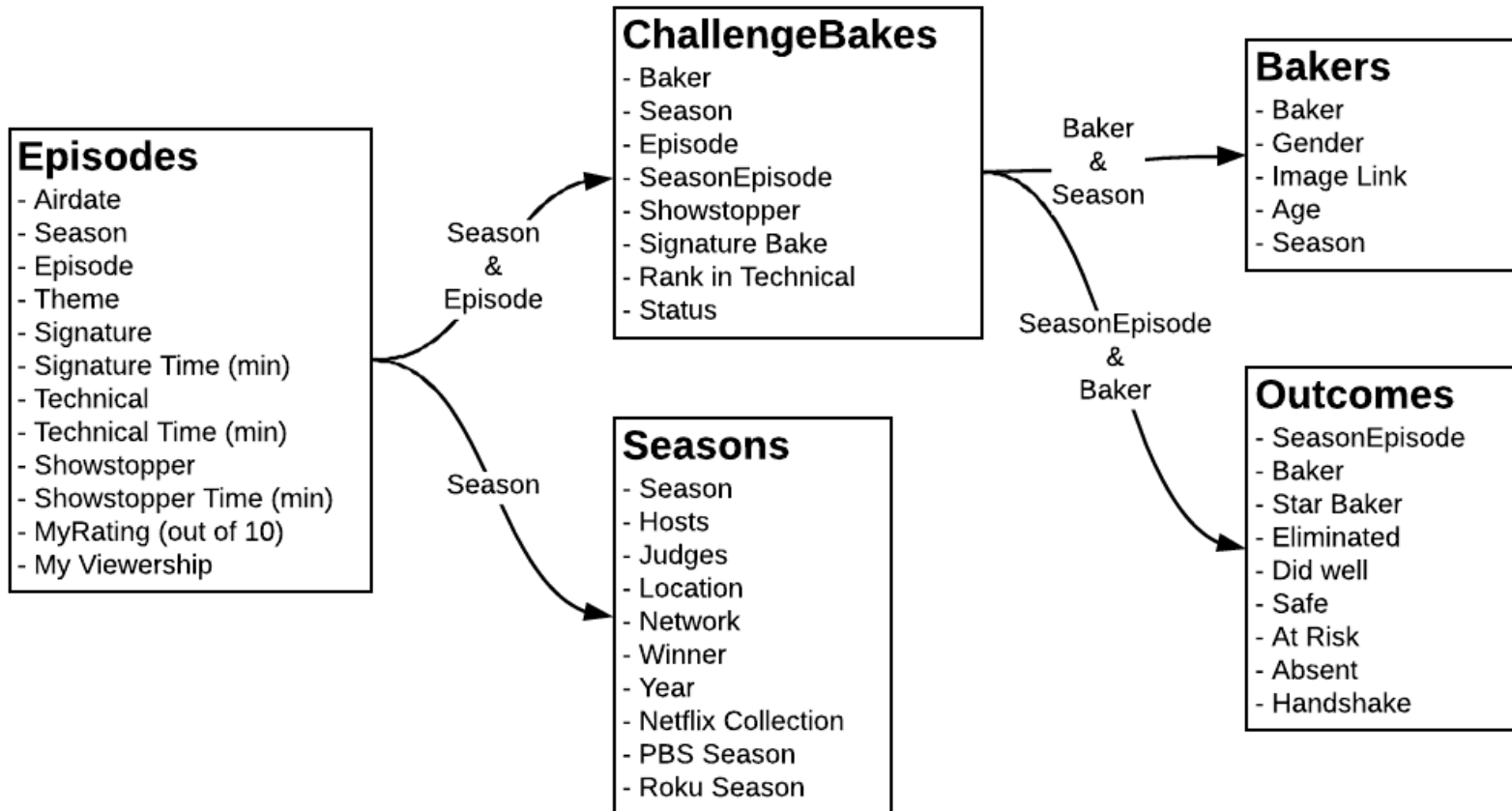


# Data Dictionary for the Great British Bake Off data set

This data dictionary contains information about the five csv files for the GBO data set. The data is for educational purposes only. It does not indicate Love Productions, Ltd., relating to The Great British Bake Off, endorses Tableau or its services.

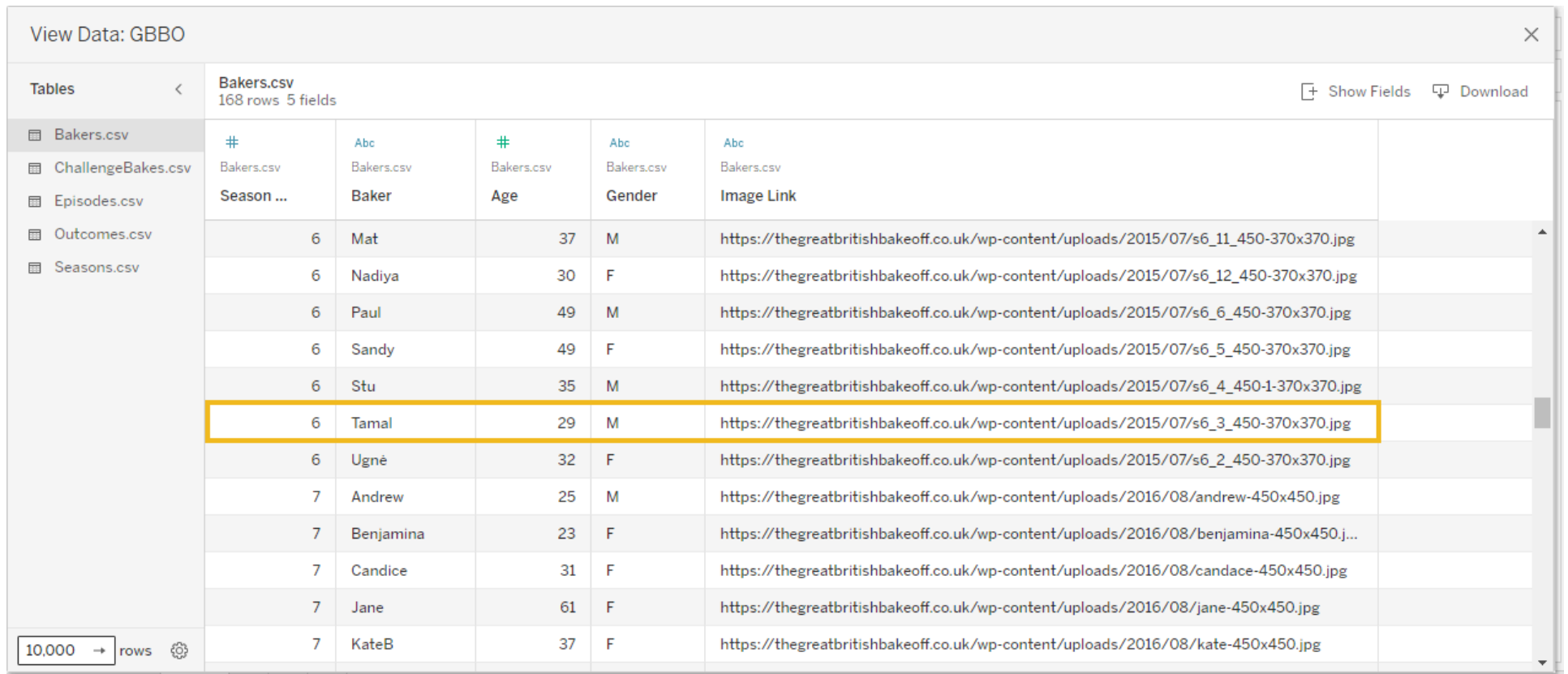


*A schematic showing the relationships between the tables in the data set. Episodes is to the far left, with two branches. One branch is a relationship to Seasons on the Season field. The other branch from Episodes is to ChallengeBakes on the Season and Episode fields. ChallengeBakes also has two branches. One is a relationship to the Bakers table on both the Baker and Season fields. The other is to Outcomes on the Baker field and SeasonEpisode field.*

# Bakers

- **Season:** which season they were on
- **Baker:** the name used to refer to the contestant during the show.
- **Age:** age when they were a contestant
- **Gender:** M or F
- **Image Link:** a URL to a hosted image of the contestant from thegreatbritishbakeoff.co.uk

The Bakers table has 5 fields and 168 rows. **What is a row?** A single contestant.



View Data: GBBO

Tables < Bakers.csv 168 rows 5 fields Show Fields Download

#	Baker	Age	Gender	Image Link
6	Mat	37	M	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_11_450-370x370.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_11_450-370x370.jpg</a>
6	Nadiya	30	F	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_12_450-370x370.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_12_450-370x370.jpg</a>
6	Paul	49	M	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_6_450-370x370.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_6_450-370x370.jpg</a>
6	Sandy	49	F	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_5_450-370x370.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_5_450-370x370.jpg</a>
6	Stu	35	M	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_4_450-1-370x370.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_4_450-1-370x370.jpg</a>
6	Tamal	29	M	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_3_450-370x370.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_3_450-370x370.jpg</a>
6	Ugné	32	F	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_2_450-370x370.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2015/07/s6_2_450-370x370.jpg</a>
7	Andrew	25	M	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/andrew-450x450.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/andrew-450x450.jpg</a>
7	Benamina	23	F	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/benamina-450x450.j...">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/benamina-450x450.j...</a>
7	Candice	31	F	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/candace-450x450.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/candace-450x450.jpg</a>
7	Jane	61	F	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/jane-450x450.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/jane-450x450.jpg</a>
7	KateB	37	F	<a href="https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/kate-450x450.jpg">https://thegreatbritishbakeoff.co.uk/wp-content/uploads/2016/08/kate-450x450.jpg</a>

10,000 → rows ⚙️

The view data screen from Tableau Desktop, showing the Bakers.csv table with a representative row highlight to show the level of detail of the table is a single contestant.

# ChallengeBakes

- **Baker:** The name used to refer to the contestant during the show
- **Episode:** The episode number
- **Season:** The season number for that episode
- **Season Episode:** Season and episode information formatted as *s#e#*, such as *s4e8* for Season 4, Episode 8
- **Showstopper:** What the contestant made for the Showstopper challenge in that episode
- **Signature:** What the contestant made for the Signature challenge in that episode
- **Rank in technical:** Their rank in the Technical challenge for that episode, with 1 as the winner.
- **Status:** The final outcome for the contestant for that episode. Values are Eliminated, At risk, Safe, Did well, Star Baker, Absent, Runner-up, and Winner. More explanation of these values can be found in the Outcomes table section.

The ChallengeBakes table has 8 fields and 1003 rows. **What is a row?** A record for each combination of season, episode, and contestant.

View Data: GBB0

ChallengeBakes.csv  
1,003 rows 8 fields

Show Fields Download

#	#	Abc	Abc	Abc	#	Abc	Abc
Season ...	Episode ...	Baker	Se...	Signature (ChallengeB...	Rank in Te...	Showstopper (ChallengeBak...	Status
6	8	Tamal	s6e8	Lime & Mascarpone and Ma...	4	Passionfruit & Mango and Pistachi...	Did well
6	8	Flora	s6e8	Peach & Lemon Thyme and ...	3	Lime & Basil and Coconut & White...	At risk
6	8	Paul	s6e8	Café Tipple & Banana Crun...	5	Religieuse à l'ancienne	Eliminated
6	9	Nadiya	s6e9	Peanut Salted Caramel and ...	4	Peacock in Nan's Door	Star Baker
6	9	IanC	s6e9	Chocolate and Bay Tart	3	Chocolate Well	At risk
6	9	Tamal	s6e9	Chocolate New York Pie	2	Chocolate Bell Tower	Did well
6	9	Flora	s6e9	Passion Fruit and Chocolate...	1	Cocoa Carousel	Eliminated
6	10	Nadiya	s6e10	Cardamom and Almond Bu...	1	My Big Fat British Wedding Cake	Winner
6	10	IanC	s6e10	Elderflower and Lemon & S...	2	Colossal Curvy Carrot Cake	Runner-up
6	10	Tamal	s6e10	Cinnamon and Apple & Toff...	3	Sticky Toffee Pudding Cake	Runner-up
7	1	Candice	s7e1	Raspberry and Rhubarb Dri...	5	Mirror Mirror On The Wall, Who Is ...	At risk
7	1	Andrew	s7e1	Lemon and Rosemary Drizz...	12	Ultimate Indulgence' Mirror Glaze ...	Safe

10,000 rows

Season 6 episode 9 has four rows for the semifinal. Season 6 episode 10 has three rows because one contestant was eliminated and only three made it to the final.

# Episodes

- **Airdate:** The date the episode originally aired on British TV
- **Season:** The season number for that episode
- **Episode:** The episode number
- **Theme:** The week's theme
- **Signature:** What the judges asked for in the Signature challenge for that episode
- **Signature Time (min):** How many minutes the contestants had for the Signature challenge in that episode, in minutes
- **Technical:** What the judges asked for in the Technical challenge for that episode
- **Technical Time (min):** How many minutes the contestants had for the Technical challenge in that episode, in minutes
- **Showstopper:** What the judges asked for in the Showstopper challenge for that episode
- **Showstopper Time (min):** How many minutes the contestants had for the Showstopper challenge in that episode, in minutes
- **My Rating (out of 10):** *Artificial data.* A rating for that episode. Due to licensing restrictions, ratings is artificial (not real-world) data.
- **My Viewership:** *Artificial data.* A popularity metric for the episode. Due to licensing restrictions, viewership is artificial (not real-world) data.

The Episodes table has 12 fields and 134 rows. **What is a row?** An episode (technically, a season-episode).

View Data: GBBO

Tables < Episodes.csv 134 rows 12 fields Show Fields Download

	Airdate	S...	E...	Theme	Signa...	S...	Techn...	T...	Showst...	S...	MyRating ...	My Viewership
Bakers.csv	10/1/2014	5	9	Pâtisserie	24 Baklava	210	Schichttorte	120	24 Entremets	300	7.90000	37.3300
ChallengeBakes.csv	10/8/2014	5	10	Final	Viennoiserie	210	12 Mini Vic...	120	Pièce Montée	300	8.30000	46.6800
Episodes.csv	8/5/2015	6	1	Cake	Madeira C...	120	Frosted W...	105	Black Forest ...	210	7.80000	43.1400
Outcomes.csv	8/12/2015	6	2	Biscuit	24 Biscotti	120	8 Arlettes	150	36 Biscuits ...	240	8.00000	43.0300
Seasons.csv	8/19/2015	6	3	Bread	2 Quick Br...	90	4 Baguettes	150	3D Bread Sc...	300	7.80000	42.8100
	8/26/2015	6	4	Dessert	12 Crème ...	120	Spanische ...	240	Three-tiered...	270	8.20000	44.1300
	9/2/2015	6	5	Alternativ...	Sugar-free...	150	12 Gluten-...	120	Dairy-free Ic...	270	8.00000	44.2300
	9/9/2015	6	6	Pastry	Frangipan...	120	12 Flaounes	120	48 Vol-au-Ve...	225	7.80000	44.1100
	9/16/2015	6	7	Victorian	Raised Ga...	180	Tennis Cake	180	Charlotte Ru...	330	7.70000	43.2100
	9/23/2015	6	8	Pâtisserie	24 Cream ...	210	9 Mokatines	120	Religieuse à ...	240	8.00000	42.1200
	9/30/2015	6	9	Chocolate	Chocolate ...	150	Chocolate ...	75	Chocolate C...	240	8.30000	46.9700
10,000 → rows	10/7/2015	6	10	Final	16 Iced Bu...	180	Mille-Feuille	120	Classic Briti...	240	8.60000	52.5800

# Outcomes

- **SeasonEpisode:** Season and episode information formatted as *s#e#*, such as *s4e8* for Season 4, Episode 8
- **Baker:** The name used to refer to the contestant during the show
- **Star Baker:** A binary field. 1 indicates the contestant was named Star Baker for that episode.
- **Eliminated:** A binary field. 1 indicates the contestant was Eliminated that episode.
- **Did well:** A binary field. 1 indicates the contestant was called out by the judges as performing well that episode.
- **Safe:** A binary field. 1 indicates the contestant wasn't called out as particularly high or low performing for that episode.
- **At risk:** A binary field. 1 indicates the contestant was called out by the judges as in danger of being eliminated that episode.
- **Absent:** A binary field. 1 indicates the contestant was missing for some or all of that episode but returned later.
- **Handshake:** A binary field. 1 indicates the contestant was given a handshake by judge Paul Hollywood.

The Outcomes table has 9 fields and 964 rows. **What is a row?** A record for each combination of contestant, season, and episode—excluding the final episode of each season.

View Data: GBBO

Tables < Outcomes.csv  
964 rows 9 fields Show Fields Download

SeasonEpisode	Baker (O...)	Star Baker	Eliminated	Did well	Safe	At risk	Absent	Handshake
s6e7	Tamal	1	0	0	0	0	0	1
s6e8	Flora	0	0	0	0	1	0	0
s6e8	IanC	0	0	0	1	0	0	0
s6e8	Nadiya	1	0	0	0	0	0	0
s6e8	Paul	0	1	0	0	0	0	0
s6e8	Tamal	0	0	1	0	0	0	0
s6e9	Flora	0	1	0	0	0	0	0
s6e9	IanC	0	0	0	0	1	0	0
s6e9	Nadiya	1	0	0	0	0	0	1
s6e9	Tamal	0	0	1	0	0	0	0
s7e1	Andrew	0	0	0	1	0	0	0
s7e1	Benamina	0	0	1	0	0	0	0

10,000 rows

# Seasons

- **Hosts:** Who hosted that season (two per season)
- **Judges:** Who judged that season (two per season)
- **Location:** Where the season was filmed. This field is formatted as [name of the site], [the nearest mappable town], [UK].
- **Network:** What British network originally aired the season
- **Winner:** The winning baker
- **Year:** The four-digit year that the season originally aired on British TV
- **Netflix Collection:** What Netflix labels that season
- **PBS Season:** What PBS labels that season
- **Roku Season:** What Roku labels that season
- **Season:** The season number. Also synonymous with "Series".

The Seasons table has 10 fields and 56 rows. **What is a row?** The row definition isn't clear. Because each season has two hosts and two judges, each season gets four rows (one per Host-Judge); therefore, a single row doesn't contain all the information for any single season..

View Data: GBBO

Seasons.csv  
56 rows 10 fields

Show Fields Download

#	Abc	Abc	Abc	Abc	#	Abc	#	#	#
Season ...	Hosts	Judges	Network	Winner	Year	Location	Netflix ...	PBS ...	Roku ...
6	Mel Giedroyc	Mary Berry	BBC One	Nadiya	2015	Welford Park, Newb...	Null	3	6
6	Mel Giedroyc	Paul Hollywood	BBC One	Nadiya	2015	Welford Park, Newb...	Null	3	6
6	Sue Perkins	Mary Berry	BBC One	Nadiya	2015	Welford Park, Newb...	Null	3	6
6	Sue Perkins	Paul Hollywood	BBC One	Nadiya	2015	Welford Park, Newb...	Null	3	6
7	Mel Giedroyc	Mary Berry	BBC One	Candice	2016	Welford Park, Newb...	Null	4	7
7	Mel Giedroyc	Paul Hollywood	BBC One	Candice	2016	Welford Park, Newb...	Null	4	7
7	Sue Perkins	Mary Berry	BBC One	Candice	2016	Welford Park, Newb...	Null	4	7
7	Sue Perkins	Paul Hollywood	BBC One	Candice	2016	Welford Park, Newb...	Null	4	7
8	Noel Fielding	Prue Leith	Channel 4	Sophie	2017	Welford Park, Newb...	5	Null	Null
8	Noel Fielding	Paul Hollywood	Channel 4	Sophie	2017	Welford Park, Newb...	5	Null	Null
8	Sandi Toks...	Prue Leith	Channel 4	Sophie	2017	Welford Park, Newb...	5	Null	Null
8	Sandi Toks...	Paul Hollywood	Channel 4	Sophie	2017	Welford Park, Newb...	5	Null	Null

10,000 → rows

# In the Data pane

For a tidy Data pane, after creating your data model with the appropriate relationships as outlined in the diagram above, we recommend the following clean-up steps:

## Bakers table

- **Image Link:** Image Role > URL
- **Season:** Hide

## Challenges table

- **Baker:** Hide
- **Episode:** Hide
- **Season:** Hide
- **Showstopper** → **Individual Showstopper**
- **Signature** → **Individual Signature**

## Episodes table

- **Episode:** Convert to dimension
- **Season:** Hide
- **MyRating (out of 10)** → **My Rating**

## Outcomes table

- **Baker:** Hide
- **SeasonEpisode:** Hide

## Seasons table

- **Netflix Collection:** Convert to dimension
- **PBS Season:** Convert to dimension
- **Roku Season:** Convert to dimension
- **Season:** Convert to dimension

