

## Dataset: [Customers EDA + Stats](#)

**Dataset Description:** This dataset contains information about an online electronic store. The store has **3 warehouses** from which goods are delivered to customers.

### Columns Description

- **order\_id:** A unique id for each order
- **customer\_id:** A unique id for each customer
- **date:** The date the order was made, given in YYYY-MM-DD format
- **nearest\_warehouse:** A string denoting the name of the nearest warehouse to the customer
- **shopping\_cart:** A list of tuples representing the order items: the first element of the tuple is the item ordered, and the second element is the quantity ordered for such item.
- **order\_price:** A float denoting the order price in USD. The order price is the price of items before any discounts and/or delivery charges are applied.
- **delivery\_charges:** A float representing the delivery charges of the order
- **customer\_lat:** Latitude of the customer's location
- **customer\_long:** Longitude of the customer's location
- **coupon\_discount:** An integer denoting the percentage discount to be applied to the order\_price.
- **order\_total:** A float denoting the total of the order in USD after all discounts and/or delivery charges are applied.
- **season:** A string denoting the season in which the order was placed.
- **is\_expedited\_delivery:** A boolean denoting whether the customer has requested an expedited delivery
- **distance\_to\_nearest\_warehouse:** A float representing the arc distance, in kilometres, between the customer and the nearest warehouse to him/her.
- **latest\_customer\_review:** A string representing the latest customer review on his/her most recent order
- **is\_happy\_customer:** A boolean denoting whether the customer is a happy customer or had an issue with his/her last order.

**Requirements:**

- **Data Cleaning & Trans**
- **EDA**
- **Business Insights**
- **Feature Engineering**
- **Hypothesis testing (Clarify your assumptions)**
- **Informative Dashboard (Bonus) >> plotly dashboards accepted & any other visualizing tools.**
  - **Informative means:** It can be represented in front of business manager for example.