	NOT APPLICABLE

Merge Columns

Choose how to merge the selected columns.

Separator Space

New column name (optional) Author 1

OK Cancel

PREVIEW DOWN

Search

ENG IN



Now again we have to split author 2 column (by lower case to uppercase) and merge them based on space and trim them.

SPLIT COLUMN

FORMAT abc Parse

By Delimiter

By Number of Characters

By Positions

By Lowercase to Uppercase

Split Column by Lowercase to Uppercase

Split values in the selected column based on transitions from a lowercase letter to an uppercase letter.

Author	Author.2	Narrator
Geronimo Stilton	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Jeff Kinney	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Suzanne Collins	NOT APPLICABLE	NOT APPLICABLE
Winter Morgan	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Mary Pope Osborne	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Mary Pope Osborne	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Mary Pope Osborne	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Philip Pullman	NOT APPLICABLE	NOT APPLICABLE
Geronimo Stilton	NOT APPLICABLE	NOT APPLICABLE
Mary Pope Osborne	NOT APPLICABLE	NOT APPLICABLE
Shannon Messenger	NOT APPLICABLE	NOT APPLICABLE
Mary Pope Osborne	NOT APPLICABLE	NOT APPLICABLE
Shannon Messenger	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
Christopher Paolini	NOT APPLICABLE	NOT APPLICABLE
Eoin Colfer	NOT APPLICABLE	NOT APPLICABLE
Geronimo Stilton	NOT APPLICABLE	NOT APPLICABLE
Rick Riordan	NOT APPLICABLE	NOT APPLICABLE
A.A.Milne	NOT APPLICABLE	NOT APPLICABLE
Mary Pope Osborne	NOT APPLICABLE	NOT APPLICABLE

Replace Values Unpivot Columns Move ABC Extract Merge Columns Statistics Standard Scientific 10² Trigonometry Fill Pivot Column Convert to List Format abc Parse Any Column Text Column Number Column Date & Time Column Structured Column

Query Settings

PROPERTIES

Name: audible_uncleaned

APPLIED STEPS

- Source
- Promoted Headers
- Extracted Text After Delimiter
- Split Column by Delimiter
- Changed Type
- Filtered Rows
- Removed Columns
- Replaced Value
- Split Column by Character Transition
- Merged Columns
- Trimmed Text
- Split Column by Character Transition

Author	author.2.1	author.2.2	author.3	narrator
Geronimo Stilton	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:BillLobley
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:RobbieDaymond
Jeff Kinney	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:DanRussell
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:SoneelaNankani
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:JesseBernstein
Suzanne Collins	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:TatianaMaslany
Winter Morgan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:LukeDaniels
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:RobbieDaymond
Mary Pope Osborne	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:MaryPopeOsborne
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:RobbieDaymond
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:JesseBernstein
Mary Pope Osborne	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:MaryPopeOsborne
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:MichaelCrouch
Philip Pullman	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:PhilipPullman,fulcast,Ruth
Geronimo Stilton	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:BillLobley
Mary Pope Osborne	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:MaryPopeOsborne
Shannon Messenger	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:CaitlinKelly
Mary Pope Osborne	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:MaryPopeOsborne
Shannon Messenger	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:CaitlinKelly
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:RobbieDaymond
Christopher Paolini	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:GerrardDoyle
Eoin Colfer	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:GerryO'Brien
Geronimo Stilton	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:BillLobley
Rick Riordan	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:JesseBernstein
A.A.Milne	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:PeterDennis
Mary Pope Osborne	NOT APPLICABLE	null	NOT APPLICABLE	Narratedby:MaryPopeOsborne



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

- ***Now again we have to split author 3 column (by lower case to uppercase) and merge them based on space and trim them.***
- ***Repeat the same on narrator as we did for author name.***
- ***WE HAVE TO SPLIT narrator COLUMN BASED ON COMMAS.***





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

The screenshot shows a Microsoft Excel spreadsheet with a Power Query editor overlay. The spreadsheet contains a table with columns: releasedate, language, stars, price, time, and Hours. The 'time' column contains text values like "43 mins" and "3 hrs and 18 mins". The 'Hours' column contains numerical values ranging from 0 to 13. The Power Query editor is open, showing a 'Custom Column' dialog. In the 'Custom column formula' field, the following M code is written:

```
= if Text.Contains([time],"hr")
then Number.FromText(Text.BeforeDelimiter([time], "hr"))

else if Text.Contains([time],"hrs")
then Number.FromText(Text.BeforeDelimiter([time], "hrs"))

else 0
```

The 'Available columns' list includes: Narrator 2, Narrator 3, releasedate, language, stars, price, and time. The 'Custom Column' dialog also displays a note: "No syntax errors have been detected." At the bottom right of the dialog are 'OK' and 'Cancel' buttons.

rice	time	Hours
	43 mins	0
	3 hrs and 18 mins	3
	6 hrs and 9 mins	6
	6 hrs and 50 mins	6
	3 hrs	3
	8 hrs and 59 mins	8
	1 hr and 34 mins	1
	3 hrs and 5 mins	3
	26 mins	0
	1 hr and 13 mins	1
	1 hr and 2 mins	1
	8 hrs and 25 mins	8
	13 hrs and 18 mins	13
	2 hrs and 15 mins	2
	7 mins	0
	2 hrs and 12 mins	2
	59 mins	0
	5 hrs and 37 mins	5
	4 hrs and 2 mins	4
	1 hr and 13 mins	1
	5 hrs and 51 mins	5



we are multiplying hour with 60 and adding minute column

From Number From Date & Time

column(#"Changed Type1", "Hours", each if Text.Contains([time],"hr"))

	language	stars	price	time	Hours
2-2021	french	Not rated yet	679	43 mins	0
9-2020	English	Not rated yet	585	3 hrs and 18 mins	3
1-2016	English	4.5 out of 5 stars3 ratings	322	6 hrs and 9 mins	6
8-2021	English	Not rated yet	615	6 hrs and 50 mins	6
0-2021	English	Not rated yet	595	2 hrs and 12 mins	2
9-2021					0
3-2021					0
2-2021					0
2-2021					0
3-2021					0
5-2021					0
8-2020					0
3-2021					0
1-2020	English	Not rated yet	410	2 hrs and 12 mins	2
7-2021	german	Not rated yet	300	59 mins	0
0-2021	English	5 out of 5 stars81 ratings	799	5 hrs and 37 mins	5
0-2019	English	4.5 out of 5 stars20 ratings	754	4 hrs and 2 mins	4
7-2020	English	4.5 out of 5 stars3 ratings	323	1 hr and 13 mins	1
5-2012	English	3.5 out of 5 stars3 ratings	615	5 hrs and 51 mins	5
4-2022	English	Not rated yet	335	2 hrs and 33 mins	2
2-2022	English	Not rated yet	234	7 mins	0
8-2021	English	Not rated yet	703	3 hrs and 24 mins	3
3-2022	English	Not rated yet	422	1 hr and 33 mins	1
8-2021	English	Not rated yet	703	3 hrs and 32 mins	3
1-2020	English	Not rated yet	493	4 hrs and 58 mins	4

Multiply

Enter a number by which to multiply each value in the column.

Value

OK Cancel

QI

View

P N A

ABC 123

Merge Columns ABC Extract 10² Trigonometry

Format abc Parse Statistics Standard Scientific Rounding

From Text From Number From Date & Time

= Table.TransformColumns(#"Removed Columns2", {"Hours", each _ * 60, type number})

releasedate

Custom Column

Add a column that is computed from the other columns.

New column name Minutes

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
```

Available columns

Narrator 2 Narrator 3 releasedate language stars price time Hours

Learn about Power Query formulas

No syntax errors have been detected.

OK Cancel

xt

Search

Windows Start Google Chrome Microsoft Edge Microsoft Word Microsoft Excel



In the table we extracted hours column and minutes column, but we added both the columns to get minutes column in one format.

Screenshot of Microsoft Power Query Editor showing the transformation of time data.

The source data table has columns: price, time, Hours, Minutes, and Addition.

The query settings show:

- PROPERTIES**: Name is set to "audible_uncleaned".
- APPLIED STEPS**: A list of 35 steps including:
 - Merged Columns4
 - Split Column by Character ...
 - Merged Columns5
 - Trimmed Text2
 - Split Column by Delimiter2
 - Changed Type3
 - Removed Columns2
 - Split Column by Delimiter3
 - Changed Type4
 - Removed Columns3
 - Split Column by Delimiter4
 - Changed Type5
 - Removed Columns4
 - Changed Type6
 - Changed Type7
 - Added Custom
 - Reordered Columns
 - Added Custom1
 - Renamed Columns
 - Multiplied Column
 - Changed Type8
 - Inserted Addition
 - Removed Columns5

The preview shows the transformed data with the "Addition" column removed, resulting in a single "time" column.

Row	Original Row	Transformed Row
1	679 43 mins	43 mins
2	585 3 hrs and 18 mins	198
3	322 6 hrs and 9 mins	369
4	615 6 hrs and 50 mins	410
5	585 3 hrs	180
6	1005 8 hrs and 59 mins	539
7	397 1 hr and 34 mins	94
8	200 3 hrs and 5 mins	185
9	251 26 mins	26
10	233 1 hr and 13 mins	73
11	756 1 hr and 2 mins	62
12	1005 8 hrs and 25 mins	505
13	820 13 hrs and 18 mins	798
14	410 2 hrs and 15 mins	135
15	376 7 mins	7
16	410 2 hrs and 12 mins	132
17	300 59 mins	59
18	799 5 hrs and 37 mins	337
19	754 4 hrs and 2 mins	242
20	323 1 hr and 13 mins	73
21	615 5 hrs and 51 mins	351
22	335 2 hrs and 33 mins	153
23	234 7 mins	7
24	703 3 hrs and 24 mins	204
25	422 1 hr and 33 mins	93
26	703 3 hrs and 32 mins	212
27	493 4 hrs and 58 mins	298



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

audible_uncleaned - Power Query Editor

File Home Transform Add Column View

Data Type: Currency Replace Values Unpivot Columns Merge Columns

Transpose Reverse Rows Detect Data Type Fill Move

Group Use First Row By as Headers Count Rows Rename Pivot Column Convert to List

Split Column Format Extract 10² Trigonometry

Statistics Standard Scientific .00 Rounding

Parse Date Time Duration

Any Column Text Column Number Column Date & Time Column Structured Column

Queries [1] audible_uncleaned

= Table.SelectRows(#"Changed Type11", each true)

	releasedate	language	stars	\$ price	time	1.2 Duration in Mins
1	11-02-2022	french	Not rated yet	679.00	43 mins	43
2	29-09-2020	English	Not rated yet	585.00	3 hrs and 18 mins	198
3	15-11-2016	English	4.5 out of 5 stars3 ratings	322.00	6 hrs and 9 mins	369
4	05-08-2021	English	Not rated yet	615.00	6 hrs and 50 mins	410
5	12-10-2021	English	Not rated yet	585.00	3 hrs	180
6	28-09-2021	English	Not rated yet	1,005.00	8 hrs and 59 mins	539
7	22-03-2013	spanish	Not rated yet	397.00	1 hr and 34 mins	94
8	23-12-2021	russian	Not rated yet	200.00	3 hrs and 5 mins	185
9	25-01-2022	English	Not rated yet	251.00	26 mins	26
10	23-02-2022	german	Not rated yet	233.00	1 hr and 13 mins	73
11	15-02-2022	french	Not rated yet	756.00	1 hr and 2 mins	62
12	01-03-2022	English	Not rated yet	1,005.00	8 hrs and 25 mins	505
13	03-05-2018	English	5 out of 5 stars1 rating	820.00	13 hrs and 18 mins	798
14	05-08-2008	English	Not rated yet	410.00	2 hrs and 15 mins	135
15	17-03-2022	french	Not rated yet	376.00	7 mins	7
16	05-11-2020	English	Not rated yet	410.00	2 hrs and 12 mins	132
17	23-07-2021	german	Not rated yet	300.00	59 mins	59
18	12-10-2021	English	5 out of 5 stars81 ratings	799.00	5 hrs and 37 mins	337
19	29-10-2019	English	4.5 out of 5 stars20 ratings	754.00	4 hrs and 2 mins	242
20	09-07-2020	English	4.5 out of 5 stars3 ratings	323.00	1 hr and 13 mins	73
21	01-05-2012	English	3.5 out of 5 stars3 ratings	615.00	5 hrs and 51 mins	351
22	01-04-2022	English	Not rated yet	335.00	2 hrs and 33 mins	153
23	01-02-2022	English	Not rated yet	234.00	7 mins	7
24	31-08-2021	English	Not rated yet	703.00	3 hrs and 24 mins	204
25	01-03-2022	English	Not rated yet	422.00	1 hr and 33 mins	93
26	17-08-2021	English	Not rated yet	703.00	3 hrs and 32 mins	212
27	12-11-2020	English	Not rated yet	493.00	4 hrs and 58 mins	298
28						

13 COLUMNS, 199+ ROWS Column profiling based on entire data set

PREVIEW DOWNLOADED AT 11:59 PM

29°C Partly sunny

Search

ENG IN 12:00 AM 11-09-2024

Query Settings

PROPERTIES Name audible_uncleaned All Properties

APPLIED STEPS

- Changed Type3
- Removed Columns2
- Split Column by Delimiter3
- Changed Type4
- Removed Columns3
- Split Column by Delimiter4
- Changed Type5
- Removed Columns4
- Changed Type6
- Changed Type7
- Added Custom
- Reordered Columns
- Added Custom1
- Renamed Columns
- Multiplied Column
- Changed Type8
- Inserted Addition
- Removed Columns5
- Renamed Columns1
- Changed Type9
- Filtered Rows2
- Changed Type11

Filtered Rows3



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

Queries [1]

audible_uncleaned

= Table.SelectRows(#"Changed Type11",

		releasedate	language
1		11-02-2022	french
2		29-09-2020	English
3		15-11-2016	English
4		05-08-2021	English
5		12-10-2021	English
6		28-09-2021	English
7		22-03-2013	spanish
8	„Ñ€Ð³ÐµÐ…	23-12-2021	russian
9		25-01-2022	English
10		23-02-2022	german
11		15-02-2022	french
12		01-03-2022	English
13		03-05-2018	English
14		05-08-2008	English

Thank you