Data Analysis Report: E-Commerce Shipping Data

1. General Overview
The dataset contains 10,999 entries related to orders, customers, shipment details, and delivery status.
The basic statistics show:

- Average product cost: varies across the dataset but no extreme values observed.

- Average customer rating: around 3.0, which is relatively low.

- On-time delivery rate: only 40.3% of orders were delivered on time, meaning 59.7% were delayed, which is quite high.

2. Shipping Mode Analysis
- The most commonly used shipping mode is Ship, followed by Road and Flight.

- Average customer ratings are almost identical across all shipping modes.

- No clear relationship was found between the type of shipment and on-time delivery performance.

3. Warehouse Analysis
- The warehouse block with the highest number of orders is F.

- Other warehouses (A, B, C, D) have similar order volumes.

- Consequently, warehouse F also has the highest number of delayed deliveries.

- However, the differences across warehouses are minimal, which seems unusual for real business data.

4. Customer Ratings
- Delivery performance (on-time vs late) is very similar across all rating levels (1 to 5), which is unrealistic.

- The average customer rating per warehouse is also nearly the same, with the highest (Block D) being just 3.2, only 0.6 points higher than the lowest.

- This uniformity raises questions about the authenticity of the dataset.

5. Product Importance
- Surprisingly, high-importance products are delivered late more frequently than medium or low-importance products.

- Again, customer ratings across different product importance levels are nearly identical.

- This contradicts expected real-world behavior, where more important products would typically receive better service.

Conclusion

The analysis indicates that the dataset is likely synthetic or generated, rather than collected from real business operations.
The patterns are too uniform across different categories, and there is no significant variation in key metrics (such as ratings, delivery timing, or customer behavior).
This makes it difficult to draw meaningful insights. Nevertheless, the dataset is still useful for learning analytical methods, building dashboards, and practicing data storytelling.