21st July 2024

Report By: Vinita Jain, Shreya Biswas, Evan Cady, Lakshmialekhya Tatampudi

IS6480 – Data Warehousing

Group 9 – Summer 2024

DVD RENTAL DATA WAREHOUSE

Table of Contents

EXECUTIVE SUMMARY

For the purposes of this project, our team has decided to pursue the topic of DVD rentals. We worked with a dataset that contains detailed information from several years of records pertaining to customer rental activities, staff, information about each location, and related information.

When working to provide specific insights for our fictitious business, we focused our efforts on the most important functions that a business of this nature would benefit from. In designing our data warehouse, it was important that we do it in the most efficient way possible. We started with the main requirements and which business questions we wanted to answer. Then, we designed dimensional models that would allow us to more accurately implement the data warehouse to suit our needs listed in requirements and matrix. From there, we set up a Power BI workspace which was then shared with all team members. Within this, we created our dataflows and lakehouse.

When we completed construction of our data warehouse, we worked as a team to create visualizations that answer our business questions and requirements. When we finished our visualizations, we had a report that could be effectively turned into an app and then shared with an audience of our choice. Through completing this step, we were able to provide answers to our business questions in a way that someone who has little knowledge of this fictitious company would rapidly gain insight into its functions and performance.

In conclusion, we were able to provide clear answers to our questions relating to rental activities, details of rentals, business performance, and related items and present this in a way that the average person would be able to make sense of.

BACKGROUND INFORMATION

a. Services Provided

Our data warehouse is to be conceived with the knowledge that our primary dataset is representative of a chain of stores that rent DVDs. With that being said, this is a company that provides video rental services through physical media for a fee to consumers, who are then obligated to return such media to a store within a given timeframe or be billed for late fees.

b. Organizational Objectives

Our company wants to build a solution that will allow it to more easily make sense of data that is based on its day-to-day operations. It wants to better understand its overarching trends to better meet the needs of its consumers by understanding what their needs are. To help with this, the company also wants to examine performance of employees and see which employees are completing the most transactions to presumably offer programs that motivate employees to perform better and to allocate talent more efficiently. The company will also want to take a look at performance of each location to see where more resources are needed or where there is an opportunity for expansion. Late returns are a key factor impacting profitability, and understanding the costs associated with customers returning items late is important. Among others, the company also wants to make sense of customer spending to possibly offer a rewards program that might boost revenue.

c. Why a Data Warehouse is Useful?

Creating a data warehouse will help the company organize the vast amount of data that it has collected from all areas of its business. This will make it easier and more efficient to analyze the data that has been collected. This is especially true based on the historical data that is in the database and the large amount of data that we have. We also can have a more central data system that pulls from several tables and includes only the items that we find necessary to analyze when seeking to provide answers to our business questions.

REQUIREMENTS

SUMMARY OF MAIN REQUIREMENTS

Requirement Name	Short Description	Status
Film Rental		
Films under each Category	Which film category has highest number of Films	Completed
Most rented films	Which films have been most rented and least rented	Completed
Track stock levels	Check the availability of DVD titles and ensure that popular titles are always in stock.	Incomplete
Language analysis	Which language films are rented the most at every store	Incomplete
Customer Analysis		
Customers Preference	What film category are the customers most interested and least interested in?	Completed
Customer demographics	What is the average age of customer interests according to film genre?	Incomplete
Loyalty programs	Analyze customer spending over time. This information can be used to give some special offers to customers frequently renting	Incomplete
Management		
Store rental	What is the average number of rentals at each store per year?	Incomplete
Staff Analysis	Which staff member processed the highest total payment amount in each store in each year?	Incomplete
Financial Analysis		
Income Analysis	What is the trend of the income that we have made over the time period?	Completed
Gross Revenue	Which film category has made highest Income and lowest Income?	Completed

Sales Peak analysis	What are the sales peak hours and days at particular store	Incomplete
Returns	Analyze late returns. This insight can be used to calculate revenue impact.	Incomplete

BUS MATRIX

	Dimensions								
Business processes	**DimFilm**	**DimStore**	**DimLanguage**	**DimCustomer**	**DimStaff**	**DimInventory**	**DimDate**	**DimFilmCategory**	**DimCustomerAge**
Film Rental									
Films under each Category	x								
Most rented films	x			x			x		
Track stock levels	x	x				x	x		
Language analysis	x	x	x				x		
Customer Analysis									
Customers Preference	x			x			x	x	
Customer demographics	x	x					x	x	x
Loyalty programs	x	x		x			x		
Management									
Store rental	x	x					x		
Staff Analysis	x	x			x		x		
Financial Analysis									

Income Analysis	x			x			x		
Gross Revenue	x			x			x	x	
Sales Peak analysis	x	x		x			x		
Returns	x	x					x		

DIMENSIONAL MODEL

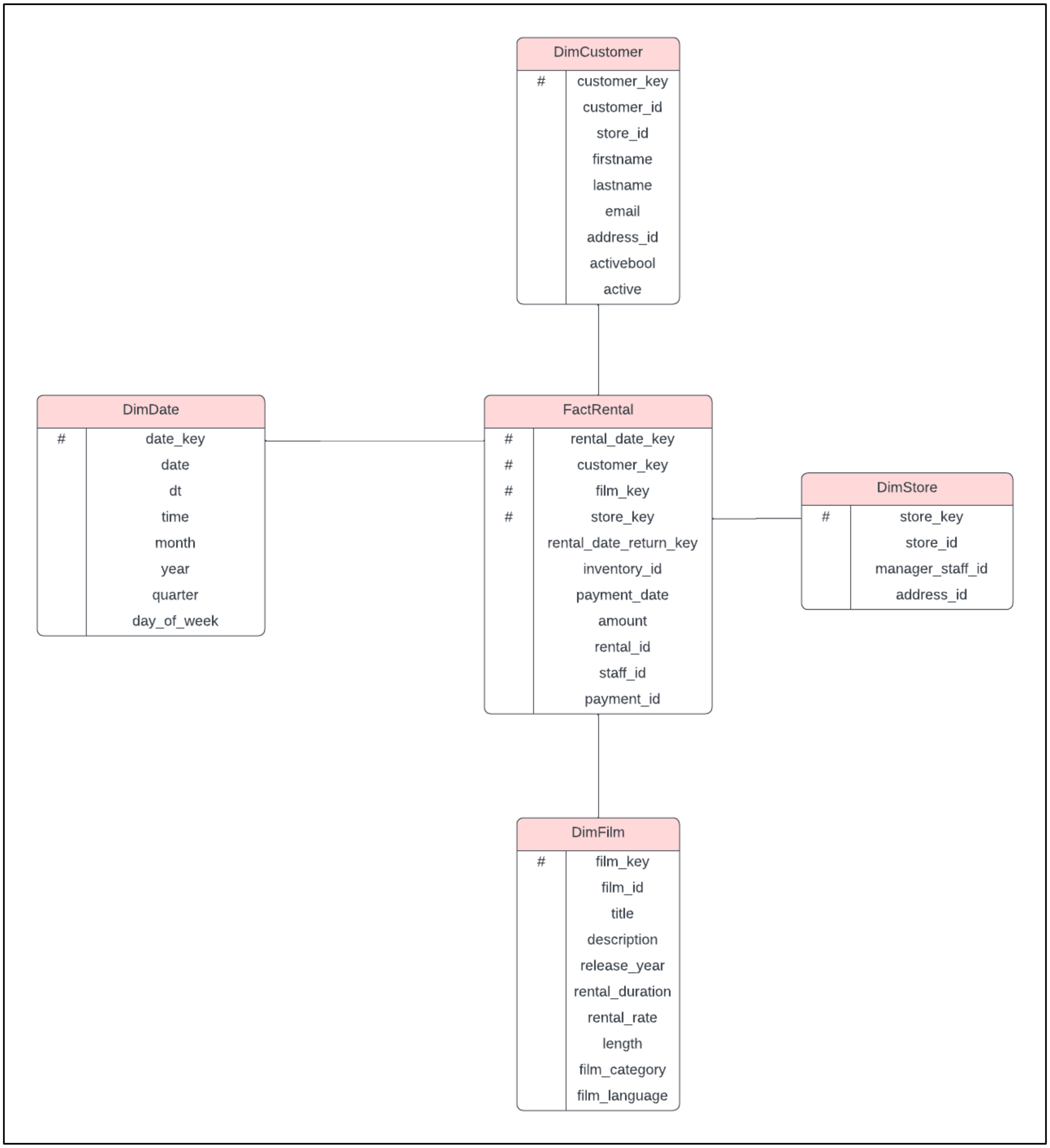


Figure 1: This model is structured to support querying and reporting on various aspects of the DVD rental business, such as customer activities, rental patterns, and store performance.

DIMENSIONAL MODEL DESCRIPTION

Fact Table

FactRental:

a. rental_date_key: Foreign key to DimDate.
b. customer_key: Foreign key to DimCustomer.
c. film_key: Foreign key to DimFilm.
d. store_key: Foreign key to DimStore.
e. rental_date_return_key: Another foreign key to DimDate for the return date.
f. inventory_id: Identifier for the inventory.
g. payment_date: Date of payment.
h. amount: Amount paid.
i. rental_id: Identifier for the rental.
j. staff_id: Identifier for the staff.
k. payment_id: Identifier for the payment.

Dimension Tables

DimDate:

a. date_key: Primary key for the date dimension.
b. date: Full date value.
c. dt: Date and time.
d. time: Time portion of the date.
e. month: Month portion of the date.
f. year: Year portion of the date.
g. quarter: Quarter portion of the year.
h. day_of_week: Day of the week.

DimCustomer:

a. customer_key: Primary key for the customer dimension.
b. customer_id: Identifier for the customer.
c. store_id: Identifier for the store.
d. firstname: First name of the customer.
e. lastname: Last name of the customer.
f. email: Email address of the customer.
g. address_id: Identifier for the address.
h. activebool: Boolean indicating if the customer is active.
i. active: Another indicator for the active status of the customer.

DimStore:

a. store_key: Primary key for the store dimension.
b. store_id: Identifier for the store.
c. manager_staff_id: Identifier for the manager staff.
d. address_id: Identifier for the address.

DimFilm:

a. film_key: Primary key for the film dimension.
b. film_id: Identifier for the film.
c. title: Title of the film.
d. description: Description of the film.
e. release_year: Year the film was released.
f. rental_duration: Duration for which the film can be rented.
g. rental_rate: Rate at which the film can be rented.
h. length: Length of the film.
i. film_category: Category of the film.
j. film_language: Language of the film.

Relationships

a. fact_most_rented is connected to DimDate via rental_date_key and rental_date_return_key.
b. fact_most_rented is connected to DimCustomer via customer_key.
c. fact_most_rented is connected to DimFilm via film_key.
d. fact_most_rented is connected to DimStore via store_key and staff_key.

ARCHITECTURE

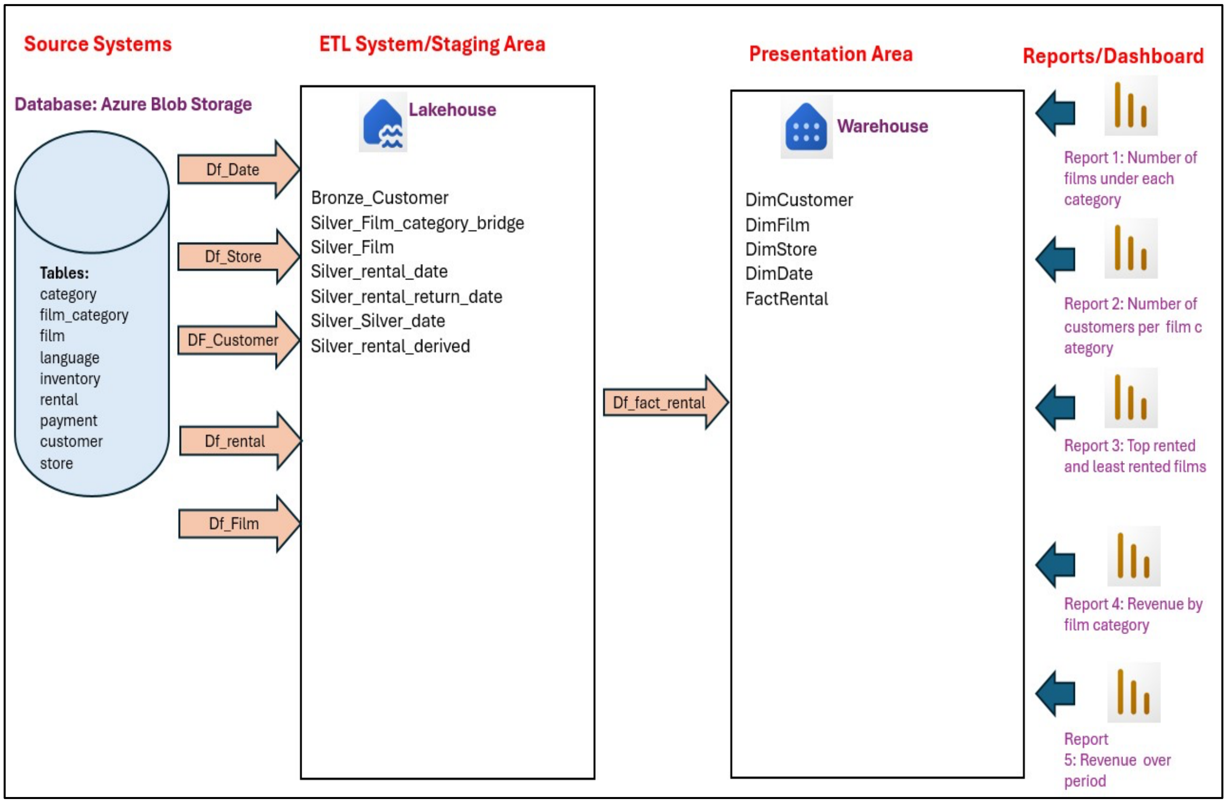
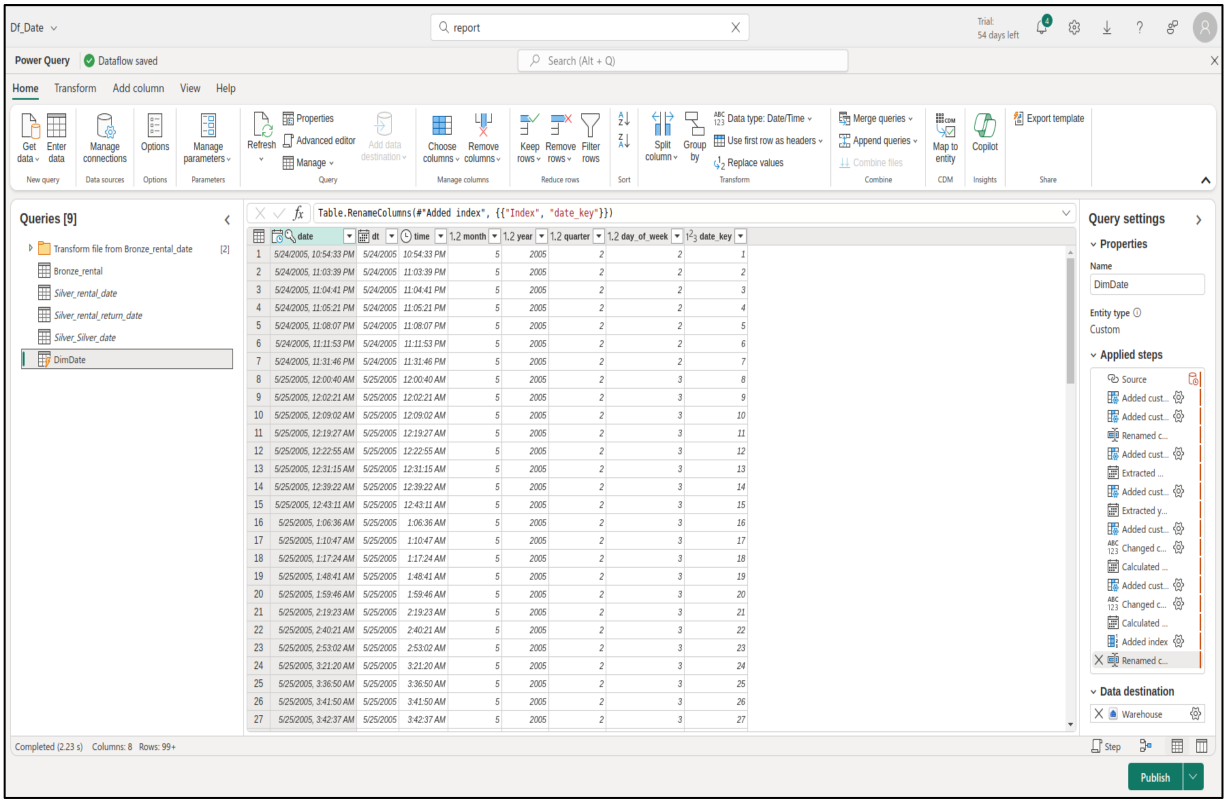


Figure 2: Data flows from the Azure Blob Storage (Source Systems) into the Lakehouse (ETL System/Staging Area), where various transformations and stages are managed. From there, the data moves to the Warehouse (Presentation Area) and is used for generating report

DATA PIPELINES



Df_Date

Queries [9]

- Transform file from Bronze_rental_date [2]
- Bronze_rental
- *Silver_rental_date*
- *Silver_rental_return_date*
- *Silver_Silver_date*
- DimDate

```
Table.RenameColumns(#"Added index", {{"Index", "date_key"}})
```

	date	dt	time	month	year	quarter	day_of_week	date_key
1	5/24/2005, 10:54:33 PM	5/24/2005	10:54:33 PM	5	2005	2	2	1
2	5/24/2005, 11:03:39 PM	5/24/2005	11:03:39 PM	5	2005	2	2	2
3	5/24/2005, 11:04:41 PM	5/24/2005	11:04:41 PM	5	2005	2	2	3
4	5/24/2005, 11:05:21 PM	5/24/2005	11:05:21 PM	5	2005	2	2	4
5	5/24/2005, 11:08:07 PM	5/24/2005	11:08:07 PM	5	2005	2	2	5
6	5/24/2005, 11:11:53 PM	5/24/2005	11:11:53 PM	5	2005	2	2	6
7	5/24/2005, 11:31:46 PM	5/24/2005	11:31:46 PM	5	2005	2	2	7
8	5/25/2005, 12:00:40 AM	5/25/2005	12:00:40 AM	5	2005	2	3	8
9	5/25/2005, 12:02:21 AM	5/25/2005	12:02:21 AM	5	2005	2	3	9
10	5/25/2005, 12:09:02 AM	5/25/2005	12:09:02 AM	5	2005	2	3	10
11	5/25/2005, 12:19:27 AM	5/25/2005	12:19:27 AM	5	2005	2	3	11
12	5/25/2005, 12:22:55 AM	5/25/2005	12:22:55 AM	5	2005	2	3	12
13	5/25/2005, 12:31:15 AM	5/25/2005	12:31:15 AM	5	2005	2	3	13
14	5/25/2005, 12:39:22 AM	5/25/2005	12:39:22 AM	5	2005	2	3	14
15	5/25/2005, 12:43:11 AM	5/25/2005	12:43:11 AM	5	2005	2	3	15
16	5/25/2005, 1:06:36 AM	5/25/2005	1:06:36 AM	5	2005	2	3	16
17	5/25/2005, 1:10:47 AM	5/25/2005	1:10:47 AM	5	2005	2	3	17
18	5/25/2005, 1:17:24 AM	5/25/2005	1:17:24 AM	5	2005	2	3	18
19	5/25/2005, 1:48:41 AM	5/25/2005	1:48:41 AM	5	2005	2	3	19
20	5/25/2005, 1:59:46 AM	5/25/2005	1:59:46 AM	5	2005	2	3	20
21	5/25/2005, 2:19:23 AM	5/25/2005	2:19:23 AM	5	2005	2	3	21
22	5/25/2005, 2:40:21 AM	5/25/2005	2:40:21 AM	5	2005	2	3	22
23	5/25/2005, 2:53:02 AM	5/25/2005	2:53:02 AM	5	2005	2	3	23
24	5/25/2005, 3:21:20 AM	5/25/2005	3:21:20 AM	5	2005	2	3	24
25	5/25/2005, 3:36:50 AM	5/25/2005	3:36:50 AM	5	2005	2	3	25
26	5/25/2005, 3:41:50 AM	5/25/2005	3:41:50 AM	5	2005	2	3	26
27	5/25/2005, 3:42:37 AM	5/25/2005	3:42:37 AM	5	2005	2	3	27

Query settings

Properties

Name: DimDate

Entity type: Custom

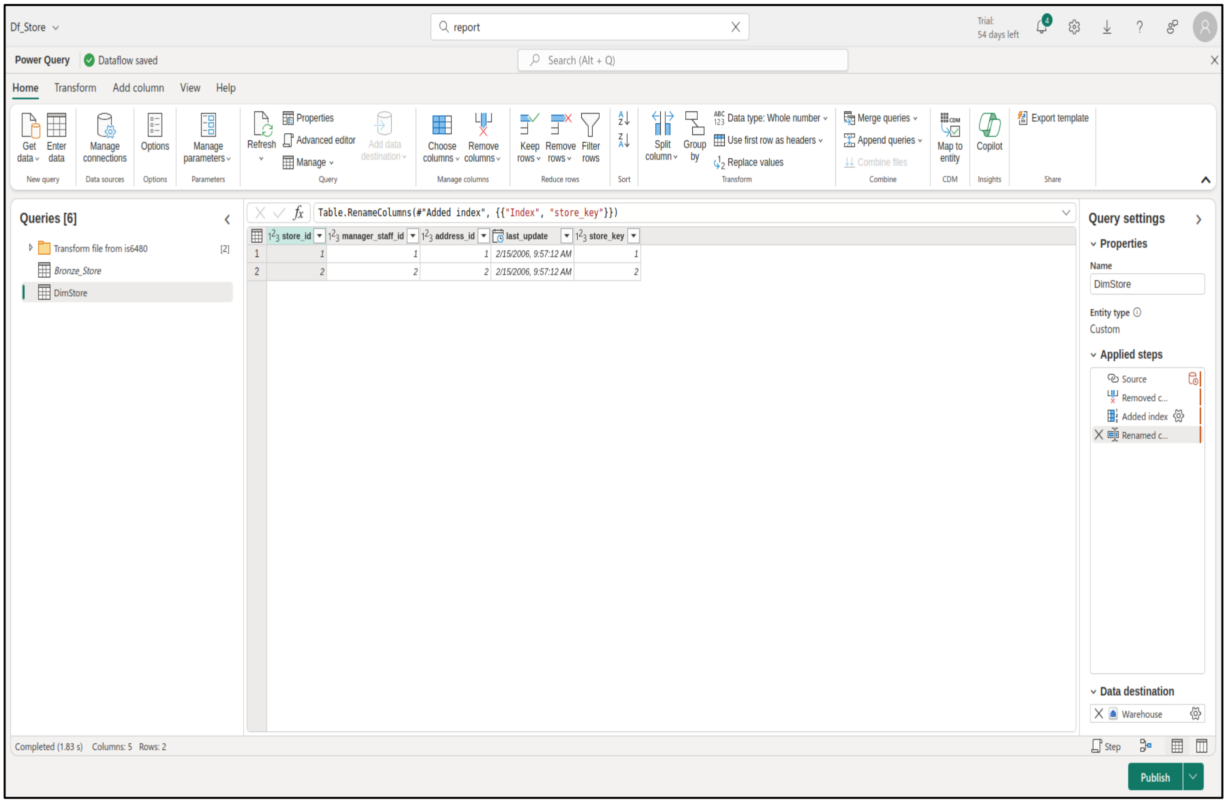
Applied steps

- Source
- Added cust...
- Added cust...
- Renamed c...
- Added cust...
- Extracted ...
- Added cust...
- Extracted y...
- Added cust...
- Changed c...
- Calculated ...
- Added cust...
- Changed c...
- Calculated ...
- Added index
- Renamed c...

Data destination: Warehouse

Completed (2.23 s) Columns: 8 Rows: 99+

Figure 3: Df_Date : rental table is loaded from azure blob. It is referenced as Silver_rental_date and Silver_rental_return_date. Both the referenced tables are merged into Silver_Silver_date and duplicates are removed. Then DimDate is created from Silver_Silver_Date



```
Table.RenameColumns(#"Added index", {{"Index", "store_key"}})
```

	store_id	manager_staff_id	address_id	last_update	store_key
1	1	1	1	2/15/2006, 9:57:12 AM	1
2	2	2	2	2/15/2006, 9:57:12 AM	2

Completed (1.83 s) Columns: 5 Rows: 2

Query settings

Properties

Name: DimStore

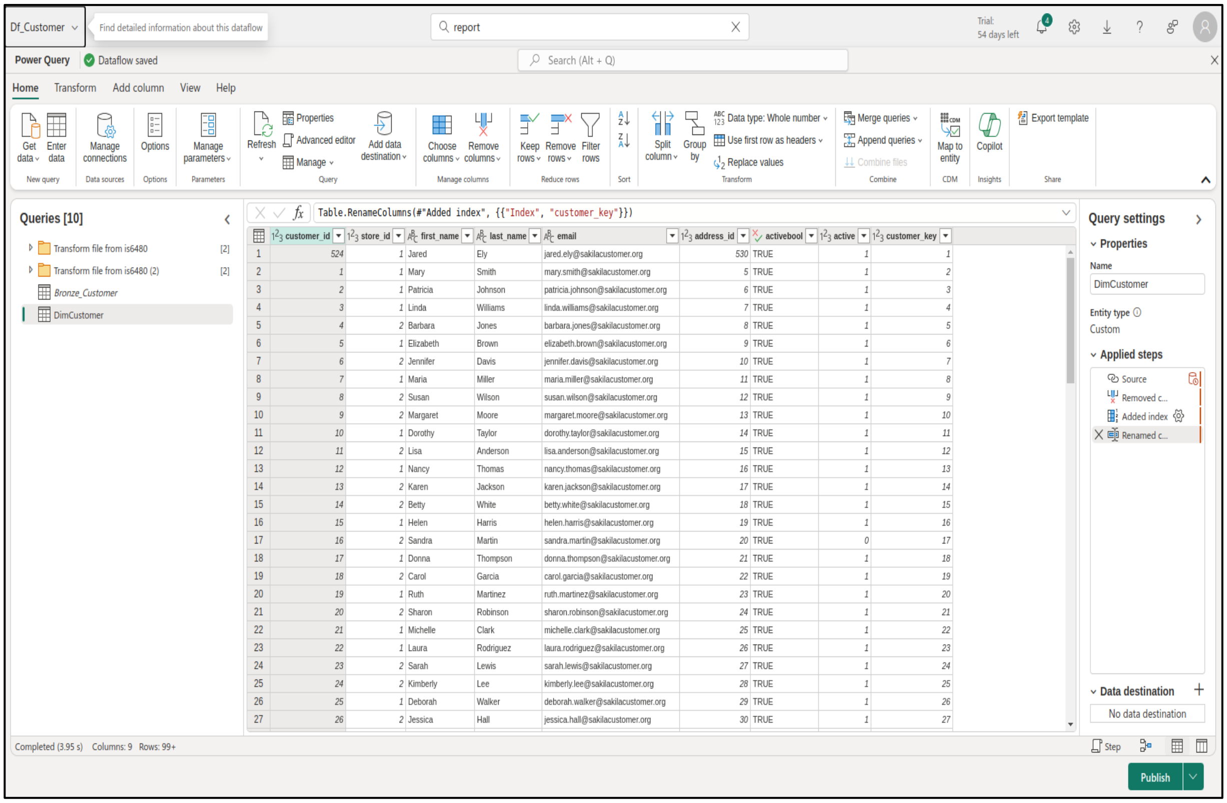
Entity type: Custom

Applied steps

- Source
- Removed c...
- Added index
- Renamed c...

Data destination: Warehouse

Figure 4: Df_Store: store table is loaded from azure blob storage and referenced to create DimStore.



```
Table.RenameColumns(#"Added index", {{"Index", "customer_key"}})
```

	customer_id	store_id	first_name	last_name	email	address_id	activebool	active	customer_key
1	524	1	Jared	Ely	jared.ely@sakilacustomer.org	530	TRUE	1	1
2	1	1	Mary	Smith	mary.smith@sakilacustomer.org	5	TRUE	1	2
3	2	1	Patricia	Johnson	patricia.johnson@sakilacustomer.org	6	TRUE	1	3
4	3	1	Linda	Williams	linda.williams@sakilacustomer.org	7	TRUE	1	4
5	4	2	Barbara	Jones	barbara.jones@sakilacustomer.org	8	TRUE	1	5
6	5	1	Elizabeth	Brown	elizabeth.brown@sakilacustomer.org	9	TRUE	1	6
7	6	2	Jennifer	Davis	jennifer.davis@sakilacustomer.org	10	TRUE	1	7
8	7	1	Maria	Miller	maria.miller@sakilacustomer.org	11	TRUE	1	8
9	8	2	Susan	Wilson	susan.wilson@sakilacustomer.org	12	TRUE	1	9
10	9	2	Margaret	Moore	margaret.moore@sakilacustomer.org	13	TRUE	1	10
11	10	1	Dorothy	Taylor	dorothy.taylor@sakilacustomer.org	14	TRUE	1	11
12	11	2	Lisa	Anderson	lisa.anderson@sakilacustomer.org	15	TRUE	1	12
13	12	1	Nancy	Thomas	nancy.thomas@sakilacustomer.org	16	TRUE	1	13
14	13	2	Karen	Jackson	karen.jackson@sakilacustomer.org	17	TRUE	1	14
15	14	2	Betty	White	betty.white@sakilacustomer.org	18	TRUE	1	15
16	15	1	Helen	Harris	helen.harris@sakilacustomer.org	19	TRUE	1	16
17	16	2	Sandra	Martin	sandra.martin@sakilacustomer.org	20	TRUE	0	17
18	17	1	Donna	Thompson	donna.thompson@sakilacustomer.org	21	TRUE	1	18
19	18	2	Carol	Garcia	carol.garcia@sakilacustomer.org	22	TRUE	1	19
20	19	1	Ruth	Martinez	ruth.martinez@sakilacustomer.org	23	TRUE	1	20
21	20	2	Sharon	Robinson	sharon.robinson@sakilacustomer.org	24	TRUE	1	21
22	21	1	Michelle	Clark	michelle.clark@sakilacustomer.org	25	TRUE	1	22
23	22	1	Laura	Rodriguez	laura.rodriguez@sakilacustomer.org	26	TRUE	1	23
24	23	2	Sarah	Lewis	sarah.lewis@sakilacustomer.org	27	TRUE	1	24
25	24	2	Kimberly	Lee	kimberly.lee@sakilacustomer.org	28	TRUE	1	25
26	25	1	Deborah	Walker	deborah.walker@sakilacustomer.org	29	TRUE	1	26
27	26	2	Jessica	Hall	jessica.hall@sakilacustomer.org	30	TRUE	1	27

Name: DimCustomer

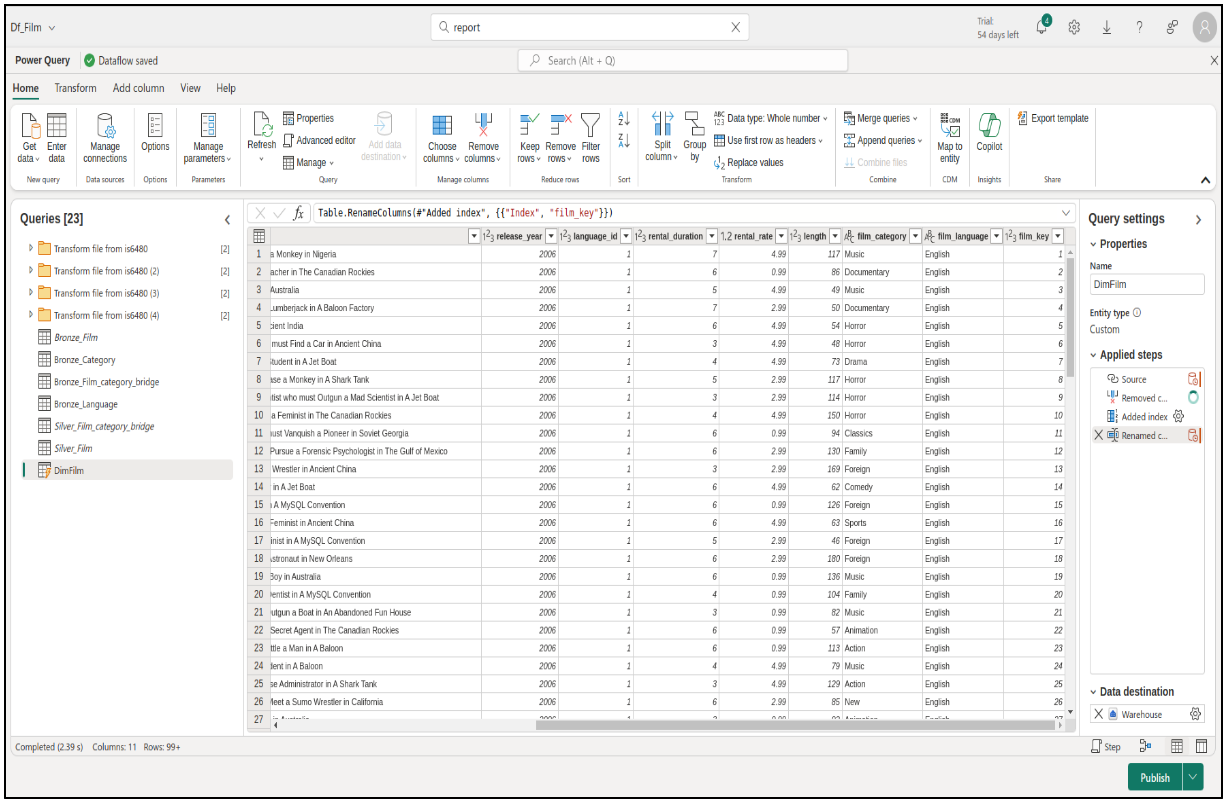
Entity type: Custom

Applied steps:
- Source
- Removed c...
- Added index
- Renamed c...

Data destination: No data destination

Completed (3.95 s) Columns: 9 Rows: 99+

Figure 5: Df_Customer: customer table is loaded from azure blob storage and referenced to create DimCustomer.



Queries [23]

- Transform file from is6480 [2]
- Transform file from is6480 (2) [2]
- Transform file from is6480 (3) [2]
- Transform file from is6480 (4) [2]
- Bronze_Film
- Bronze_Category
- Bronze_Film_category_bridge
- Bronze_Language
- Silver_Film_category_bridge
- Silver_Film
- DimFilm

```
Table.RenameColumns(#"Added index", {{"Index", "film_key"}})
```

		release_year	language_id	rental_duration	rental_rate	length	film_category	film_language	film_key
1	a Monkey in Nigeria	2006	1	7	4.99	117	Music	English	1
2	acher in The Canadian Rockies	2006	1	6	0.99	86	Documentary	English	2
3	Australia	2006	1	5	4.99	49	Music	English	3
4	.umberjack in A Baloon Factory	2006	1	7	2.99	50	Documentary	English	4
5	:ient India	2006	1	6	4.99	54	Horror	English	5
6	must Find a Car in Ancient China	2006	1	3	4.99	48	Horror	English	6
7	;tudent in A Jet Boat	2006	1	4	4.99	73	Drama	English	7
8	ase a Monkey in A Shark Tank	2006	1	5	2.99	117	Horror	English	8
9	itist who must Outgun a Mad Scientist in A Jet Boat	2006	1	3	2.99	114	Horror	English	9
10	a Feminist in The Canadian Rockies	2006	1	4	4.99	150	Horror	English	10
11	ıust Vanquish a Pioneer in Soviet Georgia	2006	1	6	0.99	94	Classics	English	11
12	Pursue a Forensic Psychologist in The Gulf of Mexico	2006	1	6	2.99	130	Family	English	12
13	Wrestler in Ancient China	2006	1	3	2.99	169	Foreign	English	13
14	in A Jet Boat	2006	1	6	4.99	62	Comedy	English	14
15	ı A MySQL Convention	2006	1	6	0.99	126	Foreign	English	15
16	Feminist in Ancient China	2006	1	6	4.99	63	Sports	English	16
17	inist in A MySQL Convention	2006	1	5	2.99	46	Foreign	English	17
18	stronaut in New Orleans	2006	1	6	2.99	180	Foreign	English	18
19	Boy in Australia	2006	1	6	0.99	136	Music	English	19
20	)entist in A MySQL Convention	2006	1	4	0.99	104	Family	English	20
21	utgun a Boat in An Abandoned Fun House	2006	1	3	0.99	82	Music	English	21
22	Secret Agent in The Canadian Rockies	2006	1	6	0.99	57	Animation	English	22
23	ttle a Man in A Baloon	2006	1	6	0.99	113	Action	English	23
24	dent in A Baloon	2006	1	4	4.99	79	Music	English	24
25	se Administrator in A Shark Tank	2006	1	3	4.99	129	Action	English	25
26	Meet a Sumo Wrestler in California	2006	1	6	2.99	85	New	English	26
27	in Australia	2006	1	3	0.99	92	Animation	English	27

Completed (2.39 s) Columns: 11 Rows: 99+

Query settings

Properties

Name: DimFilm

Entity type: Custom

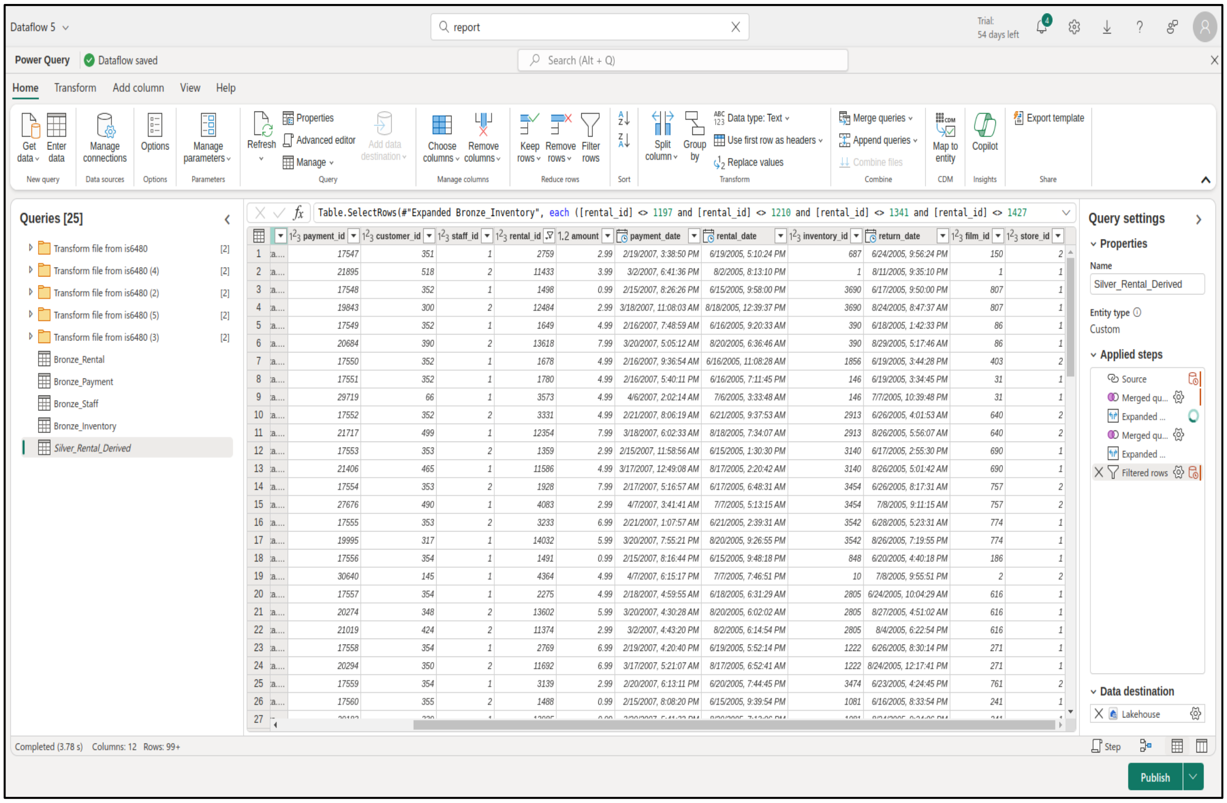
Applied steps

- Source
- Removed c...
- Added index
- Renamed c...

Data destination

- Warehouse

Figure 6: Df_Film: film, category, film_category and language tables are loaded from azure blob storage. Film_category and category tables are left joined on category_id to get category name in film_category table(Silver_film_category_bridge).Then film and Silver_film_category_bridge tables are left joined on film_id to get film category name in film table(Silver_Film). DimFilm is then created from Silver_Film table.



```
Table.SelectRows(#"Expanded Bronze_Inventory", each ([rental_id] <> 1197 and [rental_id] <> 1210 and [rental_id] <> 1341 and [rental_id] <> 1427
```

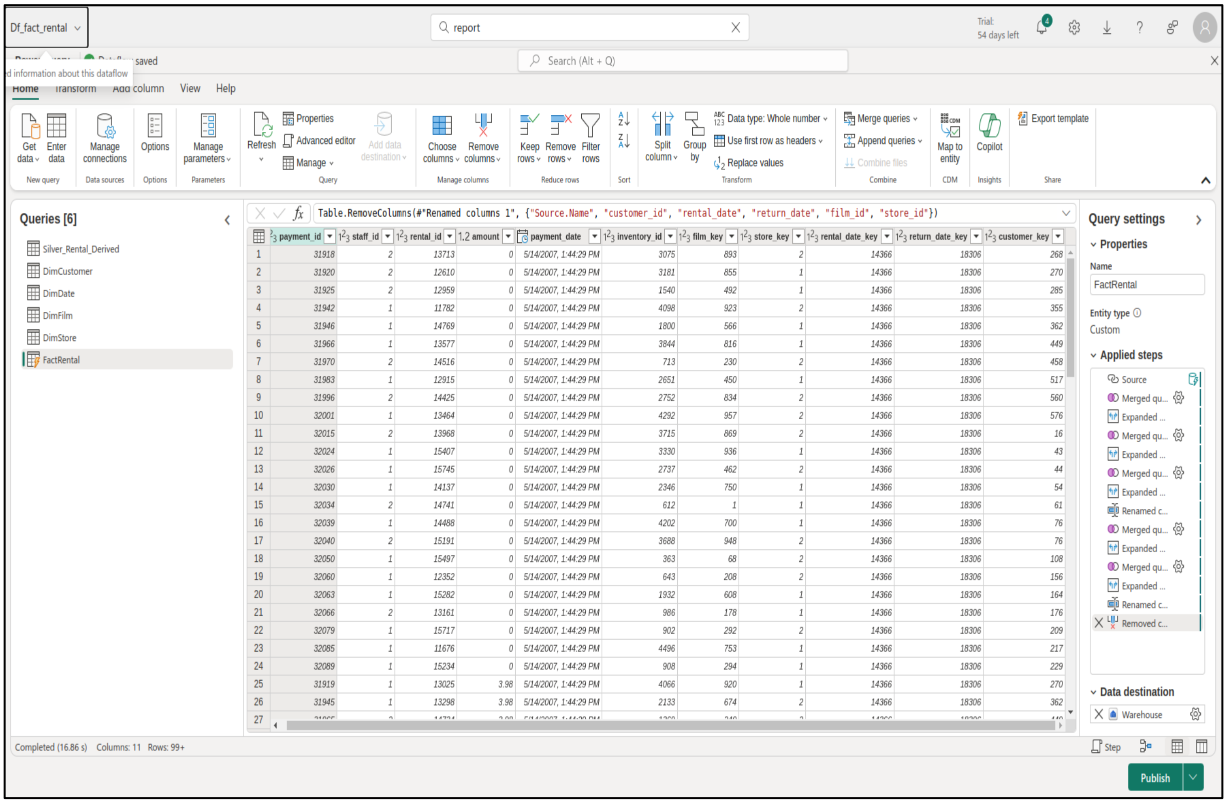
		payment_id	customer_id	staff_id	rental_id	amount	payment_date	rental_date	inventory_id	return_date	film_id	store_id
1	:a....	17547	351	1	2759	2.99	2/19/2007, 3:38:50 PM	6/19/2005, 5:10:24 PM	687	6/24/2005, 9:56:24 PM	150	2
2	:a....	21895	518	2	11433	3.99	3/2/2007, 6:41:36 PM	8/2/2005, 8:13:10 PM	1	8/11/2005, 9:35:10 PM	1	1
3	:a....	17548	352	1	1498	0.99	2/15/2007, 8:26:26 PM	6/15/2005, 9:58:00 PM	3690	6/17/2005, 9:50:00 PM	807	1
4	:a....	19843	300	2	12484	2.99	3/18/2007, 11:08:03 AM	8/18/2005, 12:39:37 PM	3690	8/24/2005, 8:47:37 AM	807	1
5	:a....	17549	352	1	1649	4.99	2/16/2007, 7:48:59 AM	6/16/2005, 9:20:33 AM	390	6/18/2005, 1:42:33 PM	86	1
6	:a....	20684	390	2	13618	7.99	3/20/2007, 5:05:12 AM	8/20/2005, 6:36:46 AM	390	8/29/2005, 5:17:46 AM	86	1
7	:a....	17550	352	1	1678	4.99	2/16/2007, 9:36:54 AM	6/16/2005, 11:08:28 AM	1856	6/19/2005, 3:44:28 PM	403	2
8	:a....	17551	352	1	1780	4.99	2/16/2007, 5:40:11 PM	6/16/2005, 7:11:45 PM	146	6/19/2005, 3:34:45 PM	31	1
9	:a....	29719	66	1	3573	4.99	4/6/2007, 2:02:14 AM	7/6/2005, 3:33:48 AM	146	7/7/2005, 10:39:48 PM	31	1
10	:a....	17552	352	2	3331	4.99	2/21/2007, 8:06:19 AM	6/21/2005, 9:37:53 AM	2913	6/26/2005, 4:01:53 AM	640	2
11	:a....	21717	499	1	12354	7.99	3/18/2007, 6:02:33 AM	8/18/2005, 7:34:07 AM	2913	8/26/2005, 5:56:07 AM	640	2
12	:a....	17553	353	2	1359	2.99	2/15/2007, 11:58:56 AM	6/15/2005, 1:30:30 PM	3140	6/17/2005, 2:55:30 PM	690	1
13	:a....	21406	465	1	11586	4.99	3/17/2007, 12:49:08 AM	8/17/2005, 2:20:42 AM	3140	8/26/2005, 5:01:42 AM	690	1
14	:a....	17554	353	2	1928	7.99	2/17/2007, 5:16:57 AM	6/17/2005, 6:48:31 AM	3454	6/26/2005, 8:17:31 AM	757	2
15	:a....	27676	490	1	4083	2.99	4/7/2007, 3:41:41 AM	7/7/2005, 5:13:15 AM	3454	7/8/2005, 9:11:15 AM	757	2
16	:a....	17555	353	2	3233	6.99	2/21/2007, 1:07:57 AM	6/21/2005, 2:39:31 AM	3542	6/28/2005, 5:23:31 AM	774	1
17	:a....	19995	317	1	14032	5.99	3/20/2007, 7:55:21 PM	8/20/2005, 9:26:55 PM	3542	8/26/2005, 7:19:55 PM	774	1
18	:a....	17556	354	1	1491	0.99	2/15/2007, 8:16:44 PM	6/15/2005, 9:48:18 PM	848	6/20/2005, 4:40:18 PM	186	1
19	:a....	30640	145	1	4364	4.99	4/7/2007, 6:15:17 PM	7/7/2005, 7:46:51 PM	10	7/8/2005, 9:55:51 PM	2	2
20	:a....	17557	354	1	2275	4.99	2/18/2007, 4:59:55 AM	6/18/2005, 6:31:29 AM	2805	6/24/2005, 10:04:29 AM	616	1
21	:a....	20274	348	2	13602	5.99	3/20/2007, 4:30:28 AM	8/20/2005, 6:02:02 AM	2805	8/27/2005, 4:51:02 AM	616	1
22	:a....	21019	424	2	11374	2.99	3/2/2007, 4:43:20 PM	8/2/2005, 6:14:54 PM	2805	8/4/2005, 6:22:54 PM	616	1
23	:a....	17558	354	1	2769	6.99	2/19/2007, 4:20:40 PM	6/19/2005, 5:52:14 PM	1222	6/26/2005, 8:30:14 PM	271	1
24	:a....	20294	350	2	11692	6.99	3/17/2007, 5:21:07 AM	8/17/2005, 6:52:41 AM	1222	8/24/2005, 12:17:41 PM	271	1
25	:a....	17559	354	1	3139	2.99	2/20/2007, 6:13:11 PM	6/20/2005, 7:44:45 PM	3474	6/23/2005, 4:24:45 PM	761	2
26	:a....	17560	355	2	1488	0.99	2/15/2007, 8:08:20 PM	6/15/2005, 9:39:54 PM	1081	6/16/2005, 8:33:54 PM	241	1

Name: Silver_Rental_Derived

Entity type: Custom

Completed (3.78 s)   Columns: 12   Rows: 99+

Figure 7: Df_rental_derived: Silver_rental_derived is created from rental, payment, staff and inventory tables.



```
Table.RemoveColumns(#"Renamed columns 1", {"Source.Name", "customer_id", "rental_date", "return_date", "film_id", "store_id"})
```

	payment_id	staff_id	rental_id	amount	payment_date	inventory_id	film_key	store_key	rental_date_key	return_date_key	customer_key
1	31918	2	13713	0	5/14/2007, 1:44:29 PM	3075	893	2	14366	18306	268
2	31920	2	12610	0	5/14/2007, 1:44:29 PM	3181	855	1	14366	18306	270
3	31925	2	12959	0	5/14/2007, 1:44:29 PM	1540	492	1	14366	18306	285
4	31942	1	11782	0	5/14/2007, 1:44:29 PM	4098	923	2	14366	18306	355
5	31946	1	14769	0	5/14/2007, 1:44:29 PM	1800	566	1	14366	18306	362
6	31966	1	13577	0	5/14/2007, 1:44:29 PM	3844	816	1	14366	18306	449
7	31970	2	14516	0	5/14/2007, 1:44:29 PM	713	230	2	14366	18306	458
8	31983	1	12915	0	5/14/2007, 1:44:29 PM	2651	450	1	14366	18306	517
9	31996	2	14425	0	5/14/2007, 1:44:29 PM	2752	834	2	14366	18306	560
10	32001	1	13464	0	5/14/2007, 1:44:29 PM	4292	957	2	14366	18306	576
11	32015	2	13968	0	5/14/2007, 1:44:29 PM	3715	869	2	14366	18306	16
12	32024	1	15407	0	5/14/2007, 1:44:29 PM	3330	936	1	14366	18306	43
13	32026	1	15745	0	5/14/2007, 1:44:29 PM	2737	462	2	14366	18306	44
14	32030	1	14137	0	5/14/2007, 1:44:29 PM	2346	750	1	14366	18306	54
15	32034	2	14741	0	5/14/2007, 1:44:29 PM	612	1	1	14366	18306	61
16	32039	1	14488	0	5/14/2007, 1:44:29 PM	4202	700	1	14366	18306	76
17	32040	2	15191	0	5/14/2007, 1:44:29 PM	3688	948	2	14366	18306	76
18	32050	1	15497	0	5/14/2007, 1:44:29 PM	363	68	2	14366	18306	108
19	32060	1	12352	0	5/14/2007, 1:44:29 PM	643	208	2	14366	18306	156
20	32063	1	15282	0	5/14/2007, 1:44:29 PM	1932	608	1	14366	18306	164
21	32066	2	13161	0	5/14/2007, 1:44:29 PM	986	178	1	14366	18306	176
22	32079	1	15717	0	5/14/2007, 1:44:29 PM	902	292	2	14366	18306	209
23	32085	1	11676	0	5/14/2007, 1:44:29 PM	4496	753	1	14366	18306	217
24	32089	1	15234	0	5/14/2007, 1:44:29 PM	908	294	1	14366	18306	229
25	31919	1	13025	3.98	5/14/2007, 1:44:29 PM	4066	920	1	14366	18306	270
26	31945	1	13298	3.98	5/14/2007, 1:44:29 PM	2133	674	2	14366	18306	362
27	31965	2	14724	3.98	5/14/2007, 1:44:29 PM	1369	240	2	14366	18306	449

Figure 8: Df_fact_rental: FactRental table is created from the Silver_Rental_Derived and all the dimension tables created.

DASHBOARD AND REPORTS

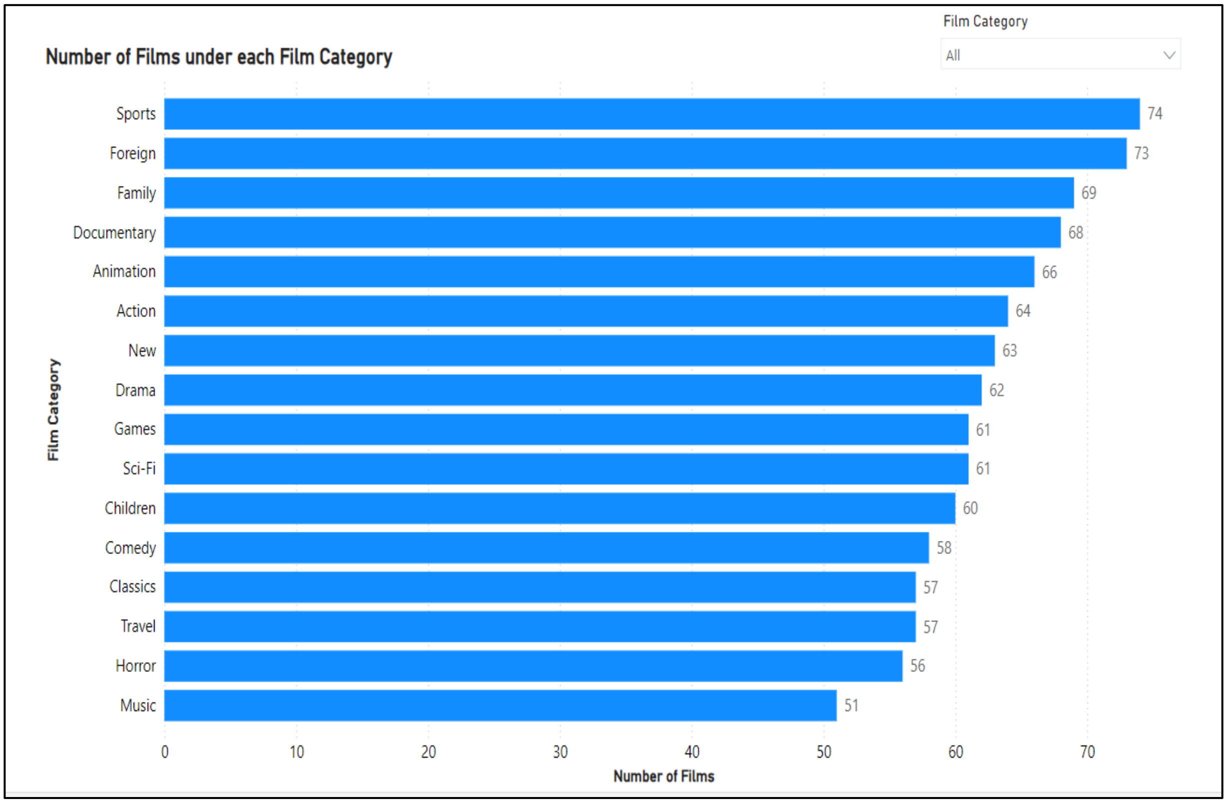
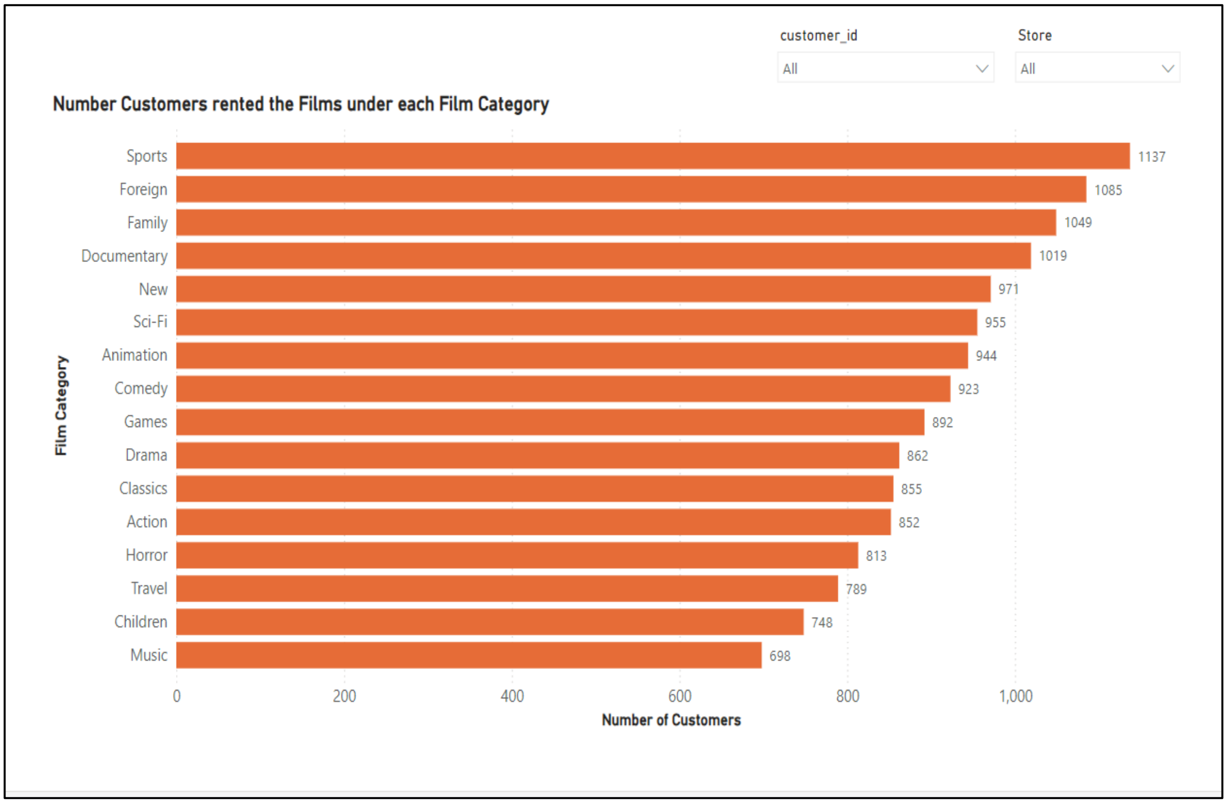


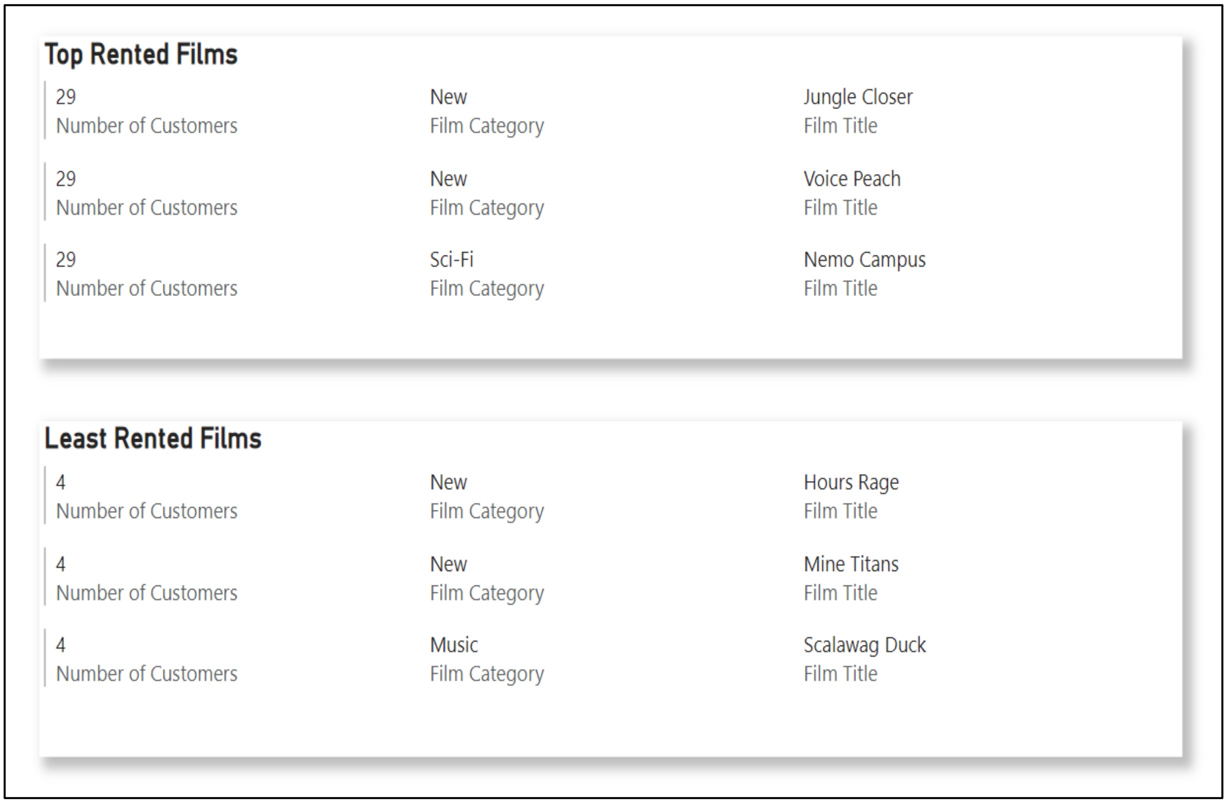
Figure 9: This Visual shows the number of films for each film Category. The sports Category has the highest number of Films.



customer_id
All

Store
All

Figure 10: This Visual shows Number of Customers that have rented the films by Film Category. The highest rented Film category is Sports.



Top Rented Films

29 Number of Customers	New Film Category	Jungle Closer Film Title
29 Number of Customers	New Film Category	Voice Peach Film Title
29 Number of Customers	Sci-Fi Film Category	Nemo Campus Film Title

Least Rented Films

4 Number of Customers	New Film Category	Hours Rage Film Title
4 Number of Customers	New Film Category	Mine Titans Film Title
4 Number of Customers	Music Film Category	Scalawag Duck Film Title

Figure 11: This visual shows the highest and least rented films with their Film Categories and Titles.

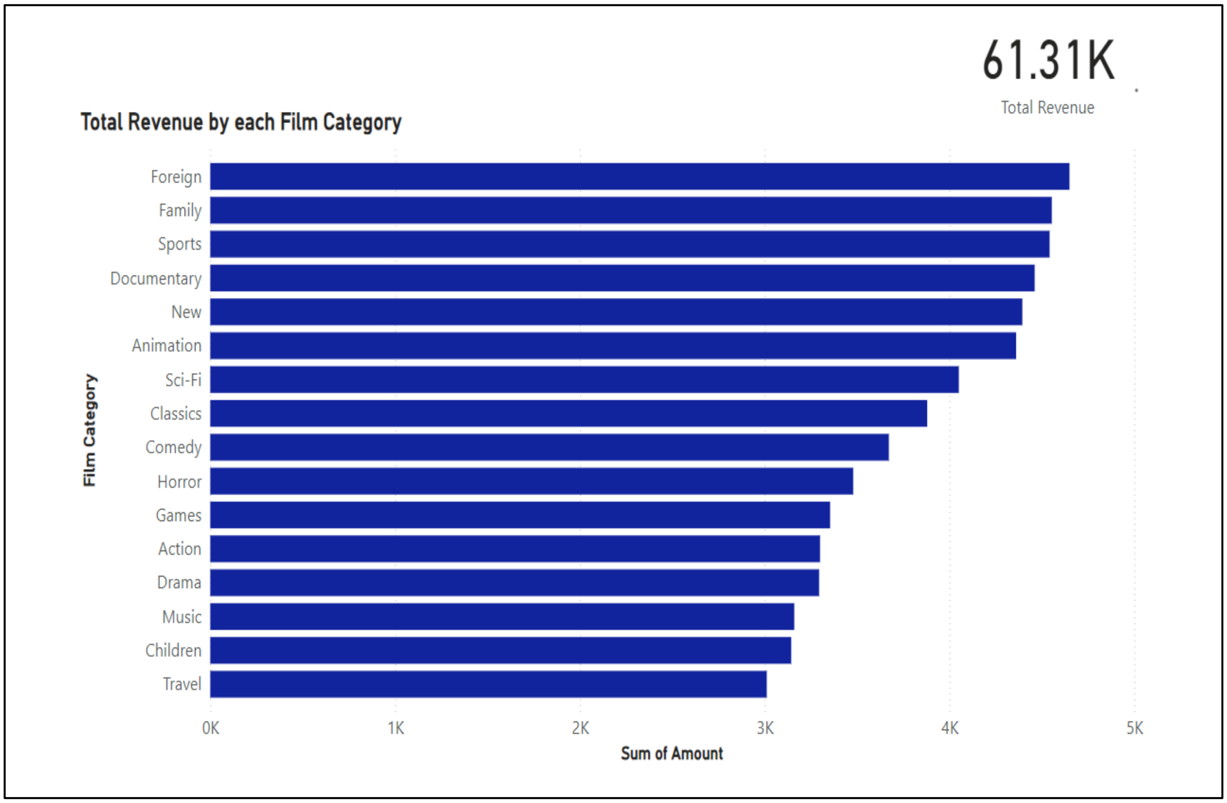


Figure 12: This Visual shows the Total Revenue each Film Category has made.

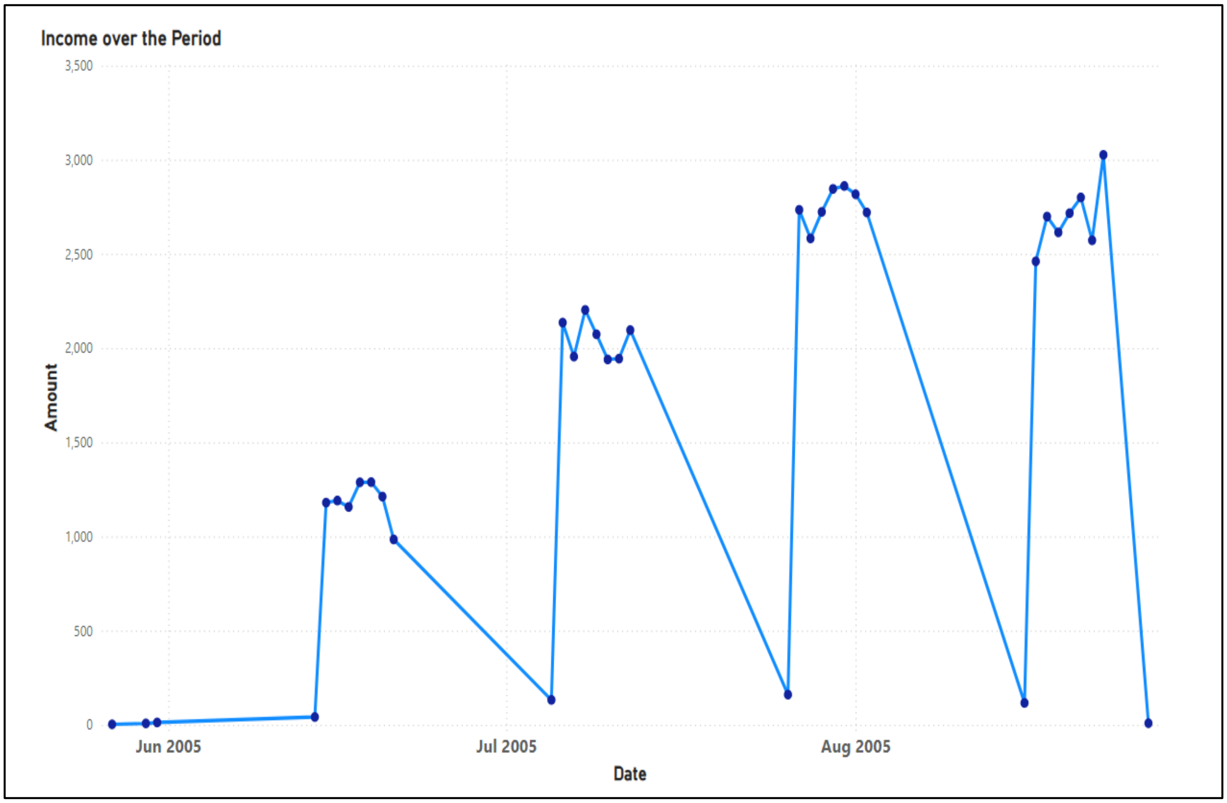


Figure 13: This visual shows the Revenue performance over the period of time.

PLANNING AHEAD

1. Dashboards/reports/analyses that can be supported by the current data warehouse but have yet to be developed

BELOW MENTIONED ADDITIONAL REPORTS WE CAN CREATE IN FUTURE USING THE CURRENT DATA WAREHOUSE

Customer Segmentation Analysis:

a. Segment customers based on rental behavior, spending patterns, and preferences.
b. Identify high-value customers and target them with personalized offers.

Rental Trends and Patterns:

a. Analyze daily, weekly, and monthly rental patterns.
b. Identify peak rental times and seasons.

2. Dashboards/reports/analyses that would need updates to the dimensional model in order to support

SUGGESTED DIMENSIONAL MODEL UPDATES:

1. **DimCustomer**:
   - Add columns for age, gender, and location.
2. **DimFilm**:
   - Add columns for rating, director, and actors.
3. **DimStore**:
   - Add detailed address attributes such as city, state, country, and size.
4. **FactRental**:
   - Include detailed payment information with attributes like payment_method and payment_status.

APPENDIX

a. Link to Presentation: https://drive.google.com/file/d/1lBdQdjvb7vzmf1G6WCouHVg4QKYj72Lm/view?usp=sharing
b. Power BI Reporting: https://app.powerbi.com/Redirect?action=OpenApp&appId=3443816e-f440-43a8-8a7a-108cf100c571&ctid=5217e0e7-539d-4563-b1bf-7c6dcf074f91
c. Data Source and ERD: https://www.postgresqltutorial.com/postgresql-getting-started/postgresql-sample-database/