

CLEANING DATASET USING POWER QUERY EDITOR

AUDIBLE



#1

Uncleaned Dataset



Below is the uncleaned dataset

audible_uncleaned - Power Query Editor

File Home Transform Add Column View

Close & Load Refresh Advanced Editor Properties Choose Columns Remove Columns Keep Rows Remove Rows Split Column Group By Data Type: Text Use First Row as Headers Replace Values Merge Queries Append Queries Combine Files Manage Parameters Data source settings New Source Recent Sources

Queries [1] audible_uncleaned

Table.TransformColumnTypes(#"Promoted Headers",{{"name", type text}, {"author", type text}, {"narrator", type text}, {"time", type text}, {"releasedate", type text}, {"language", type text}, {"stars", type text}, {"price", type text}})

	author	narrator	time	releasedate	language	stars	price
1	Writtenby:GeronimoStilton	Narratedby:BillLobley	2 hrs and 20 mins	04-08-2008	English	5 out of 5 stars34 ratings	468
2	Writtenby:RickRiordan	Narratedby:RobbieDaymond	13 hrs and 8 mins	01-05-2018	English	4.5 out of 5 stars41 ratings	820
3	Writtenby:JeffKinney	Narratedby:DanRussell	2 hrs and 3 mins	06-11-2020	English	4.5 out of 5 stars38 ratings	410
4	Writtenby:RickRiordan	Narratedby:SoneelaNankani	11 hrs and 16 mins	05-10-2021	English	4.5 out of 5 stars12 ratings	615
5	Writtenby:RickRiordan	Narratedby:JesseBernstein	10 hrs	13-01-2010	English	4.5 out of 5 stars181 ratings	820
6	Writtenby:SuzanneCollins	Narratedby:TatianaMaslany	10 hrs and 35 mins	30-10-2018	English	5 out of 5 stars72 ratings	656
7	Writtenby:WinterMorgan	Narratedby:LukeDaniels	2 hrs and 23 mins	25-11-2014	English	5 out of 5 stars11 ratings	233
8	Writtenby:RickRiordan	Narratedby:RobbieDaymond	12 hrs and 32 mins	02-05-2017	English	5 out of 5 stars50 ratings	820
9	Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne	10 hrs and 56 mins	02-05-2017	English	5 out of 5 stars5 ratings	1256
10	Writtenby:RickRiordan	Narratedby:RobbieDaymond	13 hrs and 22 mins	24-09-2019	English	5 out of 5 stars58 ratings	820
11	Writtenby:RickRiordan	Narratedby:JesseBernstein	8 hrs and 48 mins	14-01-2010	English	4.5 out of 5 stars130 ratings	820
12	Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne	5 hrs and 23 mins	24-08-2011	English	5 out of 5 stars6 ratings	1206
13	Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne	6 hrs and 1 min	27-09-2011	English	5 out of 5 stars7 ratings	1206
14	Writtenby:RickRiordan	Narratedby:MichaelCrouch	12 hrs and 58 mins	03-10-2017	English	5 out of 5 stars41 ratings	820
15	Writtenby:PhilipPullman	Narratedby:PhilipPullman,fullcast,RuthWilson	11 hrs and 55 mins	24-06-2021	English	4 out of 5 stars2 ratings	1093
16	Writtenby:GeronimoStilton	Narratedby:BillLobley	2 hrs and 25 mins	08-02-2008	English	4.5 out of 5 stars33 ratings	467
17	Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne	5 hrs and 4 mins	26-12-2004	English	5 out of 5 stars7 ratings	1206
18	Writtenby:ShannonMessenger	Narratedby:CaitlinKelly	14 hrs and 41 mins	06-11-2018	English	5 out of 5 stars20 ratings	836
19	Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne	10 hrs and 18 mins	02-05-2017	English	5 out of 5 stars11 ratings	1256
20	Writtenby:ShannonMessenger	Narratedby:CaitlinKelly	16 hrs and 42 mins	06-11-2018	English	5 out of 5 stars13 ratings	1003
21	Writtenby:RickRiordan	Narratedby:RobbieDaymond	12 hrs and 12 mins	06-10-2020	English	5 out of 5 stars79 ratings	615
22	Writtenby:ChristopherPaolini	Narratedby:GerrardDoyle	23 hrs and 29 mins	14-04-2009	English	4.5 out of 5 stars47 ratings	957
23	Writtenby:EoinColfer	Narratedby:GerryO'Brien	6 hrs and 46 mins	11-04-2013	English	5 out of 5 stars27 ratings	683
24	Writtenby:GeronimoStilton	Narratedby:BillLobley	2 hrs and 37 mins	01-02-2011	English	4.5 out of 5 stars25 ratings	469
25	Writtenby:RickRiordan	Narratedby:JesseBernstein	10 hrs and 32 mins	21-01-2010	English	5 out of 5 stars101 ratings	820
26	Writtenby:A.A.Milne	Narratedby:PeterDennis	2 hrs and 46 mins	28-12-2003	English	5 out of 5 stars15 ratings	468
27	Writtenby:MaryPopeOsborne	Narratedby:MaryPopeOsborne	7 hrs and 9 mins	01-09-2020	English	5 out of 5 stars3 ratings	1256
28	Writtenby:GeronimoStilton	Narratedby:BillLobley	2 hrs and 32 mins	01-10-2010	English	5 out of 5 stars18 ratings	469

8 COLUMNS, 999+ ROWS

PREVIEW DOWNLOADED AT 23:15

23:16 04-08-2024

24°C Mostly cloudy

ENG

Query Properties

Name: audible_uncleaned

APPLIED STEPS

Source Promoted... Change...

#2

Applied Steps



The steps applied to clean and standardize the dataset are below

APPLIED STEPS

Source

Promoted Headers

Capitalized Each Word

Extracted Text After Delimiter

Split Column by Delimiter

Removed Columns

Replaced Value

Split Column by Character Transition

Merged Columns(Author 1)

Trimmed Text

Split Column by Character Transition1

Merged Columns(Author 2)

Trimmed Text1

Split Column by Character Transition2

Merged Columns(Author 3)

Trimmed Text2

Extracted Text After Delimiter1

Split Column by Delimiter1

Changed Type1

Removed Columns1

Replaced Value1

Split Column by Character Transition3

Merged Columns

Trimmed Text3

Split Column by Character Transition4

Merged Columns1

Trimmed Text4

Split Column by Character Transition5

Merged Columns2

Trimmed Text5

Extracting Hours

Multiplied Column

Renamed Columns

Extracting Minutes

Inserted Addition

Removed Columns2

Renamed Columns1

Replaced Value2

Changed Type(price, releasedate)

Merged Columns3

Rounded Off

Extracting No Of Ratings

Extracted Text Before Delimiter

Replaced Value3

Changed Type

Removed Columns3

Reordered Columns

#3

Custom Formulas



The formula below extracts the “Hours” value from the time column

Custom Column

Add a column that is computed from the other columns.

New column name

Custom column formula

```
= if Text.Contains([time], "hr")  
then Number.FromText(Text.BeforeDelimiter([time], "hr"))  
  
else 0
```

[Learn about Power Query formulas](#)

✓ No syntax errors have been detected.

OK Cancel

Available columns

name
Author 1
Author 2
Author 3
Narrator 1
Narrator 2
Narrator 3
...

<< Insert

#4

Custom Formulas



The formula below extracts the “Minutes” value from the time column

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

name
Author 1
Author 2
Author 3
Narrator 1
Narrator 2
Narrator 3
..

<< Insert

[Learn about Power Query formulas](#)

✓ No syntax errors have been detected.

OK Cancel

#5

Custom Formulas



The formula below extracts the “Number of Ratings” value from the stars column

Custom Column

Add a column that is computed from the other columns.

New column name

Number Of Ratings

Custom column formula

```
= if Text.Contains([stars], "rating")
then Number.FromText(Text.BeforeDelimiter(Text.AfterDelimiter
([stars], "stars"), " rating"))
else 0
```

[Learn about Power Query formulas](#)

✓ No syntax errors have been detected.

OKCancel

Available columns

name

Author 1

Author 2

Author 3

Narrator 1

Narrator 2

Narrator 3

...

<< Insert

#6

Cleaned Dataset



Below is the final cleaned dataset

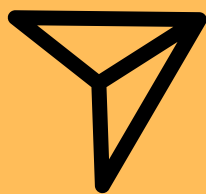
[Drive Link](#)

	A	B	C	D	E	F	G	H	I	J	K	L
	name	Author 1	Author 2	Author 3	Narrator 1	Narrator 2	Narrator 3	releaseinfo	Duration in mins	price	stars	Number Of Ratings
1	Atomic Habits	James Clear			James Clear			18-10-2018, English	335	820	5	12573
2	Atomic Habits	James Clear			James Clear			18-10-2018, English	335	820	5	12569
3	Ikigai	Héctor García	Francesc Miralles		Naoko Mori			18-09-2017, English	203	615	4.5	6981
4	The Psychology Of Money	Morgan Housel			Chris Hill			08-09-2020, English	348	668	4.5	5973
5	The Psychology Of Money	Morgan Housel			Chris Hill			08-09-2020, English	348	668	4.5	5972
6	Sapiens	Yuval Noah Harari			Derek Perkins			30-04-2015, English	918	957	5	5500
7	Rich Dad Poor Dad	Robert T. Kiyosaki			Tim Wheeler			15-06-2012, English	369	837	4.5	5324
8	Life's Amazing Secrets	Gaur Gopal Das			Gaur Gopal Das			10-12-2019, English	385	1005	4.5	4184
9	The Subtle Art Of Not Giving A F	Mark Manson			Roger Wayne			13-09-2016, English	317	1181	4.5	3461
10	Becoming	Michelle Obama			Michelle Obama			13-11-2018, English	1143	957	4.5	2924
11	How To Win Friends And Influence	Dale Carnegie			Shernaz Patel			08-11-2018, English	485	155	4.5	2494
12	Can't Hurt Me	David Goggins			David Goggins	Adam Skolnick		28-11-2018, English	817	836	5	2277
13	21 Lessons For The 21st Century	Yuval Noah Harari			Derek Perkins			30-08-2018, English	701	957	4.5	2236
14	Elon Musk	Ashlee Vance			Fred Sanders			28-04-2016, English	803	820	4.5	2161
15	The Power Of Habit	Charles Duhigg			Mike Chamberlain			05-04-2012, English	653	820	4.5	1869
16	The Monk Who Sold His Ferrari	Robin Sharma			Robin Sharma			05-07-2007, English	160	351	4.5	1838
17	Wings Of Fire	APJ Abdul Kalam	Arun Tiwari		Girish Karnad			01-04-2020, English	75	75	5	1827
18	Wings Of Fire	APJ Abdul Kalam	Arun Tiwari		Girish Karnad			01-04-2020, English	75	75	5	1823
19	Zero To One	Peter Thiel	Blake Masters		Blake Masters			04-06-2015, English	290	615	4.5	1807
20	A Promised Land	Barack Obama			Barack Obama			17-11-2020, English	1750	1708	4.5	1636
21	The Compound Effect	Darren Hardy			Darren Hardy			22-01-2019, English	284	608	4.5	1564
22	The Compound Effect	Darren Hardy			Darren Hardy			22-01-2019, English	284	608	4.5	1563
23	Think Like A Monk	Jay Shetty			Jay Shetty			08-09-2020, English	654	569	4.5	1539
24	Raavan	Amish Tripathi			Sagar Arya			04-10-2019, English	692	342	4.5	1491
25	Homo Deus	Yuval Noah Harari			Derek Perkins			08-09-2016, English	893	957	4.5	1490
26	Harry Potter And The Philosopher's	J.K. Rowling			Stephen Fry			20-11-2015, English	573	999	5	1425

LIKED THE
PROJECT?



LIKE



SHARE



COMMENT



SAVE

www.linkedin.com/in/ethendcosta