


Power BI 101

Star Schema
Practice



Challenge

- We Have Two Fact Tables Related To Movies, And We Need To Create A DIMENSION Table For Movie Titles And Release Years.
 - This Dimension Table Will Be Connected To Our Fact Tables, Allowing Us To Build A Star Schema Model.
- 
- The bottom right corner of the slide features a decorative graphic consisting of several overlapping, semi-transparent geometric shapes in shades of teal and blue, arranged in a cluster.

Rank	Title	Year	Imdb Rating
273	Hello I Must Be Going	1972	9
219	Invitation to a Gunfighter	1957	6
4	George Washington Slept Here	2000	6
53	Torrente 4: Lethal Crisis	1966	6
113	Ator, the Fighting Eagle (Ator l'invincibile)	1993	5
71	Better Luck Tomorrow	2022	5
32	Presence, The	1984	8
237	Zero Effect	2014	9
251	Fearless	2012	6
54	Vermont Is For Lovers	1953	9
265	Chorus, The (Choristes, Les)	1987	9
107	Hunt For Gollum, The	1971	5
2	Recollections of the Yellow House (RecordaÃ§Ãµes da Casa Amarela	2012	7
298	Don't Bother to Knock	1997	6
220	Story of Mankind, The	1951	8
209	Start the Revolution Without Me	1959	7
178	Gun Woman	1960	6
1	State Property	1984	5
107	Barton Fink	1973	7
66	Atlas Shrugged: Who Is John Galt? (Atlas Shrugged: Part III)	1992	5
144	Third World Cop	1955	7
53	Keeping Up with the Steins	1996	8
256	Possessed	1955	6
285	Five-Year Engagement, The	1980	5
20	Chronicles of Narnia: The Voyage of the Dawn Treader, The	1971	9
221	La Putain du roi	2006	6
271	Children of the Corn IV: The Gathering	2017	9
77	Passing Fancy (Dekigokoro)	1981	6
22	Alphabet City	1984	8
40	Mozart's Sister (Nannerl, la soeur de Mozart)	1965	7
94	Adam's Rib	2020	5
141	Eve of the Storm, The	2006	7


Rank	Title	Year	Lifetime Gross
273	Hello I Must Be Going	1972	14662565
219	Invitation to a Gunfighter	1957	10795451
4	George Washington Slept Here	2000	7354888
53	Torrente 4: Lethal Crisis	1966	13137894
113	Ator, the Fighting Eagle (Ator l'invincible)	1993	6493159
71	Better Luck Tomorrow	2022	4674122
32	Presence, The	1984	9118128
237	Zero Effect	2014	7937432
251	Fearless	2012	911360
54	Vermont Is For Lovers	1953	13369625
265	Chorus, The (Choristes, Les)	1987	3085672
107	Hunt For Gollum, The	1971	11897298
2	Recollections of the Yellow House (RecordaÃ§Ãµes da Casa Amarela	2012	2561670
298	Don't Bother to Knock	1997	6887742
220	Story of Mankind, The	1951	4020856
209	Start the Revolution Without Me	1959	14595041
178	Gun Woman	1960	4690875
1	State Property	1984	2487039
107	Barton Fink	1973	7234817
66	Atlas Shrugged: Who Is John Galt? (Atlas Shrugged: Part III)	1992	11005772
144	Third World Cop	1955	1298407
53	Keeping Up with the Steins	1996	8729769
256	Possessed	1955	12420849
285	Five-Year Engagement, The	1980	11855008
20	Chronicles of Narnia: The Voyage of the Dawn Treader, The	1971	3940942
221	La Putain du roi	2006	11898886
271	Children of the Corn IV: The Gathering	2017	3843401
77	Passing Fancy (Dekigokoro)	1981	11897186
22	Alphabet City	1984	10017368
40	Mozart's Sister (Nannerl, la soeur de Mozart)	1965	12287440
94	Adam's Rib	2020	4345994
141	Eve of the Storm, The	2006	791232

❑ HOW TO

Set Up A Star Schema Using Two Fact Tables—movie Sales And Movie Ranking—along With A Shared Movie Dimension Table.

❑ BEST PRACTICES

For Maintaining A Clean And Efficient Data Model.



Movie Dimension Table

- The Movie Dimension Table Will Be Your Master List Of Movies, Containing Essential Details Like The Movie Title And The Year It Was Produced.

Create Movie Dimension Table

- Include Only The Two Important Columns: Movie Title And Production Year From Facts Tables

Remove Any Other Unnecessary Columns To Keep The Table Simple.

The bottom right corner of the slide features a decorative graphic consisting of several overlapping, semi-transparent geometric shapes in various shades of teal and blue, arranged in a dynamic, abstract pattern.

Remove Duplicate Entries

- Ensure Each Movie-year Combination Is **Unique** To Avoid Confusion And Errors.
 - 1. Check For Duplicates:**
Look For Rows Where The Movie Title And Production Year Are The Same.
 - 2. Eliminate Duplicates:**
Remove Any Duplicate Entries So That Each Movie For Each Year Appears Only Once In The Movie Dimension Table.

Add a Surrogate Key

- A Surrogate Key **Uniquely** Identifies Each Movie In The Dimension Table, Making It Easier To Link To Fact Tables.
- **Create A Unique Identifier:**
Add A New Column Called Moviekey.
Use An **Index Column** To Assign A Unique Number To Each Movie Entry.
This Moviekey Will Act As The Surrogate Key.

Link Movie Dimension with FactMovieSales

Connect Your Movie Sales Data To The Movie Dimension Table Using The Movie Key.

1. Merge Tables
 2. Add Movie Key
 3. Clean Up
- 
- A series of overlapping teal and light blue geometric shapes, primarily squares and rectangles, arranged in a cluster in the bottom right corner of the slide.

Check for Missing MovieKeys

- Ensure That Every Sale Record Has A Corresponding Moviekey To Maintain Data Integrity.


1. Identify Missing Keys:

Look For Any **Blank Or Null Values** In The Moviekey Column In Factmoviesales.

2. Investigate Issues:

If You Find Missing Keys, Check Why Certain Movies Didn't Match Between The Fact Table And The Dimension Table. This Might Involve Correcting Data Inconsistencies.

Repeat for FactMovieRanking

- Load Tables into Power BI
 - Establish Relationships in Power BI
 - Hide MovieKey
 - Build Your Visualizations
 - Create Reports
 - Expand Your Model
- 
- The bottom right corner of the slide features a decorative graphic consisting of several overlapping, semi-transparent geometric shapes in shades of teal and cyan, arranged in a dynamic, abstract pattern.

Expand Your Model

- Keep Your Data Model Flexible And Scalable For Future Additions.

Add More Tables:

You Can Introduce Additional Dimension Or Fact Tables.

Connect Them To The Existing Movie Dimension Or Other Shared Dimensions To **Maintain The Star Schema Structure.**


Best Practices



Keep It Simple

- Stick To The **Star Schema** Layout Where Multiple Fact Tables Connect To Centralized Dimension Tables. This Makes Your Model Easier To Understand And Maintain.

Use Surrogate Keys

- Always Include A Unique Surrogate Key (Like Moviekey) In Your Dimension Tables. This Helps Prevent Issues With Duplicate Or Changing Natural Keys.
- 
- The bottom right corner of the slide features a decorative graphic consisting of several overlapping, semi-transparent geometric shapes in shades of teal and cyan, arranged in a dynamic, abstract pattern.

Maintain Consistency

- Ensure That The Fields Used For Merging (E.G., Movie Title, Production Year) Are Consistent And Formatted The Same Way In Both Fact And Dimension Tables.

Remove Duplicates

- Always Eliminate Duplicate Records In Your Dimension Tables To Ensure Each Combination Of Attributes Is **Unique**.

Optimize Storage

- Use **Numeric Surrogate** Keys Instead Of Concatenated Text Fields To Save Space And Improve Performance, Especially With Large Datasets.

Clean Your Data

- Before Loading Data Into Power BI, Ensure It's **Free Of Errors And Null** Values, Particularly In Key Columns Used For Merging.

Hide Technical Fields

- Hide Columns Like Moviekey In Your Power BI Reports To Keep The User Interface Clean And Focused On Relevant Data.

Maintain a Master List

- Treat Your Movie Dimension Table As The Authoritative Source For Movie Information, Ensuring It Includes All Unique Movie-year Combinations.

Validate Relationships

- After Setting Up Relationships, **Test** Them With Simple Visualizations To Confirm They Work Correctly Before Building More Complex Reports.
- 
- The bottom right corner of the slide features a decorative graphic consisting of several overlapping, semi-transparent geometric shapes, primarily squares and rectangles, in various shades of teal and light blue.

Plan for Scalability

- Design Your Star Schema To Easily Accommodate Additional Dimensions Or Fact Tables As Your Data Needs Grow.



Thank You

Walid Hussein

Several overlapping squares and rectangles in various shades of teal and cyan are positioned in the bottom right corner of the slide, creating a modern, abstract design element.