# SQL Outer Joins

Inner Join:

- Returns all the rows where the keys match

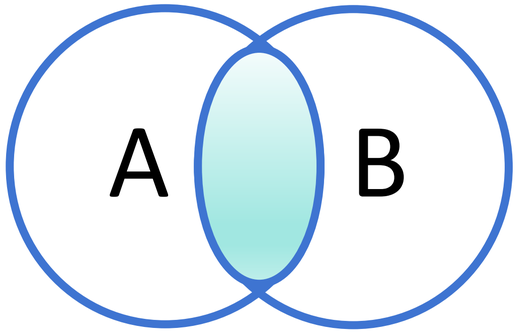
Left Outer Join:

- Returns all the rows where the keys match PLUS all the rows from the LEFT table

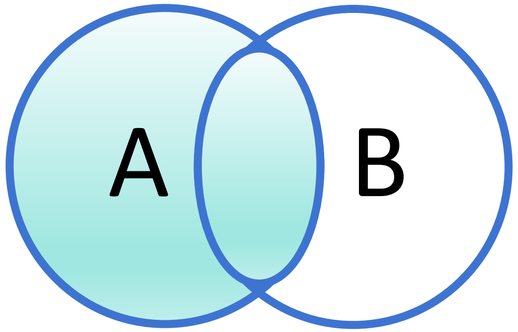
Right Outer Join:

- Returns all the rows where the keys match PLUS all the rows from the RIGHT table

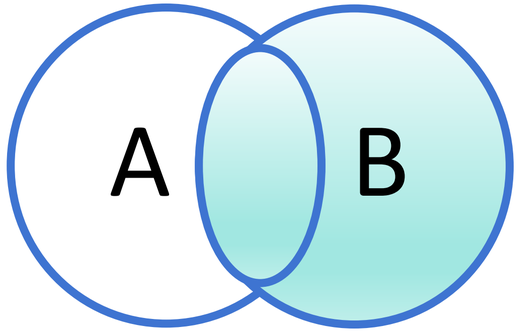
# SQL Outer Joins




Inner Join




Left Outer Join




Right Outer Join

# SQL Outer Joins

## Left Outer Join

```
SELECT LastName, Firstname, count(OrderID) as "Order Total"
   FROM "alanparadise/nw"."employees" E LEFT OUTER JOIN
        "alanparadise/nw"."orders" O
   on E.employeeid  =  O.employeeid
   GROUP BY LastName, FirstName
   Order By 3 desc;
```

Returns matching rows from the orders table, and ALL the rows from the employees table

Will show any employees with no orders.

# SQL Outer Joins

## Right Outer Join

```
SELECT LastName, Firstname, count(OrderID) as "Order Total"
   FROM "alanparadise/nw"."employees" E RIGHT OUTER JOIN
        "alanparadise/nw"."orders" O
   on E.employeeid  =  O.employeeid
   GROUP BY LastName, FirstName
   Order By 3 desc;
```

Returns matching rows from the employees table, and ALL the rows from the orders table

Will show any orders with no matching employee.

# SQL Outer Joins

Data Analysis using Outer Joins

Are there orders in the nwOrders table that have an invalid reference to a Northwinds customer?

# SQL Outer Joins

Analyzing Orders and Customers.

```
SELECT COUNT(customerid)FROM "alanparadise/nw"."customers"
```

- There are 87 customers in the customers table

```
SELECT COUNT(distinct customerid)FROM "alanparadise/nw"."orders"
```

- There are 88 distinct customers in the orders table

Is my data corrupt?

What's going on here…?

# SQL Outer Joins

Find any orders in the orders table whose customerID is NOT in the customers table

Method One:  use a subquery

```
SELECT DISTINCT customerid
      FROM "alanparadise/nw"."orders"
      WHERE customerid NOT IN (
            SELECT customerID FROM "alanparadise/nw"."customers");
```

This shows us FOUR customers who have orders in the Orders table that have no matching row in the Customers table

```
BONAP   BSBEV LACOR LAMAI
```

# SQL Outer Joins

Find any orders in the orders table whose customerID is NOT in the customers table

Method Two:  use an Outer Join

```
SELECT DISTINCT O.customerid
        FROM "alanparadise/nw"."orders" O LEFT OUTER JOIN
                "alanparadise/nw"."customers" C on
                C.customerid = O.customerid
        WHERE C.customerid is NULL
```

This shows us FOUR customers who have orders in the Orders table that have no matching row in  the Customers table

```
BONAP   BSBEV LACOR   LAMAI
```

# SQL Outer Joins

So what is the impact of this?

(That is, what is the impact of having several orders with  bad customerID's?)

# SQL Outer Joins

Inner Join:

```
SELECT C.customerid, CompanyName, COUNT(orderid)
          FROM "alanparadise/nw"."orders" O
                  JOIN "alanparadise/nw"."customers" C ON O.customerID = C.customerID
          GROUP BY C.customerid, CompanyName


SELECT COUNT(orderid)
     FROM "alanparadise/nw"."orders" O
          JOIN "alanparadise/nw"."customers" C ON O.customerid = C.customerid
```

This shows us a total of 763 orders

# SQL Outer Joins

Outer Join:

```
SELECT C.customerid, CompanyName, COUNT(orderid)
            FROM "alanparadise/nw"."orders" O
                    LEFT OUTER JOIN "alanparadise/nw"."customers" C
                       ON O.customerID = C.customerID
            GROUP BY C.customerid, CompanyName


SELECT COUNT(orderid)
    FROM "alanparadise/nw"."orders" O
            LEFT OUTER JOIN "alanparadise/nw"."customers" C ON O.customerid = C.customerid
```

This shows us a grand total of 808 orders

# SQL Outer Joins

So what is the impact of several orders having bad or missing customerID's?

This shows a discrepancy in the number of orders

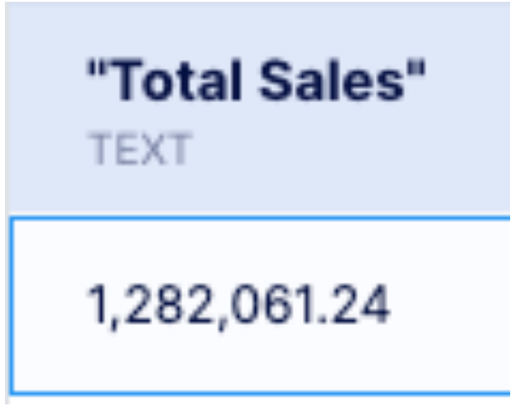
- 763 orders versus 808 orders

Let's do a similar analysis comparing the dollar amounts (joining Orders, OrderDetails and Customers)

# SQL Joins – Analysis

```
SELECT  to_char(sum(unitprice * quantity),'999,999.99') as "Total Sales"
   from "alanparadise/nw"."customers" C
        JOIN
        "alanparadise/nw"."orders" O ON C.customerid  =  O.customerid
           JOIN
           "alanparadise/nw"."orderdetails" D ON O.orderid  =  D.orderid
```

This shows total sales:



# SQL Joins – Analysis

```
SELECT  to_char(sum(unitprice * quantity),'999,999,999.99') as "Total Sales"
   from "alanparadise/nw"."customers" C
        RIGHT OUTER JOIN
        "alanparadise/nw"."orders" O ON C.customerid  =  O.customerid
           JOIN
           "alanparadise/nw"."orderdetails" D ON O.orderid  =  D.orderid
```

This shows total sales:

| "Total Sales" |
| --- |
| TEXT |
| 1,324,266.49 |

# SQL Outer Joins

We know there are some orders with missing or bad customer ids.

$1,282,021.64  VERSUS  $1,324,266.49

Such discrepancies would fail an audit !!!