/mnt/sdb3/ITI/ITI/1-July/Data warehouse and data lake/Menna/The project

.venvpro ❯

"/mnt/sdb3/ITI/ITI/1-July/Data warehouse and data lake/Menna/The project/Olist\_ETL\_Project/.ven

vpro/bin/python" "/mnt/sdb3/ITI/ITI/1-July/Data warehouse and data lake/Menna/The project/Olist

\_ETL\_Project/scripts/Check data report.py"

======================================================================

Loading Raw Data Files...

======================================================================

✅ Customers: 99,441 rows

✅ Orders: 99,441 rows

✅ Order Items: 112,650 rows

✅ Products: 32,951 rows

✅ Sellers: 3,095 rows

======================================================================

ANALYSIS 1: Missing Values Detection

======================================================================

Missing Values in Orders Table:

--------------------------------------------------

order\_approved\_at | 160 (0.16%)

order\_delivered\_carrier\_date | 1,783 (1.79%)

order\_delivered\_customer\_date | 2,965 (2.98%)

======================================================================

ANALYSIS 2: Price Distribution & Outliers

======================================================================

Price Statistics:

--------------------------------------------------

count 112650.000000

mean 120.653739

std 183.633928

min 0.850000

25% 39.900000

50% 74.990000

75% 134.900000

max 6735.000000

Name: price, dtype: float64

📊 Total Products: 112,650

⚠️ Price Outliers Detected: 8,427 (7.48%)

💰 Most Expensive Product: R$ 6,735.00

======================================================================

ANALYSIS 3: Geographic Distribution

======================================================================

Top 10 States by Customer Count:

--------------------------------------------------

customer\_state

SP 41746

RJ 12852

MG 11635

RS 5466

PR 5045

SC 3637

BA 3380

DF 2140

ES 2033

GO 2020

Name: count, dtype: int64

======================================================================

ANALYSIS 4: Order Status Distribution

======================================================================

Order Status Breakdown:

--------------------------------------------------

order\_status

delivered 96478

shipped 1107

canceled 625

unavailable 609

invoiced 314

processing 301

created 5

approved 2

Name: count, dtype: int64

======================================================================

ANALYSIS 5: Time Series - Orders Over Time

======================================================================

📅 Date Range: 2016-09-04 21:15:19 to 2018-10-17 17:30:18

📦 Peak Month: 2017-11 with 7,544 orders

======================================================================

ANALYSIS 6: Delivery Time Analysis

======================================================================

Delivery Time Statistics:

--------------------------------------------------

count 96463.000000

mean 12.095715

std 9.551358

min 1.000000

25% 6.000000

50% 10.000000

75% 15.000000

max 209.000000

Name: delivery\_days, dtype: float64

======================================================================

ANALYSIS 7: Product Category Analysis

======================================================================

Top 5 Product Categories:

--------------------------------------------------

product\_category\_name

cama\_mesa\_banho 11115

beleza\_saude 9670

esporte\_lazer 8641

moveis\_decoracao 8334

informatica\_acessorios 7827

Name: count, dtype: int64

======================================================================

ANALYSIS 8: Revenue Analysis

======================================================================

💰 Total Revenue: R$ 15,843,553.24

📦 Total Orders: 112,650

💵 Average Order Value: R$ 140.64

======================================================================

📊 FINAL SUMMARY REPORT

======================================================================

Data Quality Summary:

==================================================

✅ Total Customers: 99,441

✅ Total Orders: 99,441

✅ Total Products Sold: 112,650

✅ Unique Products: 32,951

✅ Active Sellers: 3,095

Missing Data Issues:

==================================================

⚠️ Orders with missing delivery dates: 2,965

⚠️ Orders with missing approval: 160

Business Insights:

==================================================

💰 Total Revenue: R$ 15,843,553.24

📊 Average Order Value: R$ 140.64

⏱️ Average Delivery Time: 12.1 days

🏆 Top State: SP

🎯 Most Sold Category: cama\_mesa\_banho

✅ Analysis Complete! All charts saved successfully.

======================================================================

/mnt/sdb3/ITI/ITI/1-July/Data warehouse and data lake/Menna/The project 18s

.venvpro ❯