

# Pizza Sales Analysis Project Summary

## Project Overview

This Python-based analysis explores a comprehensive pizza sales dataset containing 48,620 transactions. Using powerful data science libraries, we uncover key business insights to optimize sales strategies and understand customer preferences.

## Tools & Libraries Used

- Pandas & NumPy: For data loading, manipulation, and statistical analysis
- Matplotlib & Seaborn: For visualizations (not fully implemented in the provided snippet)
- Jupyter Notebook: Interactive analysis environment

## Data Processing Steps

1. Initial Exploration:
  - Loaded 48,620 records with 14 original features
  - Identified sales metrics, pizza categories, sizes, and timestamps
2. Data Cleaning:
  - Removed redundant columns (Unnamed: 12, pizza\_ingredients)
  - Handled missing values (48579 nulls in total column)
  - Verified no duplicate records exist
  - Standardized categorical data (e.g., "Classic" → "classic")
3. Feature Engineering:
  - Calculated key metrics:
    - Total revenue: \$817,860.05
    - Average transaction value: \$16.82
  - Prepared time-based features for trend analysis

## Key Insights

1. Sales Performance:
  - Highest daily sales observed on November 26, 2015 (261 transactions)
  - Peak order time: 12:32 PM (26 orders at this timestamp)
2. Product Analysis:

- Most popular pizza: Classic Deluxe (M) (2,416 orders)
  - Category distribution: Classic (28%), Veggie (24%), Supreme (19%), Chicken (29%)
  - Size preference: Large (38%) dominates sales
3. Pricing Analysis:
- Price range: \$9.75 - \$35.95 per pizza
  - Average pizza price: \$16.49

#### Potential Visualizations (Implicit)

- Revenue trends by day/month
- Category sales distribution pie charts
- Heatmaps of order time frequency
- Price distribution histograms
- Top-selling pizza rankings

#### Business Value

This analysis provides actionable insights for:

- Menu optimization (highlight bestsellers)
- Inventory management (size/category demand)
- Promotional timing (peak hours/days)
- Pricing strategy refinement
- Resource allocation planning

The project demonstrates professional-grade data wrangling skills and lays the foundation for impactful data-driven decision-making in the food industry. The clean, documented workflow ensures reproducibility and scalability for future analyses.

Impressive Note: The project handles real-world data complexities, including missing values, inconsistent casing, and irrelevant columns, showcasing robust analytical skills that would benefit any business intelligence team.