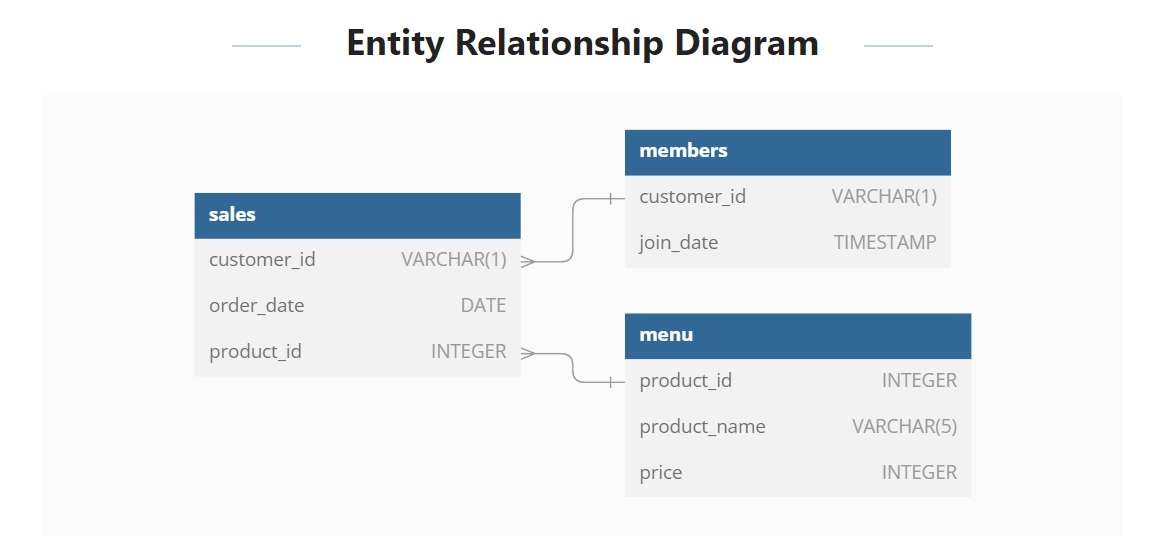
**Case Study #1 - Danny's Diner**

[**https://8weeksqlchallenge.com/case-study-1/**](https://8weeksqlchallenge.com/case-study-1/)



**Problem Statement**

Danny wants to use the data to answer a few simple questions about his customers, especially about their visiting patterns, how much money they’ve spent and also which menu items are their favourite. Having this deeper connection with his customers will help him deliver a better and more personalised experience for his loyal customers.



**Example Datasets:**

**Table1: Sales**

| **customer\_id** | **order\_date** | **product\_id** |
| --- | --- | --- |
| **A** | **2021-01-01** | **1** |
| **A** | **2021-01-01** | **2** |
| **A** | **2021-01-07** | **2** |
| **A** | **2021-01-10** | **3** |
| **A** | **2021-01-11** | **3** |
| **A** | **2021-01-11** | **3** |
| **B** | **2021-01-01** | **2** |
| **B** | **2021-01-02** | **2** |
| **B** | **2021-01-04** | **1** |
| **B** | **2021-01-11** | **1** |
| **B** | **2021-01-16** | **3** |
| **B** | **2021-02-01** | **3** |
| **C** | **2021-01-01** | **3** |
| **C** | **2021-01-01** | **3** |
| **C** | **2021-01-07** | **3** |

**Table 2: Menu**

| **product\_id** | **product\_name** | **price** |
| --- | --- | --- |
| **1** | **sushi** | **10** |
| **2** | **curry** | **15** |
| **3** | **ramen** | **12** |

**Table 3: Members**

| **customer\_id** | **join\_date** |
| --- | --- |
| **A** | **2021-01-07** |
| **B** | **2021-01-09** |

**Case Study Questions**

Each of the following case study questions can be answered using a single SQL statement:

1. What is the total amount each customer spent at the restaurant?
2. How many days has each customer visited the restaurant?
3. What was the first item from the menu purchased by each customer?
4. What is the most purchased item on the menu and how many times was it purchased by all customers?
5. Which item was the most popular for each customer?
6. Which item was purchased first by the customer after they became a member?
7. Which item was purchased just before the customer became a member?
8. What are the total items and amount spent for each member before they became a member?
9. If each $1 spent equates to 10 points and sushi has a 2x points multiplier - how many points would each customer have?
10. In the first week after a customer joins the program (including their join date) they earn 2x points on all items, not just sushi - how many points do customer A and B have at the end of January?

**Schemas and Data**

schema\_sales = StructType(

    [

        StructField("Customer\_ID", StringType(), True),

        StructField("Order\_Date", DateType(), True),

        StructField("Product\_ID", IntegerType(), True),

    ]

)

schema\_menu = StructType(

    [

        StructField("Product\_ID", IntegerType(), True),

        StructField("Product\_Name", StringType(), True),

        StructField("Price", IntegerType(), True),

    ]

)

schema\_members = StructType(

    [

        StructField("Customer\_ID", StringType(), True),

        StructField("Join\_Date", DateType(), True),

    ]

)

sales\_data=[('A',datetime.date(2021,1,1),1),

            ('A',datetime.date(2021,1,1),2),

            ('A',datetime.date(2021,1,7),2),

            ('A',datetime.date(2021,1,10),3),

            ('A',datetime.date(2021,1,11),3),

            ('A',datetime.date(2021,1,11),3),

            ('B',datetime.date(2021,1,1),2),

            ('B',datetime.date(2021,1,2),2),

            ('B',datetime.date(2021,1,4),1),

            ('B',datetime.date(2021,1,11),1),

            ('B',datetime.date(2021,1,16),3),

            ('B',datetime.date(2021,2,1),3),

            ('C',datetime.date(2021,1,1),3),

            ('C',datetime.date(2021,1,1),3),

            ('C',datetime.date(2021,1,7),3)]

menu\_data=([1,'Sushi',10],[2,'Curry',15],[3,'Ramen',12])

members\_data=[('A',datetime.date(2021,1,7)),('B',datetime.date(2021,1,9))]

**Creating the DataFrames**

df\_sales=spark.createDataFrame(sales\_data,schema\_sales)

df\_menu=spark.createDataFrame(menu\_data,schema\_menu)

df\_members=spark.createDataFrame(members\_data,schema\_members)

**Import**

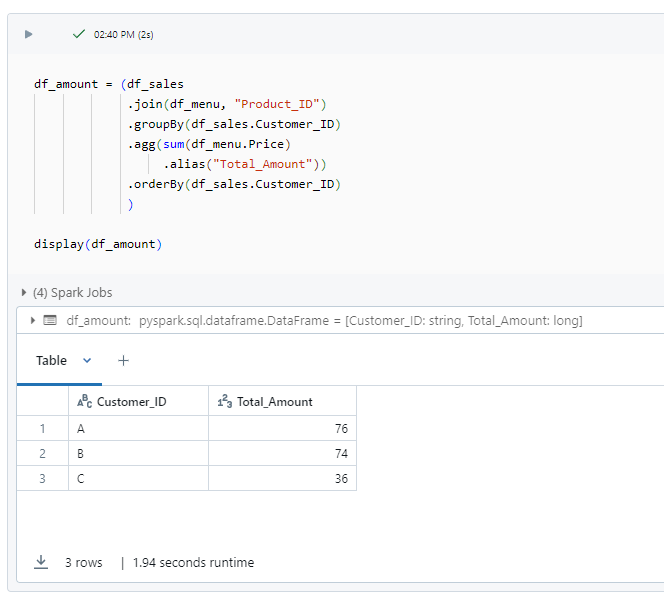
from pyspark.sql.types import \*

from pyspark.sql.functions import \*

from pyspark.sql.window import \*

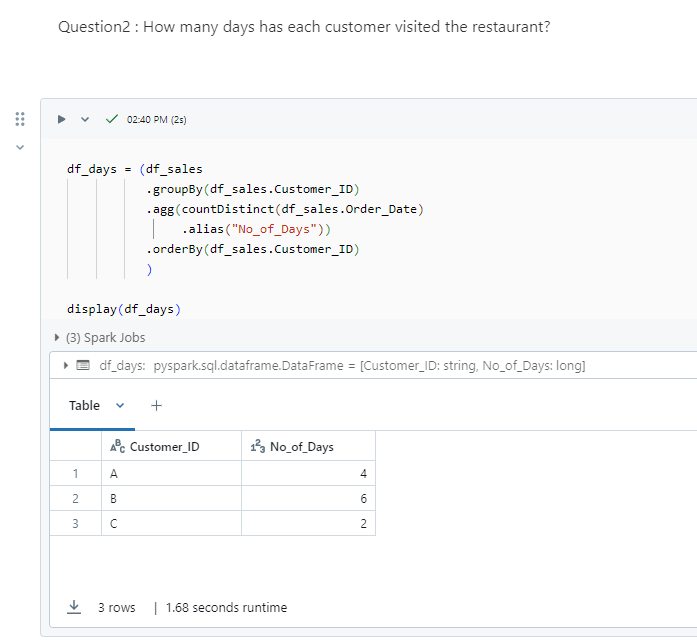
import datetime

**Question1 : What is the total amount each customer spent at the restaurant?**

****

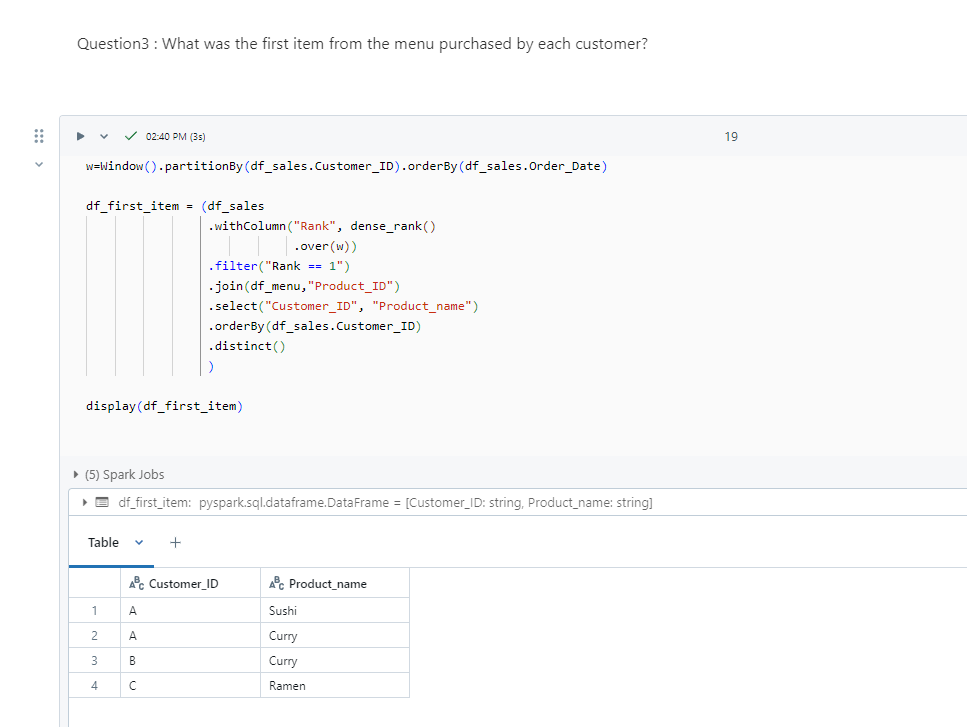
* Customer A spent 76.
* Customer B spent 74.
* Customer C spent 36.

**Question2 : How many days has each customer visited the restaurant?**

****

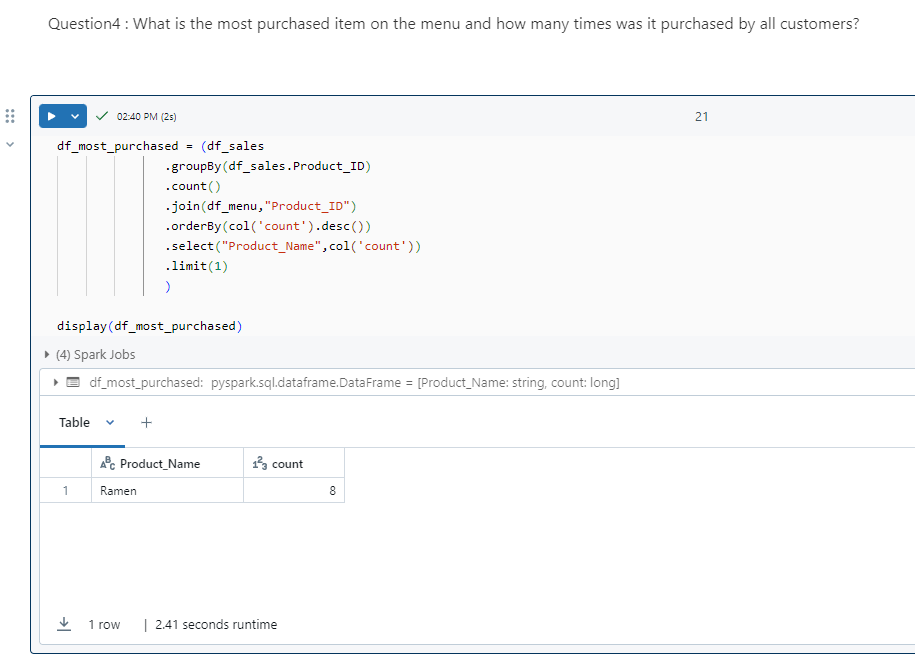
* Customer A visited 4 times.
* Customer B visited 6 times.
* Customer C visited 2 times.

**Question3 : What was the first item from the menu purchased by each customer?**

****

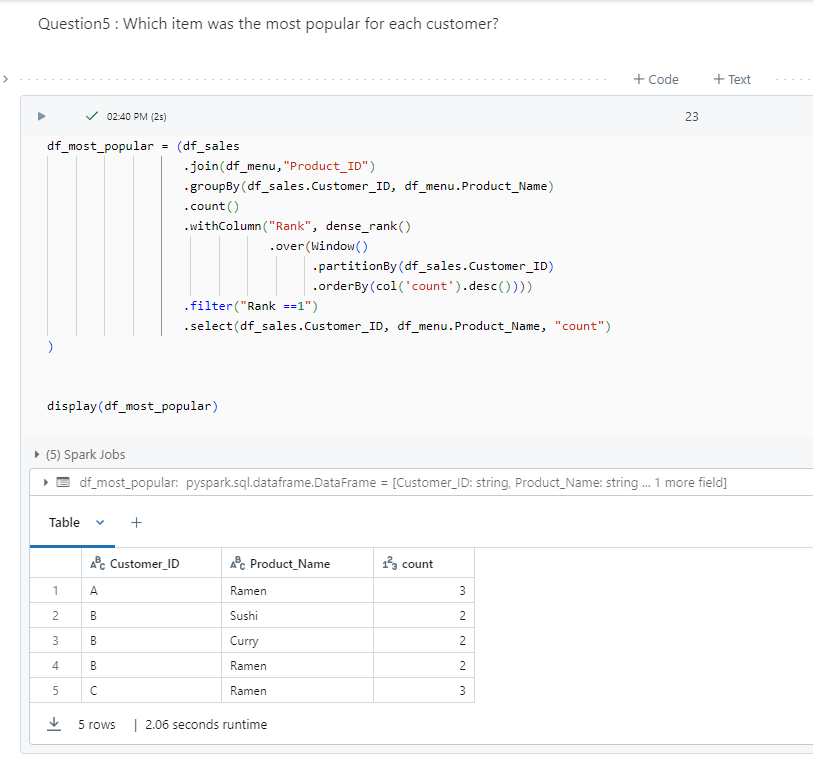
* Customer A's first orders are Curry and Sushi
* Customer B's first order is Curry.
* Customer C's first order is Ramen

**Question4 : What is the most purchased item on the menu and how many times was it purchased by all customers?**

****

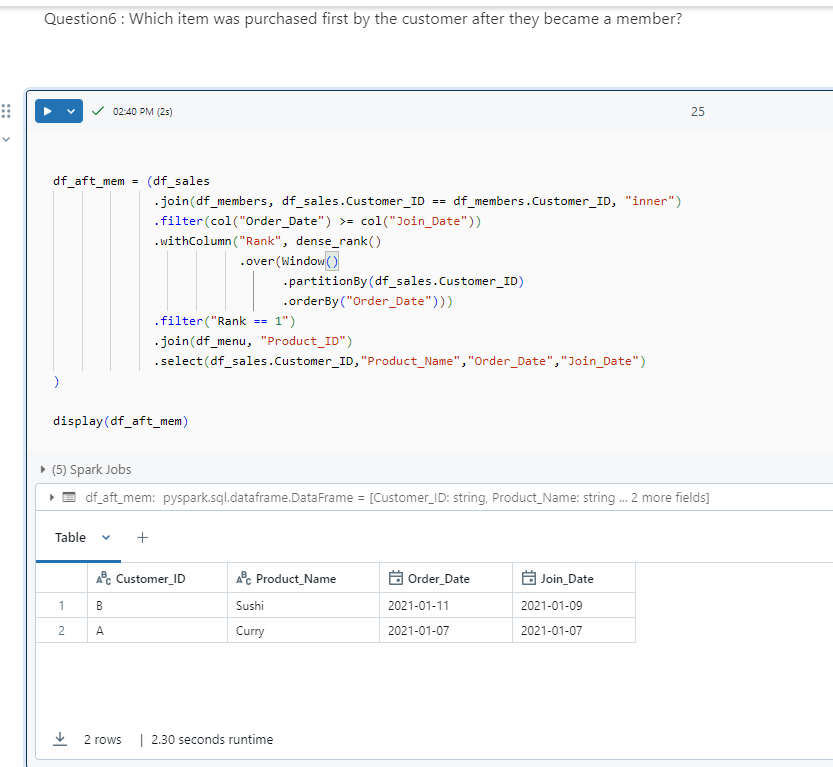
Most purchased item on the menu is Ramen which is **8 times**.

**Question5 : Which item was the most popular for each customer?**

****

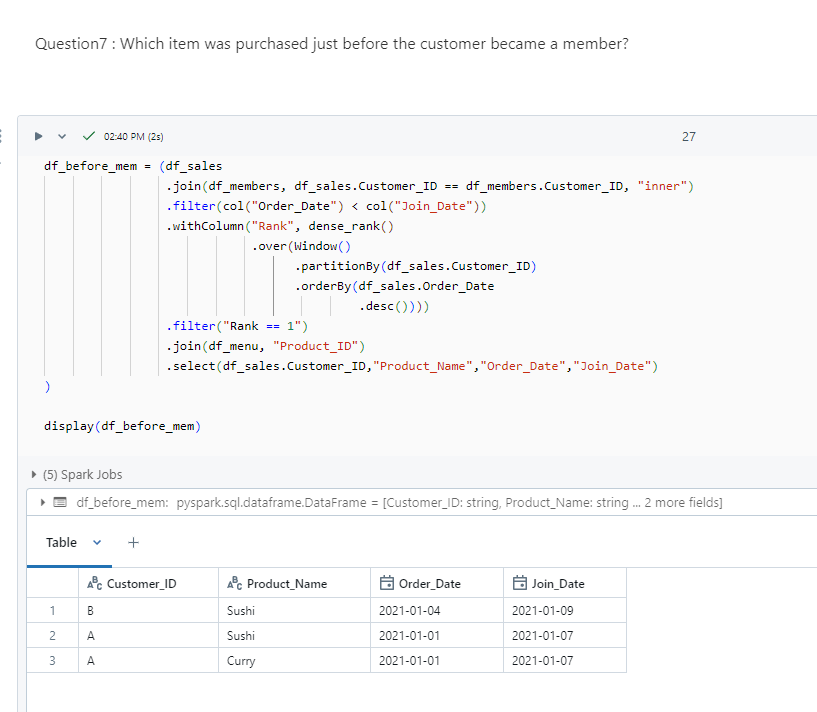
* Customer A and C's favourite item is Ramen.
* Customer B enjoys all items on the menu.

**Question6 : Which item was purchased first by the customer after they became a member?**

****

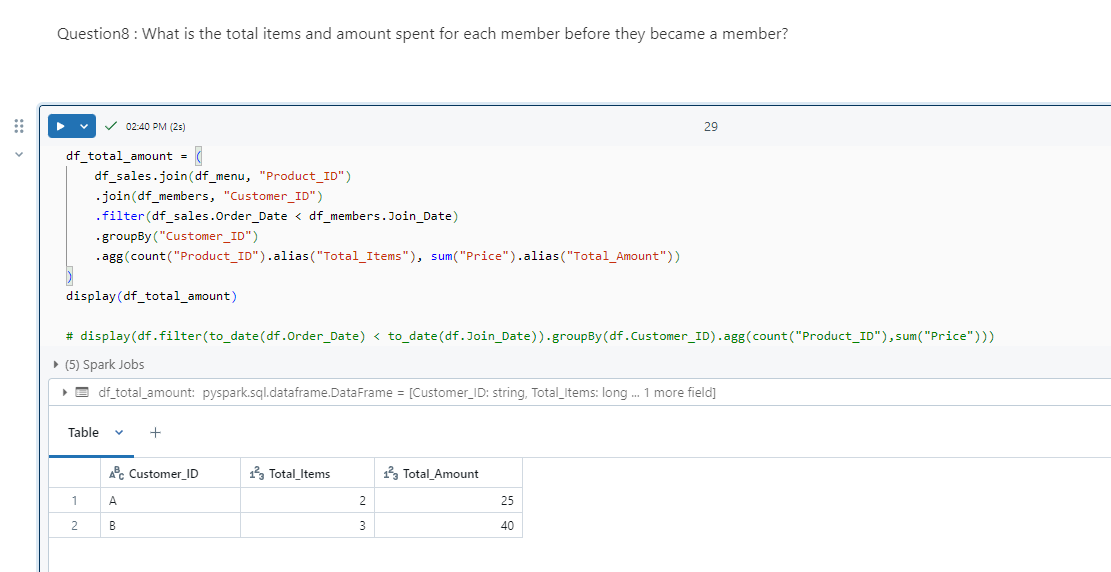
* Customer A's first order as member is Curry.
* Customer B's first order as member is Sushi.

**Question7 : Which item was purchased just before the customer became a member?**

****

* Customer A’s last order before becoming a member is Sushi and Curry.
* However, for Customer B, it's Sushi.

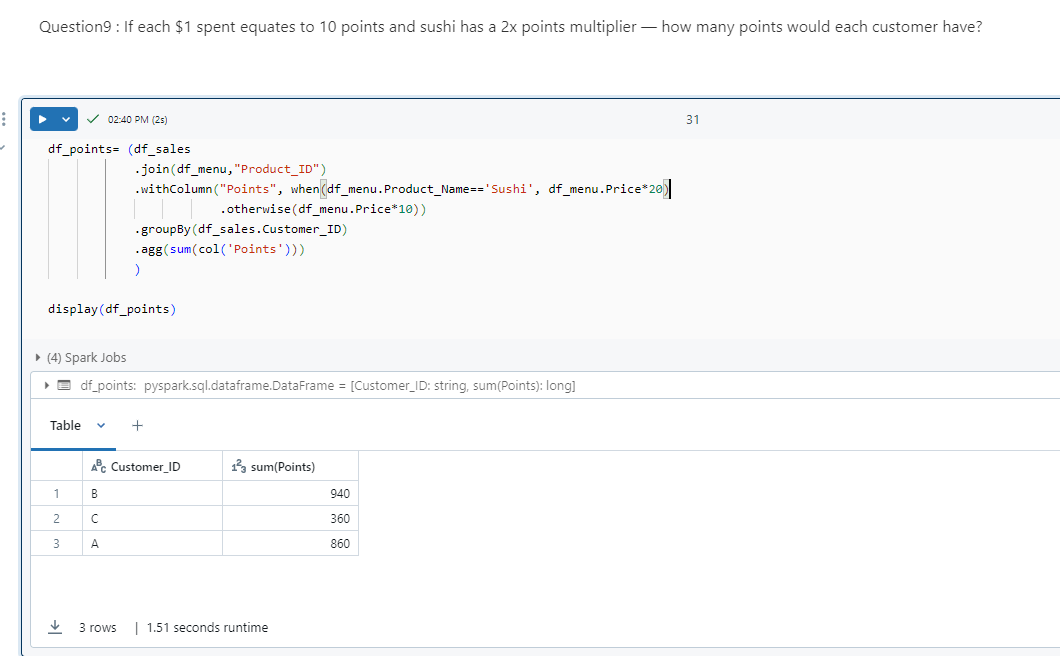
**Question8 : What is the total items and amount spent for each member before they became a member?**

****

Before becoming members,

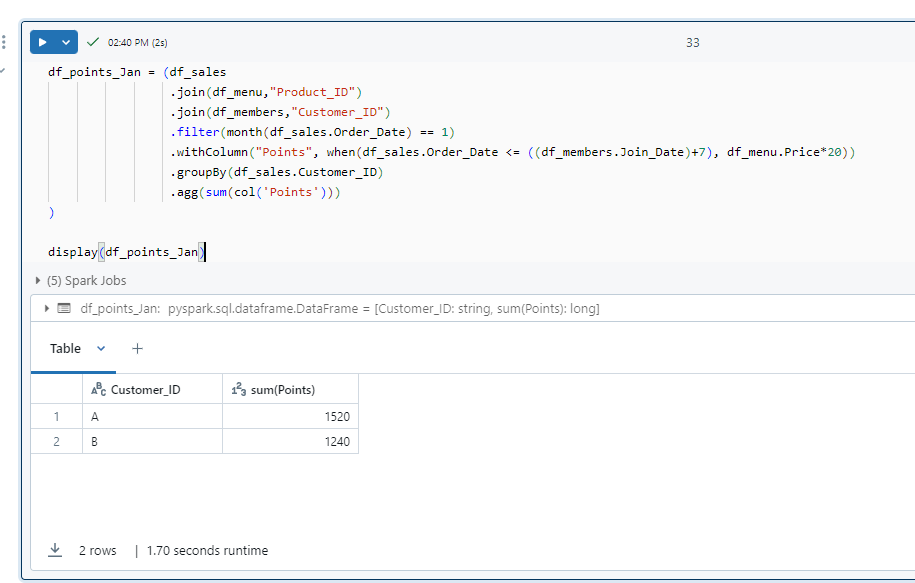
* Customer A spent 25 on 2 items.
* Customer B spent 40 on 3 items

**Question9 : If each $1 spent equates to 10 points and sushi has a 2x points multiplier — how many points would each customer have?**

****

* Total points for Customer A is 860.
* Total points for Customer B is 940.
* Total points for Customer C is 360.

**Question10 : In the first week after a customer joins the program (including their join date) they earn 2x points on all items, not just sushi — how many points do customer A and B have at the end of January?**

****

* Total points for Customer A is 1520.
* Total points for Customer B is 1240.