Grocery Recommendation System

Members:

Shruti Milind Randive (002740632),
Amretasre Rengarajan Thiruvengadam (002762670),
Sraddha Pedda Gangireddy Gari (002743943),
Dharma Thanishq Nimmala (002709754).

Objective:

This project aims to assist users in buying groceries at the most affordable price or in the most convenient location possible, depending on their needs. The system will retrieve the desired product from the database and compare the pricing of the item from multiple retailers in accordance with the price stated. It also enables us to find the product details from the closest retailers based on the user's location. This system will also keep records of each employee of respective stores and these employees are allowed to add and update products for their specific store location.

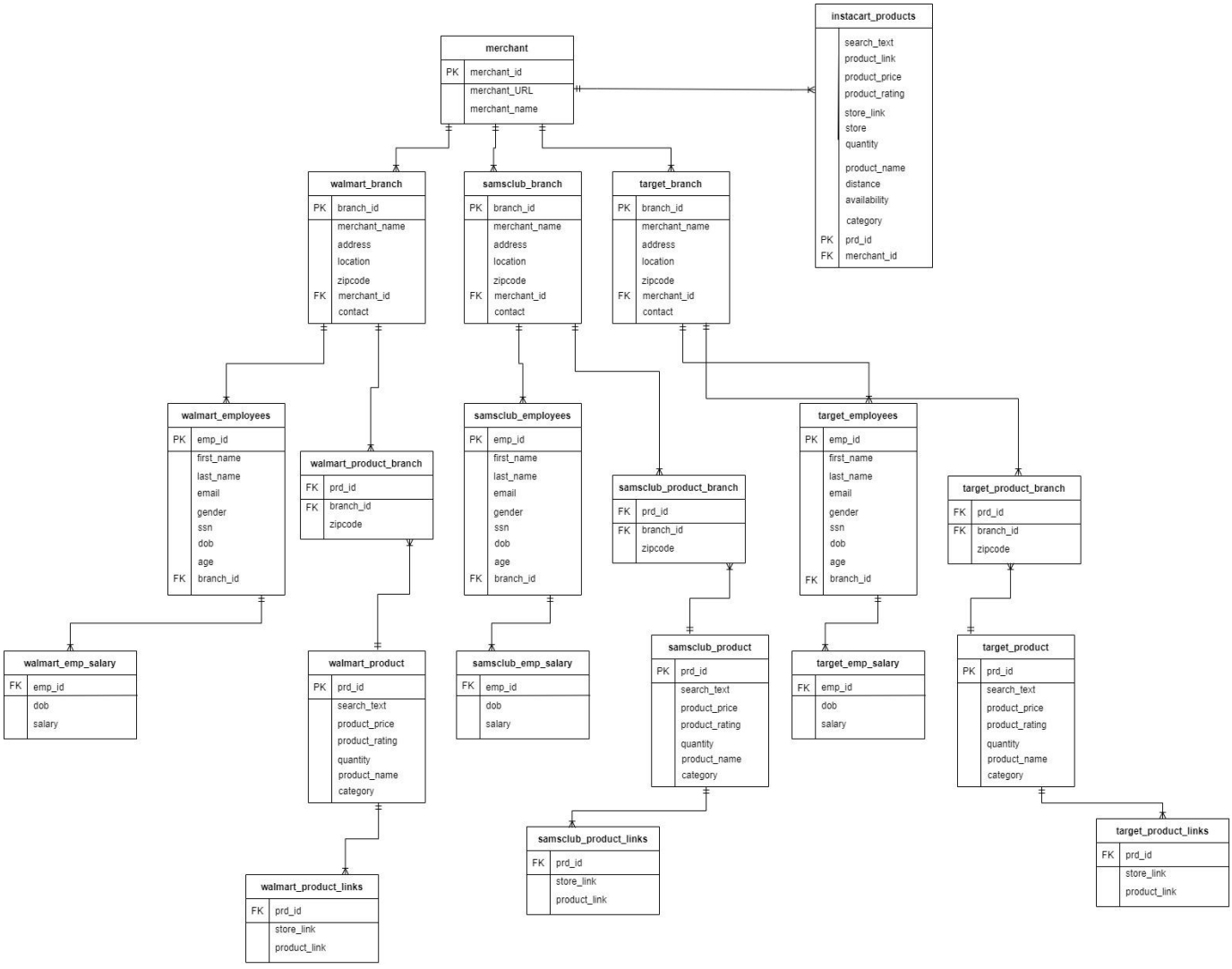
Description:

The Grocery_Recommendation_System database contains data on all the products available in merchants such as Walmart, Target, Samsclub, and the online delivery application Instacart. This data has been web scrapped from the **Google Shopping** web platform which includes millions of products from various merchants. The merchant table contains these merchant details along with their web URLs. The product tables contain columns such as product name, price, product inks, ratings, location in terms of zip code, etc. The Branch table data have been scrapped from the **Yellow Pages** website which contains the address and zip codes of all the merchant locations that are available in Boston City. The generated Employee tables contain dummy data concerning employee data such as employee id, first_name, last_name, and age for the selected merchants.

GIT HUB REPOSITORY LINK:

https://github.com/SraddhaP/Grocery_Recommendation_System

Entity Relation Diagram Of Grocery Recommendation System:



CREATE TABLE QUERIES:

1. **Merchant Table**

```
CREATE TABLE `merchant` ( `merchant_id` bigint NOT NULL AUTO_INCREMENT,
`mechant_name` text, `merchant_URL` text );
```

2. **Walmart Branch Table**

```
CREATE TABLE `walmart_branch` ( `merchant_name` text, `address` text, `location` text,
`zipcode` text, `contact` text, `branch_id` int NOT NULL AUTO_INCREMENT,
`merchant_id` int DEFAULT NULL,
 PRIMARY KEY (`branch_id`),
 UNIQUE KEY `Branch_id` (`branch_id`));
```

3. **Target Branch Table**

```
CREATE TABLE `target_branch` ( `merchant_name` text, `address` text, `location` text,
`zipcode` text, `contact` text, `branch_id` int NOT NULL AUTO_INCREMENT,
`merchant_id` int DEFAULT NULL,
 PRIMARY KEY (`branch_id`),
 UNIQUE KEY `Branch_id` (`branch_id`));
```

4. **Sam's Club Branch Table**

```
CREATE TABLE `samsclub_branch` ( `merchant_name` text, `address` text, `location` text,
`zipcode` text, `contact` text, `branch_id` int NOT NULL AUTO_INCREMENT,
`merchant_id` int DEFAULT NULL,
 PRIMARY KEY (`branch_id`),
 UNIQUE KEY `Branch_id` (`branch_id`));
```

5. **Walmart Employee Tables**

```
CREATE TABLE `walmart_employees` ( `emp_id` bigint NOT NULL, `first_name` text,
`last_name` text, `email` text, `gender` text, `ssn` bigint DEFAULT NULL, `age` bigint
DEFAULT NULL, `branch_id` int DEFAULT NULL,
 PRIMARY KEY (`emp_id`),
 KEY `walmart_emp_fk1` (`branch_id`),
```

CONSTRAINT `walmart_emp_fk1` FOREIGN KEY (`branch_id`) REFERENCES `walmart_branch` (`branch_id`));

6. Target Employee Tables

CREATE TABLE `target_employees` (`emp_id` bigint NOT NULL, `first_name` text, `last_name` text, `email` text, `gender` text, `ssn` bigint DEFAULT NULL, `dob` text, `age` bigint DEFAULT NULL, `branch_id` int DEFAULT NULL,
PRIMARY KEY (`emp_id`),
KEY `target_emp_fk1` (`branch_id`),
CONSTRAINT `target_emp_fk1` FOREIGN KEY (`branch_id`) REFERENCES `target_branch` (`branch_id`));

7. Sam's Club Employee Tables

CREATE TABLE `samsclub_employees` (`emp_id` bigint NOT NULL, `first_name` text, `last_name` text, `email` text, `gender` text, `ssn` bigint DEFAULT NULL, `Age` bigint DEFAULT NULL, `branch_id` int DEFAULT NULL, PRIMARY KEY (`emp_id`),
KEY `samsclub_emp_fk1` (`branch_id`),
CONSTRAINT `samsclub_emp_fk1` FOREIGN KEY (`branch_id`) REFERENCES `samsclub_branch` (`branch_id`));

8. Walmart Employee Salary Table

CREATE TABLE `walmart_emp_salary` (`emp_id` bigint DEFAULT NULL, `salary` bigint DEFAULT NULL,
KEY `walmart_emp_salary_fk1` (`emp_id`),
CONSTRAINT `walmart_emp_salary_fk1` FOREIGN KEY (`emp_id`) REFERENCES `walmart_employees` (`emp_id`));

9. Target Employee Salary Table

CREATE TABLE `target_emp_salary` (`emp_id` bigint DEFAULT NULL, `salary` bigint DEFAULT NULL,
KEY `target_emp_salary_fk1` (`emp_id`),
CONSTRAINT `target_emp_salary_fk1` FOREIGN KEY (`emp_id`) REFERENCES `target_employees` (`emp_id`));

10. Sam's Club Employee Salary Table

CREATE TABLE `samsclub_emp_salary` (`emp_id` bigint DEFAULT NULL, `salary` bigint DEFAULT NULL,

KEY `samsclub_emp_salary_fk1` (`emp_id`),
CONSTRAINT `samsclub_emp_salary_fk1` FOREIGN KEY (`emp_id`) REFERENCES `samsclub_employees` (`emp_id`));

11. Walmart Product Table

CREATE TABLE `walmart_product` (`search_text` text, `product_price` double DEFAULT NULL, `product_rating` double DEFAULT NULL, `quantity` text, `product_name` text, `category` text, `prd_id` bigint NOT NULL DEFAULT '0',
PRIMARY KEY (`prd_id`));

12. Target Product Table

CREATE TABLE `target_product` (`search_text` text,`product_price` double DEFAULT NULL,`product_rating` double DEFAULT NULL,`quantity` text,`product_name` text,
`category` text,`prd_id` bigint NOT NULL DEFAULT '0',
PRIMARY KEY (`prd_id`));

13. Sam's Club Product Table

CREATE TABLE `samsclub_product` (`search_text` text, `product_price` double DEFAULT NULL, `product_rating` double DEFAULT NULL, `quantity` text, `product_name` text,`category` text, `prd_id` bigint NOT NULL DEFAULT '0',
PRIMARY KEY (`prd_id`));

14. Instacart Product Table

CREATE TABLE `instacart_products` (`search_text` text,`product_link` text, `product_price` double DEFAULT NULL, `product_rating` double DEFAULT NULL,`store` text, `store_link` text, `quantity` text, `product_name` text,
`distance` double DEFAULT NULL, `availabilty` text, `category` text,
`prd_id` bigint NOT NULL AUTO_INCREMENT, `merchant_id` int NOT NULL,
PRIMARY KEY (`prd_id`),
UNIQUE KEY `Prd_id` (`prd_id`),
KEY `instacart_products_fk1` (`merchant_id`),
CONSTRAINT `instacart_products_fk1` FOREIGN KEY (`merchant_id`) REFERENCES `merchant` (`merchant_id`)
);

15. Walmart Product Branch Relation Table

```
CREATE TABLE `walmart_product_branch` ( `zipcode` bigint DEFAULT NULL,
 `branch_id` int DEFAULT NULL,`prd_id` bigint NOT NULL DEFAULT '0',
 KEY `walmart_product_branch_fk1` (`branch_id`),
 KEY `walmart_product_branch_fk2` (`prd_id`),
  CONSTRAINT `walmart_product_branch_fk1` FOREIGN KEY (`branch_id`) REFERENCES `walmart_branch` (`branch_id`),
   CONSTRAINT `walmart_product_branch_fk2` FOREIGN KEY (`prd_id`) REFERENCES `walmart_product` (`prd_id`)
);
```

16. Target Product Branch Relation Table

```
CREATE TABLE `target_product_branch` ( `zipcode` bigint DEFAULT NULL,
 `branch_id` int DEFAULT NULL, `prd_id` bigint NOT NULL DEFAULT '0',
 KEY `target_product_branch_fk1` (`branch_id`),
 KEY `target_product_branch_fk2` (`prd_id`),
  CONSTRAINT `target_product_branch_fk1` FOREIGN KEY (`branch_id`) REFERENCES `target_branch` (`branch_id`),
   CONSTRAINT `target_product_branch_fk2` FOREIGN KEY (`prd_id`) REFERENCES `target_product` (`prd_id`)
);
```

17. Sam's Club Product Branch Relation Table

```
CREATE TABLE `samsclub_product_branch` (`zipcode` bigint DEFAULT NULL,
 `branch_id` int DEFAULT NULL,`prd_id` bigint NOT NULL DEFAULT '0',
 KEY `samsclub_product_branch_fk1` (`branch_id`),
 KEY `samsclub_product_branch_fk2` (`prd_id`),
 CONSTRAINT `samsclub_product_branch_fk1` FOREIGN KEY (`branch_id`) REFERENCES `samsclub_branch` (`branch_id`),
  CONSTRAINT `samsclub_product_branch_fk2` FOREIGN KEY (`prd_id`) REFERENCES `samsclub_product` (`prd_id`)
);
```

18. Walmart Product Link Table

```
CREATE TABLE `walmart_product_links` (`store_link` text,`prd_id` bigint NOT NULL DEFAULT '0', `product_link` text,
```

```
 KEY `walmart_product_links_fk1` (`prd_id`),
   CONSTRAINT `walmart_product_links_fk1` FOREIGN KEY (`prd_id`) REFERENCES `walmart_product` (`prd_id`)
);
```

19. Target Product Link Table

```
CREATE TABLE `target_product_links` (`store_link` text,
 `prd_id` bigint NOT NULL DEFAULT '0',
 `product_link` text,
 KEY `target_product_links_fk1` (`prd_id`),
   CONSTRAINT `target_product_links_fk1` FOREIGN KEY (`prd_id`) REFERENCES `target_product` (`prd_id`)) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci;
```

20. Sam’s Club Product Link Table

```
CREATE TABLE `samsclub_product_links` ( `store_link` text,
 `prd_id` bigint NOT NULL DEFAULT '0',`product_link` text,
 KEY `samsclub_product_links_fk1` (`prd_id`),
CONSTRAINT `samsclub_product_links_fk1` FOREIGN KEY (`prd_id`) REFERENCES `samsclub_product` (`prd_id`));
```

CONSTRAINT QUERIES:

1. ADDING CONSTRAINTS TO MERCHANT BRANCH TABLE

```
ALTER TABLE walmart_branch
ADD CONSTRAINT `walmart_branch_fk1`
FOREIGN KEY (merchant_id) REFERENCES merchant(merchant_id);

ALTER TABLE target_branch
ADD CONSTRAINT `target_branch_fk1`
FOREIGN KEY (merchant_id) REFERENCES merchant(merchant_id);

ALTER TABLE samsclub_branch
ADD CONSTRAINT `samsclub_branch_fk1`
FOREIGN KEY (merchant_id) REFERENCES merchant(merchant_id);

ALTER TABLE instacart_products
ADD CONSTRAINT `instacart_products_fk1`
FOREIGN KEY (merchant_id) REFERENCES merchant(merchant_id);
```

2. ADDING CONSTRAINTS FOR MERCHANT EMPLOYEE TABLES

```
ALTER TABLE walmart_employees
MODIFY COLUMN branch_id INT;

ALTER TABLE target_employees
MODIFY COLUMN branch_id INT;

ALTER TABLE samsclub_employees
MODIFY COLUMN branch_id INT;

ALTER TABLE walmart_employees
ADD CONSTRAINT PRIMARY KEY (emp_id),
ADD CONSTRAINT `walmart_emp_fk1`
FOREIGN KEY (branch_id) REFERENCES walmart_branch(branch_id);

ALTER TABLE target_employees
ADD CONSTRAINT PRIMARY KEY (emp_id),
```

```
ADD CONSTRAINT `target_emp_fk1`
FOREIGN KEY (branch_id) REFERENCES target_branch(branch_id);

ALTER TABLE samsclub_employees
ADD CONSTRAINT PRIMARY KEY (emp_id),
ADD CONSTRAINT `samsclub_emp_fk1`
FOREIGN KEY (branch_id) REFERENCES samsclub_branch(branch_id);
```

3. ADDING CONSTRAINTS FOR NORMALISED TABLES

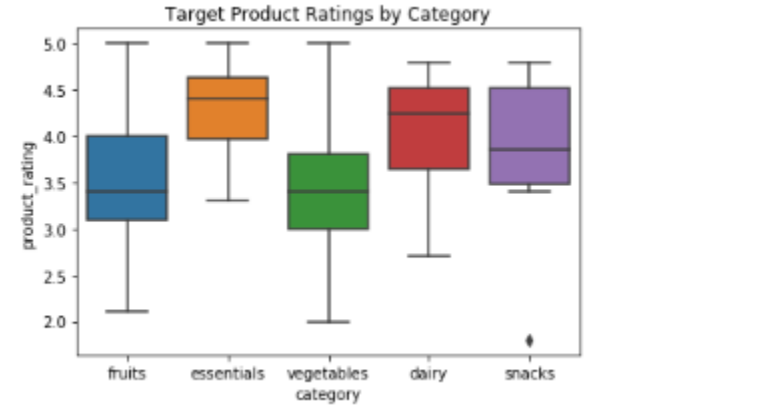
```
ALTER TABLE walmart_product_branch
ADD CONSTRAINT `walmart_product_branch_fk1`
FOREIGN KEY (branch_id) REFERENCES walmart_branch(branch_id);

ALTER TABLE target_product_branch
ADD CONSTRAINT `target_product_branch_fk1`
FOREIGN KEY (branch_id) REFERENCES target_branch(branch_id);

ALTER TABLE samsclub_product_branch
ADD CONSTRAINT `samsclub_product_branch_fk1`
FOREIGN KEY (branch_id) REFERENCES samsclub_branch(branch_id);

ALTER TABLE walmart_product_branch
ADD CONSTRAINT `walmart_product_branch_fk2`
FOREIGN KEY (prd_id) REFERENCES walmart_product(prd_id);

ALTER TABLE target_product_branch
ADD CONSTRAINT `target_product_branch_fk2`
FOREIGN KEY (prd_id) REFERENCES target_product(prd_id);

ALTER TABLE samsclub_product_branch
ADD CONSTRAINT `samsclub_product_branch_fk2`
FOREIGN KEY (prd_id) REFERENCES samsclub_product(prd_id);

ALTER TABLE walmart_product_links
ADD CONSTRAINT `walmart_product_links_fk1`
FOREIGN KEY (prd_id) REFERENCES walmart_product(prd_id);
```

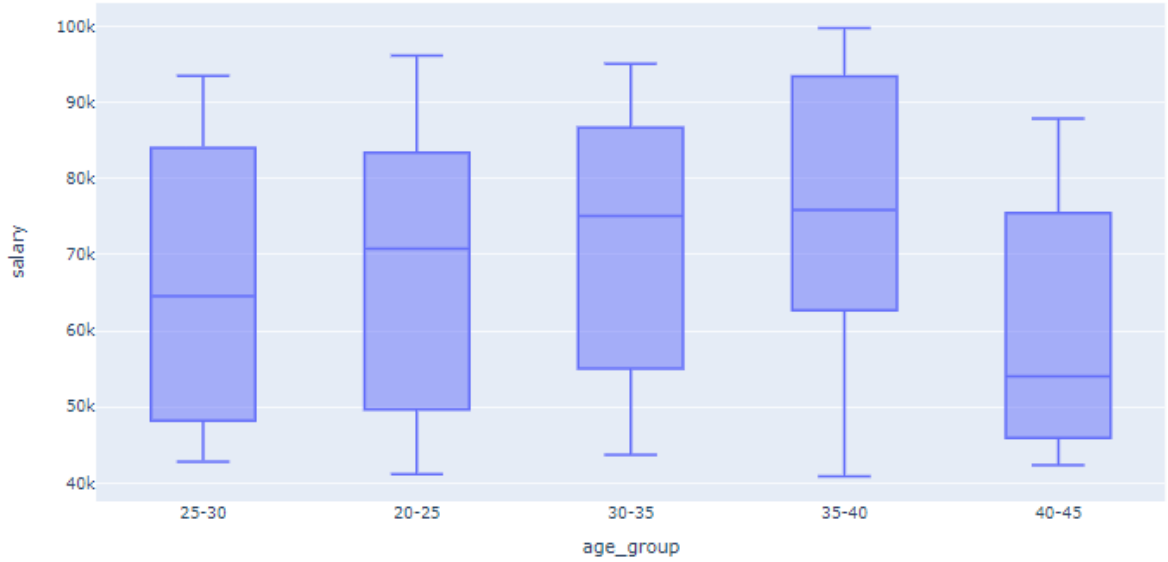
```
ALTER TABLE target_product_links
ADD CONSTRAINT `target_product_links_fk1`
FOREIGN KEY (prd_id) REFERENCES target_product(prd_id);

ALTER TABLE samsclub_product_links
ADD CONSTRAINT `samsclub_product_links_fk1`
FOREIGN KEY (prd_id) REFERENCES samsclub_product(prd_id);
```

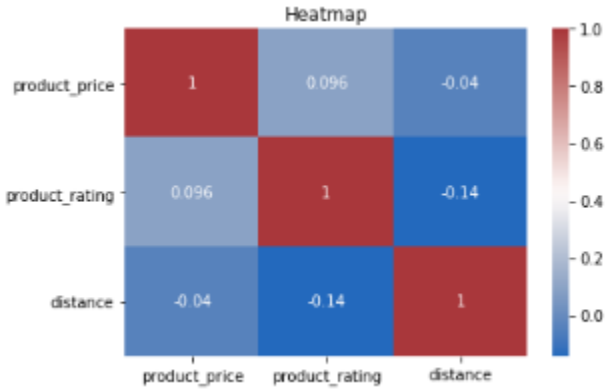
4. ADDING CONSTRAINTS FOR EMPLOYEE SALARY TABLES

```
ALTER TABLE samsclub_emp_salary
ADD CONSTRAINT `samsclub_emp_salary_fk1`
FOREIGN KEY (emp_id) REFERENCES samsclub_employees(emp_id);

ALTER TABLE target_emp_salary
ADD CONSTRAINT `target_emp_salary_fk1`
FOREIGN KEY (emp_id) REFERENCES target_employees(emp_id);

ALTER TABLE walmart_emp_salary
ADD CONSTRAINT `walmart_emp_salary_fk1`
FOREIGN KEY (emp_id) REFERENCES walmart_employees(emp_id);
```

SCREEN SHOTS OF DATA VISUALIZATION:

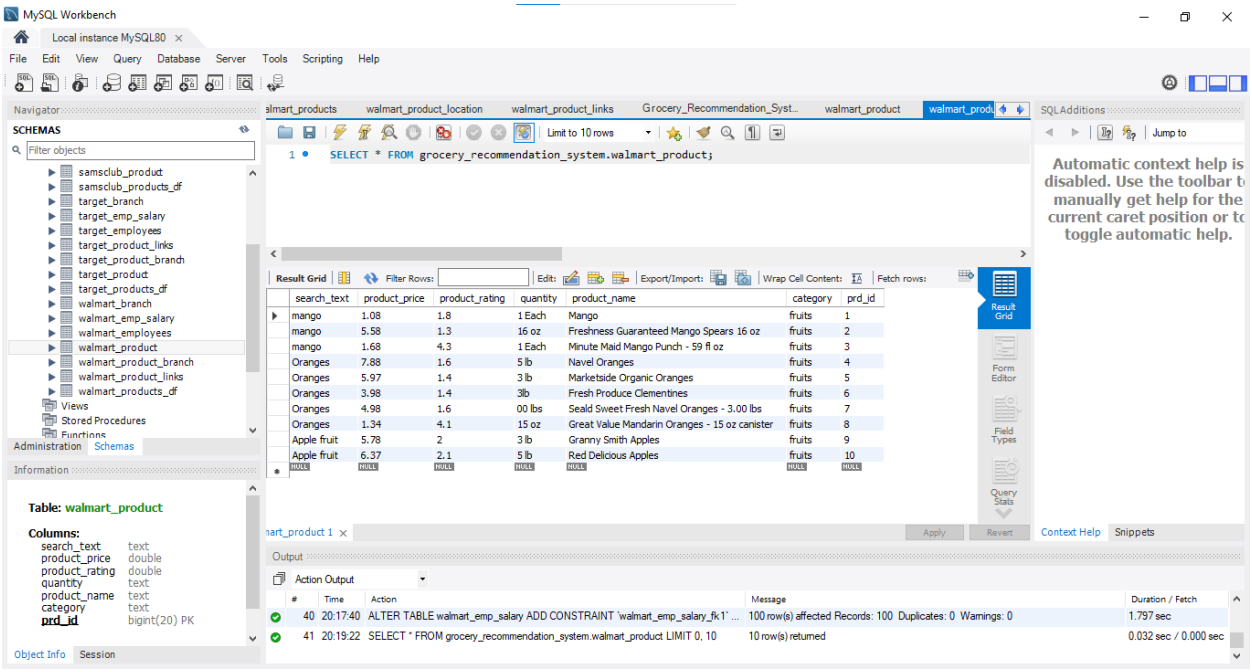



100k
90k
80k
70k
60k
50k
40k
salary
25-30
20-25
30-35
35-40
40-45
age_group

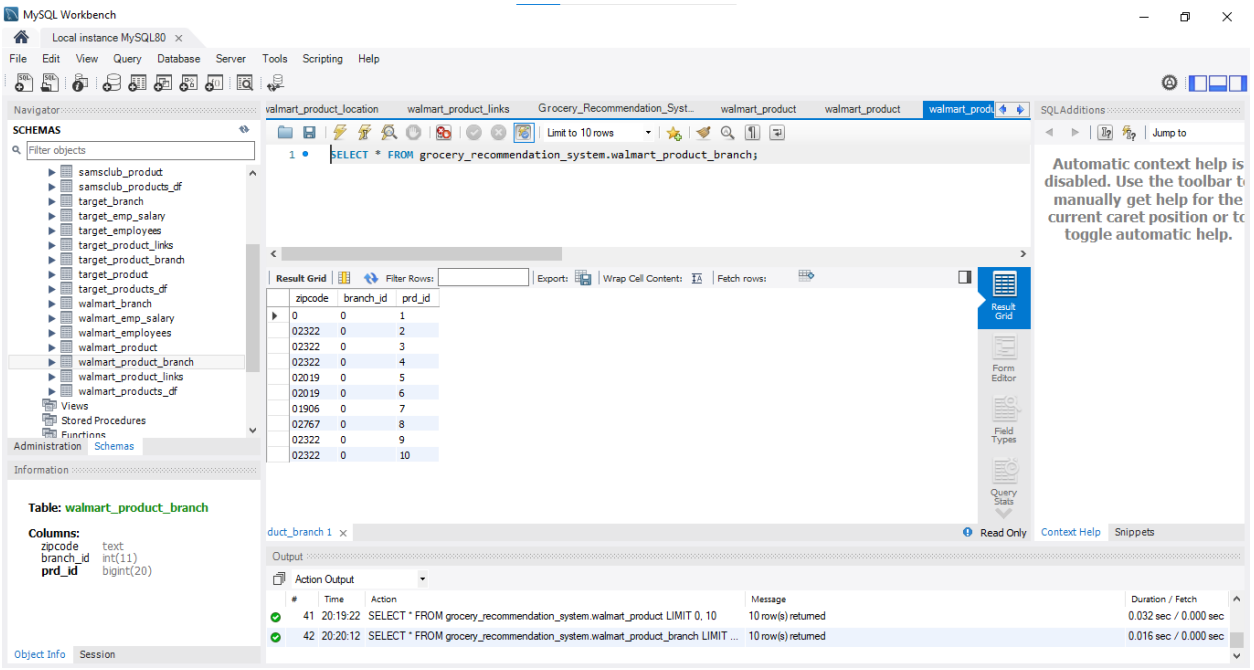

Heatmap
product_price
product_rating
distance
1
0.096
-0.04
0.096
1
-0.14
-0.04
-0.14
1
product_price
product_rating
distance
1.0
0.8
0.6
0.4
0.2
0.0

SNAP SHOTS OF TABLES FROM MYSQL WORKBENCH

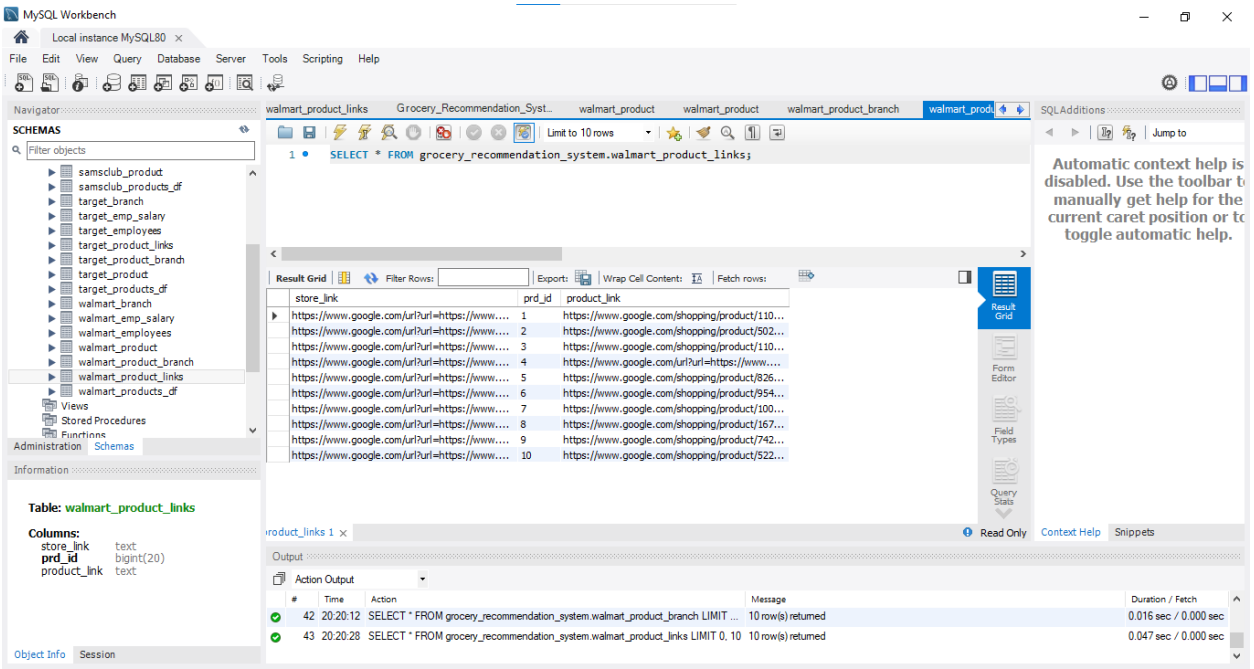
1. Walmart_Product Table



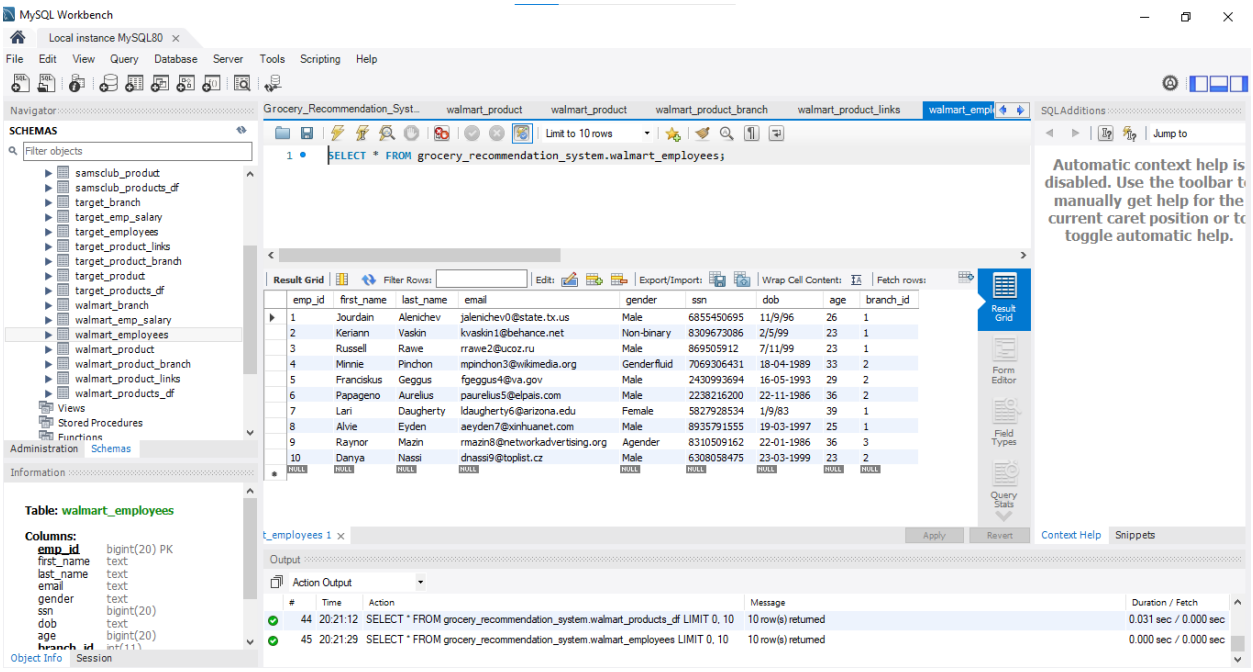
2. Walmart_Product_Branch Relation Table



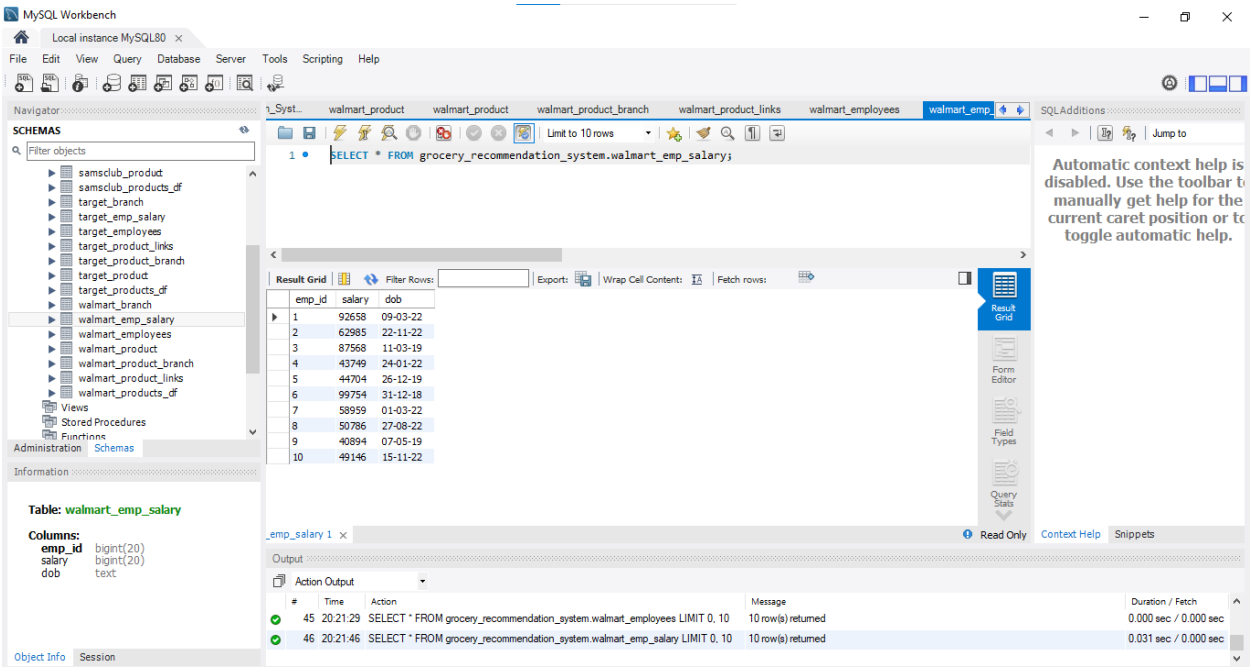
3. Walmart_Product_Links Table



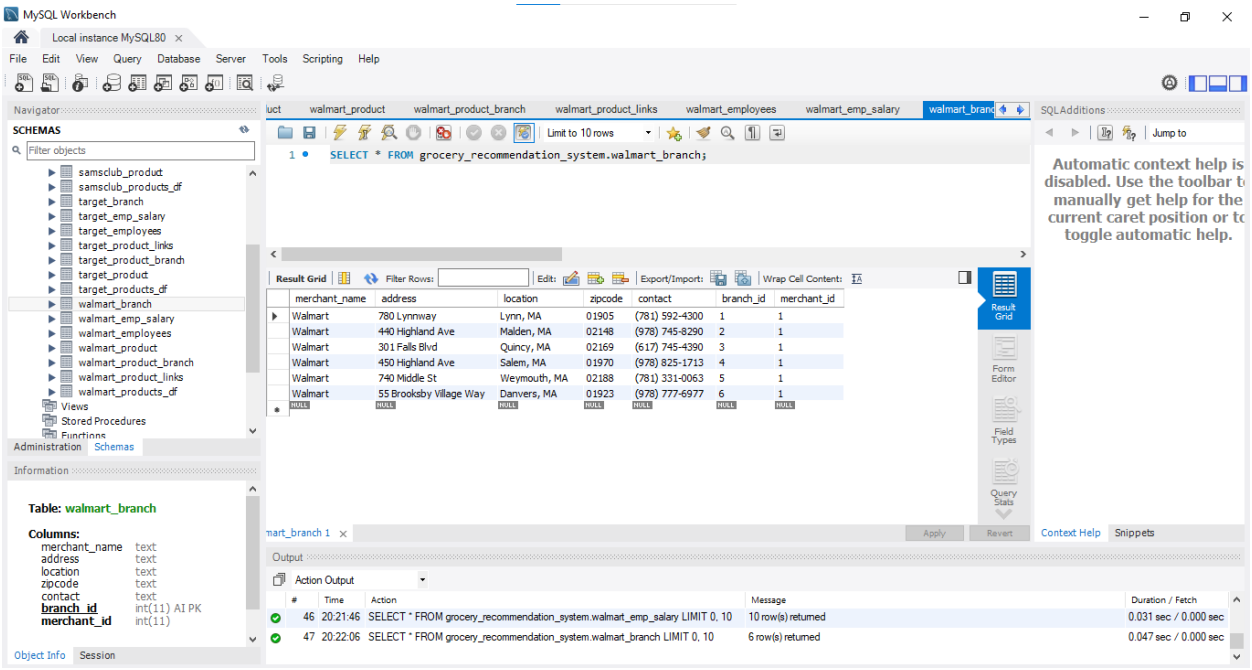
4. Walmart_Employees Table



5. Walmart_Employee_Salary Table

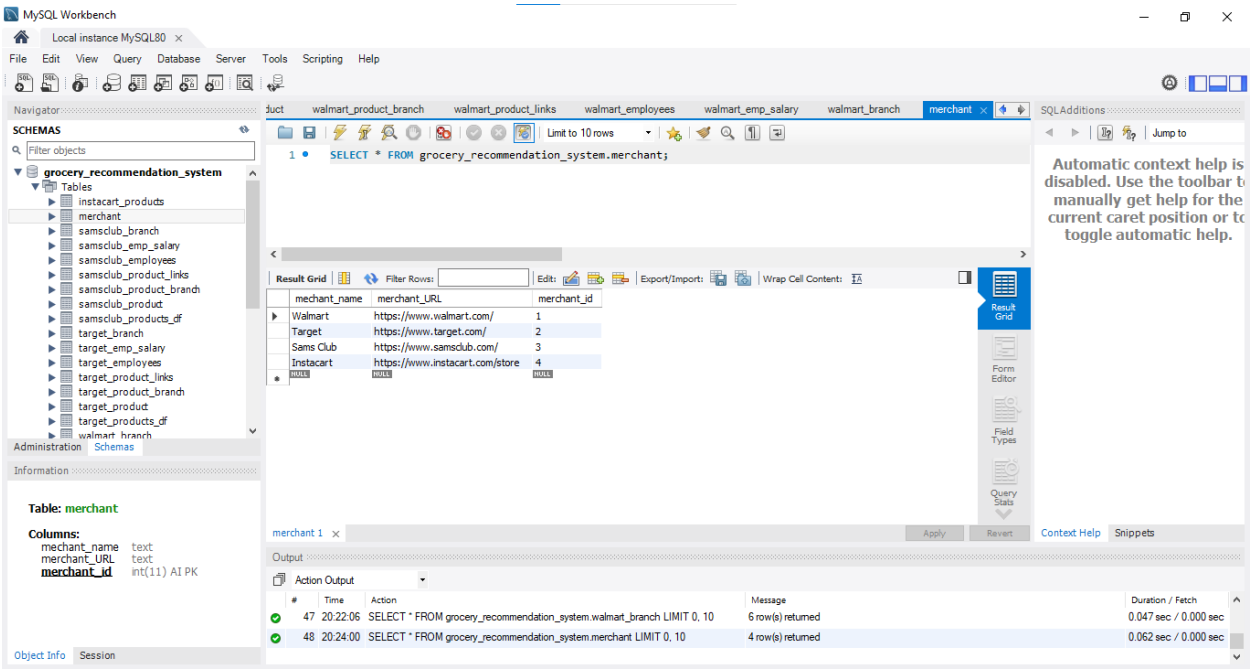


6. Walmart_Branch Table



Similarly we have tables for Sam's Club, Target and Instacart.

7. Merchant Table

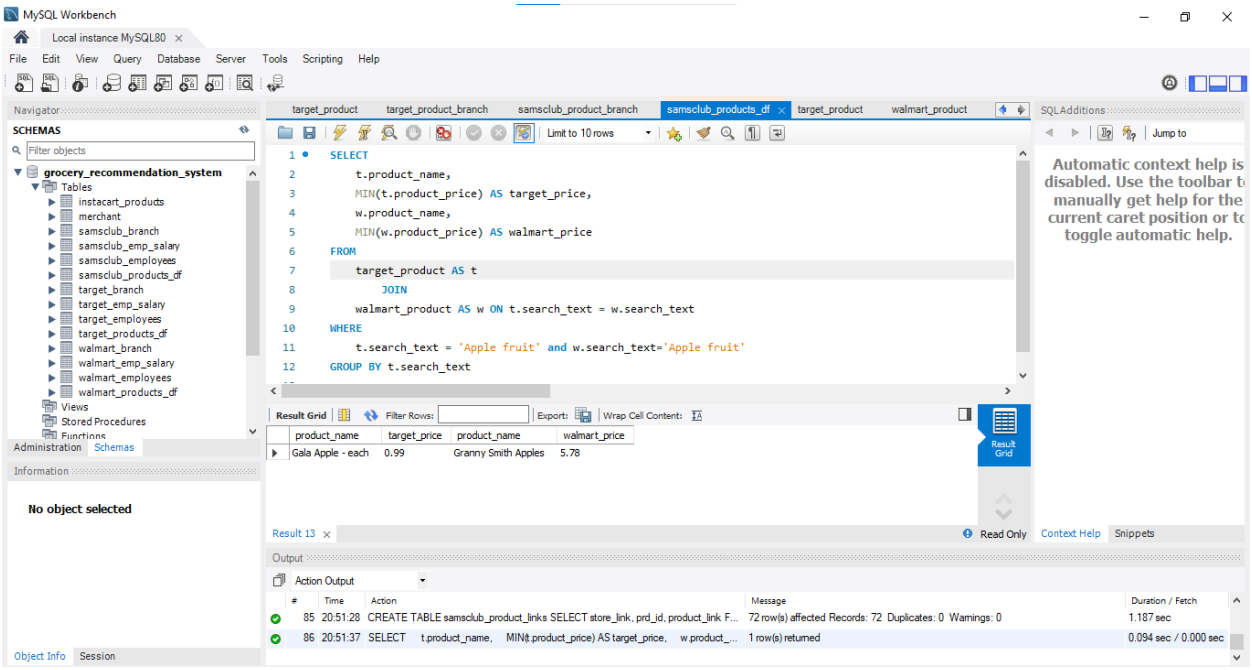


USE CASES

1. **As a customer, in which store can I get the cheapest Apples?**

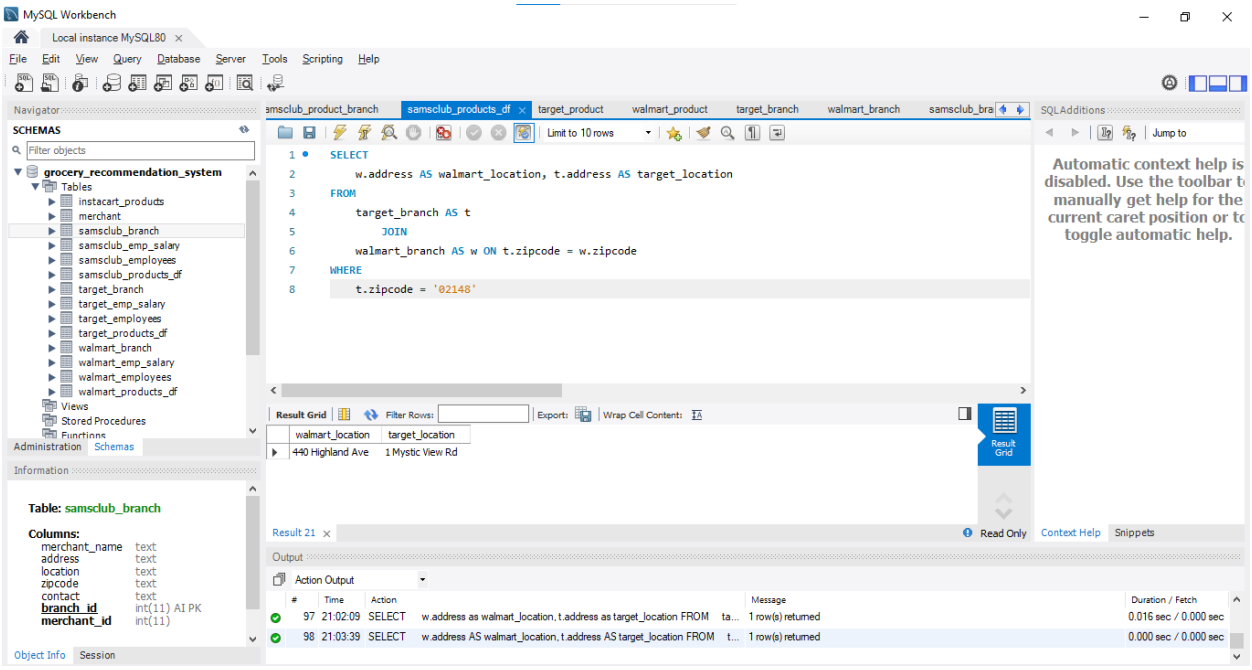
```
SELECT
   t.product_name,
   MIN(t.product_price) AS target_price,
   w.product_name,
   MIN(w.product_price) AS walmart_price
FROM
   target_product AS t
      JOIN
   walmart_product AS w ON t.search_text = w.search_text
WHERE
```

```
    t.search_text = 'Apple fruit' and w.search_text='Apple fruit'
GROUP BY t.search_text;
```



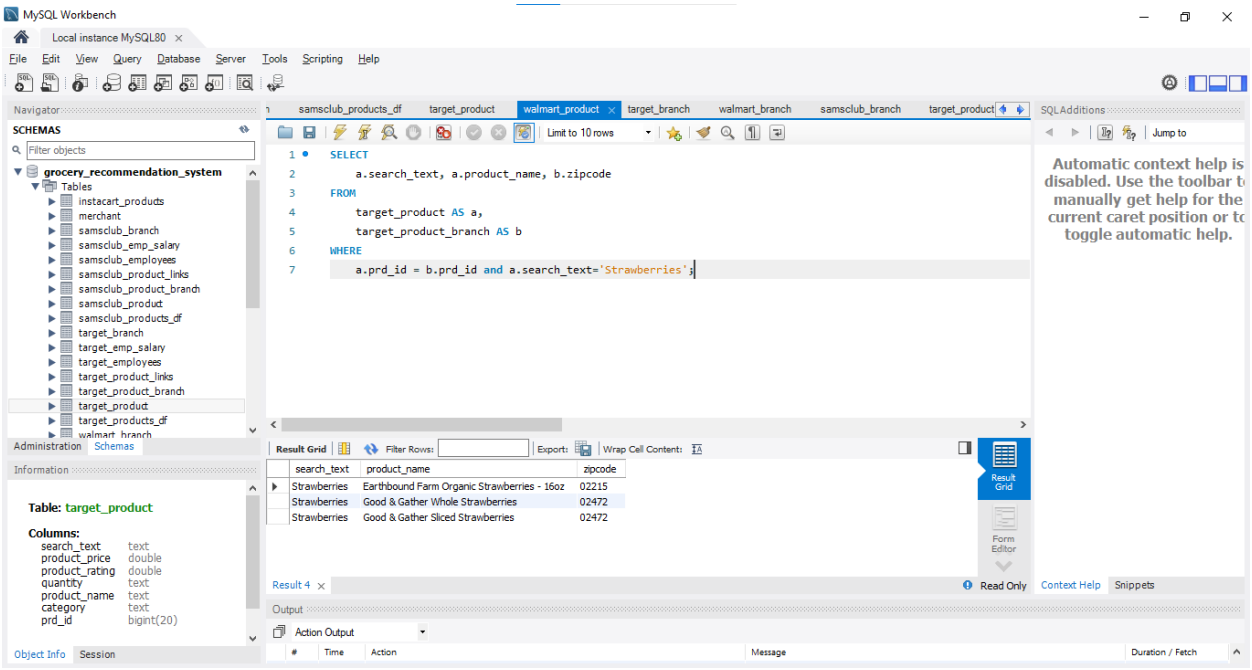
2.Which stores are located in my zipcode (02148)

```
SELECT
   w.address AS walmart_location, t.address AS target_location
FROM
   target_branch AS t
      JOIN
   walmart_branch AS w ON t.zipcode = w.zipcode
WHERE
   t.zipcode = '02148';
```



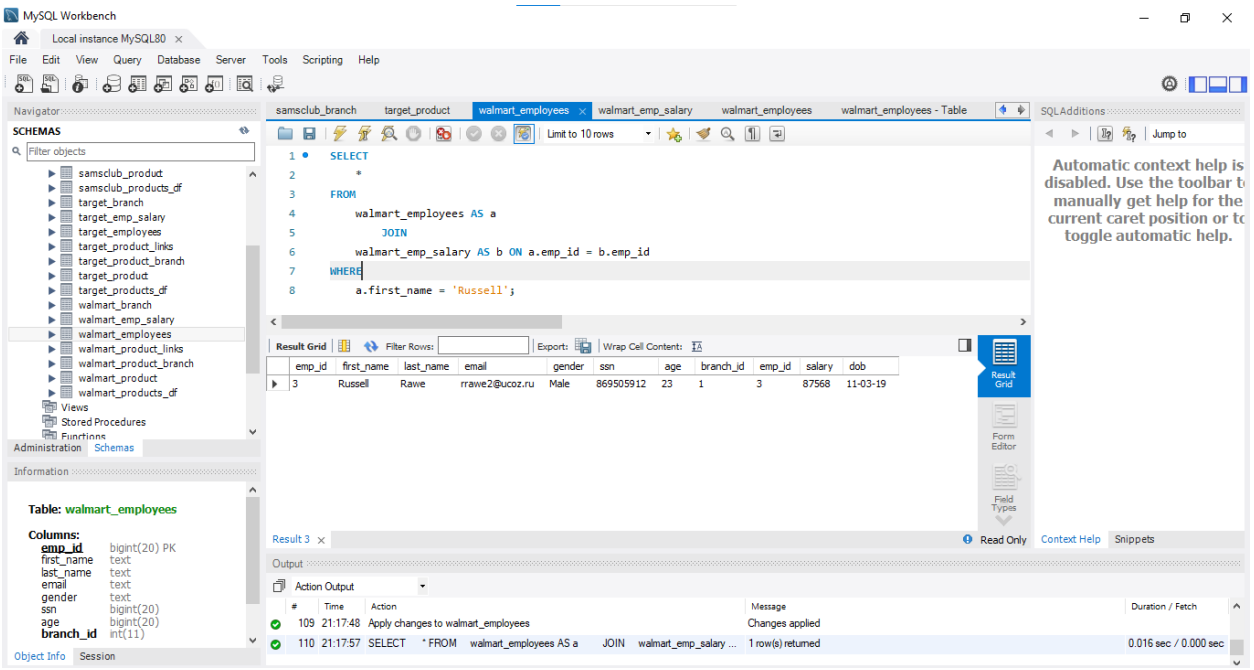
3. As a customer can I know in which Target branch is Strawberry available ?

```
SELECT
   a.search_text, a.product_name, b.zipcode
FROM
   target_product AS a,
   target_product_branch AS b
WHERE
   a.prd_id = b.prd_id and a.search_text='Strawberries';
```



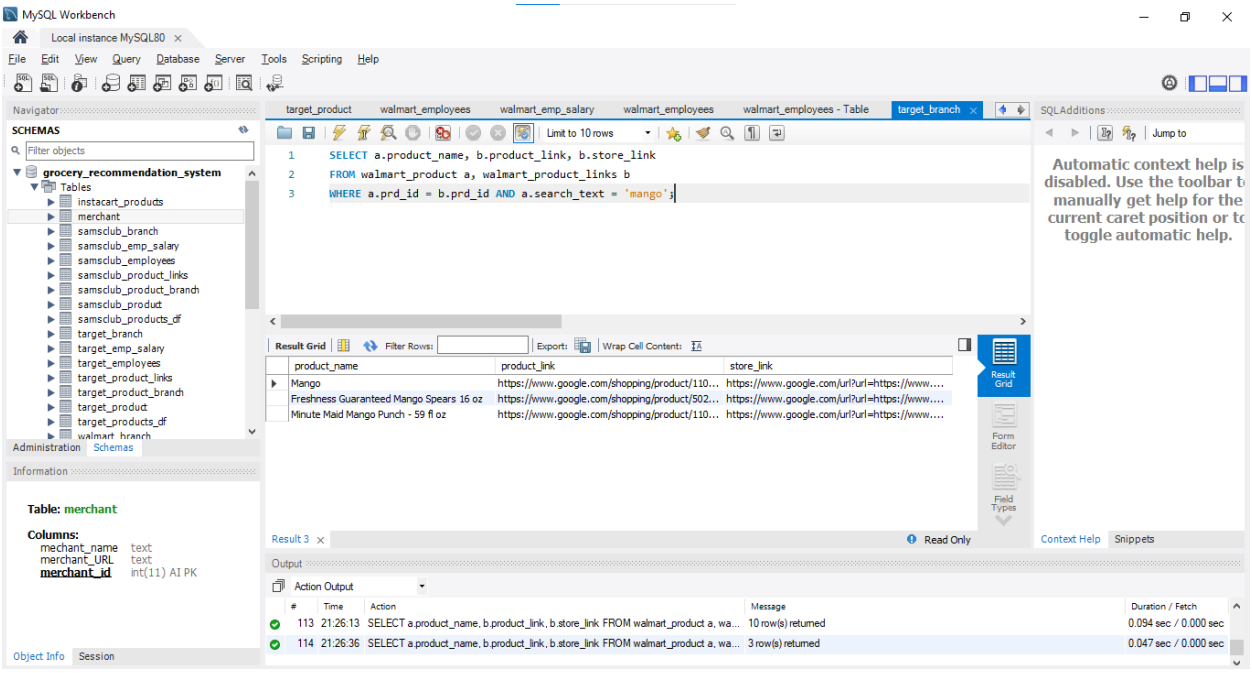
4. As Walmart employee can I see my personal details

```
SELECT
   *
FROM
   walmart_employees AS a
      JOIN
   walmart_emp_salary AS b ON a.emp_id = b.emp_id
WHERE
   a.first_name = 'Russell';
```



5. As a customer, I would like to know the links of Mangoes available in Walmart

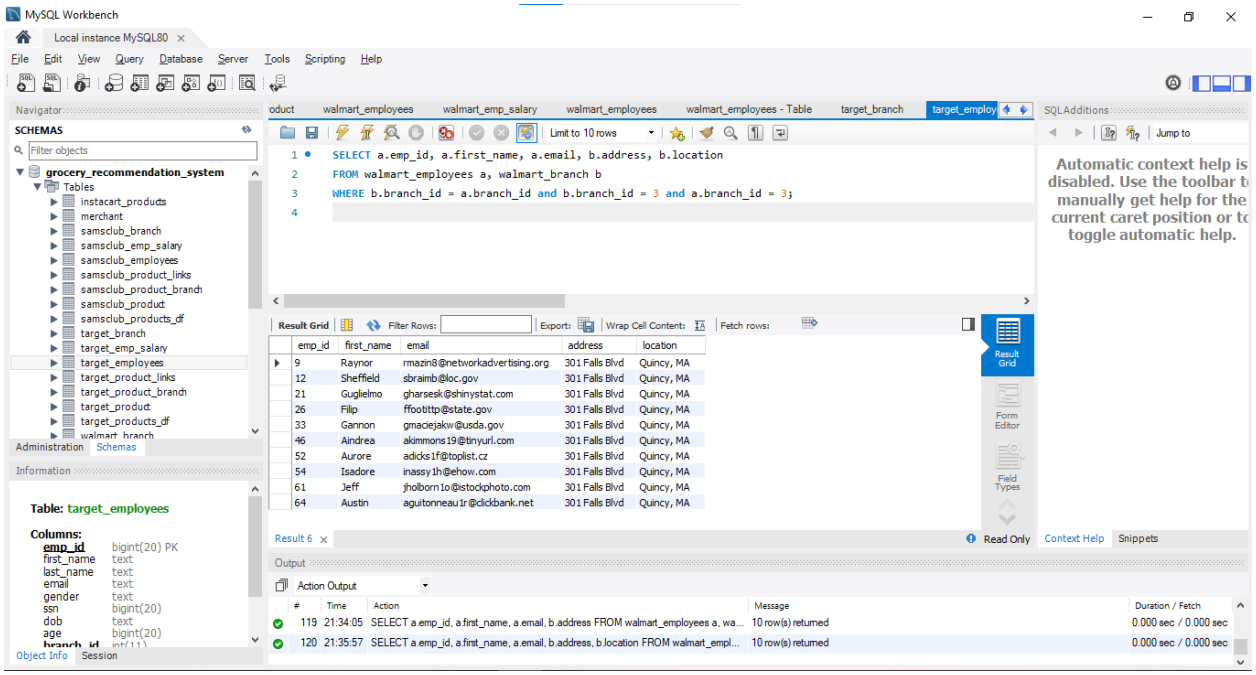
```
SELECT
   a.product_name, b.product_link, b.store_link
FROM
   walmart_product a,
   walmart_product_links b
WHERE
   a.prd_id = b.prd_id
      AND a.search_text = 'mango';
```



6. As an Employee, I would like to know all the Walmart employees who work in Quincy

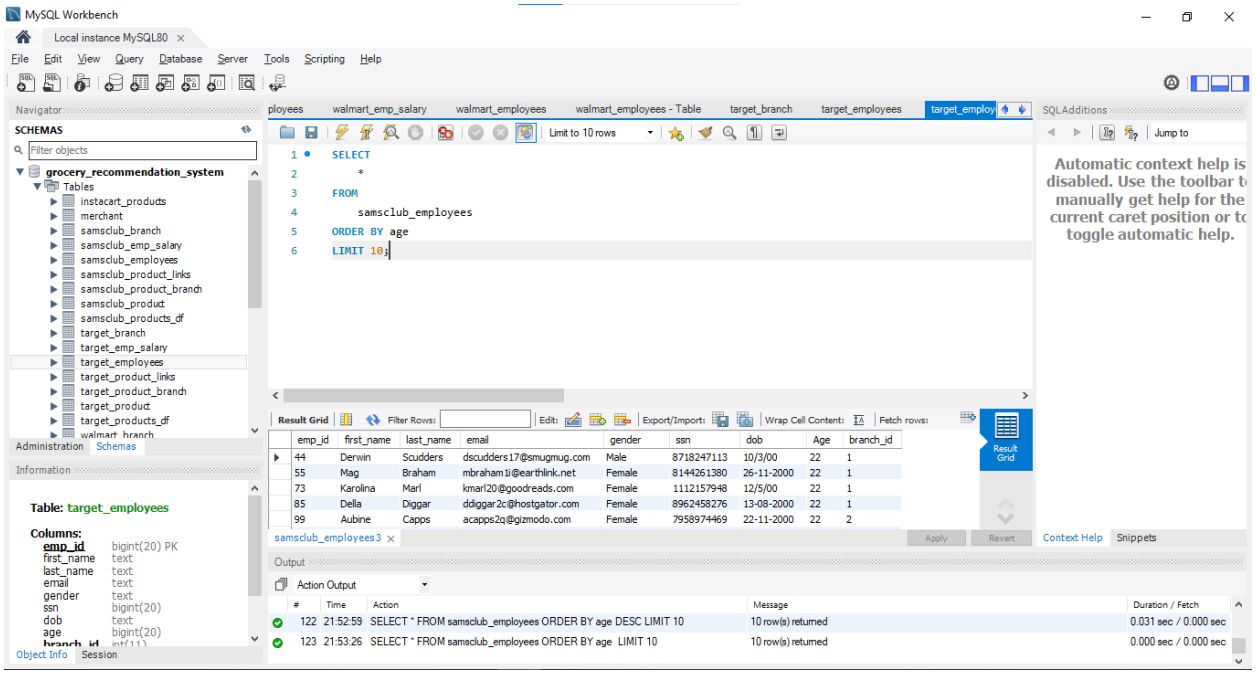
```
SELECT
   a.emp_id, a.first_name, a.email, b.address, b.location
FROM
   walmart_employees a,
   walmart_branch b
WHERE
   b.branch_id = a.branch_id
      AND b.branch_id = 3
```

AND a.branch_id = 3;



7. As Samsclub manager, get top 10 youngest employees

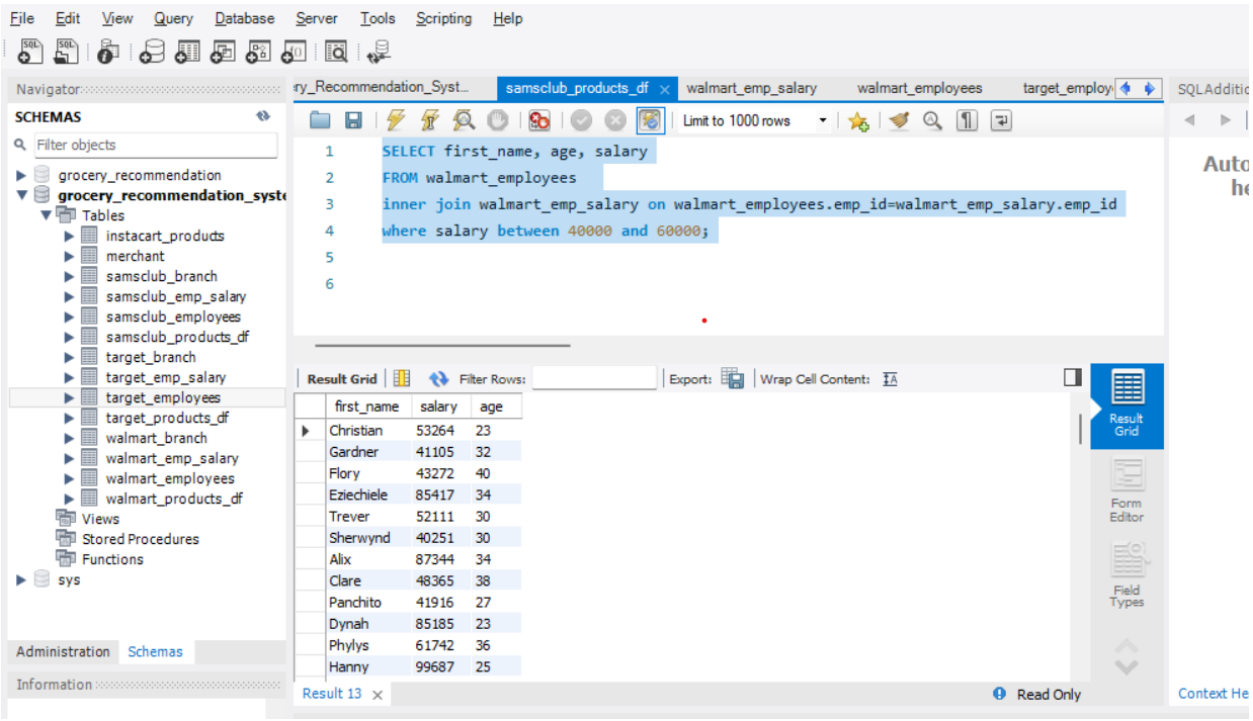
```
SELECT
   *
FROM
   samsclub_employees
ORDER BY age
LIMIT 10;
```



8. Display the first name, age and salary of the walmart employees whose salary is in between $40000 and $60000

```
SELECT
        first_name, age, salary
FROM
        walmart_employees
```

```
        inner join walmart_emp_salary on
    walmart_employees.emp_id=walmart_emp_salary.emp_id
Where
        salary between 40000 and 60000;
```

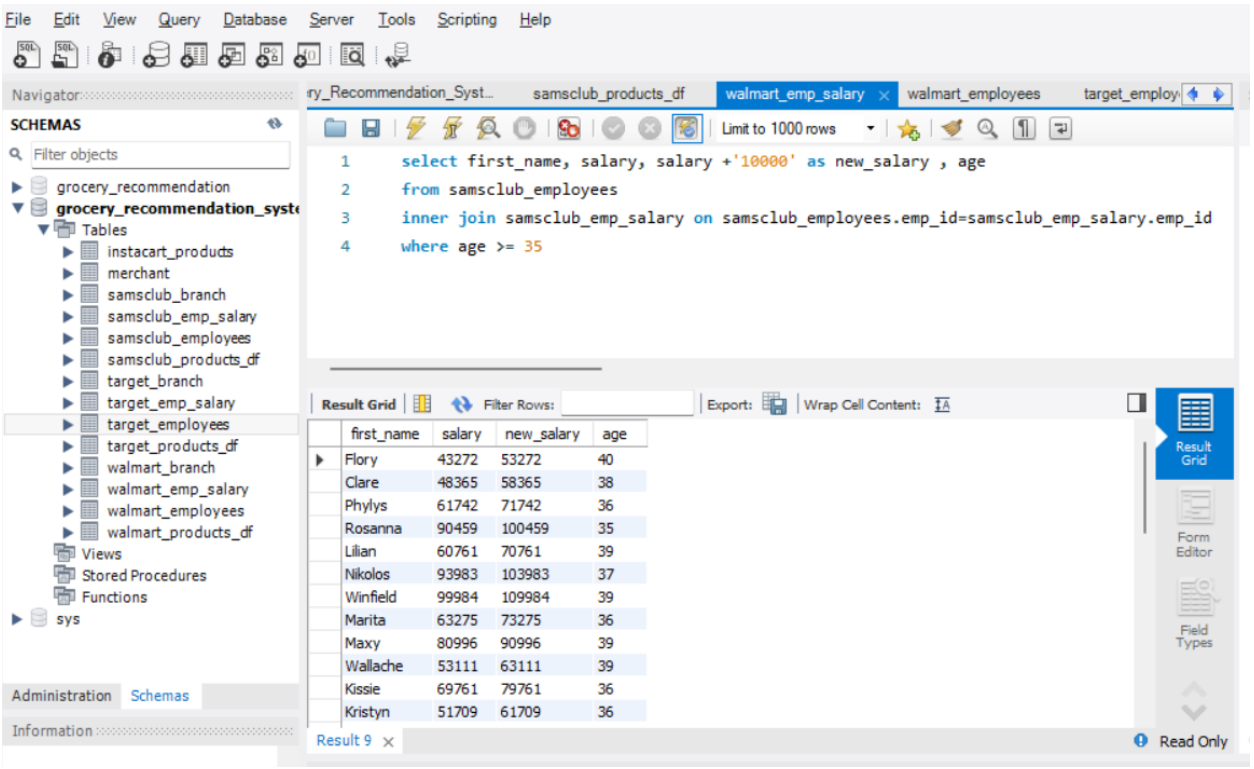


9. Increment the salaries of Sam's club employees by $10000 whose age is 35 years or older

```
Select
        first_name, salary, salary +'10000' as new_salary , age
from
        Samsclub_employees
            inner join samsclub_emp_salary on
        samsclub_employees.emp_id=samsclub_emp_salary.emp_id
```

where

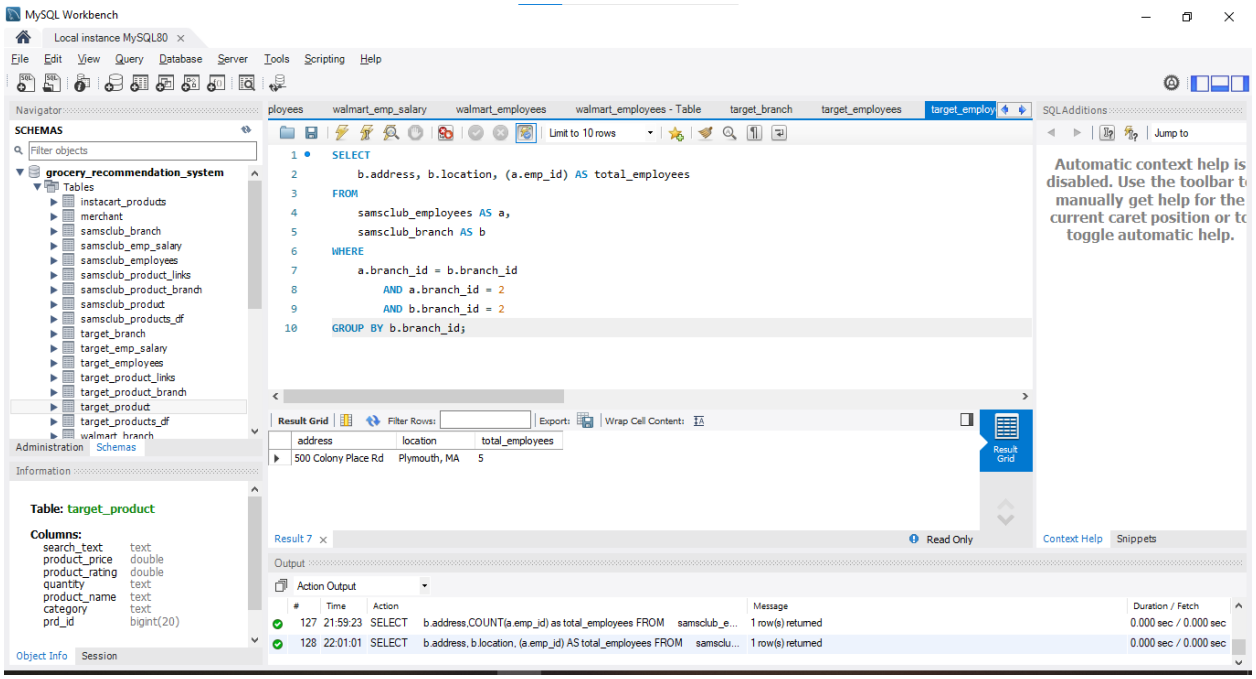
age >= 35;



10. Total number Sam's Club employees working in Plymouth, MA

```
SELECT
   b.address, COUNT(a.emp_id) AS total_employees
FROM
   samsclub_employees AS a,
   samsclub_branch AS b
WHERE
```

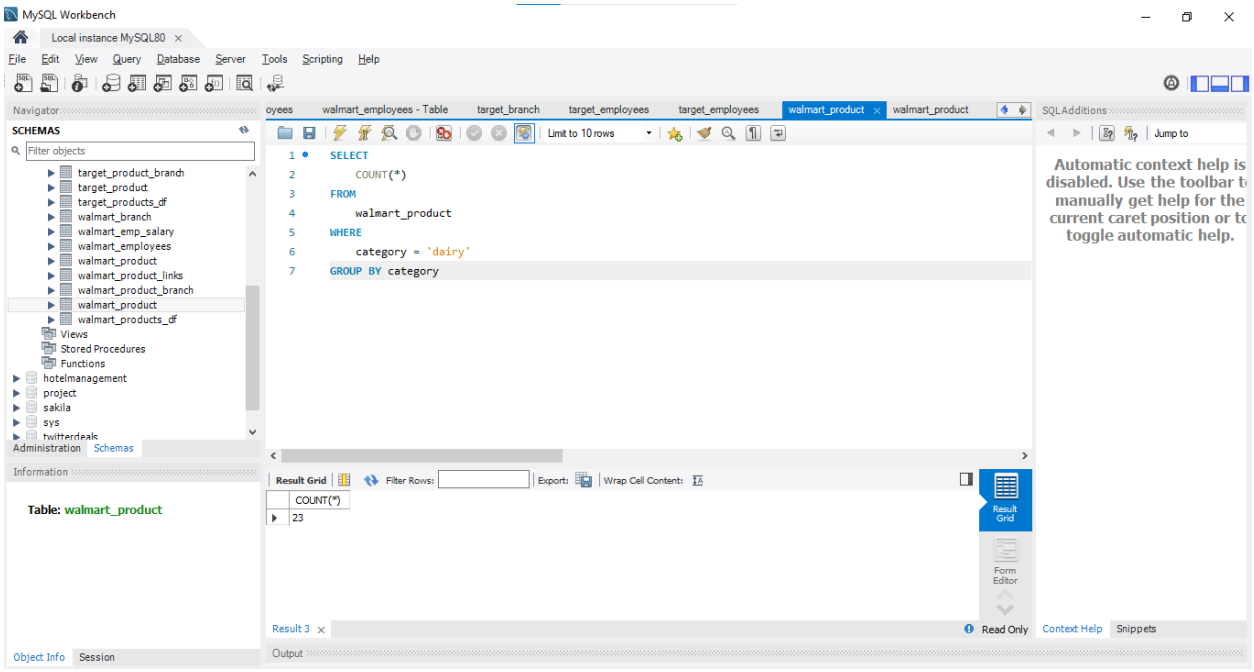
```
    a.branch_id = b.branch_id
        AND a.branch_id = 2
        AND b.branch_id = 2
GROUP BY b.branch_id;
```



11. As a manager can I know total count of Dairy products available all over walmart

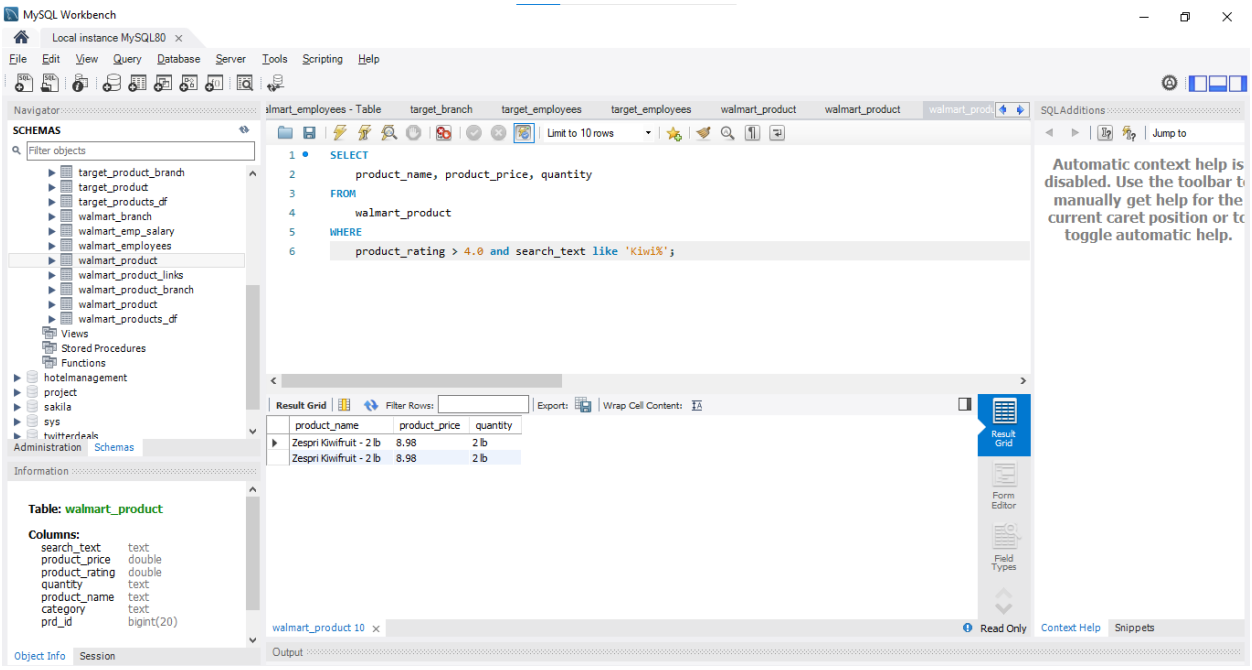
```
SELECT
    COUNT(*)
FROM
    walmart_product
WHERE
```

```
    category = 'dairy'
GROUP BY category;
```



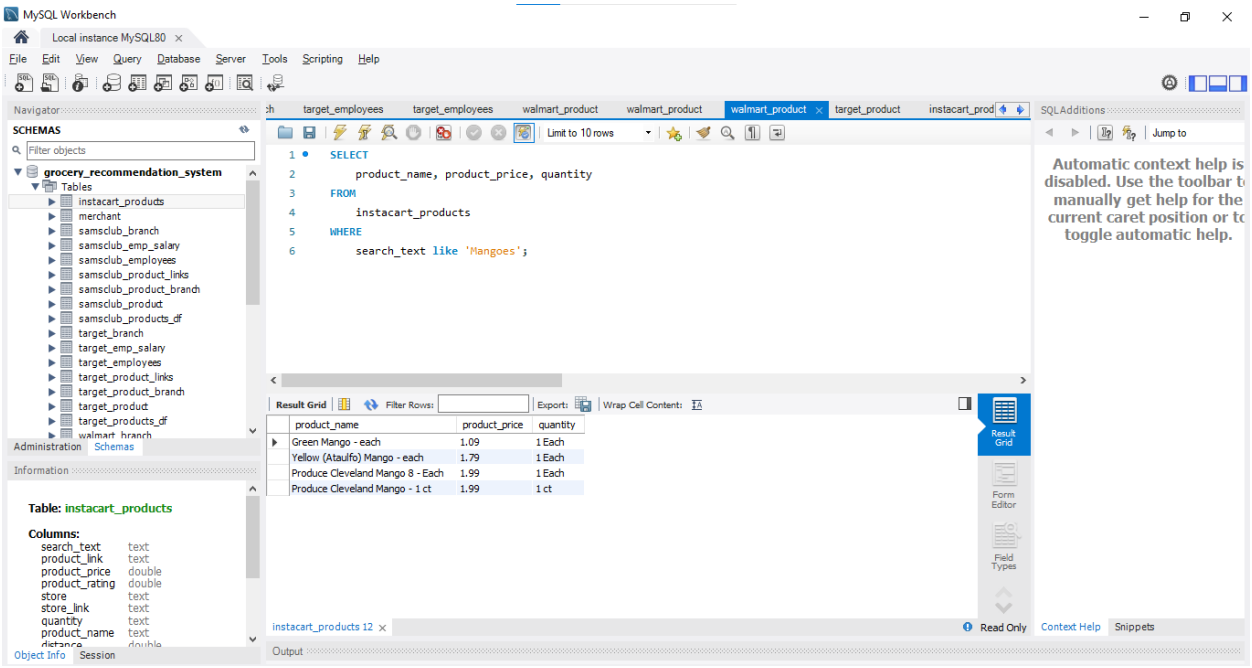
12. As a Walmart customer I would like to order Kiwi fruit having ratings more than 4.

```
SELECT
   product_name, product_price, quantity
FROM
   walmart_product
WHERE
   product_rating > 4.0 and search_text like 'Kiwi%';
```



13. As a customer, can I know what varieties of mangoes that are available in Instacart?

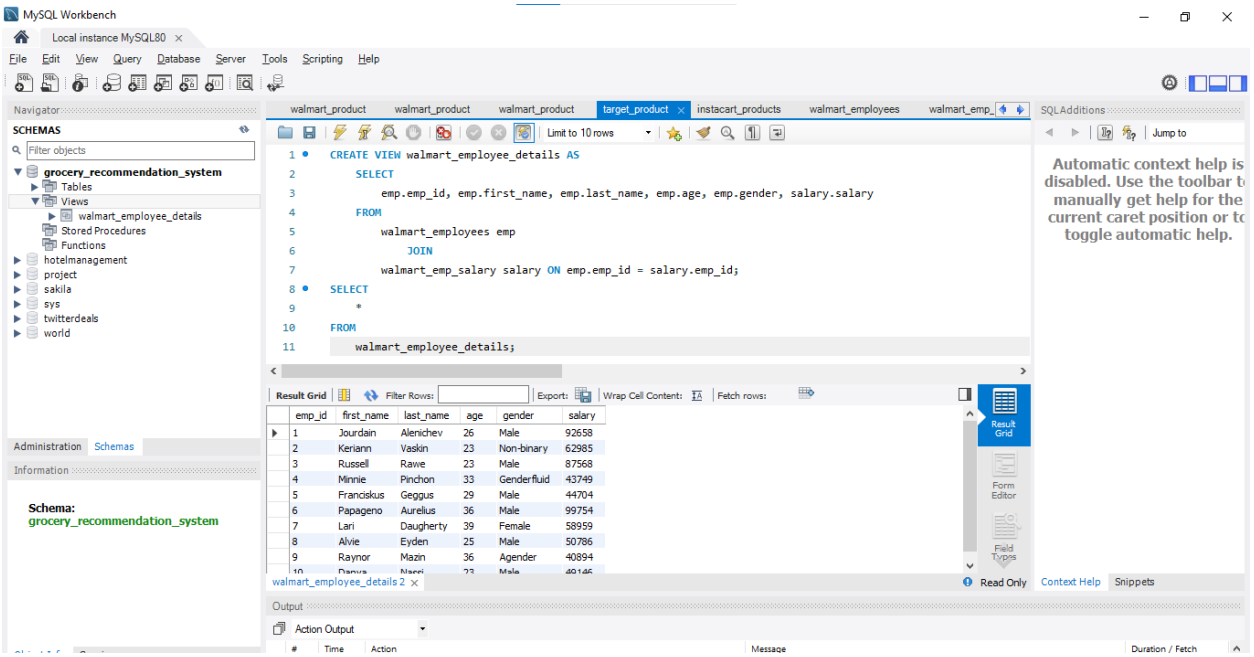
```
SELECT
        product_name, product_price, quantity
FROM
        instacart_products
WHERE
        search_text like 'Mangoes';
```



14. View of Walmart employee table that is visible to only walmart manager

```
CREATE VIEW walmart_employee_details AS
   SELECT
      emp.emp_id, emp.first_name, emp.last_name, emp.age, emp.gender,
salary.salary
   FROM
      walmart_employees emp
         JOIN
      walmart_emp_salary salary ON emp.emp_id = salary.emp_id;
```

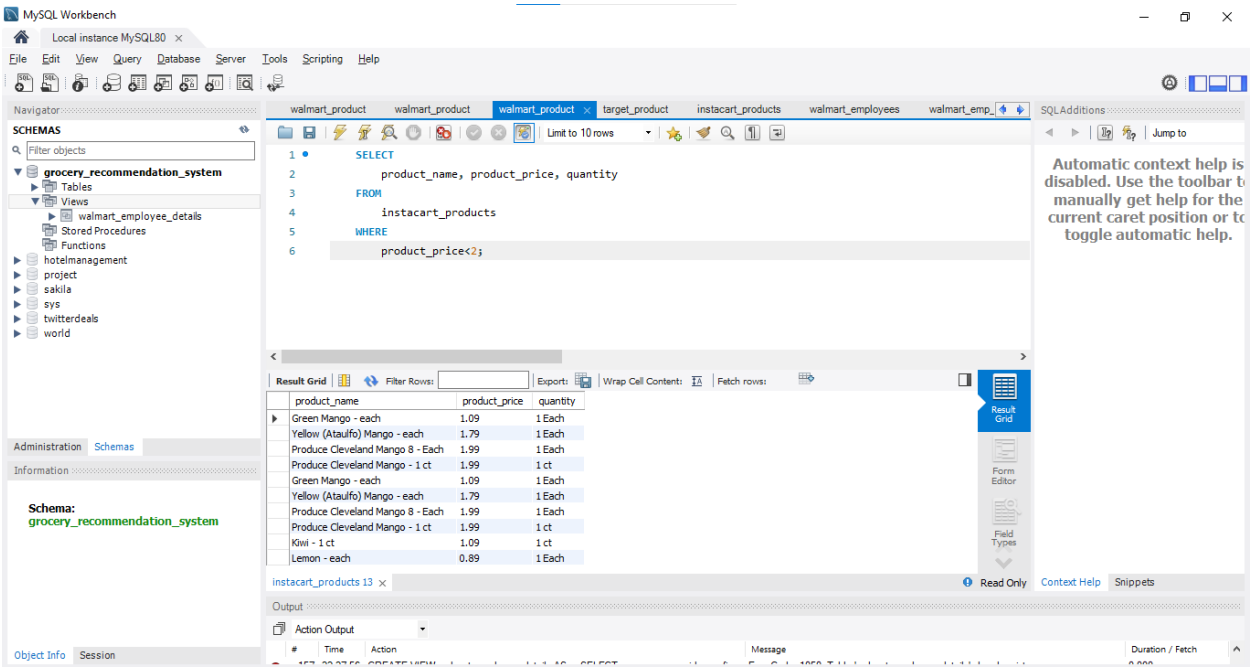
```
SELECT
    *
FROM
    Walmart_employee_details;
```



15. As a customer, can i know products under 2$ which i can order from Instacart

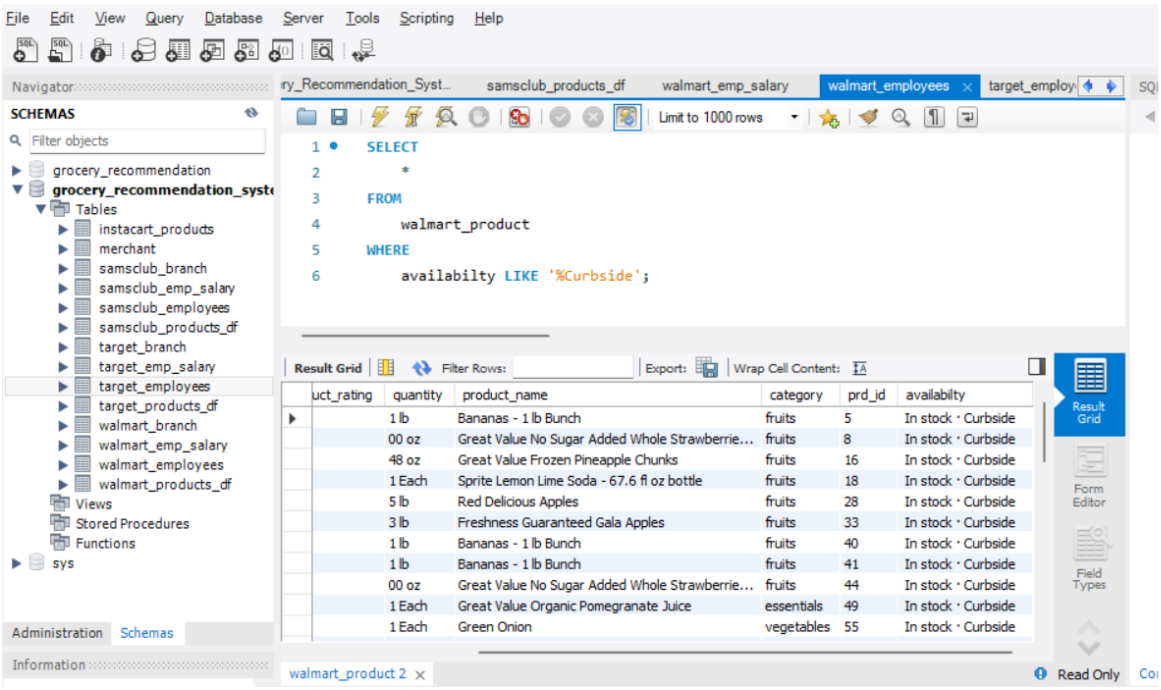
```
SELECT
        product_name, product_price, quantity
FROM
        instacart_products
```

```
    WHERE
        product_price<2;
```



16. May I know products that are available for Pickup from Walmart

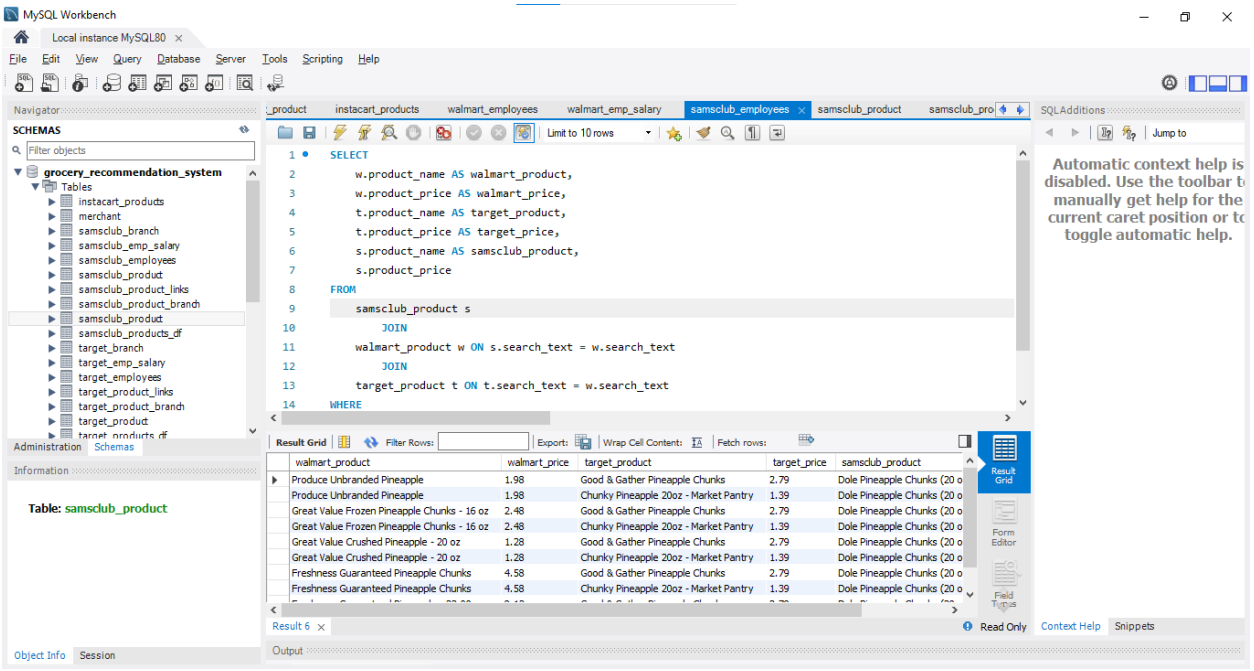
```
SELECT
   *
FROM
   walmart_product
WHERE
   availabilty LIKE '%Curbside';
```



17. As a customer can I get the price of pineapple having ratings above 4

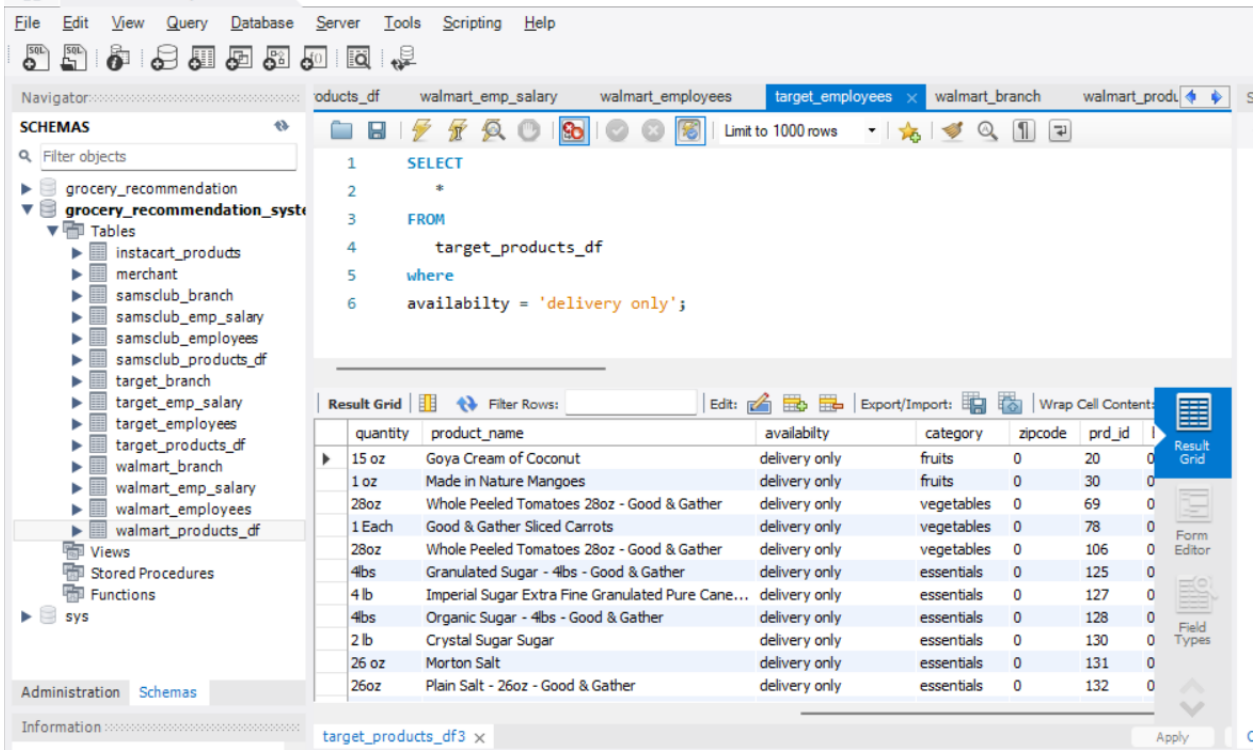
```
SELECT
   w.product_name AS walmart_product,
   w.product_price AS walmart_price,
   t.product_name AS target_product,
   t.product_price AS target_price,
   s.product_name AS samsclub_product,
   s.product_price
FROM
```

```
    samsclub_product s
        JOIN
    walmart_product w ON s.search_text = w.search_text
        JOIN
    target_product t ON t.search_text = w.search_text
WHERE
    s.search_text = 'Pineapple'
        AND s.product_rating > 4 and s.availabilty='delivery only';
```



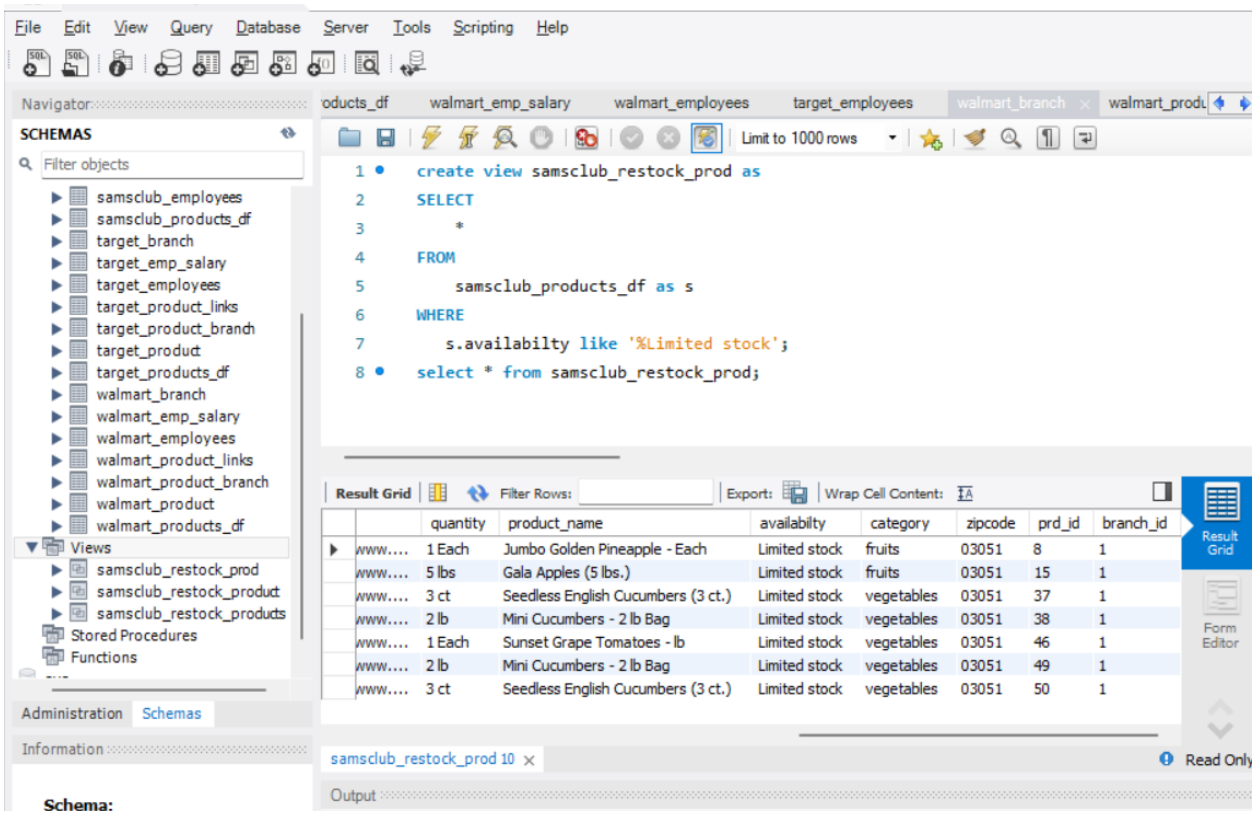
18. As a customer can I know the products available for delivery only in target

```
SELECT
  *
FROM
  target_products_df
where
availabilty = 'delivery only';
```



19. View only the products that has to be restocked for a walmart employee

```
create view samsclub_restock_prod as
SELECT
        *
FROM
   samsclub_products_df as s
WHERE
  s.availabilty like '%Limited stock';
select * from samsclub_restock_prod;
```



20. As a customer i like to know links of product having ratings 4?

```
SELECT
   p.product_name, l.product_link
FROM
   target_product AS p,
   target_product_links AS l
WHERE
   p.prd_id = l.prd_id
```

AND p.product_rating = 4;