Jack Wilburn

Assignment 5

1. What tables are in the database?
    - ActedIn
    - Directed
    - EpisodeOf
    - Genre
    - HasProfession
    - KnownFor
    - OfGenre
    - Person
    - Produced
    - Profession
    - Rating
    - Title
    - TitleType
    - Wrote
2. For each table, list the name and data type of each column. Which column(s) form the primary key? Do any appear to be foreign keys, and if so, to which table?
    - ActedIn
        - personId
            - int(11): integer with up to 11 digits
            - FK to Person
            - Part of PK
        - titleId
            - int(11): integer with up to 11 digits
            - FK to Title

        - Part of PK
- Directed
    - personId
        - int(11): integer with up to 11 digits
        - FK to Person
        - Part of PK
    - titleId
        - int(11): integer with up to 11 digits
        - FK to Title
        - Part of PK
- EpisodeOf
    - episodeId
        - int(11): integer with up to 11 digits
        - FK to Title
        - PK
    - seriesId
        - int(11): integer with up to 11 digits
        - MUL implies non-unique index
        - FK to Title
    - seasonNum
        - int(11): integer with up to 11 digits
    - episodeNum
        - int(11): integer with up to 11 digits
- Genre
    - id
        - int(11): integer with up to 11 digits
        - PK

        - text
            - varChar(16): string with up to 16 characters
            - Unique
    - HasProfession
        - personId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Person
        - professionId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Profession
    - KnownFor
        - personId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Person
        - titleId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Title
    - OfGenre
        - titleId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Title

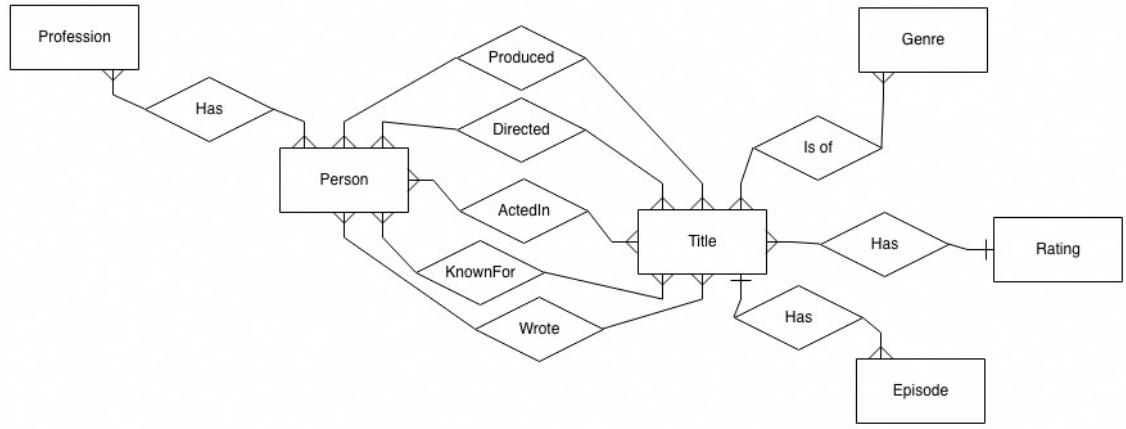
        - genreId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Genre
    - Person
        - id
            - int(11): integer with up to 11 digits
            - PK
        - name
            - varChar(128): string with up to 128 characters
        - birthYear
            - int(4): integer with up to 4 digits
        - deathYear
            - int(4): integer with up to 4 digits
    - Produced
        - titleId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Title
        - personId
            - int(11): integer with up to 11 digits
            - Part of PK
            - FK to Person
    - Profession
        - id
            - int(11): integer with up to 11 digits

        - PK
    - text
        - varChar(32): string with up to 32 characters
        - Unique
- Rating
    - titleId
        - int(11): integer with up to 11 digits
        - PK
        - FK to Title
    - avgRating
        - decimal(3,1): A number with one digit after the decimal point and up to 3 digits total
    - numVotes
        - int(11): integer with up to 11 digits
- Title
    - id
        - int(11): integer with up to 11 digits
        - PK
    - type
        - int(11): integer with up to 11 digits
        - MUL implies non-unique index
    - title
        - varChar(512): string with up to 512 characters
    - runtime
        - int(11): integer with up to 11 digits
- TitleType
    - id

            - int(11): integer with up to 11 digits
            - PK
        - text
            - varChar(16): string with up to 16 characters
            - Unique
    - Wrote
        - titleId
            - int(11): integer with up to 11 digits
            - PK
            - FK to Title
        - personId
            - int(11): integer with up to 11 digits
            - PK
            - FK to Person
3. Which tables appear to describe entities and which are relationships?
    - Entities
        - EpisodeOf
        - Genre
        - Person
        - Profession
        - Rating
        - Title
    - Relationships
        - ActedIn
        - Directed
        - HasProfession

- KnownFor
- OfGenre
- Produced
- TitleType
- Wrote

4. Using ERDPlus or similar software, make your best guess at an ERD describing the database structure.



5. How many titles are there?
    - SELECT COUNT(id) FROM Title;
        - 5152736
6. How many people are there?
    - SELECT COUNT(id) FROM Person;
        - 8891791
7. How many of these people have a name starting with K?
    - SELECT COUNT(id) FROM Person WHERE LEFT(name, 1) = 'k';
        - 403580
    - SELECT COUNT(id) FROM Person WHERE name LIKE "k%"
        - 403580

8. What is the longest title in the database? (The LENGTH() function determines the length of a string.)
    - SELECT title FROM Title ORDER BY LENGTH(title) DESC LIMIT 1;
        - What Could Possibly Be Amiss with These Women in Real Estate? And Who Pray Tell Is This Young Lady? Why Are These Women So Thrown by Such an Unassuming Young Woman? And Her Absolutely Charming Husband? The Men Certainly Seem Strange. Yes Something Seems to Be Off Here. Is That a Body? Hello?
9. How many people were born before 1600?
    - SELECT COUNT(name) FROM Person WHERE birthYear < 1600;
        - 137
10. How many people died in 2018?
    - SELECT COUNT(name) FROM Person WHERE deathYear = 2018;
        - 2855
11. Suppose that the first person added to the database got an ID number of 1, the second got ID number 2, and so on. What are the names of the first 10 people added to the database?
    - SELECT name FROM Person ORDER BY id ASC LIMIT 10;
        - Fred Astaire
        - Lauren Bacall
        - Brigitte Bardot
        - John Belushi
        - Ingmar Bergman
        - Ingrid Bergman
        - Humphrey Bogart
        - Marlon Brando
        - Richard Burton
        - James Cagney
12. How many of the titles in the database are video games?

- SELECT COUNT(title) FROM Title

  JOIN TitleType ON Title.type = TitleType.id

  WHERE text = 'videoGame';
- 21940

13.What is the largest number of ratings any title has?

- SELECT MAX(numVotes) FROM Rating;
- 2002417

14.Which title (or titles) is it and what is its average rating?

- SELECT title, avgRating FROM Rating

  JOIN Title ON Title.id = Rating.titleId

  WHERE numVotes = (SELECT MAX(numVotes) FROM Rating);

- title avgRating
- The Shawshank Redemption 9.3

15.Who acted in the movie “But I'm a Cheerleader”?

- SELECT name FROM ActedIn

  JOIN Title ON ActedIn.titleId = Title.id

  JOIN Person ON ActedIn.personId = Person.id

  WHERE title = "But I'm a Cheerleader";

- name
- Natasha Lyonne
- Clea DuVall
- Brandt Wille
- Michelle Williams

16.How many stunt people are in the database?

- SELECT COUNT(*) FROM Person

JOIN HasProfession ON Person.id = HasProfession.personId

JOIN Profession ON HasProfession.professionId = Profession.id

WHERE text = 'stunts';

- 63794

17. **[Extra credit]** What is the title, season number, and episode number of every episode of Babylon 5, sorted by season and then episode number? (**Tricky** — involves a recursive relationship)
    - SELECT seasonNum, episodeNum, Title2.title FROM Title

      JOIN EpisodeOf ON Title.id = EpisodeOf.seriesId

      JOIN Title AS Title2 ON Title2.id = EpisodeOf.episodeId

      WHERE Title.title LIKE 'Babylon 5%'

      ORDER BY seasonNum, episodeNum;

    - seasonNum episodeNum title
    - 1 1 Midnight on the Firing Line
    - 1 2 Soul Hunter
    - 1 3 Born to the Purple
    - 1 4 Infection
    - 1 5 The Parliament of Dreams
    - 1 6 Mind War
    - 1 7 The War Prayer
    - 1 8 And the Sky Full of Stars
    - 1 9 Deathwalker
    - 1 10 Believers
    - 1 11 Survivors
    - 1 12 By Any Means Necessary
    - 1 13 Signs and Portents
    - 1 14 TKO
    - 1 15 Grail
    - 1 16 Eyes

◦ 1	17	Legacies
◦ 1	18	A Voice in the Wilderness: Part 1
◦ 1	19	A Voice in the Wilderness: Part 2
◦ 1	20	Babylon Squared
◦ 1	21	The Quality of Mercy
◦ 1	22	Chrysalis
◦ 2	1	Points of Departure
◦ 2	2	Revelations
◦ 2	3	The Geometry of Shadows
◦ 2	4	A Distant Star
◦ 2	5	The Long Dark
◦ 2	6	Spider in the Web
◦ 2	7	Soul Mates
◦ 2	8	A Race Through Dark Places
◦ 2	9	The Coming of Shadows
◦ 2	10	GROPOS
◦ 2	11	All Alone in the Night
◦ 2	12	Acts of Sacrifice
◦ 2	13	Hunter, Prey
◦ 2	14	There All the Honor Lies
◦ 2	15	And Now for a Word
◦ 2	16	In the Shadow of Z'ha'dum
◦ 2	17	Knives
◦ 2	18	Confessions and Lamentations
◦ 2	19	Divided Loyalties
◦ 2	20	The Long, Twilight Struggle
◦ 2	21	Comes the Inquisitor
◦ 2	22	The Fall of Night
◦ 3	1	Matters of Honor
◦ 3	2	Convictions
◦ 3	3	A Day in the Strife
◦ 3	4	Passing Through Gethsemane
◦ 3	5	Voices of Authority

◦	3	6	Dust to Dust
◦	3	7	Exogenesis
◦	3	8	Messages from Earth
◦	3	9	Point of No Return
◦	3	10	Severed Dreams
◦	3	11	Ceremonies of Light and Dark
◦	3	12	Sic Transit Vir
◦	3	13	A Late Delivery from Avalon
◦	3	14	Ship of Tears
◦	3	15	Interludes and Examinations
◦	3	16	War Without End: Part One
◦	3	17	War Without End: Part Two
◦	3	18	Walkabout
◦	3	19	Grey 17 Is Missing
◦	3	20	And the Rock Cried Out, No Hiding Place
◦	3	21	Shadow Dancing
◦	3	22	Z'ha'dum
◦	4	1	The Hour of the Wolf
◦	4	2	Whatever Happened to Mr. Garibaldi?
◦	4	3	The Summoning
◦	4	4	Falling Toward Apotheosis
◦	4	5	The Long Night
◦	4	6	Into the Fire
◦	4	7	Epiphanies
◦	4	8	The Illusion of Truth
◦	4	9	Atonement
◦	4	10	Racing Mars
◦	4	11	Lines of Communication
◦	4	12	Conflicts of Interest
◦	4	13	Rumors, Bargains and Lies
◦	4	14	Moments of Transition
◦	4	15	No Surrender, No Retreat
◦	4	16	Exercise of Vital Powers

- 4 17 The Face of the Enemy
- 4 18 Intersections in Real Time
- 4 19 Between the Darkness and the Light
- 4 20 Endgame
- 4 21 Rising Star
- 4 22 The Deconstruction of Falling Stars
- 5 1 No Compromises
- 5 2 The Very Long Night of Londo Mollari
- 5 3 The Paragon of Animals
- 5 4 A View from the Gallery
- 5 5 Learning Curve
- 5 6 Strange Relations
- 5 7 Secrets of the Soul
- 5 8 Day of the Dead
- 5 9 In the Kingdom of the Blind
- 5 10 A Tragedy of Telepaths
- 5 11 Phoenix Rising
- 5 12 The Ragged Edge
- 5 13 The Corps Is Mother, the Corps Is Father
- 5 14 Meditations on the Abyss
- 5 15 Darkness Ascending
- 5 16 And All My Dreams, Torn Asunder
- 5 17 Movements of Fire and Shadow
- 5 18 The Fall of Centauri Prime
- 5 19 The Wheel of Fire
- 5 20 Objects in Motion
- 5 21 Objects at Rest
- 5 22 Sleeping in Light

18. For each type of title: what is the name of the type and the number of titles of this type in the database?

- SELECT text, type, COUNT(type) FROM Title

  JOIN TitleType ON Title.type = TitleType.id

  GROUP BY type;

- text COUNT(type)
- short 644893
- movie 490793
- tvMovie 126414
- tvSeries 151766
- tvEpisode 3538604
- tvShort 8715
- tvMiniSeries 24222
- tvSpecial 12417
- video 132972
- videoGame 21940

19. For each genre: what is the name of the genre and the average runtime of movies of this genre, sorted by runtime? Many of the runtimes listed in the database are zero (meaning missing information) so include in the average only nonzero runtimes.
    - SELECT text, AVG(runtime) FROM Genre

      JOIN OfGenre ON Genre.id = OfGenre.genreId

      JOIN Title ON OfGenre.titleId = Title.id

      WHERE runtime != 0

      GROUP BY genreId

      ORDER BY AVG(runtime);

    - text AVG(runtime)
    - Short 12.9371
    - Animation 19.7924
    - Family 33.8504
    - Fantasy 35.6613
    - Reality-TV 36.8992

- Sci-Fi 38.2246
- Comedy 38.8569
- Game-Show 41.0717
- Music 41.1684
- Horror 43.6675
- Adventure 44.3124
- Documentary 46.7192
- Mystery 49.4319
- Action 50.0251
- Romance 50.3294
- Drama 50.3964
- Western 50.8007
- News 51.9512
- Biography 53.4057
- Crime 55.9504
- Talk-Show 57.8853
- History 58.6634
- Thriller 59.3740
- Musical 61.6510
- Sport 66.4076
- War 69.6917
- Film-Noir 83.4458

20. Same as the previous query, but include only those genres whose average runtime is less than 45 minutes.
    - SELECT text, AVG(runtime) FROM Genre

      JOIN OfGenre ON Genre.id = OfGenre.genreId

      JOIN Title ON OfGenre.titleId = Title.id

WHERE runtime != 0

GROUP BY genreId

HAVING AVG(runtime) < 45

ORDER BY AVG(runtime);

- text AVG(runtime)
- Short 12.9371
- Animation 19.7924
- Family 33.8504
- Fantasy 35.6613
- Reality-TV 36.8992
- Sci-Fi 38.2246
- Comedy 38.8569
- Game-Show 41.0717
- Music 41.1684
- Horror 43.6675
- Adventure 44.3124

21. **[Extra credit]** What is the name and the number of episodes written of every person who has ever written at least five episodes of Babylon 5?
    - SELECT COUNT(personId), personId, name FROM Title

      JOIN EpisodeOf ON Title.id = EpisodeOf.seriesId

      JOIN Title AS Title2 ON Title2.id = EpisodeOf.episodeId

      JOIN Wrote ON Wrote.titleId = Title2.id

      JOIN Person ON Person.id = Wrote.personId

      WHERE Title.title LIKE 'Babylon 5%'

      GROUP BY personId

      HAVING COUNT(personId) >= 5;

COUNT(personId)	personId	name
7	228489	Lawrence G. DiTillio
109	833089	J. Michael Straczynski