

Load the Dataset

```
orders = pd.read_csv("olist_orders_dataset.csv")
items = pd.read_csv("olist_order_items_dataset.csv")
customers = pd.read_csv("olist_customers_dataset.csv")
```

Data Exploration

```
▶ print(orders.shape)
print(items.shape)
print(customers.shape)

...
(99441, 8)
(112650, 7)
(99441, 5)

orders.head()
```

	order_id	customer_id	order_status	order_purchase_timestamp	order_approved_at	ord
0	e481f51cbdc54678b7cc49136f2d6af7	9ef432eb6251297304e76186b10a928d	delivered	2017-10-02 10:56:33	2017-10-02 11:07:15	
1	53cdb2fc8bc7dce0b6741e2150273451	b0830fb4747a6c6d20dea0b8c802d7ef	delivered	2018-07-24 20:41:37	2018-07-26 03:24:27	
2	47770eb9100c2d0c44946d9cf07ec65d	41ce2a54c0b03bf3443c3d931a367089	delivered	2018-08-08 08:38:49	2018-08-08 08:55:23	
3	949d5b44dbf5de918fe9c16f97b45f8a	f88197465ea7920adcdbec7375364d82	delivered	2017-11-18 19:28:06	2017-11-18 19:45:59	
4	ad21c59c0840e6cb83a9ceb5573f8159	8ab97904e6daea8866dbdbc4fb7aad2c	delivered	2018-02-13 21:18:39	2018-02-13 22:20:29	

```
▶ items.head()

...
order_id  order_item_id          product_id      seller_id  shipping_lim
0 00010242fe8c5a6d1ba2dd792cb16214 1 4244733e06e7ecb4970a6e2683c13e61 48436dade18ac8b2bce089ec2a041202 2017-09-19
1 00018f77f2f0320c557190d7a144bdd3 1 e5f2d52b802189ee658865ca93d83a8f dd7ddc04e1b6c2c614352b383efe2d36 2017-05-03
2 000229ec398224ef6ca0657da4fc703e 1 c777355d18b72b67abbeef9df44fd0fd 5b51032eddd242adc84c38acab88f23d 2018-01-18
3 00024acbcdff0a6daa1e931b038114c75 1 7634da152a4610f1595efa32f14722fc 9d7a1d34a5052409006425275ba1c2b4 2018-08-15
4 00042b26cf59d7ce69dfabb4e55b4fd9 1 ac6c3623068f30de03045865e4e10089 df560393f3a51e74553ab94004ba5c87 2017-02-13
```

```
▶ customers.head()

...
customer_id  customer_unique_id  customer_zip_code_prefix  customer_city  customer_state
0 06b8999e2fba1a1fbc88172c00ba8bc7 861eff4711a542e4b93843c6dd7febb0 14409 franca SP
1 18955e83d337fd6b2def6b18a428ac77 290c77bc529b7ac935b93aa66c333dc3 9790 sao bernardo do campo SP
2 4e7b3e00288586ebd08712fdd0374a03 060e732b5b29e8181a18229c7b0b2b5e 1151 sao paulo SP
3 b2b6027bc5c5109e529d4dc6358b12c3 259dac757896d24d7702b9acbbff3f3c 8775 mogi das cruzes SP
4 4f2d8ab171c80ec8364f7c12e35b23ad 345ecd01c38d18a9036ed96c73b8d066 13056 campinas SP
```

Missing Values Check

```
Missing values in Each Column:  
order_month    0  
total_sales    0  
month_index    0  
dtype: int64  
  
Total Missing Values: 0
```

```
... Dataset Shape: (23, 3)  
  
Columns:  
Index(['order_month', 'total_sales', 'month_index'], dtype='object')  
  
Dataset Info:  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 23 entries, 0 to 22  
Data columns (total 3 columns):  
 #   Column      Non-Null Count  Dtype     
---  --          --          --          --  
 0   order_month  23 non-null    object    
 1   total_sales  23 non-null    float64  
 2   month_index  23 non-null    int64    
dtypes: float64(1), int64(1), object(1)  
memory usage: 684.0+ bytes
```

Monthly sales

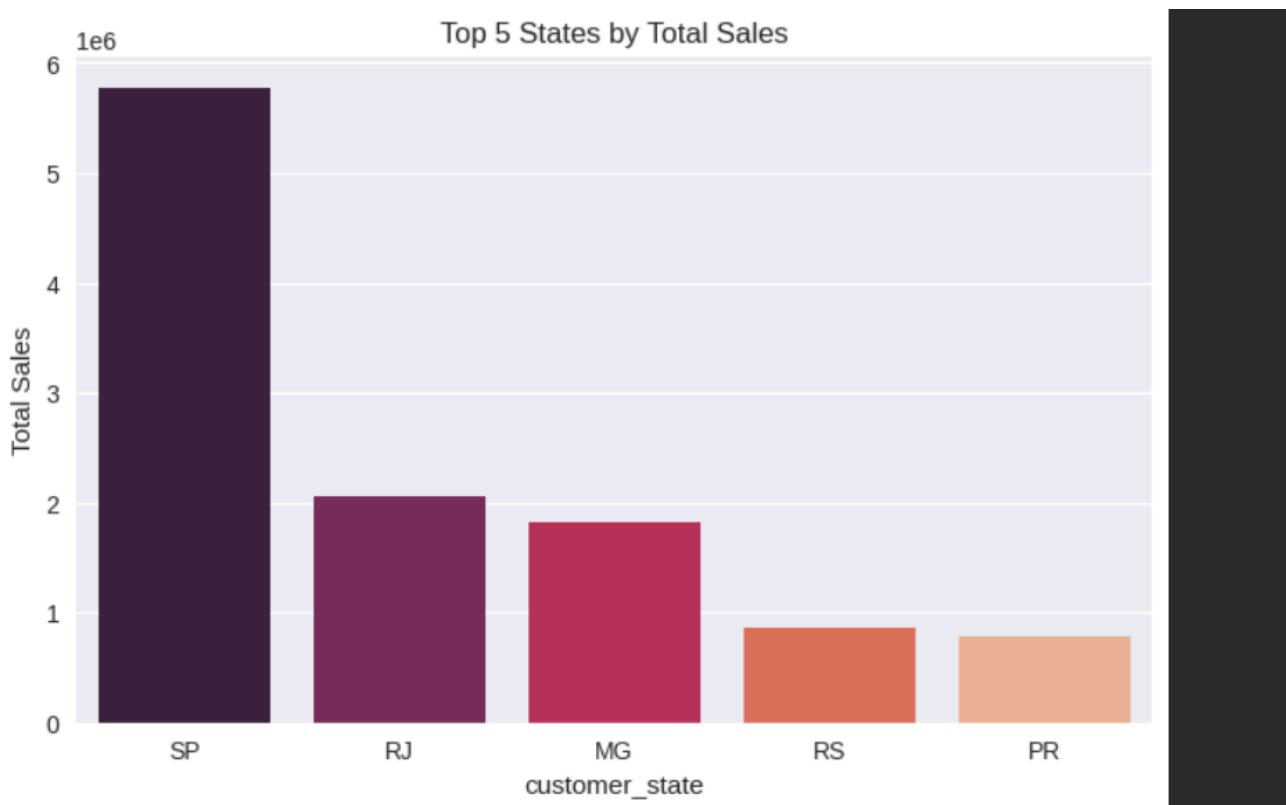
```
monthly_sales = (  
    sales_data  
    .groupby('order_month')['total_sales']  
    .sum()  
    .reset_index()  
)  
  
monthly_sales['order_month'] = monthly_sales['order_month'].astype(str)  
monthly_sales.head()
```

	order_month	total_sales
0	2016-09	143.46
1	2016-10	46490.66
2	2016-12	19.62
3	2017-01	127482.37
4	2017-02	271239.32

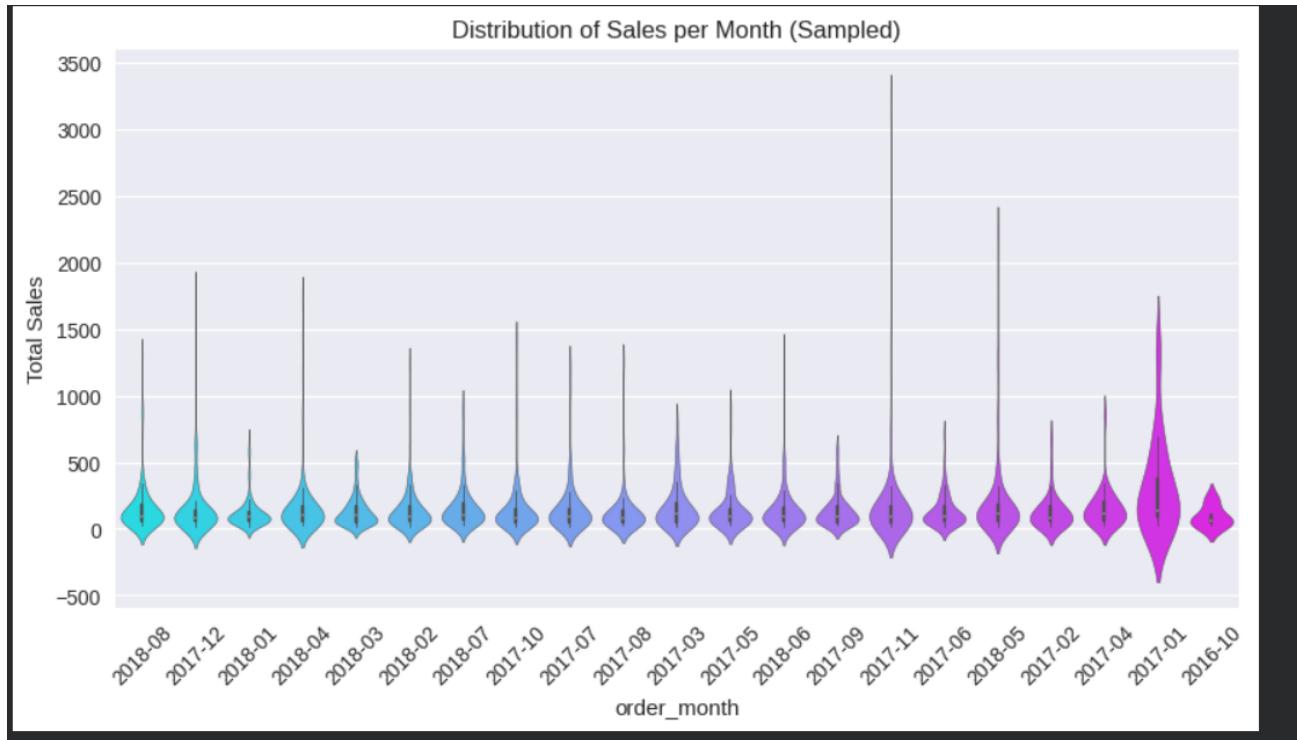
Monthly Sales Trend (LINE GRAPH)



States by Total Sales Graph



Sales Distribution Image



Evaluation Metrics

--- Linear Regression Evaluation ---

MAE: 299368.58

RMSE: 337428.93

R2 Score: -29.12

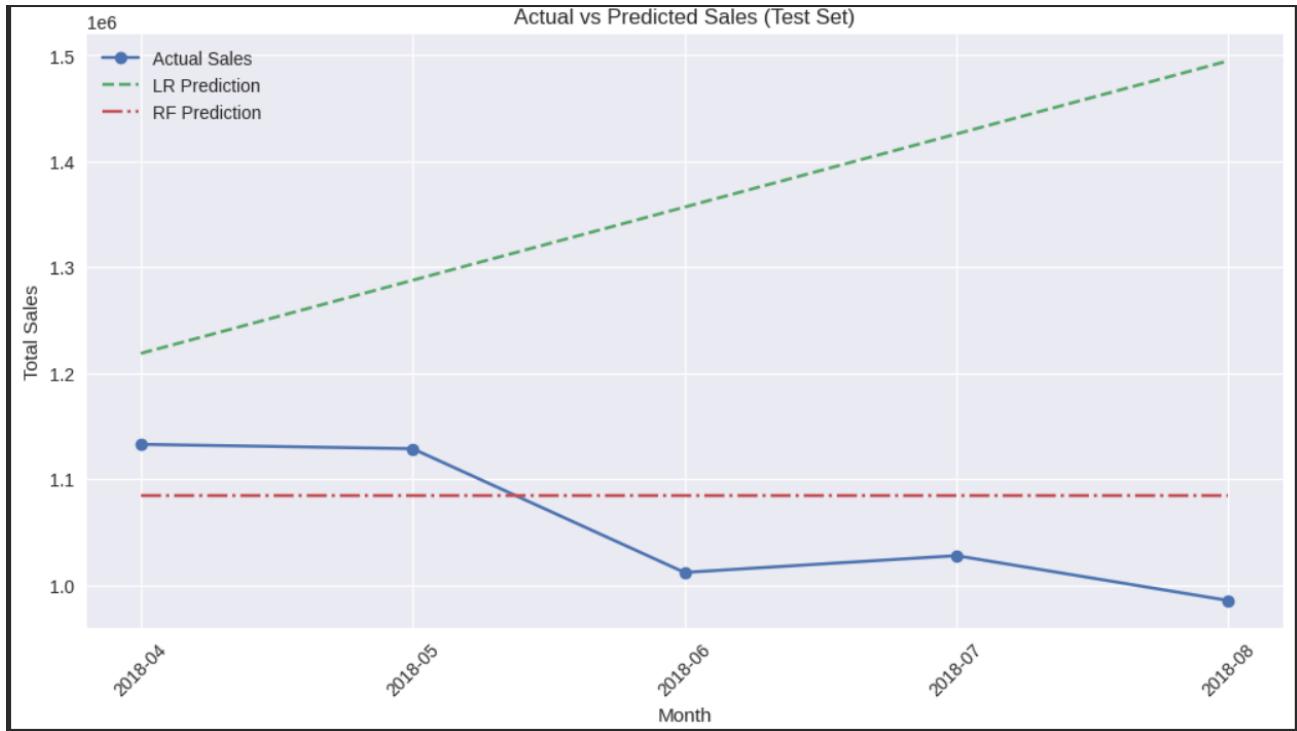
--- Random Forest Regressor Evaluation ---

MAE: 64147.02

RMSE: 67139.7

R2 Score: -0.19

Actual vs Predicted Graph



Gradio UI

```

...
Requirement already satisfied: gradio in /usr/local/lib/python3.12/dist-packages (5.50.0)
Requirement already satisfied: aiofiles<25.0,>=22.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (24.1.0)
Requirement already satisfied: anyio<5.0,>=3.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (4.12.1)
Requirement already satisfied: brotli>=1.1.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (1.2.0)
Requirement already satisfied: fastapi<1.0,>=0.115.2 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.123.10)
Requirement already satisfied: ffmpeg in /usr/local/lib/python3.12/dist-packages (from gradio) (1.0.0)
Requirement already satisfied: gradio-client==1.14.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (1.14.0)
Requirement already satisfied: groovy~0.1 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.1.2)
Requirement already satisfied: httpxx<1.0,>=0.24.1 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.28.1)
Requirement already satisfied: huggingface-hub<2.0,>=0.33.5 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.36.0)
Requirement already satisfied: jinja2<4.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (3.1.6)
Requirement already satisfied: markupsafe<4.0,>=2.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (3.0.3)
Requirement already satisfied: numpy<3.0,>=1.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (2.0.2)
Requirement already satisfied: orjson~3.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (3.11.5)
Requirement already satisfied: packaging in /usr/local/lib/python3.12/dist-packages (from gradio) (25.0)
Requirement already satisfied: pandas<3.0,>=1.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (2.2.2)
Requirement already satisfied: pillow<12.0,>=8.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (11.3.0)
Requirement already satisfied: pydantic<=2.12.3,>=2.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (2.12.3)
Requirement already satisfied: pydub in /usr/local/lib/python3.12/dist-packages (from gradio) (0.25.1)
Requirement already satisfied: python-multipart>=0.0.18 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.0.21)
Requirement already satisfied: pyyaml<7.0,>=5.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (6.0.3)
Requirement already satisfied: ruff>=0.9.3 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.14.11)
Requirement already satisfied: safehttpx<0.2.0,>=0.1.6 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.1.7)
Requirement already satisfied: semantic-version~<2.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (2.10.0)
Requirement already satisfied: starlette<1.0,>=0.40.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.50.0)
Requirement already satisfied: tomlkit<0.14.0,>=0.12.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.13.3)
Requirement already satisfied: typer<1.0,>=0.12 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.21.1)
Requirement already satisfied: typing-extensions~<4.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (4.15.0)
Requirement already satisfied: uvicorn>=0.14.0 in /usr/local/lib/python3.12/dist-packages (from gradio) (0.40.0)
Requirement already satisfied: fsspec in /usr/local/lib/python3.12/dist-packages (from gradio-client==1.14.0->gradio) (2025.3.0)
Requirement already satisfied: websockets<16.0,>=13.0 in /usr/local/lib/python3.12/dist-packages (from gradio-client==1.14.0->gradio)

```

Output

