

```
1 -- 1. SELECT with WHERE
2 SELECT customer_id, customer_city, customer_state
3 FROM olist_customers_dataset
4 WHERE customer_state = 'SP';
```

customer_id	customer_city	customer_state
06b8999e2fba1a1fbcc88172c00ba8bc7	franca	SP
18955e83d337fd6b2deff6b18a428ac77	sao bernardo do campo	SP
4e7b3e00288586ebd08712fdd0374a03	sao paulo	SP
b2b8027bc5c5109e529d4dc6358b12e3	mogi das cruzes	SP
4f2d8ab171c80ec83847c12a35b23ad	campinas	SP
1d826e7cf63160e536e0908c76c31441	sao paulo	SP
b2d1536598b73a9abd18e0d75d920a3	lencois paulista	SP
9e6b720e08816662e278caea72a	sao paulo	SP

```
5 -- 2. ORDER BY example
6 SELECT customer_id, customer_city, customer_zip_code_prefix
7 FROM olist_customers_dataset
8 ORDER BY customer_zip_code_prefix ASC
9 LIMIT 20;
```

customer_id	customer_city	customer_zip_code_prefix
7ae2a9337aa4bc799723511faa1d6830	sao paulo	1063
a09edf8c1e842e894805a206b3d73ee05	sao paulo	1064
e99b73e88afb4904ee2322fc89cf638	sao paulo	1064
5a8b64ee6ccda09ea823e6aa00e9517	sao paulo	1065
6ec2b4682814cfdac8d92bad42b3ddab	sao paulo	1065
f3e31afda80581be48ce94e7b0f3366	sao paulo	1065
c46061cd5153f35604fb9a2130df702	sao paulo	1065
15474460767452629132781071	sao paulo	1065

```
-- 3. GROUP BY (count customers by state)
1 SELECT customer_state, COUNT(*) AS count
2 FROM olist_customers_dataset
3 GROUP BY customer_state
4 ORDER BY customer_count DESC;
```

CUSTOMERS

The figure consists of a dark blue rectangular background. At the bottom, there is a horizontal row of 16 colored squares, each containing a two-letter acronym. The colors of the squares are: red-orange, yellow-orange, orange, yellow, light green, medium green, dark green, light blue, medium blue, dark blue, purple, pink, light red, medium red, dark red, and black. The acronyms are: NF, BA, SC, PR, RS, MG, QP, UN, ND, NH, NM, NW, NEM, NEMT, and APM.

▶ Run

SQLite

```
15 --Subqueries
16 SELECT customer_unique_id
17 FROM olist_customers_dataset
18 GROUP BY customer_unique_id
19 HAVING COUNT(*) > 1;
20
```

```
# customer_unique_id
```

00172711b30d52eea8b313a7f2cced02

004288347e5e88a27ded2bb23747066c

004b45ec5c64187465168251cd1c9c2f

0058f300f57d7b93c477a131a59b36c3

00a39521eb40f7012db50455bf083460

00cc12a6d8b578b8ebd21ea4e2ae8b27

011575986092c30523ecb71ff10cb473

011b13dcd51683b180c1d811250a987f

```
20 --subquery in WHERE
```

```
21 SELECT *
22 FROM olist_customers_dataset
23 WHERE customer_unique_id IN (
24     SELECT customer_unique_id
25     FROM olist_customers_dataset
26     GROUP BY customer_unique_id
27     HAVING COUNT(*) > 1
28 );
```

I

customer_id	customer_unique_id	customer_zip_code_prefix	customer_city	customer_state
879864dab9bc3047522c92c82e121...	4c93744516667ad3b8ff1fb645a3116a4	89254	jaragua do sul	SC
5adff08e34b2e993982a47070956c5c...	1175e95fb47dff9de6b2b061887e0d	81560	curitiba	PR
eabebad39a886b615b52376faec28612	295c05e81917928d76245e8427481...	5764	sao paulo	SP
2d5831cb2dff7adefba62e950ae3dc7b	e9dd12dca17352644a959d9dea133...	42890	camacari	BA
b2bed119388167a954382ca36c4777f	e079b18794454de9d2be5c12b4392...	27525	resende	RJ
a3b0fda37bae14c75a4877bed475e80c	c9158d089637ab443c78984d20da7fc0	5727	sao paulo	SP

↑ ↓ Q ⏪ ⏩ [sqliteonline.com](http://sqliteonline.com)

四

Import Export

▶ R5  
◀ SQLite

```
26 GROUP BY customer_unique_id
27 HAVING COUNT(*) > 1
28 );
29 --Aggregate functions (SUM, AVG)
30 SELECT customer_state, SUM(customer_zip_code_prefix) AS total_zip_sum
31 FROM olist_customers_dataset
32 GROUP BY customer_state
33 ORDER BY total_zip_sum DESC;
34
```

`customer_state`

四〇

四

MG

四

50

五

Type here to search



Run

SQLite

```
32 FROM olist_customers_dataset
33 GROUP BY customer_state
34 ORDER BY total_zip_sum DESC;
35 --AVG
36 SELECT customer_state, AVG(customer_zip_code_prefix) AS avg_zip
37 FROM olist_customers_dataset
38 GROUP BY customer_state
39 ORDER BY avg_zip DESC;
40
```

↓

customer_state	avg_zip
RS	94748.40468349798
SC	88818.52515809733
PR	84251.63904856294
MS	79341.07972027973
MT	78388.15876515987
TO	77426.56071428572

```
46 CREATE INDEX IF NOT EXISTS idx_customers_state
47 ON olist_customers_dataset(customer_state);
48
49 CREATE INDEX IF NOT EXISTS idx_customers_uniqueid
50 ON olist_customers_dataset(customer_unique_id);
51
52 CREATE INDEX IF NOT EXISTS idx_orders_customer
53 ON olist_customers_dataset(customer_id);
54
```

[Table](#)    [Syntax](#)    [History](#)

SQLite

```
CREATE INDEX IF NOT EXISTS idx_customers_state
ON olist_customers_dataset(customer_state);
```

[CREATE I](#)



Run

SQLite

```
4 --orders table (for JOIN)
5 DROP TABLE IF EXISTS orders;
6
7 CREATE TABLE orders (
8     order_id TEXT PRIMARY KEY,
9     customer_id TEXT,
10    order_value REAL
11 );
12
```

SQLite

```
DROP TABLE IF EXISTS orders;
```

```
CREATE TABLE orders (

```

```
    order_id TEXT PRIMARY KEY,
    customer_id TEXT,
    ...
```

```
61 ('04', '02', 71.10),
62 ('03', '01', 50.00),
63 ('04', '01', 200.00));
64 ..INNER JOIN
65
66 SELECT c.customer_id, c.customer_city, o.order_id, o.order_value
67 FROMolist.customers dataset c
68 INNER JOIN orders o
69 ON c.customer_id = o.customer_id
70
71
72
73
```

| customer\_id

| customer\_city

| order\_id

| order\_value

```

71    ON c.customer_id = o.customer_id;
72
73 --LEFT JOIN
74 SELECT c.customer_id, c.customer_city, o.order_id, o.order_value
75 FROM olist_customers_dataset c
76 LEFT JOIN orders o
77 ON c.customer_id = o.customer_id;
78
79

```

customer_id	customer_city	order_id	order_value
06b8999e2fba1a1fbcc88172c00ba8bc7	franca	NULL	NULL
18955e83d337fd6b2de6b18a428ac77	sao bernardo do campo	NULL	NULL
4e7b3e00288586eb08712fd0374a03	sao paulo	NULL	NULL
b2b6027bc5c5109e529d4d63353b12c3	mogi das cruzes	NULL	NULL
4f2b8ab71c80ec8364f7c12e35b23ad	campinas	NULL	NULL
879864da89dc30417522c92c82e1212b8	jaragua do sul	NULL	NULL

```
77    ON c.customer_id = o.customer_id;
78 --RIGHT JOIN
79 SELECT c.customer_id, c.customer_city, o.order_id, o.order_value
80 FROMolist_customers_dataset c
81 RIGHT JOIN orders o
82 ON c.customer_id = o.customer_id;
83
```

customer_id	customer_city	order_id	order_value
NULL	NULL	o1	120
NULL	NULL	o2	75.5
NULL	NULL	o3	50
NULL	NULL	o4	200