

## Nigam Pranay Mehta

### Dataset Overview

- **Total Rows (Records):** 48,620
- **Total Columns (Features):** 12
- **File Type:** CSV
- **Memory Usage:** ~4.5 MB
- **Time Period Covered:** Year 2015 (based on order\_date column)

The dataset contains transactional sales data from a pizza store, including order details, pizza types, prices, and ingredients.

### Issues Identified in Raw Data (Before Cleaning)

Although the dataset has **no missing values**, several data quality issues are present:

#### **Incorrect Data Types**

- pizza\_id, order\_id, and quantity are stored as **float** instead of integer.
- order\_date and order\_time are stored as **object (string)** instead of proper datetime format.

#### **Date Format Inconsistency**

- Some dates appear as 1/1/2015
- Some appear as 26-11-2015  
This indicates inconsistent formatting and requires standardization.

#### **Derived Column Validation Needed**

- total\_price should equal unit\_price  $\times$  quantity
- Needs verification during cleaning.

#### **Ingredient Column Structure**

- pizza\_ingredients contains multiple ingredients in a single text field.

- May require splitting for ingredient-level analysis.

## Duplicate Check Required

- Since multiple pizzas can belong to the same order\_id, duplicates need to be checked carefully (true duplicates vs valid repeated orders).

## Overall Raw Data Summary

The Pizza Sales dataset contains 48,620 transaction-level records from a pizza store in 2015. It includes detailed order-level and product-level information such as pizza category, size, ingredients, pricing, and timestamps.

Although the dataset does not contain missing values, several structural and formatting inconsistencies exist, particularly in data types and date formats. These issues must be addressed during the data cleaning process to ensure accurate analysis and reporting.

## Sample Image of Raw Data Sheet

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	pizza_id	order_id	pizza_name	quantity	order_date	order_time	unit_price	total_price	pizza_size	pizza_category	ingredients	pizza_name						
1	1	1	hawaiian_r	1	11:38:36	11:38:36	13.25	13.25	M	Classic	Sliced Ham	The Hawaiian Pizza						
2	2	2	classic_db	1	11:57:40	11:57:40	16	16	M	Classic	Pepperoni	The Classic Deluxe Pizza						
3	3	2	five_chees	1	11:57:40	11:57:40	18.5	18.5	L	Veggie	Mozzarella	The Five Cheese Pizza						
4	4	2	ital_supr_l	1	11:57:40	11:57:40	20.75	20.75	L	Supreme	Calabrese	The Italian Supreme Pizza						
5	5	2	mexicana	1	11:57:40	11:57:40	16	16	M	Veggie	Tomatoes	The Mexicana Pizza						
6	6	2	thai_ckn_l	1	11:57:40	11:57:40	20.75	20.75	L	Chicken	Chicken	The Thai Chicken Pizza						
7	7	3	ital_supr_r	1	12:12:28	12:12:28	16.5	16.5	M	Supreme	Calabrese	The Italian Supreme Pizza						
8	8	3	prsc_argla	1	12:12:28	12:12:28	20.75	20.75	L	Supreme	Prosciutto	The Prosciutto and Arugula Pizza						
9	9	4	ital_supr_r	1	12:16:31	12:16:31	16.5	16.5	M	Supreme	Calabrese	The Italian Supreme Pizza						
10	10	5	ital_supr_r	1	12:21:30	12:21:30	16.5	16.5	M	Supreme	Calabrese	The Italian Supreme Pizza						
11	11	6	bbq_ckn_s	1	12:29:36	12:29:36	12.75	12.75	S	Chicken	Barbecued	The Barbecue Chicken Pizza						
12	12	6	the_greek	1	12:29:36	12:29:36	12	12	S	Classic	Kalamata	The Greek Pizza						
13	13	7	spinach_si	1	12:50:37	12:50:37	12.5	12.5	S	Supreme	Spinach	The Spinach Supreme Pizza						
14	14	8	spinach_si	1	12:51:37	12:51:37	12.5	12.5	S	Supreme	Spinach	The Spinach Supreme Pizza						
15	15	9	classic_db	1	12:52:01	12:52:01	12	12	S	Classic	Pepperoni	The Classic Deluxe Pizza						
16	16	9	green_gar	1	12:52:01	12:52:01	12	12	S	Veggie	Spinach	The Green Garden Pizza						
17	17	9	ital_cpcllo	1	12:52:01	12:52:01	20.5	20.5	L	Classic	Capocollo	The Italian Capocollo Pizza						
18	18	9	ital_supr_l	1	12:52:01	12:52:01	20.75	20.75	L	Supreme	Calabrese	The Italian Supreme Pizza						
19	19	9	ital_supr_s	1	12:52:01	12:52:01	12.5	12.5	S	Supreme	Calabrese	The Italian Supreme Pizza						
20	20	9	mexicana	1	12:52:01	12:52:01	12	12	S	Veggie	Tomatoes	The Mexicana Pizza						
21	21	9	spicy_ital	1	12:52:01	12:52:01	20.75	20.75	L	Supreme	Capocollo	The Spicy Italian Pizza						
22	22	9	spin_pesto	1	12:52:01	12:52:01	20.75	20.75	L	Veggie	Spinach	The Spinach Pesto Pizza						

