

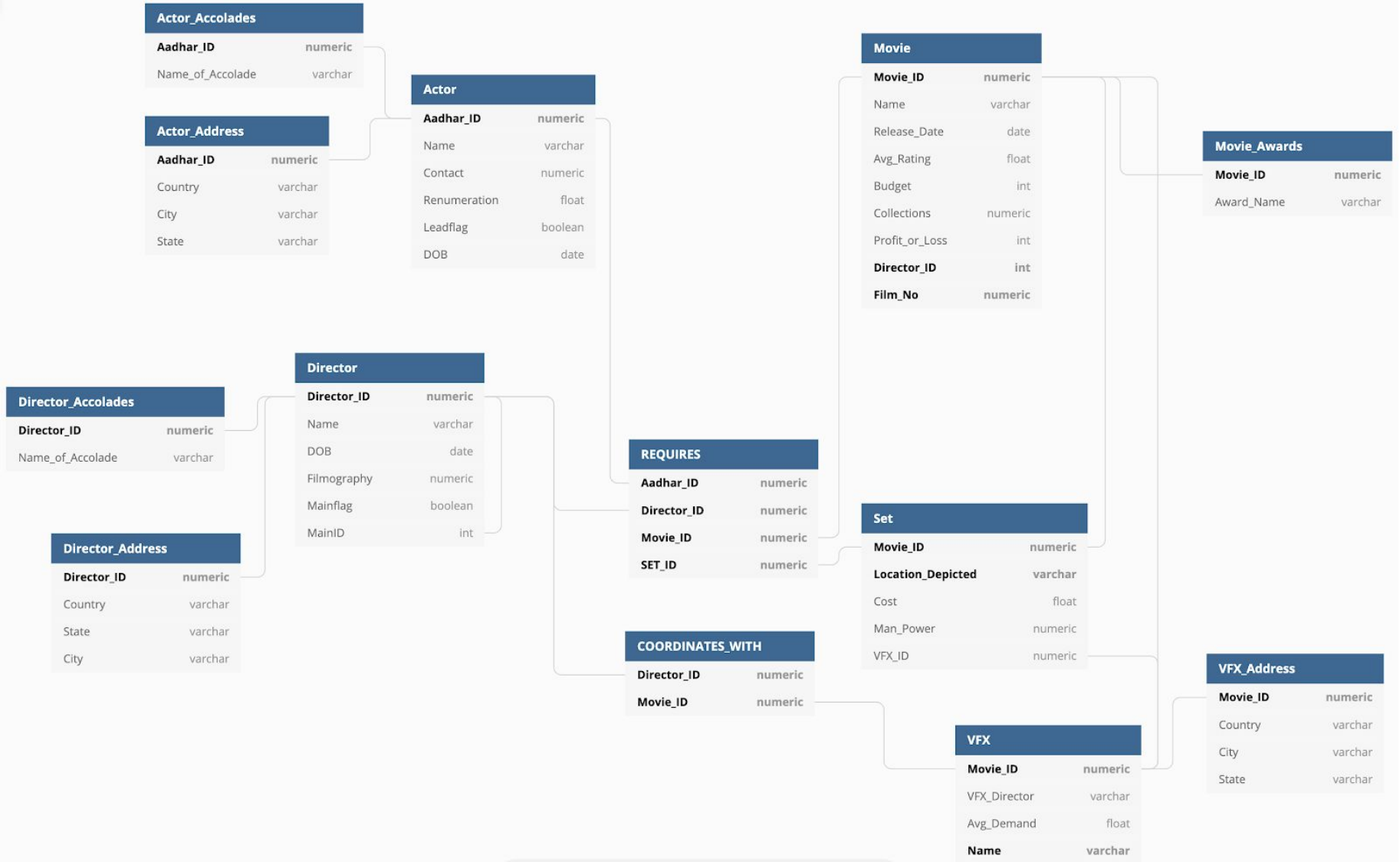
WeCookRawData

Members:
Gokul Vamsi Thota - 2019111009
T.H.Arjun - 2019111012
Mukkara Rakesh Reddy - 2019101087

Relational Model/ 1NF

This snapshot corresponds to the relational model / its conversion to first normal form (1NF). There were no changes made in converting the relational model to its first normal form, because of the highly optimised nature of the database (domains of all attributes in the relational model are atomic values).

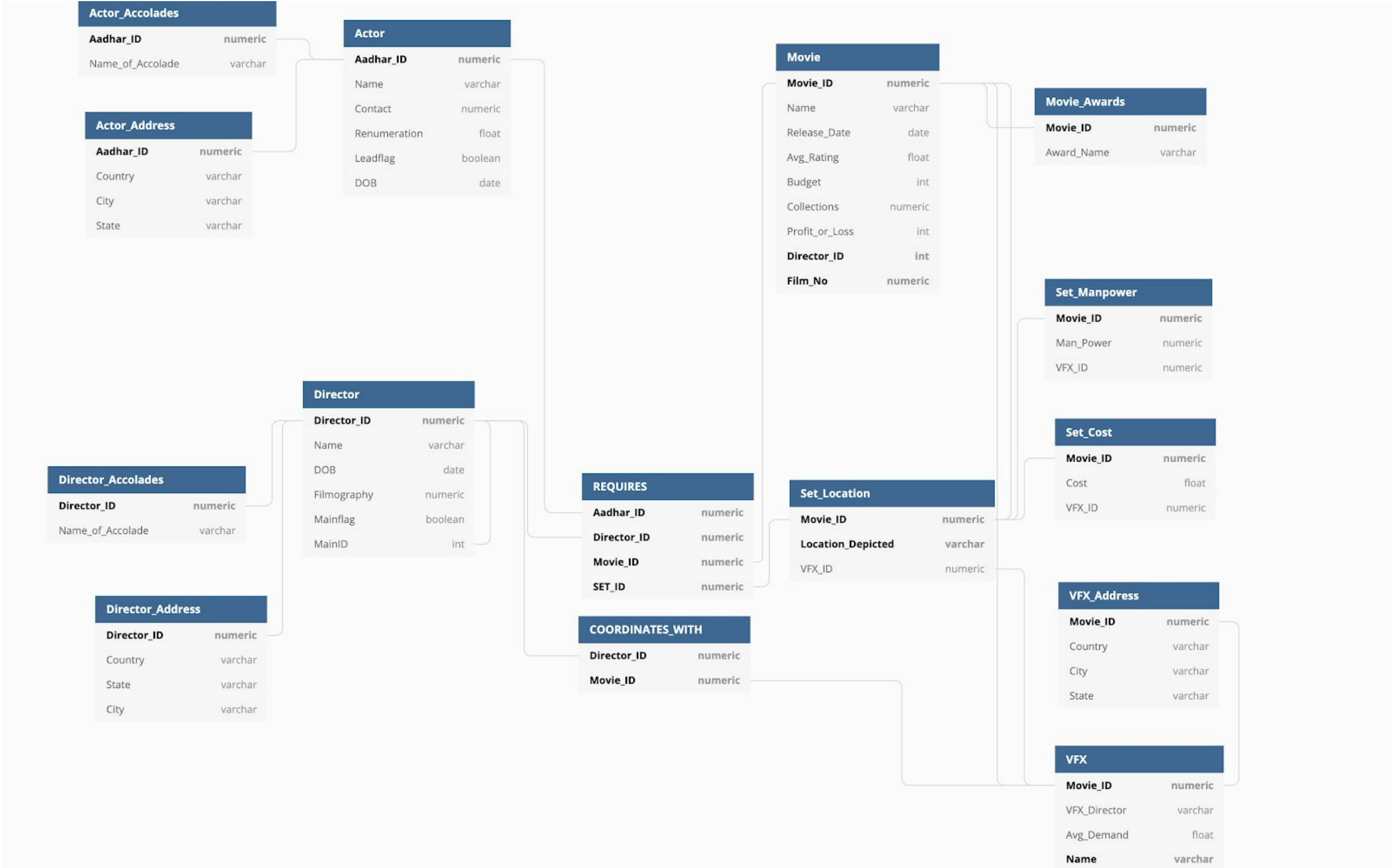
- In the 'actors' entity type, the multi-valued attributes 'Accolades' and 'Addresses' have been decomposed into respective relations. This ensures that the domain comprises atomic values. In these two relations, the actor's Aadhar_ID acts as the foreign key. It acts as the primary key for this entity type. Also, as derived attributes do not exist in the relational model, the 'age' attribute, which is a derived attribute, has been omitted from both the relations - 'actor' and 'director'.
- Similar to the 'actors' entity type, the entity type 'directors' has been broken down into 3 relations, in which 2 correspond to director's 'accolades' and 'addresses' respectively, whereas the other corresponds to the remaining attributes which contain atomic values in their domain. These two separated relations contain Director_ID as the foreign key. It acts as the primary key for this entity type.
- The entity type 'movies', majorly comprises simple attributes. The multi-valued attribute 'awards' has been separated into a different relation, with foreign key as Movie_ID. But this entity type has Movie_ID as the primary key.
- The entity type 'VFX studio' has been decomposed to two relations, in which one relation depicts the multiple-valued composite attribute 'Address', which consists of 'Movie_ID' as its foreign key. But the relation 'VFX' has 'Movie_ID' as the primary key. This is because the 'movie' entity type acts as identifying entity type for the weak entity type 'VFX studio'. Hence, its key is included as the primary key here. 'Name' acts as its partial key here.
- The weak entity type 'Set' comprises only simple attributes, hence, it is depicted by the single relation 'Set' whose primary key is 'Movie_ID' (because a set is in an identifying relationship 'REQUIRES' with the 'Movie' entity type.) Location_depicted acts as its partial key here.
- The quaternary relationship 'REQUIRES' has been depicted by a relation 'REQUIRES'. It comprises of four foreign keys, which are actually the primary keys of its participating entity types, i.e, Aadhar_ID for actors, Director_ID for directors, Movie_ID for movies, Set_ID for sets (which is 'Movie_ID', the primary key of identifying entity type of sets).
- The relationship 'COORDINATES_WITH' is depicted by the relation 'COORDINATES WITH'. It comprises two foreign keys, which are the primary keys of the entity types 'movie' (Movie_ID) and 'director' (Director_ID). It cannot be expressed as a single attribute for any of the participant entity types, because it is a many to many relationship.



2 NF

The major difference in the transition from the first normal form to the second normal form is in the decomposition of the relation 'Sets'. The major reason behind this is the existence of a partial dependency among its attributes.

- The attributes of the relation 'set' apart from its foreign keys and primary key (key of owner entity type) are - location depicted, man power and cost. The location depicted alone determines both the man power and cost of the set, it is because different kinds of locations need different amount of effort from people in creating and maintaining the particular set, and thus, it determines the cost required for making the set, which involve the money to be paid to human labour and various props etc. Hence, the attributes 'man power' and 'cost' are dependent only on a single prime attribute (location depicted, which is the partial key here), and are independent of the movie ID, leading to partial dependency.
- To accommodate this discrepancy and to normalize it to satisfy the second normal form, the relation 'set' has been decomposed into 3 different relations - Set_location, Set_man power and Set_cost. All of these contain only their corresponding attributes (attributes their name represents) and foreign keys (Movie ID and VFX ID).



3 NF

There are a significant amount of changes in the transition from second normal form to third normal form, vowing to the decent amount of inter-dependencies among non-prime attributes of various relations.

- For the 3 entity types 'Actor', 'Director' and 'VFX studio', there exists a relation describing their corresponding address. That relation consists of the attributes - country, city and state. But since all of these three are non prime attributes, but still are interdependent on each other, it is required to split these attributes into their own relation, so as to ensure that it is in third normal form. Hence, for all the three entity types mentioned above, the 'Address' relation has been decomposed into 3 relations, representing the corresponding country, city and state.
- In the relation 'Movie', we can find a few dependencies among non prime attributes - average rating, budget, collections, profit/loss, release date. The average rating of a movie decides the number of people willing to watch it, which in turn determines its collections, and those collections and budget of the movie determine the profit / loss it achieved. Also, the release date of a movie plays a key role in determining collections, as movies releasing on holidays / festivals tend to achieve decent collections (Imagine the fate of a movie which is well made, but is being released during the lockdown period). Hence, it is essential to separate the attributes 'Collections' and 'Profit or loss' from all other attributes of a 'movie'.

