**DEPA Assignment 3**

Troy Zhongyi Zhang

Netid: [zhongyiz@uchicago.edu](mailto:zhongyiz@uchicago.edu)

**SakilaSnowflakeDW-DDL:**

-- -----------------------------------------------------

-- Table `sakila\_snowflake`.`fact\_rental`

-- Write Fact table fact\_rental DDL script here

-- -----------------------------------------------------

CREATE TABLE IF NOT EXISTS `sakila\_snowflake`.`fact\_rental` (

`rental\_id` INT(10) NOT NULL,

`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\_returns` INT(10) NOT NULL,

`count\_rentals` INT(8) NOT NULL,

`rental\_duration` INT(10) NULL DEFAULT NULL,

`dollar\_amount` FLOAT,

PRIMARY KEY (`rental\_id`),

CONSTRAINT `dim\_customer\_fact\_rental\_fk`

FOREIGN KEY (`customer\_key`)

REFERENCES `sakila\_snowflake`.`dim\_customer` (`customer\_key`)

ON DELETE CASCADE

ON UPDATE CASCADE,

CONSTRAINT `dim\_film\_fact\_rental\_fk`

FOREIGN KEY (`film\_key`)

REFERENCES `sakila\_snowflake`.`dim\_film` (`film\_key`)

ON DELETE CASCADE

ON UPDATE CASCADE,

CONSTRAINT `dim\_staff\_fact\_rental\_fk`

FOREIGN KEY (`staff\_key`)

REFERENCES `sakila\_snowflake`.`dim\_staff` (`staff\_key`)

ON DELETE CASCADE

ON UPDATE CASCADE,

CONSTRAINT `dim\_store\_fact\_rental\_fk`

FOREIGN KEY (`store\_key`)

REFERENCES `sakila\_snowflake`.`dim\_store` (`store\_key`)

ON DELETE CASCADE

ON UPDATE CASCADE,

CONSTRAINT `dim\_date\_fact\_rental\_fk`

FOREIGN KEY (`rental\_date\_key`)

REFERENCES `sakila\_snowflake`.`dim\_date` (`date\_Id`)

ON DELETE CASCADE

ON UPDATE CASCADE,

CONSTRAINT `dim\_date\_fact\_rental\_fk1`

FOREIGN KEY (`return\_date\_key`)

REFERENCES `sakila\_snowflake`.`dim\_date` (`date\_Id`)

ON DELETE CASCADE

ON UPDATE CASCADE)

ENGINE = InnoDB

DEFAULT CHARACTER SET = latin1;

CREATE INDEX `dim\_store\_fact\_rental\_fk` ON `sakila\_snowflake`.`fact\_rental` (`store\_key` ASC);

CREATE INDEX `dim\_staff\_fact\_rental\_fk` ON `sakila\_snowflake`.`fact\_rental` (`staff\_key` ASC);

CREATE INDEX `dim\_film\_fact\_rental\_fk` ON `sakila\_snowflake`.`fact\_rental` (`film\_key` ASC);

CREATE INDEX `dim\_customer\_fact\_rental\_fk` ON `sakila\_snowflake`.`fact\_rental` (`customer\_key` ASC);

SET SQL\_MODE=@OLD\_SQL\_MODE;

SET FOREIGN\_KEY\_CHECKS=@OLD\_FOREIGN\_KEY\_CHECKS;

SET UNIQUE\_CHECKS=@OLD\_UNIQUE\_CHECKS;

A screenshot of a social media post

Description automatically generated

**SakilaSnowflakeDW-DML:**

# The below query might take over 30 seconds to complete and you might get an "Error Code: 2013.

# Lost connection to MySQL server during query" error

# Please follow the instructions below:

# - In the application menu, select Edit > Preferences > SQL Editor.

# - Look for the MySQL Session section and increase the DBMS connection read time out value.

# - Save the settings, quit MySQL Workbench and reopen the connection.

-- -----------------------------------------------------

-- Write Fact table fact\_rental DML script here

-- -----------------------------------------------------

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\_returns,

count\_rentals, rental\_duration, dollar\_amount

)

(SELECT

rental\_id,

r.last\_update,

customer\_key,

staff\_key,

film\_key,

store\_key,

dt.date\_id AS rental\_date\_key,

dt2.date\_id AS return\_date\_key,

sum(case when dt.date is not null then 1 else 0 end) as rentalcount,

sum(case when dt2.date is not null then 1 else 0 end) as returncount,

sum(film\_rental\_duration) as rental\_duration,

(film.film\_rental\_rate \* datediff(r.return\_date, r.rental\_date)) as dollar\_amount

FROM sakila.rental r

join dim\_customer cu on cu.customer\_id = r.customer\_id

join dim\_staff staff on r.staff\_id = staff.staff\_id

join sakila.inventory inv on r.inventory\_id = inv.inventory\_id

join dim\_film film on inv.film\_id = film.film\_id

join dim\_store store on inv.store\_id = store.store\_id

left join dim\_date dt on DATE(dt.date) = DATE(r.rental\_date)

left join dim\_date dt2 on DATE(dt2.date) = DATE(r.return\_date)

GROUP BY

rental\_id,

r.last\_update,

customer\_key,

staff\_key,

film\_key,

store\_key,

dt.date\_id,

dt2.date\_id);

A screenshot of a social media post

Description automatically generated

**To Test my SQL queries result:**

SELECT count(\*) FROM sakila\_snowflake.fact\_rental;

1 row(s) returned

A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

I got 16,044 counts for my final result. The total number should be correct.

**I will attach the full SQL scripts on Canvas.**

**To Show I didn’t use any databases other than “sakila\_snowflake”:**

A screenshot of a computer

Description automatically generated

All I have is the sakila\_snowflake database. I join ten tables in Tableau to work on the questions from part c.

**I will attach my Tableau final dashboard product on Canvas.**

**The five charts I choose from the question 2 to answer are:**

**Part b, Part d, Part e, Part f, and Part h.**

A screenshot of a cell phone

Description automatically generated

My dashboard preview⬆️

The “**Select Store Id**” filter at the upper-right corner works for **all** the dashboard sections (question parts). I added another filter “**Select range**” for just **part e**.

I will attach my dashboard on Canvas.