





## Olist

Last modified on 10/03/2025, 09:28:54

Notebook

Script

Job details

Runs

Data quality

Schedules

Version Control



Code ▾

```
•[8]: # Selecionando as colunas desejadas
df2 = df.select("product_id", "product_category_name")

s3_output_path = "s3://olist-dados/df2_olist_products.csv"

# Salvando no S3
df2.write.mode("overwrite").option("header", "true").csv(s3_output_path)
```

```
[9]: df2.show(5)
```

```
+-----+-----+
| product_id | product_category_name |
+-----+-----+
| 1e9e8ef04dbcff454... | perfumaria |
| 3aa071139cb16b67c... | artes |
| 96bd76ec8810374ed... | esporte_lazer |
| cef67bcfe19066a93... | bebes |
| 9dc1a7de274444849... | utilidades_domest... |
+-----+-----+
```



Editor



Queries



Notebooks



+ Olist x

▶ Run



Limit 100



Explain



Isolated session



Serverless: oli... ▾

dev ▾

Schedule



```
70 -- Top 10 Categorias de Produtos Mais Vendidas
71 SELECT
72     p.product_category_name,
73     COUNT(oi.order_id) AS total_vendas
74 FROM olist_order_items oi
75 JOIN olist_products p ON oi.product_id = p.product_id
76 GROUP BY p.product_category_name
77 ORDER BY total_vendas DESC
78 LIMIT 10;
```

Row 73, Col 39, Chr 5579

Result 1 (10)

Export ▾



Chart



<input type="checkbox"/>	product_category_na...	total_vendas
<input type="checkbox"/>	cama_mesa_banho	22230
<input type="checkbox"/>	beleza_saude	19340
<input type="checkbox"/>	esporte_lazer	17282

Query ID 1424690

Elapsed time: 279 ms

Total rows: 10

+ Olist x

Run



Limit 100



Explain



Isolated session



Serverless: oli...

dev

```
80 -- Ticket Médio por Categoria de Produto
81 SELECT
82     p.product_category_name,
83     ROUND(AVG(oi.price), 2) AS ticket_medio
84 FROM olist_order_items oi
85 JOIN olist_products p ON oi.product_id = p.product_id
86 GROUP BY p.product_category_name
87 ORDER BY ticket_medio DESC;
```

Result 1 (74)

<input type="checkbox"/>	product_category_na...	ticket_medio
<input type="checkbox"/>	pcs	1098.34
<input type="checkbox"/>	portateis_casa_forno_e_c...	624.29
<input type="checkbox"/>	eletrodomesticos_2	476.12
<input type="checkbox"/>	agro_industria_e_comercio	342.12
<input type="checkbox"/>	instrumentos_musicais	281.62
<input type="checkbox"/>	eletroportateis	280.78000000000003
<input type="checkbox"/>	portateis_cozinha_e_prep...	264.57
<input type="checkbox"/>	telefonica_fixa	225.69
<input type="checkbox"/>	construcao_ferramentas_...	208.99



Editor



Queries



Notebooks



+ Olist x

▶ Run



Limit 100



Explain



Isolated session



Serverless: oli...



dev



```
89 -- Top 10 Tempo Médio de Entrega por Categoria de Produto
90 SELECT
91     p.product_category_name,
92     ROUND(AVG(DATEDIFF(day, o.order_purchase, o.order_delivered_customer_date)), 2) AS tempo_medio_entrega_dias
93 FROM olist_order o
94 JOIN olist_order_items oi ON o.order_id = oi.order_id
95 JOIN olist_products p ON oi.product_id = p.product_id
96 WHERE o.order_delivered_customer_date IS NOT NULL
97 GROUP BY p.product_category_name
98 ORDER BY tempo_medio_entrega_dias DESC
99 LIMIT 10;
```

Result 1 (10)

<input type="checkbox"/>	product_category_na...	tempo_medio_entreg...
<input type="checkbox"/>	moveis_escritorio	20
<input type="checkbox"/>	artigos_de_natal	15
<input type="checkbox"/>	seguros_e_servicos	15
<input type="checkbox"/>	fashion_calcados	15



Editor



Queries



Notebooks



+ Olist x

▶ Run



Limit 100



Explain



Isolated session



Serverless: oli...



dev



```
100 -- Distribuição dos Pagamentos por Tipo
101 SELECT
102     payment_type,
103     COUNT(order_id) AS total_pedidos,
104     ROUND(SUM(payment_value), 2) AS total_valor
105 FROM olist_order_payments
106 GROUP BY payment_type
107 ORDER BY total_pedidos DESC;
```

📊 Result 1 (5)

<input type="checkbox"/>	payment type	total pedidos	total valor
<input type="checkbox"/>	credit_card	76795	12542084.190000001
<input type="checkbox"/>	boleto	19784	2869361.27
<input type="checkbox"/>	voucher	5775	379436.87
<input type="checkbox"/>	debit_card	1529	217989.78999999998
<input type="checkbox"/>	not_defined	3	0





▼

<input type="checkbox"/>	ano	mes	vendas mensais
<input checked="" type="checkbox"/>	2016	9	764
<input type="checkbox"/>	2016	10	2231827
<input type="checkbox"/>	2016	12	190
<input type="checkbox"/>	2017	1	5062841

+ Olist x

▶ Run

☒ Limit 100☐ Explain☒ Isolated session ⓘ

Serverless: oli... ▼

dev ▼

```
109  -- Identificar os Maiores Vendedores e região (Top 10)
110  WITH vendas AS (
111  SELECT
112      oi.seller_id,
113      os.seller_city,
114      os.seller_state,
115      COUNT(DISTINCT oi.order_id) AS total_vendas,
116      ROUND(SUM(oi.price)::NUMERIC, 2) AS total_faturamento
117  FROM olist_order_items oi
118  JOIN olist_sellers os ON oi.seller_id = os.seller_id
119  GROUP BY oi.seller_id, os.seller_city, os.seller_state
120  )
121  SELECT
122      ROW_NUMBER() OVER (ORDER BY total_faturamento DESC) AS vendedor,
123      seller_city,
124      seller_state,
125      total_vendas,
126      total_faturamento
127  FROM vendas
128  ORDER BY total_faturamento DESC
129  LIMIT 10;
130
```



aws

Serviços

Search

[Alt+S]

Estados Unidos

Editor

Queries

Notebooks

Olist

Run

Limit 100

Explain

Isolated session

Serverless: oli...

dev

```
109  -- Identificar os Maiores Vendedores e região (Top 10)
110  WITH vendas AS (
111      SELECT
```

Result 1 (10)

Export

	vendedor	seller_city	seller_state	total_vendas	total_faturamento
<input type="checkbox"/>	1	guariba	SP	1132	458945
<input type="checkbox"/>	2	lauro de freitas	BA	358	445552
<input type="checkbox"/>	3	ibitinga	SP	1806	400946
<input type="checkbox"/>	4	sumare	SP	585	388084
<input type="checkbox"/>	5	itaquaquecetuba	SP	982	375848
<input type="checkbox"/>	6	barueri	SP	336	352864
<input type="checkbox"/>	7	piracicaba	SP	1314	320473
<input type="checkbox"/>	8	sao paulo	SP	1160	283491
<input type="checkbox"/>	9	sao paulo	SP	915	277937
<input type="checkbox"/>	10	sao paulo	SP	1287	270343