

Description of the database

The database's purpose is to view statistics about the [Star Wars Prequels](#) movies including characters, planets, dialogue, lightsaber duels, masters and apprentices.

The assumptions are the following:

1. There are only 3 movies, each has own ID.
2. All characters have their ID, they must all be in at least one movie, at maximum all three of them.
3. Most characters have known home planets, each planet has own ID.
4. Some characters have masters (mentors, teachers).
5. Characters sometimes fight each other in lightsaber duels.

Tables

1. Movies.

Data about movies.

Column	Description	Data type	Empty value	Default	Links
PartID	Identifier of movie part	TINYINT	NO		PK
Name	Name of movie without franchise name	VARCHAR(50)	NO		
ReleaseDate	Date of release	DATE	NO		
Budget	Amount of money spent on making the movie in \$	BIGINT	NO		
BoxOffice	Amount of money earned by the movie in \$	BIGINT	NO		
WordsSpoken	Amount of words spoken in the movie (even by characters not included in Characters table)	INT	NO		
WordsSpoken ByCharacters	Amount of words spoken in the movie only by characters included in Characters table. Generated from the Dialogue table.	INT	YES		

2. Planets.

Data about planets that the characters come from.

Column	Description	Data type	Empty value	Default	Links
ID	Identifier of the planet	INT	NO	IDENTITY(0,1)	PK
Name	Name of planet	VARCHAR(20)	NO		
NumSuns	Number of suns (if known)	INT	YES	1	
NumMoons	Number of moons (if known)	INT	YES	1	
Population	Number of creatures of main planet civilization	BIGINT	YES		

3. Characters.

Data about certain chosen characters in the Star Wars Prequels trilogy.

Column	Description	Data type	Empty value	Default	Links
ID	Identifier of the character	INT	NO	IDENTITY(0,1)	PK
Name	Name of character	NVARCHAR(50)	NO		
Midichlorians	Number of midichlorians (force-sensitivity metric)	INT	YES	0	
PlayedBy	Actor's name	VARCHAR(50)	YES		
LikesSand	Whether character likes sand. NULL if unknown.	BIT	YES		
FromPlanet	Character's planet of origin. NULL if unknown.	INT	YES		FK(Planets)
Jedi	Whether the character is a Jedi(1), a Sith(0) or unknown(NULL).	BIT	YES	NULL	

4. CharactersMoviesRelations.

Instead of creating a comma separated list in the **Characters** table, this table links characters to the movies they appeared in.

Column	Description	Data type	Empty value	Default	Links
RelationID	Identifier of the relation	INT	NO	IDENTITY(0,1)	PK
CharacterID	ID of Character	INT	NO		FK(Characters)
MovieID	ID of Movie	TINYINT	NO		FK(Movies)
NumWords Spoken	# of words spoken by this character in this movie. Generated from Dialogue table.	INT	YES		

5. Dialogue.

Transactional table consisting of actual dialogue from the movie (only between characters in the **Characters** table).

Column	Description	Data type	Empty value	Default	Links
ID	Identifier of the dialogue	INT	NO	IDENTITY(0,1)	PK
Movie	ID of Movie	TINYINT	NO		FK(Movies)
FromID	ID of speaker	INT	NO		FK(Characters)
ToID	ID of listener. NULL if unknown.	INT	YES		FK(Characters)
Dialogue	What was actually spoken.	VARCHAR(586)	NO		
WordCount	# of words spoken in the Dialogue column. Generated.	INT	YES		

6. DialogueStats.

Table generated from the **Dialogue** table, this table tells us who spoke with whom and how much across all movies.

Column	Description	Data type	Empty value	Default	Links
FromID	ID of speaker	INT	NO		FK(Characters)
ToID	ID of listener. NULL if unknown.	INT	YES		FK(Characters)
NumWords	# of words spoken from FromID to ToID across all movies	INT	NO		

7. MasterApprenticeRelations.

Table specifying which character is which character's master.

Column	Description	Data type	Empty value	Default	Links
MasterID	ID of master	INT	YES		FK(Characters)
ApprenticeID	ID of apprentice	INT	YES		FK(Characters)

8. Fights.

Lightsaber duels between characters included in the Characters table.

Column	Description	Data type	Empty value	Default	Links
ID	ID of row	INT	NO		PK
FightID	ID of fight, different from ID since there can be multiple people in the same fight	INT	NO		
MovieID	ID of movie in which the fight took place	TINYINT	NO		FK(Movies)
FirstFigherID	ID of one of the fighters	INT	NO		FK(Characters)
SecondFigherID	ID of the other fighter	INT	NO		FK(Characters)
Winner	1 – first fighter won 0 – second fighter won NULL – unknown/tie	BIT	YES		

Reports

- Who is the master of the character who spoke the most per each movie (or NULL)?

View Master (or multiple separated by commas) of the character who spoke the most per each movie.

	Movie Part	Movie Name	Character With Most Words	Words	This Character's Master(s)
1	4	Phantom Menace	Qui-Gon Jinn	1085	Dooku
2	5	Attack of the Clones	Anakin Skywalker	1365	Qui-Gon Jinn, Obi-Wan Kenobi, Sheev Palpatine
3	6	Revenge of the Sith	Anakin Skywalker	1680	Qui-Gon Jinn, Obi-Wan Kenobi, Sheev Palpatine

We see that the masters who encouraged talking the most were:

- Dooku for Qui-Gon*
- Qui-Gon taught talking to Obi-Wan, they both taught Anakin alongside Palpatine*

- Which planet is the origin of the characters with the highest average of midichlorians?

Planet with highest amount of midichlorians.

	Planet	Maximum # of midichlorians	Characters Count	Top Representative
1	Tatooine	27700	1	Anakin Skywalker

We see to have the highest chance of being a Force user with large amount of midichlorians we should be born on Tatooine.

- Which character's word was worth the most \$ per each movie?

The amount of box office money in relation to percentage words spoken in all the movies combined.

	Name	Money earned by talking in millions \$
1	Anakin Skywalker	127.000848
2	Obi-Wan Kenobi	91.717366
3	Sheev Palpatine	71.431291
4	Padmé Amidala	63.338625
5	Qui-Gon Jinn	40.673638
6	Yoda	34.887044
7	Mace Windu	28.234112
8	Dooku	13.278054
9	General Grievous	3.889720
10	Maul	1.237078

We see that Anakin talking has influence on the movie making a lot of money.

4. Most talkative character grouped by movie.

Similar to the above, but we see who talked the most each movie, and also what relation that had on the movie's box office.

	Character	Movie	Worth in millions \$
1	Qui-Gon Jinn	Phantom Menace	122.020915
2	Anakin Skywalker	Attack of the Clones	111.683373
3	Anakin Skywalker	Revenge of the Sith	200.042780

5. Least talkative character grouped by movie.

Same as above, but we see who talked the least each movie, and also what relation that had on the movie's box office.

	Character	Movie	Worth in millions \$
1	Maul	Phantom Menace	3.711235
2	Sheev Palpatine	Attack of the Clones	20.454830
3	Dooku	Revenge of the Sith	5.715508

6. Which pair of character's spoke the most with each other?

View characters who spoke the most with each other from the set of the Characters table.

	Character1	Character2	Number of words	Example longest line
1	Padmé Amidala	Anakin Skywalker	3107	If Master Obi-Wan caught me doing this, he'd be v...
2	Obi-Wan Kenobi	Anakin Skywalker	2403	You are strong and wise, Anakin, and I am very pro...
3	Anakin Skywalker	Sheev Palpatine	1579	Anakin, if one is to understand the great mystery, o...
4	Yoda	Obi-Wan Kenobi	958	They say Master Sifo-Dyas placed an order for a cl...
5	Anakin Skywalker	Qui-Gon Jinn	760	Life-forms living together for mutual advantage. Wit...
6	Obi-Wan Kenobi	Qui-Gon Jinn	748	This is an odd play for the Trade Federation. We'v...
7	Padmé Amidala	Sheev Palpatine	494	There is no civility, only politics. The Republic is no...
8	Mace Windu	Anakin Skywalker	445	Anakin, we will deal with Count Dooku. The most im...
9	Dooku	Obi-Wan Kenobi	370	Don't be so sure, my young Jedi. You forget that h...
10	Padmé Amidala	Qui-Gon Jinn	287	We can enter the city using the secret passages o...

We see that Padme and Anakin were the most talkative pair, followed by Obi-Wan and Anakin.

7. Grouping characters in two groups: those who like sand and those who don't, what was the score of wins/losses when characters from opposing groups fought?

	LikesSand	Wins
1	0	2
2	1	3

3:2 for the sand liking fighters

8. Grouping characters into two groups: Sith and Jedi what was the score?

	Jedi	Wins
1	0	5
2	1	2

5:2 for the Sith

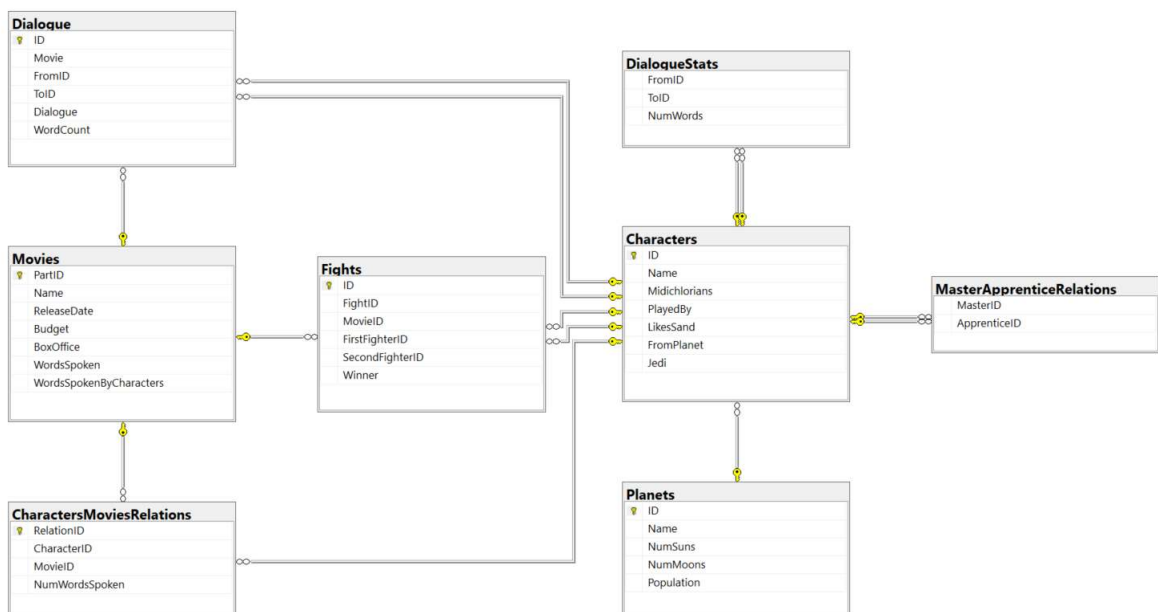
- Given a movie ID and a number N, find the character with the n-th most midichlorians in that movie (NULL if no force-sensitive users in particular movie).

Example @movie=4, @n=2;

	ID	Name	Midichlonians	PlayedBy	LikesSand	FromPlanet	Jedi	RelationID	CharacterID	MovielID	NumWordsSpoken	ROW
1	1	Sheev Palpatine	20500	Ian McDiarmid	1	1	0	1	1	4	344	2

Second most powerful force user in the first movie of the prequel trilogy is Palpatine.

ERD Diagram



Sources

The data is based on [this LA Times article](#) and also this data <https://github.com/datadesk/star-wars-analysis> published by them.

Midichlorian count is based on this source <http://starwarsuniverse2.tripod.com/id9.html>.

Movie data is sourced from Wikipedia, in-universe Star Wars information is sourced thanks to contributors of Wookieepedia.

Fights sourced from this video: <https://youtu.be/esnMDtMysHo?t=402>