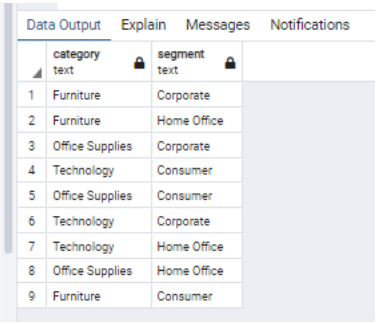
Getting familiar with categories and segments

```
select distinct p.category, c.segment
FROM products p
JOIN orders o on p.product_id = o.product_id
Join customers c on c.customer_id = o.customer_id
```

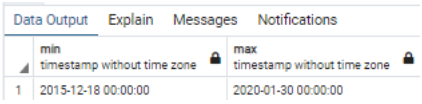


Data Output | Explain | Messages | Notifications

	category text	segment text
1	Furniture	Corporate
2	Furniture	Home Office
3	Office Supplies	Corporate
4	Technology	Consumer
5	Office Supplies	Consumer
6	Technology	Corporate
7	Technology	Home Office
8	Office Supplies	Home Office
9	Furniture	Consumer

What is the date range of the data set

```
select min(order_date), max(order_date)
from orders
```



Data Output | Explain | Messages | Notifications

	min timestamp without time zone	max timestamp without time zone
1	2015-12-18 00:00:00	2020-01-30 00:00:00

My first instinct was to want to see most of the data that I would be using to answer my problem statement. I joined columns from orders with regions, products, customers, and returns. I also only wanted orders that were not returned, selecting NULL for return quantity.

```
SELECT DISTINCT o.order_id, o.Order_date, o.sales, o.quantity, o.profit, o.discount, c.customer_name, c.segment,
p.category, p.sub_category, r2.return_quantity
from orders o
        Left join regions r on o.region_id = r.region_id
        Left join products p on o.product_id = p.product_id
        Left join customers c on o.customer_id = c.customer_id
        Left join returns r2 on o.order_id = r2.order_id
WHERE region = 'Americas' AND return_quantity IS null
limit 10;
```

Data Output | Explain | Messages | Notifications

	order_id text	order_date timestamp without time zone	sales numeric	quantity integer	profit numeric	discount numeric	customer_name text	segment text	category text	sub_category text	return_quantity integer
1	AE-2016-184765	2016-10-03 00:00:00	82.67	2	0.3	0.3	Patrick O'Donnell	Consumer	Office Supplies	Storage	[null]
2	AE-2016-1878215	2016-09-15 00:00:00	78.41	6	0.5	0.5	Patrick O'Donnell	Consumer	Technology	Machines	[null]
3	AE-2016-2721839	2016-10-28 00:00:00	82.67	2	0.3	0.3	Patrick O'Donnell	Consumer	Office Supplies	Storage	[null]
4	AE-2016-2956962	2016-09-06 00:00:00	78.41	6	0.5	0.5	Patrick O'Donnell	Consumer	Technology	Machines	[null]
5	AE-2016-3720132	2016-09-11 00:00:00	82.67	2	0.3	0.3	Patrick O'Donnell	Consumer	Office Supplies	Storage	[null]
6	AE-2016-4021172	2016-09-16 00:00:00	78.41	6	0.5	0.5	Patrick O'Donnell	Consumer	Technology	Machines	[null]
7	AE-2018-1130	2018-10-14 00:00:00	224.75	6	-232.27	0.45	Eugene Barchas	Consumer	Furniture	Bookcases	[null]
8	AE-2018-1367772	2018-12-16 00:00:00	6.97	1	0.29	0.33	Maribeth Yedwab	Corporate	Office Supplies	Storage	[null]
9	AE-2018-1369130	2018-10-13 00:00:00	224.75	6	0.44	0.48	Eugene Barchas	Consumer	Furniture	Bookcases	[null]
10	AE-2018-1530	2018-12-31 00:00:00	6.97	1	-8.6	0.3	Maribeth Yedwab	Corporate	Office Supplies	Storage	[null]

What are sales in the different segments?

```
select sum(o.sales), c.segment
from orders o
join customers c on o.customer_id = c.customer_id
Group by c.segment
```

Data Output | Explain | Messages | Notifications

	sum numeric	segment text
1	125189921.13	Consumer
2	74605957.50	Corporate
3	46087796.13	Home Office

Total Sales for Technology in Americas Region

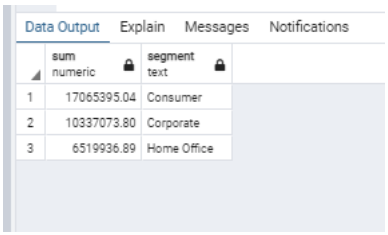
```
select sum(o.sales), c.segment
from orders o
        Left join customers c on o.customer_id = c.customer_id
        Left join products p on o.product_id = p.product_id
        Left join regions r on o.region_id = r.region_id
where category = 'Technology' AND r.region = 'Americas'
Group by c.segment
```

Data Output | Explain | Messages | Notifications

	sum numeric	segment text
1	18625232.87	Consumer
2	11354948.48	Corporate
3	7260015.28	Home Office

Total Sales for Technology in Americas Region and return quantity is NULL (No returns on order)

```
select sum(o.sales), c.segment
from orders o

    Left join customers c on o.customer_id = c.customer_id
    Left join products p on o.product_id = p.product_id
    Left join regions r on o.region_id = r.region_id
    Left join returns r2 on o.order_id = r2.order_id
where category = 'Technology' AND r.region = 'Americas' AND return_quantity IS null
Group by c.segment
```



Data Output | Explain | Messages | Notifications

	sum numeric	segment text
1	17065395.04	Consumer
2	10337073.80	Corporate
3	6519936.89	Home Office

CAN WE COMBINE THE TWO and get a difference? I only figured a union out… but can differentiate. I want separate columns.

```
select sum(o.sales) AS NET_SALES, c.segment
from orders o
    Left join customers c on o.customer_id = c.customer_id
    Left join products p on o.product_id = p.product_id
    Left join regions r on o.region_id = r.region_id
where category = 'Technology' AND r.region = 'Americas'
Group by c.segment

UNION

select sum(o.sales), c.segment
from orders o

    Left join customers c on o.customer_id = c.customer_id
```

```
        Left join products p on o.product_id = p.product_id
        Left join regions r on o.region_id = r.region_id
        Left join returns r2 on o.order_id = r2.order_id
where category = 'Technology' AND r.region = 'Americas' AND return_quantity IS null
Group by c.segment
```

Data Output | Explain | Messages | Notifications

	net_sales numeric	segment text
1	7260015.28	Home Office
2	6519936.89	Home Office
3	11354948.48	Corporate
4	10337073.80	Corporate
5	17065395.04	Consumer
6	18625232.87	Consumer

Total Sales by category in Americas Region and return quantity is NULL (No returns on order)

```
select sum(o.sales), p.category
from orders o

        Left join customers c on o.customer_id = c.customer_id
        Left join products p on o.product_id = p.product_id
        Left join regions r on o.region_id = r.region_id
        Left join returns r2 on o.order_id = r2.order_id
where r.region = 'Americas' AND return_quantity IS null
Group by p.category
```

Data Output | Explain | Messages | Notifications

	sum numeric	category text
1	28804516.40	Furniture
2	27146884.08	Office Supplies
3	33922405.73	Technology

LET'S LOOK AT SALES BY THE YEARS

```
select order_date, DATE_PART('Year', order_date)
from orders
limit 10;
```

Data Output | Explain | Messages | Notifications

	order_date timestamp without time zone	date_part double precision	sum numeric
1	2015-12-18 00:00:00	2015	31.12
2	2015-12-21 00:00:00	2015	146.64
3	2015-12-22 00:00:00	2015	572.58
4	2015-12-26 00:00:00	2015	69.68
5	2015-12-27 00:00:00	2015	171.91
6	2015-12-28 00:00:00	2015	873.78
7	2015-12-29 00:00:00	2015	187.98
8	2015-12-30 00:00:00	2015	18.52
9	2015-12-31 00:00:00	2015	41.85
10	2016-01-01 00:00:00	2016	812.53

```
SELECT
        EXTRACT(year FROM order_date) AS year, SUM(sales) as Sales
        FROM ORDERS
        Group BY Extract(year FROM order_date);
```

Data Output | Explain | Messages | Notifications

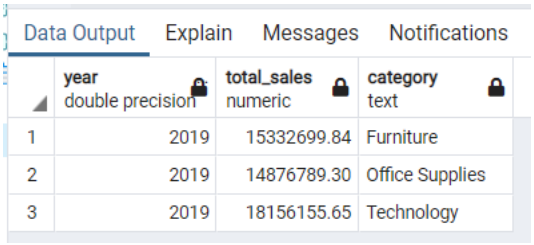
	year double precision	sales numeric
1	2015	2114.06
2	2016	9559661.49
3	2017	30623646.51
4	2018	68813298.11
5	2019	132436376.16
6	2020	4448578.43

```
SELECT
        EXTRACT(year FROM order_date) AS year, SUM(sales) as Total_Sales, p.category
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
Group BY Extract(year FROM order_date), p.category;
```

Data Output | Explain | Messages | Notifications

	year double precision	total_sales numeric	category text
1	2015	825.36	Furniture
2	2015	170.35	Office Supplies
3	2015	41.85	Technology
4	2016	1192813.91	Furniture
5	2016	1030291.66	Office Supplies
6	2016	1287650.56	Technology
7	2017	3638824.68	Furniture
8	2017	3328450.54	Office Supplies
9	2017	4345635.13	Technology
10	2018	8097243.05	Furniture
11	2018	7422151.53	Office Supplies

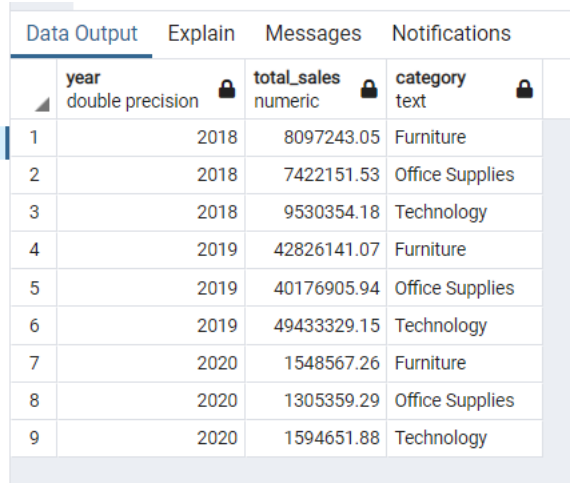
```
SELECT
        EXTRACT(year FROM order_date) AS year, SUM(sales) as Total_Sales, p.category
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null AND EXTRACT(year FROM order_date) = '2019'
Group BY Extract(year FROM order_date), p.category;
```



Data Output | Explain | Messages | Notifications

	year double precision	total_sales numeric	category text
1	2019	15332699.84	Furniture
2	2019	14876789.30	Office Supplies
3	2019	18156155.65	Technology

```
SELECT
        EXTRACT(year FROM order_date) AS year, SUM(sales) as Total_Sales, p.category
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
        AND EXTRACT(year FROM order_date) = '2018'
        or EXTRACT(year FROM order_date) = '2019'
        or EXTRACT(year FROM order_date) = '2020'
Group BY Extract(year FROM order_date), p.category;
```



Data Output | Explain | Messages | Notifications

	year double precision	total_sales numeric	category text
1	2018	8097243.05	Furniture
2	2018	7422151.53	Office Supplies
3	2018	9530354.18	Technology
4	2019	42826141.07	Furniture
5	2019	40176905.94	Office Supplies
6	2019	49433329.15	Technology
7	2020	1548567.26	Furniture
8	2020	1305359.29	Office Supplies
9	2020	1594651.88	Technology

Total technology sales w/o returns by year.

```
SELECT
        EXTRACT(year FROM order_date) AS year, SUM(sales) as Total_Sales, p.category
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
```

```
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND p.category = 'Technology'
Group BY Extract(year FROM order_date), p.category;
```

Data Output | Explain | Messages | Notifications

	year double precision	total_sales numeric	category text
1	2015	41.85	Technology
2	2016	1287650.56	Technology
3	2017	4345635.13	Technology
4	2018	9530354.18	Technology
5	2019	18156155.65	Technology
6	2020	602568.36	Technology

Total technology sales w/o returns for full years of operation.

```
SELECT
        EXTRACT(year FROM order_date) AS year, SUM(sales) as Total_Sales, p.category
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND p.category = 'Technology' AND EXTRACT(year FROM order_date) != '2015'
AND EXTRACT(year FROM order_date) != '2020'
Group BY Extract(year FROM order_date), p.category;
```

Data Output | Explain | Messages | Notifications

	year double precision	total_sales numeric	category text
1	2016	1287650.56	Technology
2	2017	4345635.13	Technology
3	2018	9530354.18	Technology
4	2019	18156155.65	Technology

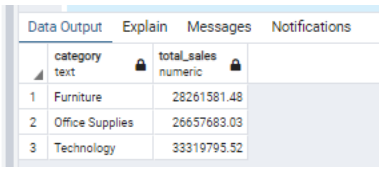
Full year sales by category

```
SELECT
        p.category, SUM(sales) as Total_Sales
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND p.category = 'Technology' AND EXTRACT(year FROM order_date) != '2015'
AND EXTRACT(year FROM order_date) != '2020'
Group by p.category

UNION

SELECT
        p.category, SUM(sales) as Total_Sales
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND p.category = 'Office Supplies' AND EXTRACT(year FROM order_date) != '2015'
AND EXTRACT(year FROM order_date) != '2020'
Group by p.category

UNION
```

```
SELECT
        p.category, SUM(sales) as Total_Sales
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND p.category = 'Furniture' AND EXTRACT(year FROM order_date) != '2015'
AND EXTRACT(year FROM order_date) != '2020'
Group by p.category;
```

Davuth's Way for the same item:

```
SELECT
        p.category, SUM(sales) as Total_Sales
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND EXTRACT(year FROM order_date) != '2015'
AND EXTRACT(year FROM order_date) != '2020'
Group by p.category;
```



Data Output | Explain | Messages | Notifications

	category text	total_sales numeric
1	Furniture	28261581.48
2	Office Supplies	26657683.03
3	Technology	33319795.52

Use the above as a template and modified to group by segments instead of category.

```
SELECT
        c.segment, SUM(sales) as Total_Sales
```

```
        FROM ORDERS o
        Left join customers c on o.customer_id = c.customer_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND c.segment = 'Consumer' AND EXTRACT(year FROM order_date) != '2015'
AND EXTRACT(year FROM order_date) != '2020'
Group by c.segment

UNION

SELECT
        c.segment, SUM(sales) as Total_Sales
        FROM ORDERS o
        Left join customers c on o.customer_id = c.customer_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND c.segment = 'Corporate' AND EXTRACT(year FROM order_date) != '2015'
AND EXTRACT(year FROM order_date) != '2020'
Group by c.segment

UNION

SELECT
        c.segment, SUM(sales) as Total_Sales
        FROM ORDERS o
        Left join customers c on o.customer_id = c.customer_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
where r.region = 'Americas' AND return_quantity IS null
AND c.segment = 'Home Office' AND EXTRACT(year FROM order_date) != '2015'
```

AND EXTRACT(year FROM order_date) != '2020'

Group by c.segment;

Data Output | Explain | Messages | Notifications

	segment text	total_sales numeric
1	Consumer	44490992.24
2	Home Office	16736939.08
3	Corporate	27011128.71

SELECT

EXTRACT(year FROM order_date) AS year, SUM(sales) as Total_Sales, c.segment

FROM ORDERS o

Left join products p on o.product_id = p.product_id

Left join returns r2 on o.order_id = r2.order_id

Left join regions r on o.region_id = r.region_id

Left join customers c on o.customer_id = c.customer_id

where r.region = 'Americas' AND return_quantity IS null

AND EXTRACT(year FROM order_date) = '2018'

or EXTRACT(year FROM order_date) = '2019'

or EXTRACT(year FROM order_date) = '2020'

Group BY Extract(year FROM order_date), c.segment;

Data Output | Explain | Messages | Notifications

	year double precision	total_sales numeric	segment text
1	2018	12751661.99	Consumer
2	2018	7970462.35	Corporate
3	2018	4327624.42	Home Office
4	2019	66124505.21	Consumer
5	2019	39821054.28	Corporate
6	2019	26490816.67	Home Office
7	2020	2442602.72	Consumer
8	2020	1282968.67	Corporate
9	2020	723007.04	Home Office

```
SELECT
	EXTRACT(year FROM order_date) AS year, c.segment, SUM(sales) as Total_Sales, sum(profit) AS Total_Profit
	FROM ORDERS o
	Left join products p on o.product_id = p.product_id
	Left join returns r2 on o.order_id = r2.order_id
	Left join regions r on o.region_id = r.region_id
	Left join customers c on o.customer_id = c.customer_id
where r.region = 'Americas' AND return_quantity IS null
	AND EXTRACT(year FROM order_date) = '2018'
	or EXTRACT(year FROM order_date) = '2019'
	or EXTRACT(year FROM order_date) = '2020'
Group BY Extract(year FROM order_date), c.segment;
```

Data Output | Explain | Messages | Notifications

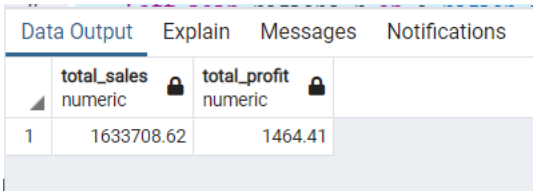
	year double precision	segment text	total_sales numeric	total_profit numeric
1	2018	Consumer	12751661.99	74362.35
2	2018	Corporate	7970462.35	42100.78
3	2018	Home Office	4327624.42	31125.86
4	2019	Consumer	66124505.21	311105.92
5	2019	Corporate	39821054.28	171301.52
6	2019	Home Office	26490816.67	126772.55
7	2020	Consumer	2442602.72	2127.68
8	2020	Corporate	1282968.67	1127.90
9	2020	Home Office	723007.04	571.26

2020 sales were only for January, and I wanted to compare them to 2019 January. I landed 2020 sales with the first quaery below but had to use the second one to be able to compare to 2019 by changing the year.

```
SELECT
	SUM(sales) as Total_Sales, sum(profit) AS Total_Profit
	FROM ORDERS o
	Left join products p on o.product_id = p.product_id
	Left join returns r2 on o.order_id = r2.order_id
	Left join regions r on o.region_id = r.region_id
	Left join customers c on o.customer_id = c.customer_id
where r.region = 'Americas' AND return_quantity IS null
```

```
        AND EXTRACT(year FROM order_date) = '2020'
```

```
SELECT
        SUM(sales) as Total_Sales, sum(profit) AS Total_Profit
        FROM ORDERS o
        Left join products p on o.product_id = p.product_id
        Left join returns r2 on o.order_id = r2.order_id
        Left join regions r on o.region_id = r.region_id
        Left join customers c on o.customer_id = c.customer_id
where r.region = 'Americas' AND return_quantity IS null AND
        '2020-01-01' <= order_date AND order_date < '2020-02-01'
```



Data Output | Explain | Messages | Notifications

	total_sales numeric	total_profit numeric
1	1633708.62	1464.41