



# OPTIFLOW

Faster • Smarter • Cheaper

# ABOUT US

- Eyad Allawy
- Amr Abdulazim
- Abdulrahman Mekky
- Mahmoud Rashidy







# PROJECT IDEA

*We analyze supply chain data to spot inefficiencies, streamline product flow, and give actionable insights that guide companies in choosing the **best supply chain strategy** for each product.*



# DATASET OVERVIEW

- **Source:** *Kaggle – DataCo Supply Chain Dataset*
- **About:** *Orders, shipping, customers & products data for global supply chain*
- **Products:** *Clothing, Sports, Electronics*
- **Shape:** *52 columns × 108,519 rows*



# DATATYPES

## Text (Categorical) Columns (33):

Type, Delivery Status, Category Name, Customer City, Customer Country, Customer Email, Customer Fname, Customer Id, Customer Lname, Customer Password, Customer Segment, Customer State, Customer Street, Customer Zipcode, Department Name, Market, Order City, Order Country, Order Customer Id, order date (DateOrders), Order Id, Order Item Cardprod Id, Order Item Id, Order Region, Order State, Order Status, Product Card Id, Product Description, Product Image, Product Name, Product Status, Shipping date (DateOrders), Shipping Mode



# DATATYPES

## <sup>1</sup><sub>3</sub><sup>2</sup><sub>4</sub> Numerical Columns (19):

Days for shipping (real), Days for shipment (scheduled),  
Benefit per order, Sales per customer, Late\_delivery\_risk,  
Category Id, Department Id, Latitude, Longitude, Order  
Item Discount, Order Item Discount Rate, Order Item  
Product Price, Order Item Profit Ratio, Order Item  
Quantity, Sales, Order Item Total, Order Profit Per Order,  
Product Category Id, Product Price

# DATA ISSUES

# Missing Values & Duplicates

OPTIFLOW

- **Customer Lname:** 8 missing → filled with "Unknown"
- **Order Zipcode:** 3 missing → filled with most frequent value (mode)
- ● **Product Description:** fully empty → column dropped
- ● **Customer Zipcode:** 150k+ missing → column dropped
- ✓ **Duplicates:** 0 found → no action needed





# Data Type Conversions

OPTIFLOW

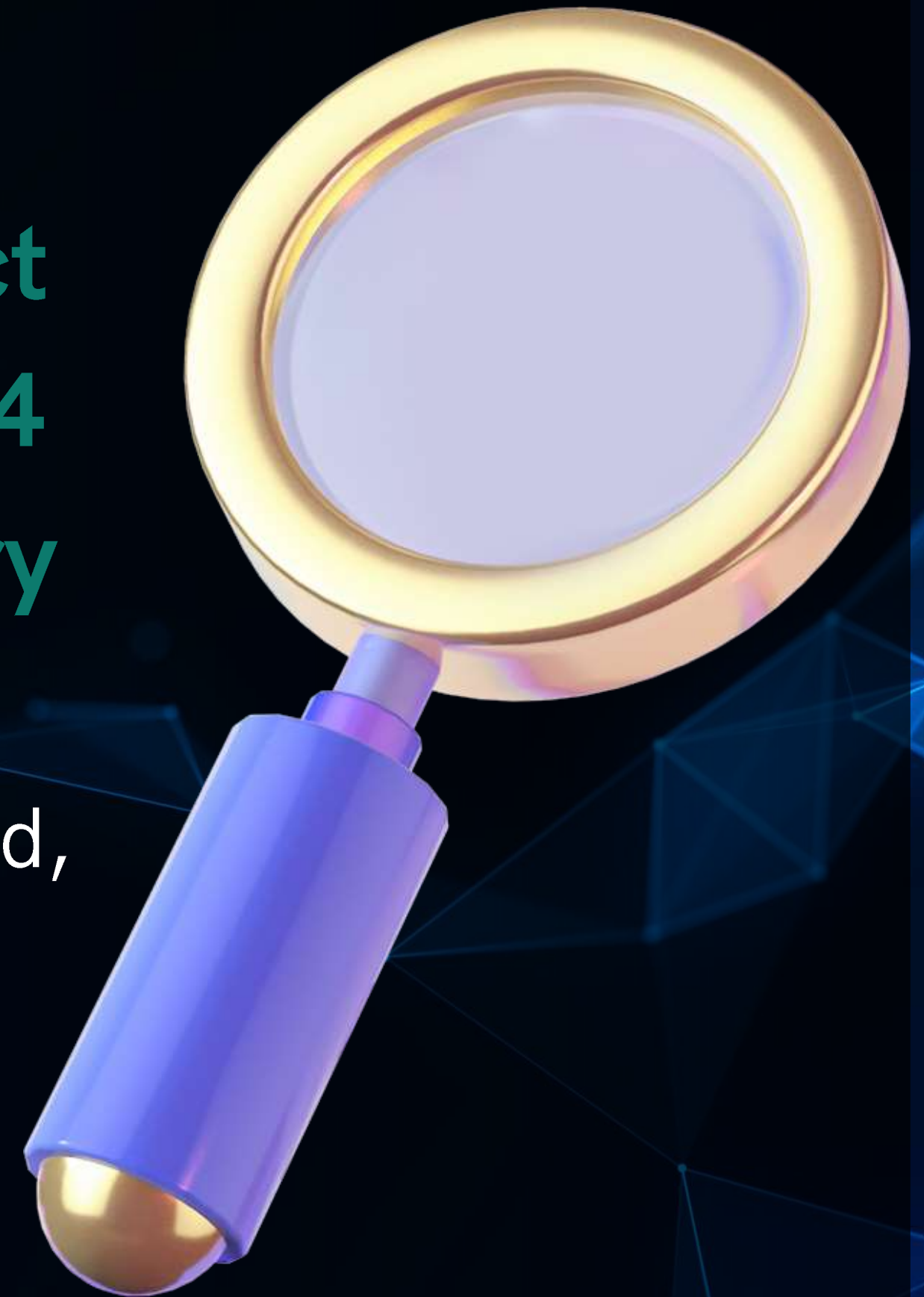
✓ ☐ Customer Zipcode: float64 → **object**

✓ ☐ Date Columns: object → **datetime64**

✓ ☐ Late\_delivery\_risk: int64 → **category**

✓ ☐ ID Columns: int64 → **category**

(Category Id, Department Id, Product Card Id,  
Product Category Id, Product Status)



*How can we help companies  
ship **faster, smarter** and  
**cheaper**?*





# 💡 INSIGHTS OVERVIEW

- **Operational** Insights (Faster)
- **Cost** Optimization (Cheaper)
- **Supplier** & Carrier Performance (Smarter)
- **Customer**-Centric Insights (Smarter)



# OPERATIONAL INSIGHTS

- **Shipping Delay** → Days for shipping (real), Days for shipment (scheduled)
- **On-Time Performance** → Delivery Status, Order Region, Shipping Mode
- **Late Delivery Risk** → Late\_delivery\_risk, order date, shipping date





# COST OPTIMIZATION

- **Shipping Cost Efficiency** → Order Item Discount, Order Profit Per Order
- **High-Cost Regions** → Customer Country, Sales per customer
- **Order Profitability** → Benefit per order





# SUPPLIER & CARRIER PERFORMANCE

- **Supplier/Carrier Ranking** → Shipping Mode, Delivery Status
- **Best Combinations** → Product Category Id, Shipping Mode
- **Lead Time Optimization** → order date, shipping date





# CUSTOMER-CENTRIC INSIGHTS

- **Customer Segmentation** → Customer Segment, Customer Country
- **High-Value Customers** → Sales per customer, Sales
- **Regional Problem Spots** → Order Region, Delivery Status



## ✂ ROLES & RESPONSIBILITIES

- **Amr Abdulazim** → Build Data Model, Data Cleaning & Preprocessing
- **Abdelrahman Mekky** → Analysis Questions Phase
- **Mahmoud Rashidy** → Forecasting Questions Phase
- **Eyad Allawy** → Visualization Dashboard & Final Presentation





*OPTIFLOW*

**THANK YOU!**