**Queries that have been used to create Tables in Redshift Cluster**

create table genres(

genre\_id integer not null,

genre\_name varchar(100)

);

create table movie\_keyword(

movie\_id integer not null,

keyword\_id integer not null

);

create table keyword(

keyword\_id integer not null,

keyword\_name varchar(100)

);

create table movie\_genre(

movie\_id integer not null,

genre\_id integer not null

);

create table production\_company(

pc\_id integer not null,

production\_name varchar(1000)

);

create table production\_datasource2(

pmovie\_id integer not null,

movie\_name varchar(1000),

country\_name varchar(100),

country\_id integer not null,

turnover\_in\_dollars integer not null

);

create table movies(

adult boolean,

budget integer,

movie\_id integer not null,

tmdb\_id integer not null,

imdb\_id integer not null,

original\_language varchar(20),

original\_title varchar(1000),

overview varchar(5000),

popularity decimal(9,6),

poster\_path varchar(200),

release\_date varchar(20),

revenue bigint,

runtime integer,

movie\_status varchar(20),

tagline varchar(1000),

title varchar(1000),

vote\_average decimal(8,2),

vote\_count integer

);

create table movie\_prodcomp(

movie\_id integer not null,

pc\_id integer not null

);

create table movie\_ratings(

movie\_id integer not null,

ratings\_id integer not null

);

create table ratings(

ratings\_id integer not null,

userid integer not null,

movie\_id integer not null,

rating decimal(8,2),

rating\_time integer

);

create table moviemetadata2(

imdb\_id integer not null,

titleType varchar(20),

Movie\_name varchar(100),

OriginalTiltle varchar(100),

isAdult varchar(20),

startYear varchar(20),

endYear varchar(20),

runtimeMinutes varchar(20),

genres varchar(20)

);

create table Ratingdatasource2(

rating\_imdb\_id integer not null,

averageRating decimal(9,6),

numVotes integer );

create table Castdatasource2(

castName varchar(100),

birthYear varchar(20),

deathYear varchar(20),

profession varchar(100),

movie\_imdb\_id integer not null);

**Queries that have been used to load the tables from the files uploaded in the S3 Bucket**

copy genres from 's3://moviesdb/movies/genres.txt'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter '|' region 'us-east-2';

copy keyword from 's3://moviesdb/movies/keyword.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2';

copy movie\_keyword from 's3://moviesdb/movies/movie\_keyword.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2';

copy movie\_rating from 's3://moviesdb/movies/movie\_ratings.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2

copy ratings from 's3://moviesdb/movies/ratings.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2';

copy movie\_genre from 's3://moviesdb/movies/movie\_genre.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2';

copy movie\_prodcomp from 's3://moviesdb/movies/movie\_prodcomp.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2'

copy production\_company from 's3://moviesdb/movies/production\_company.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2';

copy movies from 's3://moviesdb/movies/movies\_metadata\_34.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2';

copy production\_datasource2 from 's3://moviesdb/movies/pproductioin\_datasource2New.csv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter ',' region 'us-east-2';

copy moviemetadata2 from 's3://moviesdb/MovieDatasource2.tsv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter '\t' region 'us-east-2';

copy castdatasource2 from 's3://moviesdb/CastDatasource2.tsv'

credentials 'aws\_iam\_role=arn:aws:iam::813540096928:role/myredshift\_moviesrole'

delimiter '\t' region 'us-east-2';