package org.itc.com  
import org.apache.spark.SparkConf  
import org.apache.spark.sql.SparkSession  
import org.apache.spark.sql.expressions.Window  
import org.apache.spark.sql.functions.\_  
import org.apache.spark.sql.types.\_  
  
object Main extends App {  
 val *sparkconf* = new SparkConf()  
 *sparkconf*.set("spark.app.name", "firstDFDemo")  
 *sparkconf*.set("spark.master", "local[1]")  
 val *spark* = SparkSession.*builder*().config(*sparkconf*).getOrCreate()  
  
 val *salesSchema* = "customer\_id string, order\_date date, product\_id int"  
 var *salesdf* = *spark*.read.option("header", true).schema(*salesSchema*).csv(path = "D:/sales.csv")  
 //salesdf.show()  
  
 val *membersSchema* = "customer\_id string, join\_date date"  
 var *membersdf* = *spark*.read.option("header", true).schema(*membersSchema*).csv(path = "D:/members.csv")  
 //membersdf.show()  
  
 val *menuSchema* = "product\_id int, product\_name string,price int"  
 var *menudf* = *spark*.read.option("header", true).schema(*menuSchema*).csv(path = "D:/menu.csv")  
 //menudf.show()  
  
 *salesdf*.createOrReplaceTempView("tblSales")  
 *membersdf*.createOrReplaceTempView("tblmembers")  
 *menudf*.createOrReplaceTempView("tblmenu")  
  
  
 //7. Which item was purchased just before the customer became a member?  
 val *DF1* = *spark*.sql("""SELECT s.customer\_id,men.product\_name AS product\_name,MAX(s.order\_date) AS last\_purchase\_date\_before\_join  
 FROM tblSales s  
 INNER JOIN tblmembers m ON s.customer\_id = m.customer\_id  
 INNER JOIN tblmenu men ON men.product\_id = s.product\_id  
 WHERE s.order\_date < m.join\_date  
 GROUP BY s.customer\_id, men.product\_name  
 ORDER BY customer\_id""".stripMargin)  
 *DF1*.show()  
  
 //8.What is the total items and amount spent for each member before they became a member?  
 val *DF2* = *spark*.sql("""SELECT s.customer\_id,COUNT(\*) AS total\_items,SUM(m.price) AS total\_amount\_spent  
 FROM tblSales s  
 INNER JOIN tblmenu m ON s.product\_id = m.product\_id  
 INNER JOIN tblmembers mb ON s.customer\_id = mb.customer\_id  
 WHERE s.order\_date < mb.join\_date  
 GROUP BY s.customer\_id""".stripMargin)  
 *DF2*.show()  
  
 //9 If each $1 spent equates to 10 points and sushi has a 2x points multiplier - how many points would each customer have?  
 val *DF3* = *spark*.sql("""WITH PurchasePoints AS (  
 | SELECT s.customer\_id,  
 | SUM(  
 | CASE  
 | WHEN m.product\_name = 'sushi' THEN 20 \* m.price  
 | ELSE 10 \* m.price  
 | END  
 | ) AS total\_points  
 | FROM tblSales s  
 | INNER JOIN tblmenu m ON s.product\_id = m.product\_id  
 | GROUP BY s.customer\_id  
 |)  
 |SELECT customer\_id,COALESCE(total\_points, 0) AS total\_points  
 |FROM  
 | PurchasePoints  
 |""".stripMargin)  
*DF3*.show()  
  
  
  
 //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?  
 val *DF4* = *spark*.sql(  
 """SELECT s.customer\_id,  
 SUM(  
 CASE  
 WHEN s.order\_date >= mb.join\_date AND s.order\_date < DATE\_ADD(mb.join\_date, 7) THEN 20 \* m.price -- First week after joining  
 ELSE 10 \* m.price  
 END  
 ) AS total\_points  
 FROM tblSales s  
 INNER JOIN tblMenu m ON s.product\_id = m.product\_id  
 INNER JOIN tblmembers mb ON s.customer\_id = mb.customer\_id  
 WHERE s.order\_date <= '2021-01-31' -- End of January  
 GROUP BY s.customer\_id  
 """  
 )  
 *DF4*.show()  
  
  
}

+-----------+------------+------------------------------+

|customer\_id|product\_name|last\_purchase\_date\_before\_join|

+-----------+------------+------------------------------+

| A| curry| 2021-01-01|

| A| sushi| 2021-01-01|

| B| sushi| 2021-01-04|

| B| curry| 2021-01-02|

+-----------+------------+------------------------------+

+-----------+-----------+------------------+

|customer\_id|total\_items|total\_amount\_spent|

+-----------+-----------+------------------+

| B| 3| 40|

| A| 2| 25|

+-----------+-----------+------------------+

+-----------+------------+

|customer\_id|total\_points|

+-----------+------------+

| B| 940|

| C| 360|

| A| 860|

+-----------+------------+

+-----------+------------+

|customer\_id|total\_points|

+-----------+------------+

| B| 720|

| A| 1270|

+-----------+------------+