Team Project #1

Database Management (BUS-315)

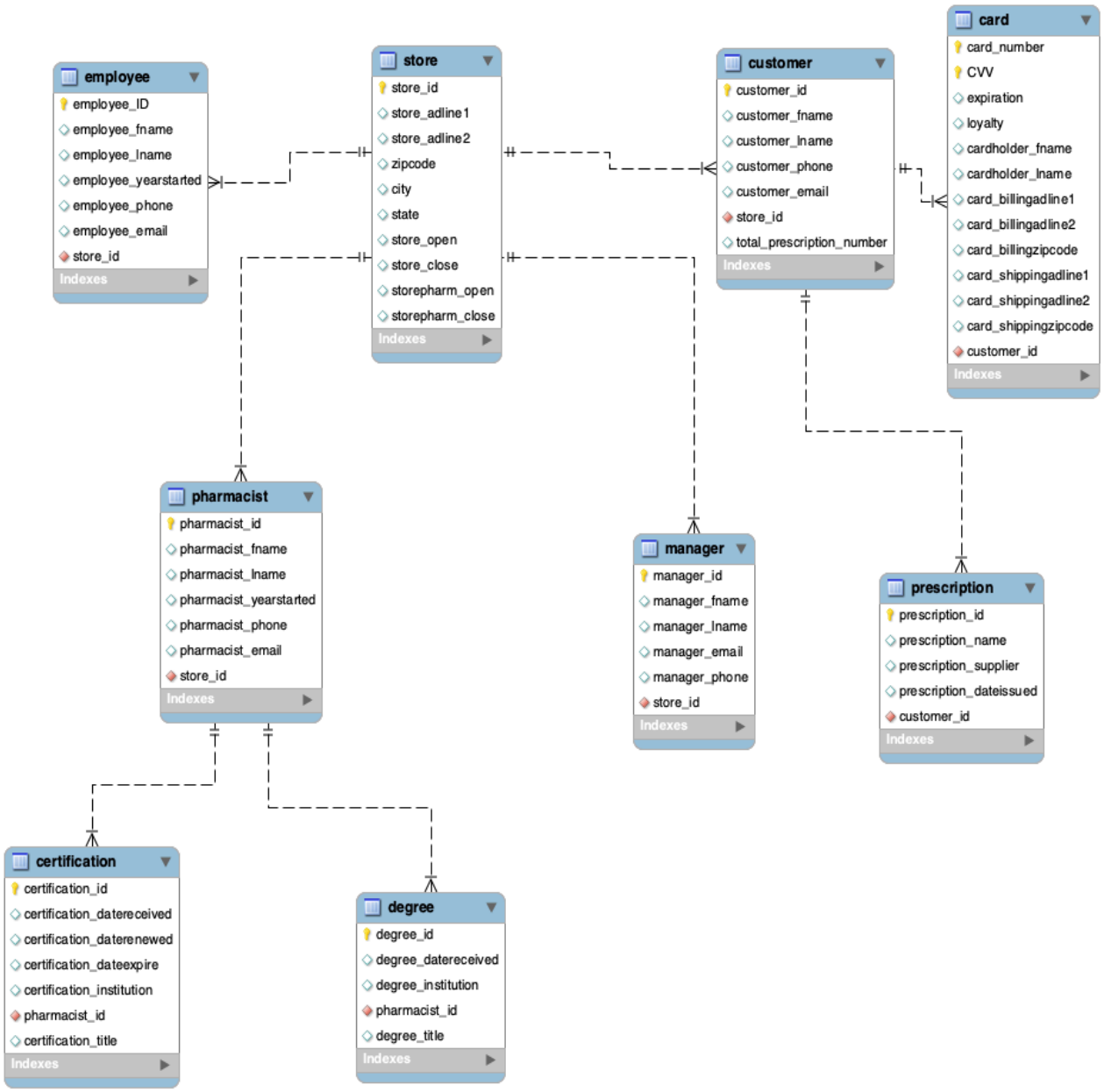
Prof. Larson

Thomas Haller

Avery Myers

Aidan Lorsson

Data Model


employee
employee_ID
employee_fname
employee_lname
employee_yearstarted
employee_phone
employee_email
store_id
Indexes
store
store_id
store_adline1
store_adline2
zipcode
city
state
store_open
store_close
storepharm_open
storepharm_close
Indexes
customer
customer_id
customer_fname
customer_lname
customer_phone
customer_email
store_id
total_prescription_number
Indexes
card
card_number
CVV
expiration
loyalty
cardholder_fname
cardholder_lname
card_billingadline1
card_billingadline2
card_billingzipcode
card_shippingadline1
card_shippingadline2
card_shippingzipcode
customer_id
Indexes
pharmacist
pharmacist_id
pharmacist_fname
pharmacist_lname
pharmacist_yearstarted
pharmacist_phone
pharmacist_email
store_id
Indexes
manager
manager_id
manager_fname
manager_lname
manager_email
manager_phone
store_id
Indexes
prescription
prescription_id
prescription_name
prescription_supplier
prescription_dateissued
customer_id
Indexes
certification
certification_id
certification_datereceived
certification_daterenewed
certification_dateexpire
certification_institution
pharmacist_id
certification_title
Indexes
degree
degree_id
degree_datereceived
degree_institution
pharmacist_id
degree_title
Indexes

The data model we ultimately decided on was based on the project 1 requirements, in which we came up with nine different tables. CVS has many stores throughout the United States, and these stores have many different pharmacists, managers, employees, and customers. Store is our one table, having a one-to-many relationship with the pharmacist, manager, employee, and customer tables. Subsequently, pharmacists have many different certifications and degrees. This constitutes two one-to-many relationships between pharmacist and degree as well as pharmacist and certification. Moreover, customers have many different credit cards and prescriptions. Therefore, these one-to-many relationships are laid out in the above data model and display the hierarchical relationships among each and every table. We were initially going to include other subsequent tables such as employee shift, pharmacist shift, manager shift, etc., but decided against it as to not take away or distract from the main operations of the business. Each "many" table includes a foreign key from its parent table allowing the two tables to be joined together in queries.

In regards to the tables and their keys, we decided that the "card" table would have a composite key, in which it included "card_number," and "CVV." Originally, we thought the card number would be sufficient, but after further deliberation, we came to the conclusion that the CVV would create a much greater distinction, as we thought the card number might not be entirely unique from card to card. To distinguish whether a card was a CVS loyalty card or from an independent company, we added a loyalty column. In this column, cards are checked "Y" for yes and "N" for no. A card checked "Y" is a loyalty card, and a card checked "N" is from an independent company.

We added a column in the customer table called total_prescription_number that serves a record of all historical prescriptions that a customer has had with CVS. This column is separate from the prescription table because the prescription table shows current prescriptions for a customer rather than historical data.

Our data model covers a small subset of CVS storefronts, as there are only 30 rows of data in the store table. Therefore, the pharmacists, employees, customers, and managers only belong to 30 stores nationwide. Moreover, we utilized ChatGPT for its ability to produce fictitious data, in which we then exported into a shared Google Sheet. Then, we imported the data from the Google Sheet (downloaded as a CSV file) into the shared team schema (we also copied and pasted values into the tables themselves).

We assumed that one store has many pharmacists, managers, and other employees. We added a table for employees to encapsulate CVS workers that don't fall under the pharmacist and manager titles, because CVS likey needs to keep track of all its workers. We also assumed that pharmacists can obtain multiple certifications for conducting pharmaceutical practices as well as multiple educational degrees, but the certifications and degrees are different. We didn't make a third table for credentials, as we assumed that a pharmacist's certifications and degrees comprise the necessary job credentials.

Manager Data Dictionary

	Field	Data Type	Key Type	Description	Constraints	Allowable Values	Example
1	manager_id	INT	Primary Key	Manager's ID	Not Null, Auto-increment	Integer, Unique	102
2	manager_fname	VARCHAR(45)		Manager's first name	Not Null	String, Max Length 50 chars	John
3	manager_lname	VARCHAR(45)		Manager's last name	Not Null	String, Max Length 50 chars	Doe
4	manager_email	VARCHAR(45)		Manager's email address	Not Null, Unique	String, Email format	johndoe@cvs.com
5	manager_phone	VARCHAR(15)		Manager's phone number	Not Null	String, Phone number format	123-456-7891
6	store_id	INT	Foreign key	Store's ID	Not Null	Integer, unique	4

Queries

1. How many employees and customers have the first name John? For CVS, being able to write a query to identify the number of employees and customers with the first name "John" is crucial for targeted communication and engagement. This information allows CVS to personalize interactions, whether it's recognizing employees with the name John for internal initiatives or tailoring promotions and marketing strategies specifically for customers sharing the same name. Additionally, periodic checks on the count contribute to data quality assurance, ensuring accurate and reliable information within the CVS database.

Query

(SELECT COUNT(*) FROM employee WHERE employee_fname = 'John') + (SELECT COUNT(*)FROM customer WHERE customer_fname = 'John') AS total_johns;

Result

total_johns
2

2. List of customers with number of prescriptions between 5 and 10. Being able to generate a list of customers with a specific number of prescriptions, in this case, between 5 and 10, is important for CVS for various reasons. This information is valuable for targeted customer engagement and marketing efforts. CVS can use this query to identify a segment of customers who fall within a particular prescription range, allowing them to tailor promotions, loyalty programs, or health-related communications specifically for this group. Additionally, it aids in understanding customer behavior and preferences, enabling CVS to enhance its services and offerings based on the prescription patterns of different customer segments.

Query

SELECT customer_id, customer_fname, customer_lname, total_prescription_number
FROM customer
WHERE total_prescription_number BETWEEN 5 AND 10;

Result

customer_id	customer_fname	customer_lname	total_prescription_number
1	John	Doe	10
4	Peter	Brown	7
10	Richard	Anderson	6
12	Martinez	Gomez	8
14	Nguyen	Tran	5
16	Khan	Ali	10
18	Williams	Davis	6
23	Allen	Cooper	9
26	Lopez	Martinez	6
27	Campbell	Wilson	7
28	Murphy	Anderson	8
30	Davis	Johnson	10
NULL	NULL	NULL	NULL

3. List of pharmacists in order of first hired to most recently hired. The ability to generate a list of pharmacists ordered by their hire date, from the earliest to the most recent, is essential for CVS for effective workforce management and operational efficiency. This query allows CVS to gain insights into the chronological order of pharmacist hires, facilitating the identification of experienced staff and newer additions. This information

can be valuable for training, scheduling, and resource allocation purposes. Additionally, it aids in assessing workforce dynamics and turnover rates, helping CVS make informed decisions regarding staffing levels and optimizing the overall efficiency of its pharmacy operations.

Query

```
SELECT pharmacist_id, pharmacist_fname, pharmacist_lname, pharmacist_datestarted
FROM pharmacist
ORDER BY pharmacist_datestarted;
```

Result

pharmacist_id	pharmacist_fname	pharmacist_lname	pharmacist_datestarted
29	Penelope	Hernandez	2010-05-24
30	John	King	2010-06-29
28	Zoe	Young	2011-04-19
24	Aria	Lee	2013-12-28
17	Avery	Thompson	2014-05-23
22	Victoria	Rodriguez	2014-10-18
26	Joseph	Hall	2015-02-08
16	Sofia	Martin	2015-04-18
7	Sophia	Miller	2015-07-30
15	Elizabeth	Harris	2016-03-13
13	Abigail	Jackson	2018-01-02
5	Ava	Brown	2018-05-20
25	Lily	Walker	2019-01-03
14	Thomas	White	2019-02-07
4	Noah	Jones	2019-04-15
1	Emma	Smith	2020-01-01
2	Liam	Johnson	2020-02-05
3	Olivia	Williams	2020-03-10
27	Riley	Allen	2020-03-14
21	Justin	Clark	2021-09-13
23	Madison	Lewis	2021-11-23
18	Ella	Garcia	2022-06-28
20	Grace	Robinson	2022-08-08
6	Isabella	Davis	2023-06-25
19	Scarlett	Martinez	2023-07-03
8	Peter	Wilson	2023-08-02
9	Charlotte	Moore	2023-09-11
10	Amelia	Taylor	2023-10-17

4. List the stores that are not located in Anytown, Bigtown, Foxtown, Iguana, Cattown, Dogtown. The capability to generate a list of stores that are not located in specific cities is crucial for CVS for targeted geographical analysis and decision-making. This query helps CVS identify stores situated outside these specified cities, providing insights into the distribution and coverage of its stores. The information is valuable for strategic planning, market expansion considerations, and optimizing store placement based on regional preferences and demographics. Additionally, it aids in ensuring that marketing campaigns or initiatives intended for specific city-based audiences do not include irrelevant store locations.

Query

```
SELECT *
FROM store
WHERE city NOT IN ('Anytown', 'Bigtown', 'Foxtown', 'Iguana', 'Cattown', 'Dogtown');
```

Result

store_id	store_adline1	store_adline2	zipcode	city	state	store_open	store_close	storepharm_open	storepharm_close
6	1415 Poplar Street	Suite 606	98765	Goattown	DE	11:00	21:00	10:00	20:00
7	1617 Birch Street	Suite 707	23456	Horsetown	FL	8:00	18:00	7:00	17:00
9	2021 Elm Street	Suite 909	67890	Lizardtown	GA	11:00	21:00	10:00	20:00
10	2223 Oak Street	Suite 100	32109	Monkeytown	HI	8:00	18:00	7:00	17:00
11	2425 Poplar Street	Suite 111	54321	Parrottown	IA	9:00	19:00	8:00	18:00
12	2627 Birch Street	Suite 122	76543	Pigeontown	ID	10:00	20:00	9:00	19:00
13	2829 Cedar Street	Suite 133	98765	Rabbittown	IL	11:00	21:00	10:00	20:00
14	3031 Elm Street	Suite 144	23456	Seagulltown	IN	8:00	18:00	7:00	17:00
15	3233 Oak Street	Suite 155	45678	Snaketown	KS	9:00	19:00	8:00	18:00
16	3435 Poplar Street	Suite 166	67890	Sparrowtown	KY	10:00	20:00	9:00	19:00
17	3637 Birch Street	Suite 177	32109	Swantown	LA	11:00	21:00	10:00	20:00
18	3839 Cedar Street	Suite 188	54321	Turtletown	MA	8:00	18:00	7:00	17:00
19	4041 Elm Street	Suite 199	76543	Zebratown	MD	9:00	19:00	8:00	18:00
20	4243 Oak Street	Suite 210	98765	Wolftown	ME	10:00	20:00	7:00	17:00
21	4445 Poplar Street	Suite 221	23456	Aardvarktown	MI	11:00	21:00	10:00	20:00
22	4647 Birch Street	Suite 232	45678	Badgertown	MN	8:00	18:00	7:00	17:00
23	4849 Cedar Street	Suite 243	67890	Beartown	MS	9:00	19:00	8:00	18:00
24	5051 Elm Street	Suite 254	32109	Bisontown	MO	10:00	20:00	9:00	19:00
25	5253 Oak Street	Suite 265	54321	Catamountt...	MT	11:00	21:00	10:00	20:00
26	5455 Poplar Street	Suite 276	76543	Deertown	NE	8:00	18:00	7:00	17:00
27	5657 Birch Street	Suite 287	98765	Elktown	NH	9:00	19:00	8:00	18:00
29	6061 Elm Street	Suite 309	45678	Goattown	NM	11:00	21:00	10:00	20:00
30	6263 Oak Street	Suite 310	67890	Horsetown	NY	8:00	18:00	7:00	17:00
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

5. List the number of pharmacists who have different degree titles. The ability to determine the number of pharmacists with distinct degree titles is important for CVS to assess the diversity and specialization within its pharmacy staff. This query allows CVS to identify the variety of educational backgrounds among pharmacists, which can be critical for optimizing pharmacy services, addressing specific healthcare needs, and ensuring a well-rounded skill set within the workforce. Understanding the distribution of degree titles among pharmacists is valuable for workforce planning, training programs, and aligning the expertise of pharmacy staff with CVS's service goals and customer requirements.

Query

```
SELECT degree_title, COUNT(DISTINCT pharmacist.pharmacist_id) AS pharmacist_count
FROM pharmacist
JOIN degree ON pharmacist.pharmacist_id = degree.pharmacist_id
GROUP BY degree_title;
```

Result

degree_title	pharmacist_count
Bachelor of Science in Pharmacy	10
Doctor of Pharmacy	10
Master of Science in Pharmacy	10

6. Calculate the average number of prescriptions per customer for each store. The capability to calculate the average number of prescriptions per customer for each store is crucial for CVS to assess the pharmacy performance and customer engagement at a local level. This query enables CVS to understand the prescription patterns and demands specific to each store, facilitating targeted marketing strategies and personalized customer interactions. The average prescriptions per customer metric is valuable for evaluating the effectiveness of pharmacy services, optimizing staffing levels, and tailoring promotions or health

initiatives based on local customer behavior. This information contributes to strategic decision-making, enhancing the overall customer experience and operational efficiency at each CVS store.

Query

```
SELECT store.store_id,
    SUM(customer.total_prescription_number) AS total_prescriptions,
    COUNT(customer.customer_id) AS total_customers,
    SUM(customer.total_prescription_number) / COUNT(customer.customer_id) AS
average_prescriptions_per_customer
FROM store
JOIN customer ON store.store_id = customer.store_id
GROUP BY store.store_id;
```

Result

store_id	total_prescriptions	total_customers	average_prescriptions_per_customer
1	10	1	10.0000
2	4	1	4.0000
3	3	1	3.0000
4	7	1	7.0000
5	12	1	12.0000
6	20	1	20.0000
7	12	1	12.0000
8	3	1	3.0000
9	2	1	2.0000
10	6	1	6.0000
11	2	1	2.0000
12	8	1	8.0000
13	1	1	1.0000
14	5	1	5.0000
15	NULL	1	NULL
16	10	1	10.0000
17	3	1	3.0000
18	6	1	6.0000
19	11	1	11.0000
20	1	1	1.0000
21	4	1	4.0000
22	1	1	1.0000
23	9	1	9.0000
24	3	1	3.0000
25	2	1	2.0000

7. List the customers who have purchased more than the average prescription amount. The ability to list customers who have purchased more than the average prescription amount is crucial for CVS to identify high-value customers and tailor engagement strategies. This query helps CVS pinpoint customers whose prescription spending exceeds the average, enabling targeted marketing campaigns, loyalty programs, or personalized communications. By understanding and catering to the needs of these higher-spending customers, CVS can enhance customer satisfaction, loyalty, and potentially increase revenue. This information is valuable for customer relationship management, allowing CVS to optimize its services based on the prescription spending behavior of different customer segments.

Query

```
SELECT customer.customer_id, customer.customer_fname, customer.customer_lname,
customer.total_prescription_number
FROM customer
WHERE customer.total_prescription_number > (
   SELECT AVG(total_prescription_number)
   FROM customer);
```

Result

customer_id	customer_fname	customer_lname	total_prescription_number
1	John	Doe	10
4	Peter	Brown	7
5	David	Jones	12
6	Sarah	Williams	20
7	Thomas	Miller	12
12	Martinez	Gomez	8
16	Khan	Ali	10
19	Brown	Miller	11
23	Allen	Cooper	9
27	Campbell	Wilson	7
28	Murphy	Anderson	8
30	Davis	Johnson	10
NULL	NULL	NULL	NULL

8. List the number of customers who have a total prescription count greater than 10. The capability to list the number of customers with a total prescription count greater than 10 is important for CVS to identify a specific segment of customers who may have higher prescription needs. This query allows CVS to focus on customers with potentially ongoing health requirements and tailor its services accordingly. The information is valuable for targeted health-related communications, loyalty programs, and ensuring that CVS can effectively meet the healthcare needs of customers who have a higher prescription count. By understanding and addressing the needs of this customer segment, CVS can enhance customer satisfaction and loyalty within this particular group.

Query

```
SELECT customer_id, customer_fname, customer_lname, SUM(total_prescription_number) AS
total_prescriptions
FROM customer
GROUP BY customer_id, customer_fname, customer_lname
HAVING SUM(total_prescription_number) > 10;
```

Result

customer_id	customer_fname	customer_lname	total_prescriptions
5	David	Jones	12
6	Sarah	Williams	20
7	Thomas	Miller	12
19	Brown	Miller	11

9. List the employees and their start years.Being able to list employees along with their start years is essential for CVS to understand the tenure of its workforce and make informed decisions regarding employee management. This query provides CVS with insights into the distribution of start years among employees, facilitating workforce planning, recognition programs, and the identification of long-term or recently onboarded staff. Additionally, this information contributes to strategic decision-making related to staffing levels, training initiatives, and employee engagement efforts, ensuring a well-balanced and experienced workforce at CVS.

Query

```
SELECT employee_ID, employee_fname, employee_lname,
DATE_FORMAT(employee_datestarted, '%Y') AS start_year
FROM employee;
```

Result

employee_ID	employee_fname	employee_lname	start_year
1	Alex	Johnson	2022
2	William	Brown	2022
3	Sarah	Miller	2022
4	David	Taylor	2022
5	Michael	Williams	2022
6	Jane	Anderson	2022
7	John	Wilson	2022
8	Peter	Smith	2022
9	Mary	Jones	2022
10	Richard	Doe	2022
11	Emily	Thomas	2023
12	Kevin	Davis	2023
13	Laura	Miller	2023
14	Michael	Taylor	2023
15	Jessica	Williams	2023
16	Mark	Anderson	2023
17	Susan	Wilson	2023
18	David	Smith	2023
19	Elizabeth	Jones	2023
20	Robert	Doe	2023
21	Steven	Thomas	2022
22	Karen	Davis	2022
23	Paul	Miller	2022
24	Anthony	Taylor	2022
25	Donna	Williams	2022
26	Mary	Anderson	2022
27	George	Wilson	2022
28	Kenneth	Smith	2022

10. List all pharmacists who have obtained degrees from military institutions. The capability to list all pharmacists who have obtained degrees from military institutions is valuable for CVS to recognize and potentially leverage the unique skills and experiences of these pharmacists. This query allows CVS to identify pharmacists with military education backgrounds, facilitating targeted internal initiatives, specialized training programs, or initiatives that may benefit from their distinct qualifications. Understanding the educational diversity among pharmacists contributes to workforce planning and the optimization of pharmacy services to cater to a variety of customer needs. This information aids CVS in recognizing the valuable contributions of pharmacists with military degrees and aligning their expertise with the overall goals of the organization.

Query

```
SELECT pharmacist.pharmacist_id, pharmacist.pharmacist_fname,
pharmacist.pharmacist_lname, degree_institution
FROM pharmacist
JOIN degree ON pharmacist.pharmacist_id = degree.pharmacist_id
WHERE degree.degree_institution REGEXP 'United States';
```

Result

pharmacist_id	pharmacist_fname	pharmacist_lname	degree_institution
22	Victoria	Rodriguez	United States Air Force Academy
23	Madison	Lewis	United States Naval Academy
25	Lily	Walker	United States Military Academy at West Point
27	Riley	Allen	United States Coast Guard Institute

	Query1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1. Subquery	x						x			
2. BETWEEN		x								
3. GROUP BY					x	x		x		
4. HAVING								x		
5. ORDER BY			x							
6.				x						

IN/NOT IN										
7. Aggregate function	**x**				**x**	**x**	**x**	**x**		
8. REGEXP										**x**
9. Date function									**x**	
10. Calculation	**x**					**x**				