**VARIABLE INFO**

The dataset contains **180,519 entries** with **53 columns**. Here's an overview of the key features and data types:

### Key Columns:

1. **Type**: Type of order (e.g., product category).
2. **Days\_for\_shipping\_real**: Actual number of days taken for shipping.
3. **Days\_for\_shipment\_scheduled**: Scheduled number of days for shipment.
4. **Benefit\_per\_order**: Benefit earned per order.
5. **Sales\_per\_customer**: Sales amount per customer.
6. **Delivery\_Status**: Status of the delivery (e.g., on time, delayed).
7. **Late\_delivery\_risk**: Indicator of late delivery risk (0 or 1).
8. **Category\_Id**: Product category identifier.
9. **Category\_Name**: Name of the product category.
10. **Customer\_City, Customer\_Country, Customer\_Email, Customer\_Id, etc.**: Customer-related features, including their city, country, and email.
11. **Customer\_Segment**: Customer segmentation (e.g., retail, wholesale).
12. **Order\_City, Order\_Country, Order\_Status, etc.**: Order-related features, such as city, country, and status.
13. **Order\_Profit\_Per\_Order**: Profit earned per order.
14. **Order\_Region**: Region of the order.
15. **Order\_Zipcode**: Zipcode related to the order (with many null values).
16. **Product\_Card\_Id, Product\_Description, Product\_Name, Product\_Price, etc.**: Product details, including card ID, description, price, and status.
17. **Shipping\_Mode**: Mode of shipping (e.g., air, ground).
18. **Order\_Item\_Quantity**: Quantity of products ordered.
19. **Order\_Item\_Discount**: Discount applied to the product.
20. **Order\_Item\_Total**: Total value of the order item.

### Data Quality:

* The dataset contains some **null values**, especially in columns like **Customer\_Lname**, **Customer\_Zipcode**, and **Product\_Description** (with 0 non-null).
* Most columns, including customer, order, and product-related features, are **non-null**.