Big Data Midterm (Fall’23)

Q3: Cosmos DB

Look at the format of the json file (ecommerce.json in Q3 directory). Use the same format and create at least 5 such customers for your DB with different combinations of orders and cart items. Then answer the following questions using Cosmos DB queries. (Make sure that the json file is created accordingly so that each of the following questions will have an answer).

JSON File:

[

{

"customerId": "customer456",

"firstName": "John",

"lastName": "Doe",

"email": "john.doe@example.com",

"address": {

"street": "123 Main St",

"city": "New York",

"state": "NY",

"zipCode": "10001"

},

"orders": [

{

"orderId": "order123",

"orderDate": "2022-10-06T10:00:00Z",

"items": [

{

"productId": "product789",

"productName": "Widget",

"quantity": 2,

"price": 10.99

},

{

"productId": "product101",

"productName": "Gadget",

"quantity": 1,

"price": 29.99

}

]

},

{

"orderId": "order234",

"orderDate": "2023-10-06T10:00:00Z",

"items": [

{

"productId": "product303",

"productName": "Dress",

"quantity": 3,

"price": 70.99

},

{

"productId": "product404",

"productName": "Shoes",

"quantity": 1,

"price": 140

},

{

"productId": "product788",

"productName": "Widget",

"quantity": 2,

"price": 30.99

}

]

}

],

"cart": [

{

"productId": "product789",

"productName": "Widget",

"quantity": 2,

"price": 10.99

},

{

"productId": "product101",

"productName": "Gadget",

"quantity": 1,

"price": 29.99

},

{

"productId": "product303",

"productName": "Dress",

"quantity": 3,

"price": 70.99

},

{

"productId": "product404",

"productName": "Shoes",

"quantity": 1,

"price": 140

},

{

"productId": "product505",

"productName": "Banana",

"quantity": 1,

"price": 1

}

],

"id": "21ff33e3-f1ab-4d3e-803e-6b99f757a289",

"\_rid": "Hyo5AKaGRdoBAAAAAAAAAA==",

"\_self": "dbs/Hyo5AA==/colls/Hyo5AKaGRdo=/docs/Hyo5AKaGRdoBAAAAAAAAAA==/",

"\_etag": "\"d800f949-0000-0700-0000-653d63970000\"",

"\_attachments": "attachments/",

"\_ts": 1698522007

},

{

"customerId": "customer1",

"firstName": "Alice",

"lastName": "Smith",

"email": "alice.smith@example.com",

"address": {

"street": "456 Elm St",

"city": "Los Angeles",

"state": "CA",

"zipCode": "90001"

},

"orders": [

{

"orderId": "order1",

"orderDate": "2022-05-15T14:30:00Z",

"items": [

{

"productId": "product101",

"productName": "Gadget",

"quantity": 2,

"price": 25.99

},

{

"productId": "product202",

"productName": "Widget",

"quantity": 3,

"price": 15.99

}

]

},

{

"orderId": "order2",

"orderDate": "2023-03-20T09:45:00Z",

"items": [

{

"productId": "product303",

"productName": "Dress",

"quantity": 2,

"price": 60.99

},

{

"productId": "product404",

"productName": "Shoes",

"quantity": 1,

"price": 120

}

]

}

],

"cart": [

{

"productId": "product303",

"productName": "Dress",

"quantity": 1,

"price": 60.99

},

{

"productId": "product505",

"productName": "Banana",

"quantity": 2,

"price": 1

}

],

"id": "598f2442-c096-4183-9f1e-ad6610ca42cc",

"\_rid": "Hyo5AKaGRdoCAAAAAAAAAA==",

"\_self": "dbs/Hyo5AA==/colls/Hyo5AKaGRdo=/docs/Hyo5AKaGRdoCAAAAAAAAAA==/",

"\_etag": "\"d800134b-0000-0700-0000-653d63a00000\"",

"\_attachments": "attachments/",

"\_ts": 1698522016

},

{

"customerId": "customer2",

"firstName": "Bob",

"lastName": "Johnson",

"email": "bob.johnson@example.com",

"address": {

"street": "789 Oak St",

"city": "Chicago",

"state": "IL",

"zipCode": "60601"

},

"orders": [

{

"orderId": "order3",

"orderDate": "2022-08-10T11:15:00Z",

"items": [

{

"productId": "product404",

"productName": "Shoes",

"quantity": 1,

"price": 120

}

]

},

{

"orderId": "order4",

"orderDate": "2023-07-05T16:20:00Z",

"items": [

{

"productId": "product101",

"productName": "Gadget",

"quantity": 1,

"price": 25.99

},

{

"productId": "product303",

"productName": "Dress",

"quantity": 3,

"price": 70.99

}

]

}

],

"cart": [

{

"productId": "product101",

"productName": "Gadget",

"quantity": 2,

"price": 25.99

}

],

"id": "bb1bb271-58a6-4cdd-91c9-f4c12090104a",

"\_rid": "Hyo5AKaGRdoDAAAAAAAAAA==",

"\_self": "dbs/Hyo5AA==/colls/Hyo5AKaGRdo=/docs/Hyo5AKaGRdoDAAAAAAAAAA==/",

"\_etag": "\"d800a14c-0000-0700-0000-653d63ac0000\"",

"\_attachments": "attachments/",

"\_ts": 1698522028

},

{

"customerId": "customer3",

"firstName": "Charlie",

"lastName": "Brown",

"email": "charlie.brown@example.com",

"address": {

"street": "567 Maple St",

"city": "San Francisco",

"state": "CA",

"zipCode": "94101"

},

"orders": [

{

"orderId": "order5",

"orderDate": "2023-01-25T13:45:00Z",

"items": [

{

"productId": "product505",

"productName": "Banana",

"quantity": 3,

"price": 1

}

]

}

],

"cart": [

{

"productId": "product101",

"productName": "Gadget",

"quantity": 1,

"price": 25.99

}

],

"id": "a8e576f0-698d-4f68-bc61-b18295db2e08",

"\_rid": "Hyo5AKaGRdoEAAAAAAAAAA==",

"\_self": "dbs/Hyo5AA==/colls/Hyo5AKaGRdo=/docs/Hyo5AKaGRdoEAAAAAAAAAA==/",

"\_etag": "\"d800f04f-0000-0700-0000-653d63c80000\"",

"\_attachments": "attachments/",

"\_ts": 1698522056

},

{

"customerId": "customer4",

"firstName": "David",

"lastName": "Lee",

"email": "david.lee@example.com",

"address": {

"street": "321 Pine St",

"city": "Houston",

"state": "TX",

"zipCode": "77001"

},

"orders": [

{

"orderId": "order6",

"orderDate": "2022-12-15T10:30:00Z",

"items": [

{

"productId": "product202",

"productName": "Widget",

"quantity": 2,

"price": 55.99

}

]

},

{

"orderId": "order7",

"orderDate": "2023-10-10T08:00:00Z",

"items": [

{

"productId": "product404",

"productName": "Shoes",

"quantity": 1,

"price": 120

},

{

"productId": "product505",

"productName": "Banana",

"quantity": 2,

"price": 1

}

]

}

],

"cart": [],

"id": "ab8af7a8-3a2f-4226-a05d-eaa2b6dbc231",

"\_rid": "Hyo5AKaGRdoFAAAAAAAAAA==",

"\_self": "dbs/Hyo5AA==/colls/Hyo5AKaGRdo=/docs/Hyo5AKaGRdoFAAAAAAAAAA==/",

"\_etag": "\"e9003038-0000-0700-0000-653ea3e90000\"",

"\_attachments": "attachments/",

"\_ts": 1698604009

},

{

"customerId": "customer5",

"firstName": "Eve",

"lastName": "Adams",

"email": "eve.adams@example.com",

"address": {

"street": "234 Cedar St",

"city": "Miami",

"state": "FL",

"zipCode": "33101"

},

"orders": [

{

"orderId": "order8",

"orderDate": "2023-02-15T14:00:00Z",

"items": [

{

"productId": "product101",

"productName": "Gadget",

"quantity": 2,

"price": 25.99

},

{

"productId": "product202",

"productName": "Widget",

"quantity": 3,

"price": 15.99

}

]

}

],

"cart": [

{

"productId": "product303",

"productName": "Dress",

"quantity": 1,

"price": 70.99

},

{

"productId": "product404",

"productName": "Shoes",

"quantity": 2,

"price": 120

},

{

"productId": "product505",

"productName": "Banana",

"quantity": 3,

"price": 1

}

],

"id": "58062f88-f0ea-4281-accc-fd1fdedb7810",

"\_rid": "Hyo5AKaGRdoGAAAAAAAAAA==",

"\_self": "dbs/Hyo5AA==/colls/Hyo5AKaGRdo=/docs/Hyo5AKaGRdoGAAAAAAAAAA==/",

"\_etag": "\"d800ae56-0000-0700-0000-653d64070000\"",

"\_attachments": "attachments/",

"\_ts": 1698522119

}

]

1. For every product in any order that is a ‘Widget’ with price greater than 20, select its CustomerID, First Name, last name, orderID, orderDate, productID, productName and productPrice.

QUERY:  
  
SELECT c.customerId, c.firstName, c.lastName, o.orderId, o.orderDate, p.productId, p.productName, p.price

FROM ecom c

JOIN o IN c.orders

JOIN p IN o.items

WHERE p.productName = 'Widget' AND p.price > 20

A screenshot of a computer

Description automatically generated

Result:

[

{

"customerId": "customer456",

"firstName": "John",

"lastName": "Doe",

"orderId": "order234",

"orderDate": "2023-10-06T10:00:00Z",

"productId": "product788",

"productName": "Widget",

"price": 30.99

},

{

"customerId": "customer4",

"firstName": "David",

"lastName": "Lee",

"orderId": "order6",

"orderDate": "2022-12-15T10:30:00Z",

"productId": "product202",

"productName": "Widget",

"price": 55.99

}

]

1. Select list of products which are in the cart and in one of the orders

QUERY:  
  
S SELECT item.productId, item.productName AS cartItem

FROM c

JOIN item IN c.cart

JOIN o IN c.orders

JOIN p IN o.items

WHERE (item.productId = p.productId)

A screenshot of a computer

Description automatically generated

Result:

[

{

"productId": "product789",

"cartItem": "Widget"

},

{

"productId": "product101",

"cartItem": "Gadget"

},

{

"productId": "product303",

"cartItem": "Dress"

},

{

"productId": "product404",

"cartItem": "Shoes"

},

{

"productId": "product303",

"cartItem": "Dress"

},

{

"productId": "product101",

"cartItem": "Gadget"

}

]

1. Select list of items that are added to cart but haven't been bought by the customer.

QUERY:  
  
SELECT c.customerId, item.productName, item.productId

FROM c

JOIN item IN c.cart

WHERE (SELECT VALUE COUNT(1) FROM o IN c.orders JOIN p IN o.items WHERE p.productId = item.productId) = 0

A screenshot of a computer

Description automatically generated

Result:

[

{

"customerId": "customer456",

"productName": "Banana",

"productId": "product505"

},

{

"customerId": "customer1",

"productName": "Banana",

"productId": "product505"

},

{

"customerId": "customer3",

"productName": "Gadget",

"productId": "product101"

},

{

"customerId": "customer5",

"productName": "Dress",

"productId": "product303"

},

{

"customerId": "customer5",

"productName": "Shoes",

"productId": "product404"

},

{

"customerId": "customer5",

"productName": "Banana",

"productId": "product505"

}

]

1. Select the customers who have ordered something in 2022 as well as in 2023.

QUERY:

SELECT c.customerId, c.firstName, c.lastName

FROM customers c

WHERE ARRAY\_LENGTH(

    ARRAY(SELECT o.orderDate

          FROM o IN c.orders

          WHERE o.orderDate >= '2022-01-01T00:00:00Z' AND o.orderDate < '2023-01-01T00:00:00Z')

    ) > 0

AND ARRAY\_LENGTH(

    ARRAY(SELECT o.orderDate

          FROM o IN c.orders

          WHERE o.orderDate >= '2023-01-01T00:00:00Z' AND o.orderDate < '2024-01-01T00:00:00Z')

    ) > 0

A screenshot of a computer

Description automatically generated

Result:

[

{

"customerId": "customer456",

"firstName": "John",

"lastName": "Doe"

},

{

"customerId": "customer1",

"firstName": "Alice",

"lastName": "Smith"

},

{

"customerId": "customer2",

"firstName": "Bob",

"lastName": "Johnson"

},

{

"customerId": "customer4",

"firstName": "David",

"lastName": "Lee"

}

]

1. What is the total expenditure for each customer for orders placed in 2023.

QUERY:

SELECT c.customerId, c.firstName, c.lastName, SUM(p.price) AS totalExpenditure

FROM customers c

JOIN o IN c.orders

JOIN p IN o.items

WHERE o.orderDate >= '2023-01-01T00:00:00Z' AND o.orderDate < '2024-01-01T00:00:00Z'

GROUP BY c.customerId, c.firstName, c.lastName

A screenshot of a computer

Description automatically generated

Result:

[

{

"customerId": "customer5",

"firstName": "Eve",

"lastName": "Adams",

"totalExpenditure": 41.98

},

{

"customerId": "customer4",

"firstName": "David",

"lastName": "Lee",

"totalExpenditure": 121

},

{

"customerId": "customer3",

"firstName": "Charlie",

"lastName": "Brown",

"totalExpenditure": 1

},

{

"customerId": "customer2",

"firstName": "Bob",

"lastName": "Johnson",

"totalExpenditure": 96.97999999999999

},

{

"customerId": "customer1",

"firstName": "Alice",

"lastName": "Smith",

"totalExpenditure": 180.99

},

{

"customerId": "customer456",

"firstName": "John",

"lastName": "Doe",

"totalExpenditure": 241.98000000000002

}

]