Algorithm Explanation:

Product and Sales Pattern Analysis Algorithm is a comprehensive data analysis system designed to deeply analyze product sales patterns, customer segment preferences, and sales trends in retail business. Through multi-dimensional data analysis, it reveals product category performance, customer behavior patterns, and seasonal trends, providing data-driven insights for business decisions.

Algorithm Input and Output:

Inputs:

1. customers_2.csv- Customer basic information data
2. products_2.csv- Product catalog and pricing information
3. sales_2.csv- Sales transaction records
4. s1_customer_segmentation_results.csv- Customer segmentation results (from Stage 1)
5. s2_customer_ltv_predictions.csv- Customer lifetime value predictions (from Stage 2)
6. s2_feature_importance_analysis.csv- Feature importance analysis results (from Stage 2)

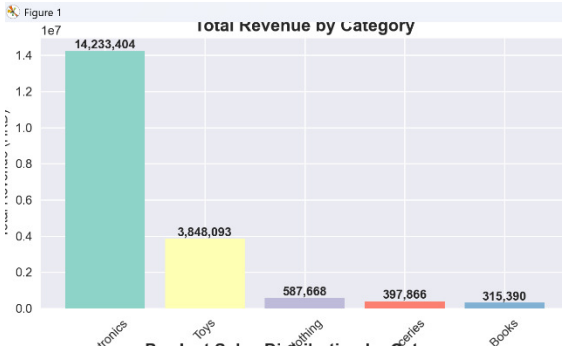
Outputs:

1. **Category Performance Analysis Report** (s3_category_performance_analysis.csv)
2. **Segment Preference Analysis** (s3_segment_preferences_analysis.csv)
3. **Sales Trend Analysis** (s3_sales_trends_analysis.csv)
4. **Visualization Charts** (direct display)
5. **Business Intelligence Recommendations**

Detailed Explanation:

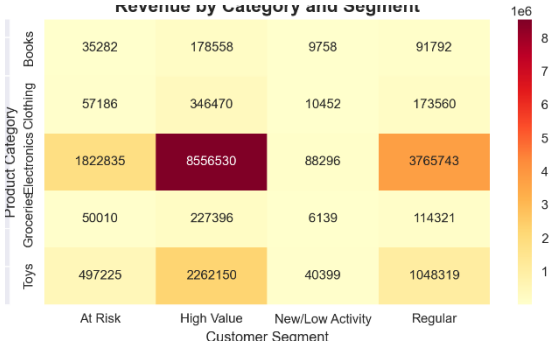
1. *Total Revenue by Category Graph*

- What it does:
  - displays the total revenue generated by each product category
- key data analysis:
  - Electronics: 14,233,404 HKD
  - Toys: 3,848,093 HKD
  - Clothing: 587,668 HKD
  - Groceries: 339,369 HKD
  - Books: 315,390 HKD
- Trend we observe:
  - suggests heavy reliance on high-value categories, with opportunities to boost low-revenue ones through promotions



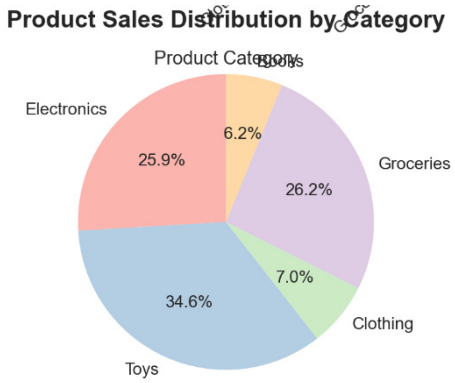
2. *Revenue by Category and Segment Graph*

- What it does:
  - revenue breakdown by category and customer segment
- key data analysis:
  - Books: At Risk 35,282; High Value 176,658; New/Low Activity 9,758; Regular 91,792 (total 313,490).
  - Clothing: At Risk 57,186; High Value 346,470; New/Low Activity 10,452; Regular 173,560 (total 587,668).
  - Electronics: At Risk 1,822,835; High Value 8,556,530; New/Low Activity 88,296; Regular 3,765,743 (total 14,233,404).
  - Groceries: At Risk 50,010; High Value 227,396; New/Low Activity 6,139; Regular 114,321 (total 397,866).
  - Toys: At Risk 497,225; High Value 2,262,150; New/Low Activity 40,399; Regular 1,048,319 (total 3,848,093).
- Trend we observe:
  - trend shows premium categories (Electronics/Toys) favored by valuable segments, while essentials (Groceries/Books) more even but low overall.



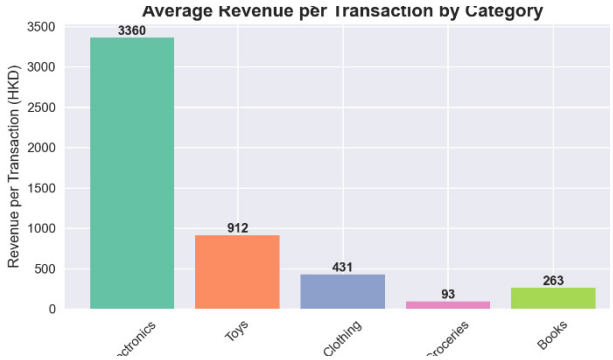
3. *Product sales Distribution by Category Graph*

- What it does:
  - represents the percentage of total product sales
- key data analysis:
  - Electronics: 34.6%.
  - Toys: 25.9%.
  - Groceries: 26.2%.
  - Clothing: 7.0%.
  - Books: 6.3%.



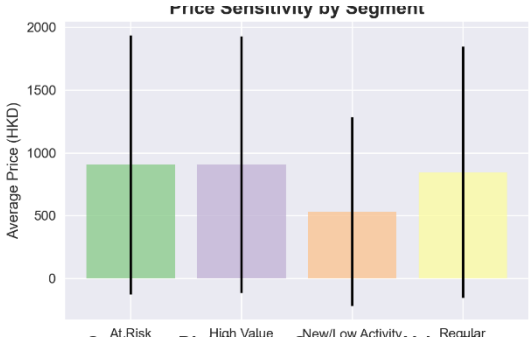
4. *Average Revenue Per Transaction by category graph*

- What it does:
  - average revenue per transaction for each category
- key data analysis:
  - Electronics: 3,360 HKD.
  - Toys: 912 HKD.
  - Clothing: 431 HKD.
  - Groceries: 93 HKD.
  - Books: 263 HKD.
- Trend we observe:
  - Electronics far outperforms



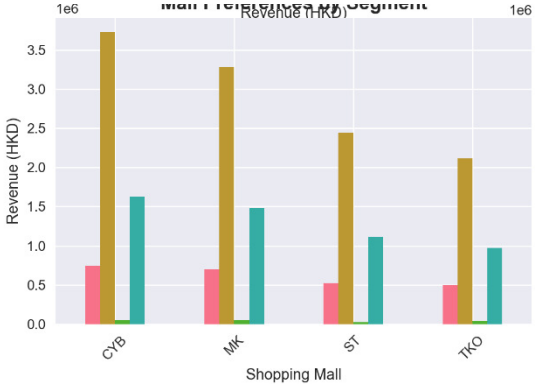
5. *Price Sensitivity by segment Graph:*

- What it does:
  - measures average price per item purchased by segment
- key data analysis:
  - At Risk: ~1,200 HKD.
  - High Value: ~1,100 HKD.
  - New/Low Activity: ~900 HKD.
  - Regular: ~800 HKD.
- Trend we observe:
  - trend shows riskier segments tolerate higher prices



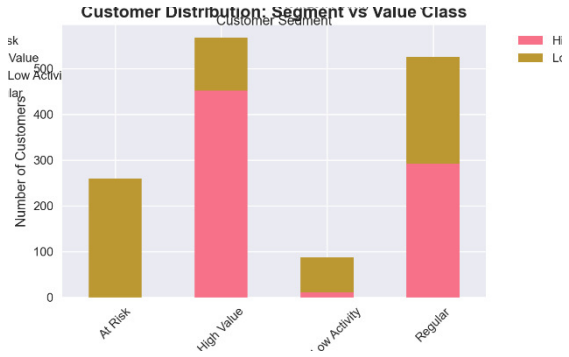
6. *mall preference by segment graph*

- what it does:
    - shows average revenue per segment at each mall
- key data analysis:
    - CYB: High Value 3.5, Regular 1.5, At Risk 0.5, New/Low 0.1.
    - MK: High Value 3.2, Regular 1.5, At Risk 0.5, New/Low 0.1.
    - ST: High Value 2.5, Regular 1.0, At Risk 0.4, New/Low 0.05.
    - TKO: High Value 2.0, Regular 1.0, At Risk 0.3, New/Low 0.05.
- Trend we observe:
    - trending toward premium malls attracting valuable customers, for location-specific campaigns



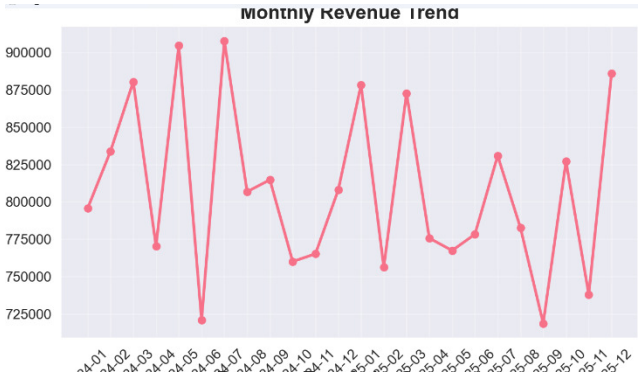
7. ***Customer Distribution: Segment VS Value Class Graph***

- What it does:
    - distributes customers by segment
- key data analysis:
    - At Risk: High ~200, Low ~59 (total 259).
    - High Value: High ~500, Low ~67 (total 567).
    - New/Low Activity: High ~20, Low ~68 (total 88).
    - Regular: High ~300, Low ~225 (total 525).
- Trend we observe:
    - trend shows strong correlation between segments and value

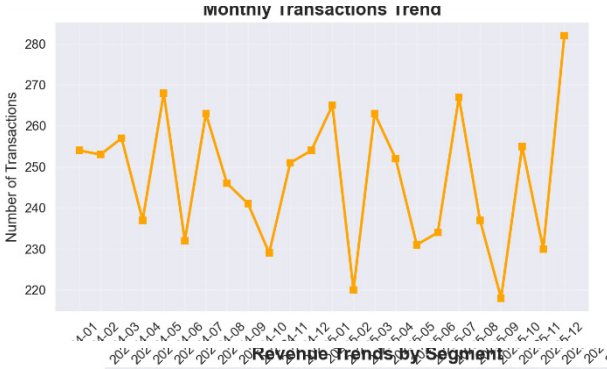


8. ***Monthly Revenue Trend Graph***

- What it is:
    - tracks total monthly revenue over time
- key data analysis:
    - Peaks: ~900,000 (2024-03, 2025-12), lows ~720,000 (2024-06, 2025-09).
    - Average: ~810,000 HKD/month.
- Trend we observe:
    - suggesting seasonal cycles (Q1/Q4 high)

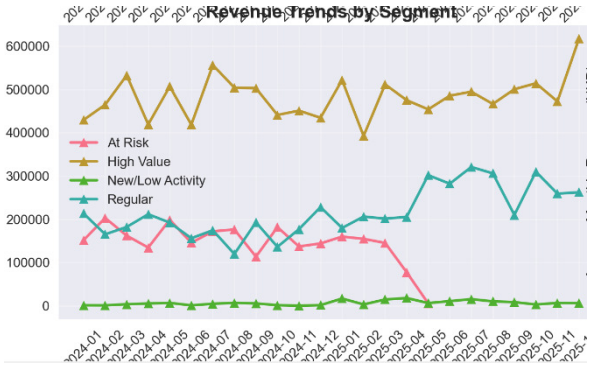


9. Monthly Transactions Trend Graph



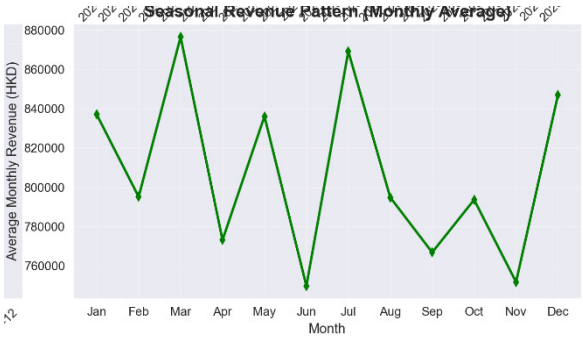
- What it is:
  - shows monthly transaction counts
- key data analysis:
  - Peaks: 280 (2025-12), lows 220 (2025-02).
  - Average: ~247 transactions/month.

10. Revenue Trends by segment Graph



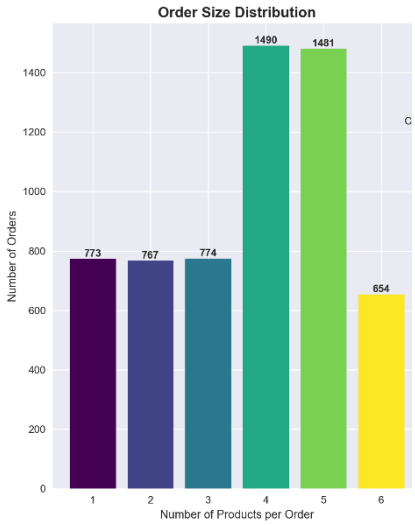
- What it is:
  - monthly revenue by segment
- key data analysis:
  - High Value: Peaks ~600,000, average ~400,000.
  - Regular: Peaks ~300,000, average ~200,000.
  - At Risk: Peaks ~150,000, average ~100,000.
  - New/Low: Low ~20,000 average.
- Trend we observe:
  - trending segment consistency in seasonality but scaled by value

11. Seasonal Revenue Pattern



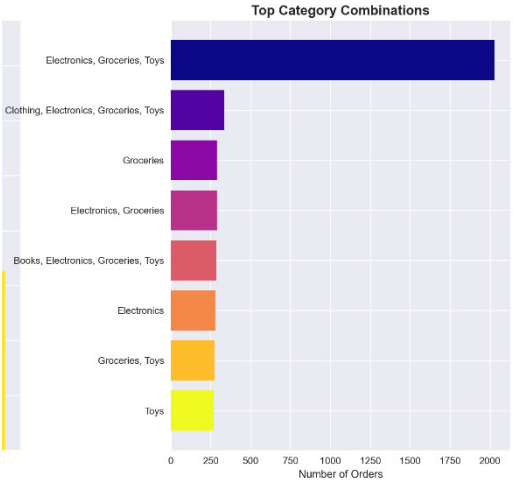
- What it is:
  - averages revenue by month number
- key data analysis:
  - Peaks: Dec ~88,000, Jul ~86,000; lows Feb ~78,000, Sep ~76,000.
  - Average: ~82,000 HKD.
- Trend we observe:
  - consistent cycle suggests holiday/weather influences, for proactive seasonal marketing.

12. Order size Distribution Graph



- What it does:
  - shows order frequency by number of products
- key data analysis:
  - 1: 773 orders.
  - 2: 767.
  - 3: 774.
  - 4: 1,490.
  - 5: 1,481.
  - 6: 654.
- Trend we observe:
  - Peak at 4-5 (~50% of orders), lower at extremes; right-skew toward larger orders, indicating multi-item purchases common, for bundling incentives.

13. Top category Combinations Graph



- What it does:
  - ranks top multi-category order combinations by frequency
- key data analysis:
  - Electronics,Groceries,Toys: ~2,000 orders.
  - Clothing,Electronics,Groceries,Toys: ~1,200.
  - Groceries: ~900.
  - Electronics,Groceries: ~800.
  - Books,Electronics,Groceries,Toys: ~700.
  - Electronics: ~500.
  - Groceries,Toys: ~400.
  - Toys: ~200.
- Trend we observe:
  - Multi-category dominant (top ~2,000 with 3+), single lower; Electronics/Groceries/Toys frequent
  - trending toward complementary bundles, suggesting cross-selling potential in essentials/premiums