

-Arpan Pradhan

### Assignment 3

#### 2. Download Sakila Snowflake dimensional data model (DDL) from canvas –{ 40 Points }

a. Run the DDL & DML provided for the Sakila dimension model

##### DDL script for fact\_rental table-

```
CREATE TABLE IF NOT EXISTS `sakila_snowflake`.`fact_rental`(  
  `rental_id` INT(10) NOT NULL AUTO_INCREMENT,  
  `rental_last_update` TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP  
ON UPDATE CURRENT_TIMESTAMP,  
  `customer_key` INT(8) NOT NULL,  
  `staff_key` INT(8) NOT NULL,  
  `film_key` INT(8) NOT NULL,  
  `store_key` INT(8) NOT NULL,  
  `rental_date_key` BIGINT(20) NOT NULL,  
  `return_date_key` BIGINT(20) DEFAULT NULL,  
  `count_rentals` INT(8) NOT NULL,  
  `count_returns` INT(10) NOT NULL,  
  `rental_duration` INT(10) NOT NULL,  
  `dollar_amount` FLOAT NOT NULL,  
  PRIMARY KEY (`rental_id`),  
  CONSTRAINT FOREIGN KEY (`customer_key`)  
    REFERENCES `sakila_snowflake`.`dim_customer` (`customer_key`)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION,  
  CONSTRAINT FOREIGN KEY (`staff_key`)  
    REFERENCES `sakila_snowflake`.`dim_staff` (`staff_key`)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION,  
  CONSTRAINT FOREIGN KEY (`film_key`)  
    REFERENCES `sakila_snowflake`.`dim_film` (`film_key`)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION,  
  CONSTRAINT FOREIGN KEY (`store_key`)  
    REFERENCES `sakila_snowflake`.`dim_store` (`store_key`)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION,  
  CONSTRAINT FOREIGN KEY (`rental_date_key`)  
    REFERENCES `sakila_snowflake`.`dim_date` (`date_id`)
```

```

        ON DELETE NO ACTION
        ON UPDATE NO ACTION,
        CONSTRAINT FOREIGN KEY (`return_date_key`)
            REFERENCES `sakila_snowflake`.`dim_date` (`date_id`)
        ON DELETE NO ACTION
        ON UPDATE NO ACTION
    )
ENGINE=INNODB DEFAULT CHARACTER SET=UTF8;

```

### **DML Script for Fact\_rental table**

```

INSERT INTO sakila_snowflake.fact_rental (
    rental_id,
    rental_last_update,
    customer_key,
    staff_key,
    film_key,
    store_key,
    rental_date_key,
    return_date_key,
    count_rentals,
    count_returns,
    rental_duration,
    dollar_amount
)
SELECT
    r.rental_id,
    r.last_update,
    ss_c.customer_key,
    ss_stf.staff_key,
    ss_f.film_key,
    ss_str.store_key,
    rent.date_Id AS rental_date_key,
    rtn.date_Id AS return_date_key,
    1,
    CASE WHEN r.return_date IS NOT NULL THEN 1 ELSE 0 END AS count_returns,
    f.rental_duration,
    p.amount AS dollar_amount
FROM
    sakila.rental r
    INNER JOIN sakila.staff stf ON r.staff_id = stf.staff_id
    INNER JOIN sakila.inventory i ON r.inventory_id = i.inventory_id
    INNER JOIN sakila.film f ON i.film_id = f.film_id
    INNER JOIN sakila.customer c ON r.customer_id = c.customer_id
    INNER JOIN sakila.store str ON c.store_id = str.store_id
    INNER JOIN sakila.payment p ON r.rental_id = p.rental_id

```

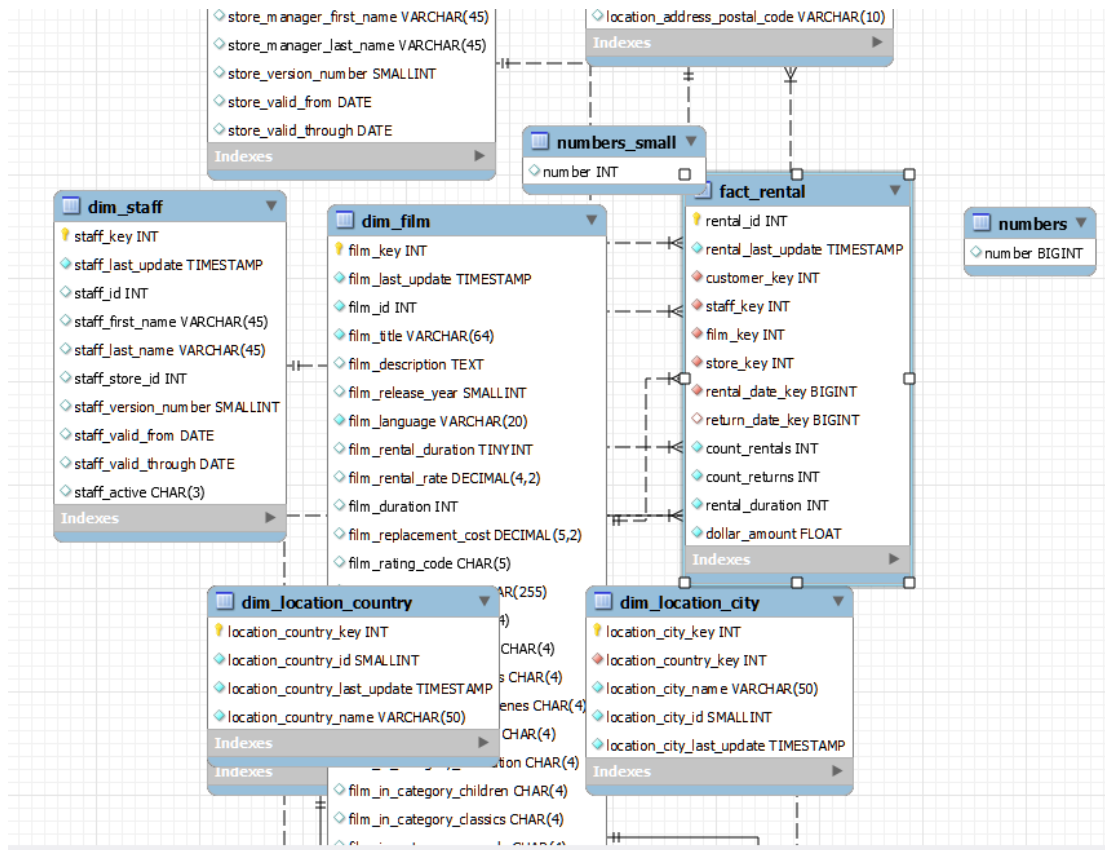
```

INNER JOIN sakila_snowflake.dim_customer ss_c ON c.customer_id =
ss_c.customer_id
INNER JOIN sakila_snowflake.dim_staff ss_stf ON stf.staff_id = ss_stf.staff_id
INNER JOIN sakila_snowflake.dim_film ss_f ON f.film_id = ss_f.film_id
INNER JOIN sakila_snowflake.dim_store ss_str ON str.store_id = ss_str.store_id
LEFT JOIN (
    SELECT r.rental_id, ss_d.date_id FROM sakila.rental r
    LEFT JOIN sakila_snowflake.dim_date ss_d ON DATE_FORMAT(r.rental_date,
'%Y-%m-%d') = ss_d.date
) rent ON r.rental_id = rent.rental_id
LEFT JOIN (
    SELECT r.rental_id, ss_d.date_id FROM sakila.rental r
    LEFT JOIN sakila_snowflake.dim_date ss_d ON DATE_FORMAT(r.return_date,
'%Y-%m-%d') = ss_d.date
) rtn ON r.rental_id = rtn.rental_id;

```

b. For the missing Fact table fact\_rental create the DDL and the DML scripts to pull data from the Sakila ER model. Modify/improve the fact table script if you need more metrics for the Tableau report below ( For example to include the dollar amount for each rental use rental rate \* number of days rented ).

c. Generate the EER diagram for the final dimensional model.



**d. Identify any scripts errors and areas of improvement in the data model.**

- There are few errors and area of improvements. More tables such as inventory can be added to perform analysis on inventory status. Naming convention was not consistent, normalization is beneficial. Also, there are tables like numbers and numbers\_small as placeholders. They could be removed to reduce unnecessary objects.

**3. Using Tableau and referencing the snowflake dimensional model ( Sakila data warehouse ) create 5 different reports that provides insights into the sakila dataset. You may use any of the below use cases or you can create your own reports, dashboards**

**– { 40 Points }**

- Create a report that helps visualize list of movies titles grouped by categories**

Rows	Name	Title
Sheet 1		
Name	Title	
Action	AMADEUS HOLY	Abc
	AMERICAN CIRCUS	Abc
	ANTITRUST TOMATOES	Abc
	ARK RIDGEMONT	Abc
	BAREFOOT MANCHURIAN	Abc
	BERETS AGENT	Abc
	BRIDE INTRIGUE	Abc
	BULL SHAWSHANK	Abc
	CADDYSHACK JEDI	Abc
	CAMPUS REMEMBER	Abc
	CASUALTIES ENCINO	Abc
	CELEBRITY HORN	Abc
	CLUELESS BUCKET	Abc
	CROW GREASE	Abc
	DANCES NONE	Abc
	DARKO DORADO	Abc
	DARN FORRESTER	Abc
	DEVIL DESIRE	Abc
	DRAGON SQUAD	Abc
	DREAM PICKUP	Abc
	DRIFTER COMMANDMEN..	Abc
	EASY GLADIATOR	Abc

Name
<input checked="" type="checkbox"/> (All)
<input checked="" type="checkbox"/> Action
<input checked="" type="checkbox"/> Animation
<input checked="" type="checkbox"/> Children
<input checked="" type="checkbox"/> Classics
<input checked="" type="checkbox"/> Comedy
<input checked="" type="checkbox"/> Documentary
<input checked="" type="checkbox"/> Drama
<input checked="" type="checkbox"/> Family
<input checked="" type="checkbox"/> Foreign
<input checked="" type="checkbox"/> Games
<input checked="" type="checkbox"/> Horror
<input checked="" type="checkbox"/> Music
<input checked="" type="checkbox"/> New
<input checked="" type="checkbox"/> Sci-Fi
<input checked="" type="checkbox"/> Sports
<input checked="" type="checkbox"/> Travel

**b. Report showing Film title and Film rating Code.**

Columns	Film Rating Code					
Rows	Film Title					

Sheet 1

Film Title		Film Rating Code				
		G	NC-17	PG	PG-13	R
ACADEMY DINOSAUR				Abc		
ACE GOLDFINGER		Abc				
ADAPTATION HOLES			Abc			
AFFAIR PREJUDICE		Abc				
AFRICAN EGG		Abc				
AGENT TRUMAN				Abc		
AIRPLANE SIERRA					Abc	
AIRPORT POLLOCK						Abc
ALABAMA DEVIL					Abc	
ALADDIN CALENDAR			Abc			
ALAMO VIDEOTAPE		Abc				
ALASKA PHANTOM				Abc		
ALI FOREVER				Abc		
ALICE FANTASIA			Abc			
ALIEN CENTER			Abc			
ALLEY EVOLUTION			Abc			
ALONE TRIP						Abc
ALTER VICTORY					Abc	
AMADEUS HOLY				Abc		
AMELIE HELLFIGHTERS						Abc
AMERICAN CIRCUS						Abc
AMISTAD MIDSUMMER		Abc				
ANACONDA CONFESSIONS						Abc
ANALYZE HOOSIERS						Abc
ANGELS LIFE		Abc				
ANNIE IDENTITY		Abc				
ANONYMOUS HUMAN			Abc			
ANTHEM LUKE					Abc	
ANTITRUST TOMATOES			Abc			
ANYTHING SAVANNAH						Abc

c. Create visualization with payments for each customer grouped by city.

## Sheet 1

First Name	Last Name	City							
AARON	SELBY	Mwene-Ditu							
ADAM	GOOCH	Adoni							
ADRIAN	CLARY	Udine							
AGNES	BISHOP	Sambhal							
ALAN	KAHN	Emeishan							
ALBERT	CROUSE	Bamenda							
ALBERTO	HENNING	Barcelona							
ALEX	GRESHAM	Uruapan							
ALEXANDER	FENNELL	Bergamo							
ALFRED	CASILLAS	Sawhaj							
ALFREDO	MCADAMS	Serpuhov							
ALICE	STEWART	Fontana							
ALICIA	MILLS	Nagaon							
ALLAN	CORNISH	Tarlac							
ALLEN	BUTTERFIELD	Hoshiarpur							
ALLISON	STANLEY	Phnom Penh							
ALMA	AUSTIN	Mannheim							
ALVIN	DELOACH	Cuauhtmoc							
AMANDA	CARTER	Nador							
AMBER	DIXON	Touliu							
AMY	LOPEZ	Jhansi							
ANA	BRADLEY	Memphis							
ANDRE	RAPP	Coquimbo							

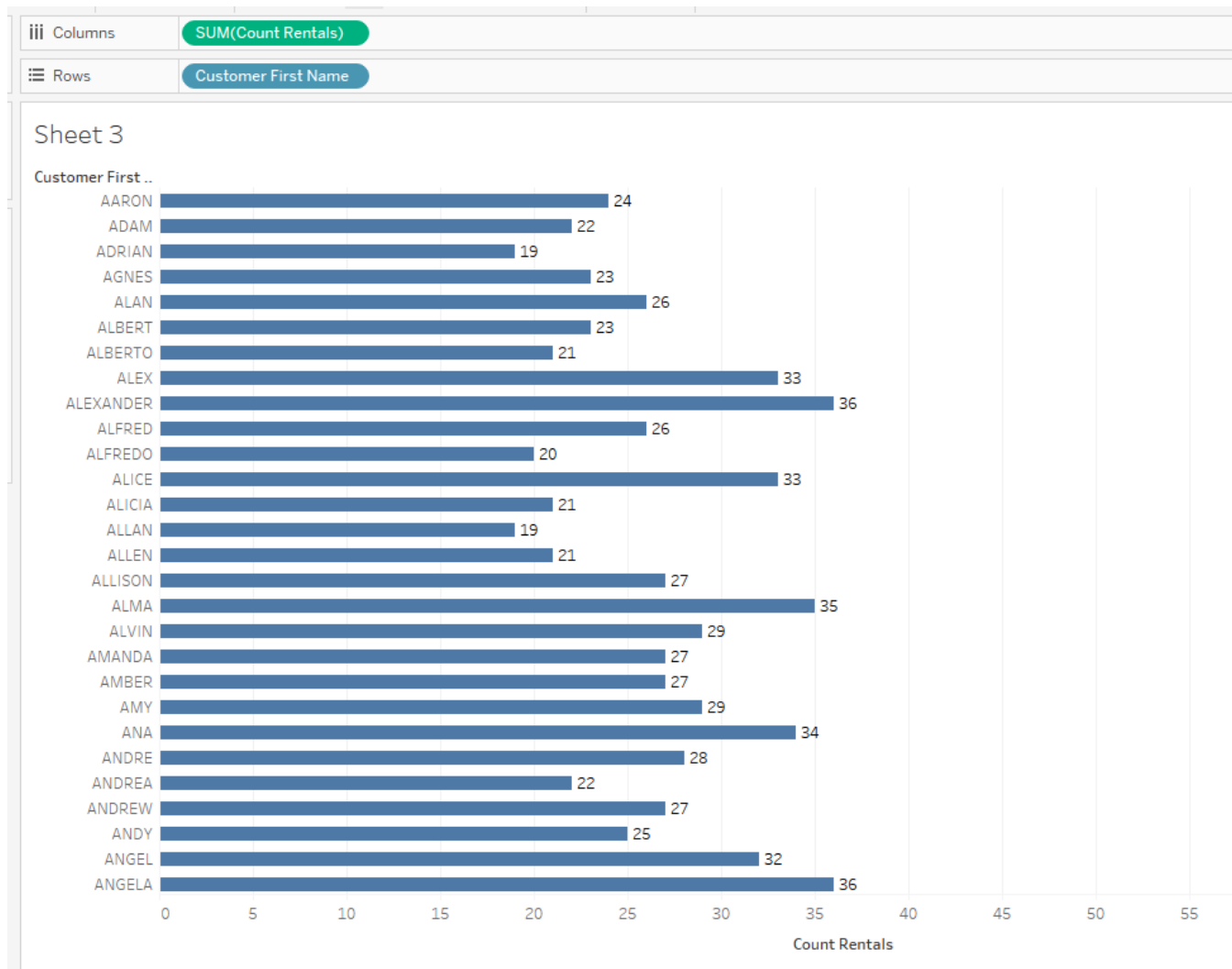
c. Create a report that list the Labels and language for all DVDs.

## Sheet 3

Film Langu..	Film Title		
English	ACADEMY DINOSAUR	Abc	^
	ACE GOLDFINGER	Abc	
	ADAPTATION HOLES	Abc	
	AFFAIR PREJUDICE	Abc	
	AFRICAN EGG	Abc	
	AGENT TRUMAN	Abc	
	AIRPLANE SIERRA	Abc	
	AIRPORT POLLOCK	Abc	
	ALABAMA DEVIL	Abc	
	ALADDIN CALENDAR	Abc	
	ALAMO VIDEOTAPE	Abc	
	ALASKA PHANTOM	Abc	
	ALI FOREVER	Abc	
	ALICE FANTASIA	Abc	
	ALIEN CENTER	Abc	
	ALLEY EVOLUTION	Abc	
	ALONE TRIP	Abc	
	ALTER VICTORY	Abc	
	AMADEUS HOLY	Abc	
	AMELIE HELLFIGHTERS	Abc	
	AMERICAN CIRCUS	Abc	
	AMISTAD MIDSUMMER	Abc	
	ANACONDA CONFESSIONS	Abc	
	ANALYZE HOOSIERS	Abc	
	ANGELS LIFE	Abc	
	ANNIE IDENTITY	Abc	
	ANONYMOUS HUMAN	Abc	
	ANTHEM LUKE	Abc	
	ANTITRUST TOMATOES	Abc	
	ANYTHING CAN HAPPEN	Abc	

e. Create charts that gives insight into the number of customers, ranged by number of rentals made





f. Create charts that gives insight into the number of rentals for each store and name of store manager.

Columns: Store Manager First .., Store Manager Last ..

### Sheet 6

Store Mana..	Store Mana..	
Jon	Stephens	30,409
Mike	Hillyer	36,998

2. Create a dashboard based on the report set that you have created .

