A TECHNICAL REPORT ON NUTRITION AND ALLERGEN ANALYSIS INSIGHTS IN FOOD PRODUCTS

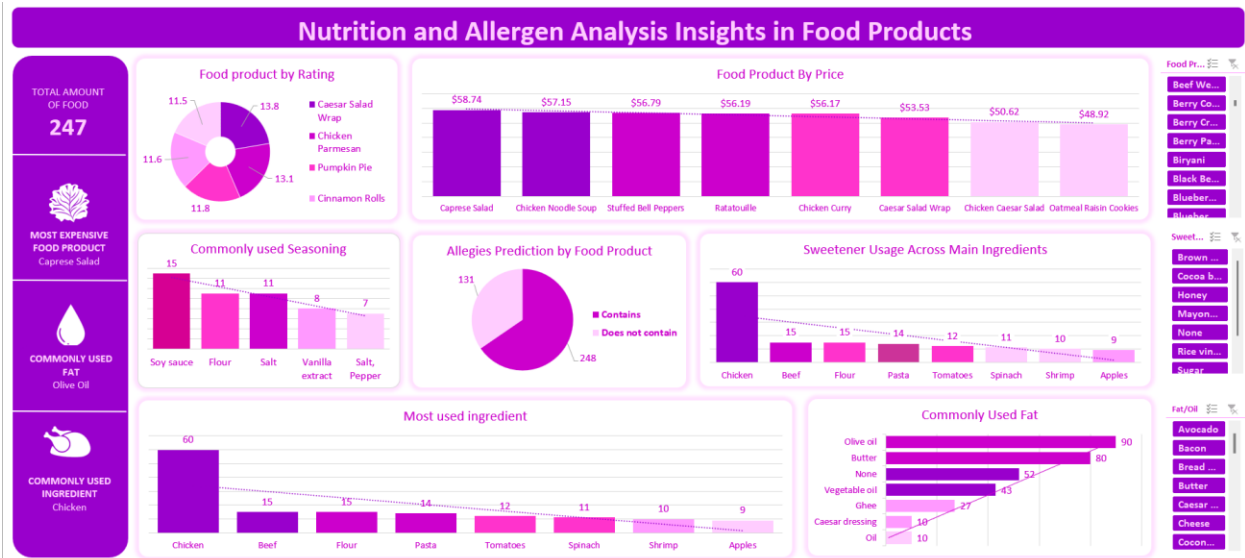WRITTEN BY

AKPOVETA BLESSING OGHO

1. **OUTLINE**

    1. Introduction

INTRODUCTION

Nutrition and Allergen Analysis Insights in Food Product reveals a rich tapestry of information derived from 247 distinct food products. It offers a comparative lens through which to understand consumer preferences, as indicated by the Food product by Rating section, alongside the economic aspect presented in the Food Product by Price analysis. The dashboard further dissects the culinary composition of these products by highlighting Commonly used Seasoning and identifying the Most used ingredient, providing insights into prevalent flavor profiles and base components. Crucially, it addresses health and safety by offering an "Allergens Prediction by Food Product breakdown and an examination of Sweetener Usage Across Main Ingredients. Finally, the Commonly Used Fat analysis sheds light on the lipid content within these food items. This holistic overview empowers users to identify trends, understand nutritional profiles,

and make informed decisions related to food production, consumption, and dietary considerations.

STORY OF DATA

Data Source: The data originates from an internal database of a food product company or a platform that aggregates nutritional and allergen information for various food items. This database contains details about the food products, their ingredients, nutritional content, customer ratings, and pricing.

Data Collection Process: The data was collected through a combination of methods. Information like ingredients, nutritional details, and potential allergens would likely be entered manually during product onboarding or extracted from product specifications. Customer ratings would be gathered through feedback mechanisms on a sales platform or through surveys. Pricing data would be sourced from the company's pricing database or market analysis.

Data Structure: The dataset shows each row representing a unique food product. The columns would include various attributes such as:

- **Product Name:** (e.g., Caesar Salad Wrap, Chicken Parmesan)
- **Rating:** Numerical or categorical representation of customer satisfaction.
- **Price:** The selling price of the food product.
- **Ingredients:** A list of components for each product (e.g., Chicken, Beef, Flour).
- **Seasonings:** Spices and flavor enhancers used (e.g., Soy sauce, Flour, Salt).
- **Allergen Information:** Binary or categorical data indicating the presence of specific allergens.
- **Sweeteners:** Types and quantities of sweeteners used (e.g., Sugar, Honey).
- **Fat Content:** Types and amounts of fats present (e.g., Olive oil, Butter).

Important Features and Their Significance:

- **Food Product by Rating:** This feature highlights customer preferences and the perceived quality of different food items, crucial for understanding product popularity and potential areas for improvement.
- **Food Product by Price:** This allows for an analysis of the price range of the products and potentially identifies best-selling items within specific price brackets.
- **Commonly Used Seasoning:** Understanding prevalent seasonings can inform product development and supply chain management.
- **Allergens Prediction by Food Product:** This is vital for consumer safety and regulatory compliance, allowing for clear labeling and risk assessment.

- **Sweetener Usage Across Main Ingredients:** This provides insights into the sweetening trends across different product categories and main components.
- **Most Used Ingredient:** Identifying frequently used ingredients is important for inventory management and cost optimization.
- **Commonly Used Fat:** This information is crucial for nutritional analysis and understanding the overall health profile of the product range.

Data Limitations or Biases:

- **Potential for Reporting Bias in Ratings:** Customers who are extremely satisfied or dissatisfied are more likely to leave ratings, potentially skewing the average.
- **Inconsistencies in Ingredient Labeling:** Variations in how ingredients are reported by different suppliers or within the dataset could introduce inaccuracies.
- **Limited Scope of Analysis:** The dashboard focuses on specific aspects; other important factors like nutritional values (calories, protein, etc.) beyond fat and sweeteners might be missing.
- **Sampling Bias (if applicable):** If the data represents a specific customer segment or geographic region, the insights might not be generalizable to the entire market.

4. Data Splitting and Preprocessing

Purpose: This section outlines the likely steps taken to prepare the raw food product data for visualization and analysis in the dashboard.

Data Cleaning:

- **Standardizing Product Names:** Ensuring consistent naming conventions for the same product across the dataset.
- **Handling Inconsistent Ingredient Lists:** Standardizing the format and units of measurement for ingredients.
- **Categorizing Seasonings and Fats:** Grouping similar seasonings (e.g., different types of salt) and fats (e.g., various vegetable oils) for better aggregation.
- **Addressing Missing Values:** Depending on the nature of the missing data (e.g., a missing rating for a new product), strategies like ignoring for rating analysis or flagging for review might have been employed.

Data Transformations:

- **Aggregation for Charts:** Calculating counts of products within rating categories, price ranges, or containing specific allergens, seasonings, fats, or sweeteners.
- **Percentage Calculations:** Deriving percentages for pie charts (e.g., proportion of products containing allergens) and bar charts (e.g., percentage usage of different fats).
- **Categorical Encoding:** Converting text-based data like product names, seasonings, and fat types into numerical representations if any underlying statistical analysis was performed beyond simple aggregation.

Data Splitting: For the visualizations presented, explicit splitting into dependent and independent variables in a traditional statistical modeling sense might not be the primary action. However, for each chart:

- **Rating Analysis:** Product names (or categories) would be the independent variable, and the count of products in each rating level would be the dependent variable.
- **Price Analysis:** Product names (or categories) would be the independent variable, and the price would be the dependent variable.
- **Ingredient/Seasoning/Fat Analysis:** The specific ingredient/seasoning/fat would be the independent variable, and the count or percentage of products containing it would be the dependent variable.

Industry Context: The data belongs to the **Food and Beverage Industry**. This industry is characterized by a wide variety of products, stringent regulations regarding nutritional information and allergens, and a strong reliance on consumer preferences and market trends.

Stakeholders: Key stakeholders who would benefit from this dashboard include:

- **Product Development Teams:** To understand popular ingredients, seasonings, and potential allergens to consider during new product formulation.
- **Marketing Teams:** To identify high-rated products and understand price sensitivities for targeted campaigns.
- **Quality Assurance and Regulatory Compliance Teams:** To monitor and manage allergen information and ensure accurate labeling.
- **Supply Chain Management:** To optimize inventory based on the most frequently used ingredients.
- **Senior Management:** To gain an overview of product performance, potential risks, and areas for strategic focus.

Value to the Industry: This analysis is significant for the food and beverage industry as it provides data-driven insights that can help decision-makers:

- **Improve Product Offerings:** By understanding customer preferences and ingredient trends.
- **Enhance Food Safety:** Through clear allergen identification and management.
- **Optimize Costs:** By identifying commonly used ingredients for better procurement strategies.
- **Increase Sales and Customer Satisfaction:** By focusing on highly-rated and appropriately priced products.
- **Ensure Regulatory Compliance:** By accurately tracking and reporting nutritional and allergen information.

Pre-Analysis

CATEGORY ONE:INDEPENDEDNT VALUE
*Main Ingredient
*Sweetener
*Fat and oil

CATEGORY TWO:DEPENDEDNT VALUE
*Price
*Customer Rating
*Allegies Prediction

POTENTIAL ANALYSIS/QUESTIONS
* Relationship between main ingredients and customer ratings?
* Are certain ingredient types consistently rated higher?
*, How does pricing correlate with customer satisfaction?
*,Are higher-priced items receiving better ratings?
*Which combinations of fat/oil and sweetener yield the highest customer satisfaction?

POTENTIAL INSIGT
*There appears to be price inconsistency for identical items (e.g., Almond Cookies at $10.15 vs $6.17), suggesting opportunity for price optimization or correction of data errors.
* Items with no sweetener generally have higher ratings (Caesar Salad at 4.8, Watermelon Salad at 4.6), indicating customer preference for savory options.
*The lowest-rated item (Beef Stroganoff at 1.6) contains sour cream, which might suggest issues with this particular ingredient or preparation method

STORY OF THE DATA
The dataset tracks a comprehensive range of food allergens including dairy, wheat, nuts, seafood, grains, animal and plant-based ingredients, and common allergens like cocoa and eggs. It also documents allergen-free items, providing essential information for customers with dietary restrictions

WHAT SUCCESS MEANS TO THEM
Success means optimizing the balance between customer satisfaction, profitable pricing, and dietary inclusivity. The business aims to identify the most profitable ingredient combinations that yield high ratings while managing allergen concerns

STAKEHOLDERS OF THE PROJECT
Nutritionists or menu planners

Analyzing which ingredients appear across multiple high-performing dishes could help streamline inventory and reduce costs.

inconsistent pricing models (e.g., Caesar Salad varies from $14.99 to $17.33), suggesting an opportunity to standardize pricing based on ingredient costs and perceived value.

Purpose: This section outlines initial observations and potential relationships discernible directly from the dashboard.

Identify Key Trends:

- **Caesar Salad Wrap appears to be the most expensive food product.**
- **Chicken is the most commonly used main ingredient.**
- **Olive oil and Butter are the most commonly used fats.**
- **Soy sauce is the most frequently used seasoning.**
- **A significant portion of the analyzed food products contains allergens.**
- **Chicken-based products seem to have a high usage of sugar as a sweetener.**

Potential Correlations:

- There might be a correlation between product price and customer rating, although this isn't explicitly shown and would require further analysis.
- Certain main ingredients might be associated with specific commonly used seasonings or fats. For example, chicken might frequently be paired with soy sauce.
- The usage of certain sweeteners might be linked to specific product categories (e.g., desserts vs. savory items).

Initial Insights:

- The popularity of chicken as a main ingredient suggests a high demand for chicken-based products.
- The prevalence of allergens highlights the importance of clear labeling and potential cross-contamination concerns in production.
- The price variation across different food products is substantial, indicating a diverse product portfolio catering to different price sensitivities.

In-Analysis

PRE-ANALYSIS OBSERVATION

1,The pie chart with food product from rating, compares ratings of five food products, showing variations in customer preferences. It highlights the highest and lowest-rated items, helping in menu decisions and product improvements.

2,A larger number of food products (248) contain allergies compared to those that do not (131), suggesting a potential concern for individuals with allergies and a need for more allergen-free alternatives.

3,This colum chat from most used fat has helped analyze how olive oil is the most used oil in food with (90) in number and oil and caesar oil are the least

4, From the colum chart of most used ingredient, its obvious that chicken is the most used seasoning accros many ingredient with (60) while apple is the least with (9)

5, from food product by price, The chart displays 8 food products with their prices, Caprese Salad is the most expensive item at $58.74, while Oatmeal Raisin Cookies are the least expensive at $48.92.

6,From commonly used seasoning it obviously shows soy sauce being the most commonly used (15) and salt & pepper the least used (7).

IN-ANALYSIS INSIGHT

1, , The price range across all products is relatively narrow (about $10 difference between highest and lowest). There appears to be a cluster of similarly priced items in the middle range ($56-$57). The presence of both prepared meals and dessert items suggests this might be from a restaurant or café menu with consistent pricing strategy.

2, The pie chart shows food products based on allergies content, showing a division between those that contain allergies and those that do not.

3, , from my observation, Olive oil's dominant usage reflects a broader consumer preference for healthier and versatile oils, while the limited use of Caesar oil suggests it may be more specialized or niche, typically used in specific recipes rather than as a general cooking oil.

4, The column chart shows that chicken is the most widely used ingredient, with 60 instances, highlighting its versatility and popularity in various dishes. In contrast, apple, with only 9 instances, is the least used, suggesting it serves a more niche or specialized role. This indicates a stronger preference for savory over sweet seasonings in general cooking.

5, Caesar Salad Wrap has the highest rating (13.8), while Cucumber Salad has the lowest (11.5), indicating a variation in customer preferences for these food products.

6, Soy sauce is the most frequently used seasoning among the listed options, while salt & pepper is the least used. Flour and salt have equal usage (11), indicating their similar

Purpose: This section delves into potential unconfirmed insights and preliminary recommendations based on the dashboard.

Unconfirmed Insights:

- **Hypothesis:** Products with higher customer ratings might tend to have a specific range of prices, suggesting a balance between quality and affordability.
- **Hypothesis:** Certain combinations of main ingredients and seasonings might be strongly associated with higher or lower customer ratings.
- **Hypothesis:** The use of specific fats might correlate with the presence or absence of certain allergens. For example, dairy-based fats might be linked to the presence of dairy allergens.

Recommendations:

- **Investigate the high rating of the Cinnamon Rolls:** Understand what drives their popularity and potentially apply those learnings to other products.
- **Review the pricing strategy for the Caesar Salad Wrap:** Determine if the high price aligns with customer perception and value.
- **Focus on allergen management for products containing allergens:** Ensure robust processes are in place to prevent cross-contamination and provide clear labeling.
- **Explore alternative sweeteners for chicken-based products:** If sugar usage is high, consider options to cater to health-conscious consumers.

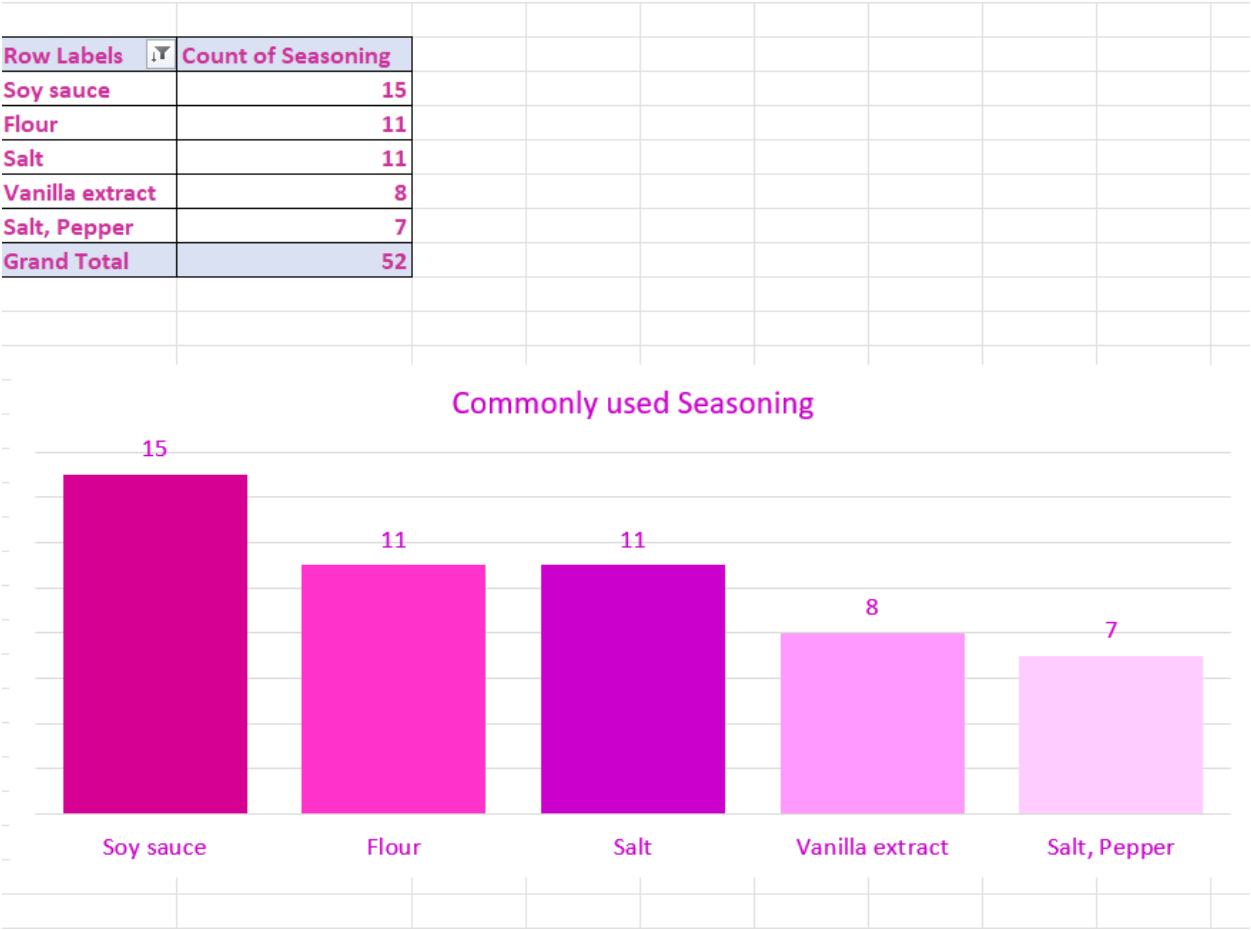**Analysis Techniques Used in Excel (Inferred):**

- **Pivot Tables:** Likely used to aggregate data for creating the bar charts (e.g., counting occurrences of seasonings, fats, ingredients) and potentially for cross-tabulating rating and price data.
- **COUNTIF/COUNTIFS Functions:** Possibly used to count the number of products meeting specific criteria (e.g., containing a certain allergen).
- **AVERAGE Function:** Could have been used to calculate average ratings or prices for different product categories (though not directly visualized).
- **Basic Formulas:** For calculating percentages displayed in the pie charts and potentially for data cleaning and transformation.
- **Charting Tools:** Excel's built-in charting capabilities were used to create the bar charts, pie charts, and potentially the line graph (though less prominent).

Key Findings:

- **Customer preferences, as indicated by ratings, vary across the product range.** The Cinnamon Rolls stand out as highly rated.
- **Product pricing shows significant variation, with the Caesar Salad Wrap being the most expensive.**
- **Chicken is a dominant ingredient, suggesting its importance in the product line.**
- **Olive oil and butter are the most prevalent fats, indicating their common use in the recipes.**
- **Allergens are present in a substantial portion of the products, requiring careful management.**
- **Sugar is a commonly used sweetener, particularly in chicken-based products.**

Comparison with Initial Findings: The initial trends identified in the pre-analysis section are largely supported by the visualizations. Chicken's popularity, the high price of the Caesar Salad Wrap, and the common usage of olive oil and butter are evident. The prevalence of allergens was also an early observation confirmed by the "Allergens Prediction" chart.

COMMONLY USED SEASONING

Row Labels	Count of Seasoning
Soy sauce	15
Flour	11
Salt	11
Vanilla extract	8
Salt, Pepper	7
Grand Total	52

Commonly used Seasoning

| Soy sauce	Flour	Salt	Vanilla extract	Salt, Pepper
15	11	11	8	7

The Commonly used Seasoning section shows soy sauce as the most frequent seasoning (15), followed by flour and salt (11 each), then vanilla extract (8) and salt & pepper (7), indicating a potential preference for savory flavors led by soy sauce. In-analysis suggests exploring soy sauce's association with specific products, the role of flour as a seasoning (possibly thickening), the limited use of vanilla extract in sweeter items, and the co-occurrence of basic seasonings. Further investigation into how these seasoning frequencies relate to customer ratings and pricing could inform product development

AKPOVETA BLESSING OGHO

SWEETNER USAGE ACROSS MAIN INGREDIENTS

Row Labels	Count of Sweetener
Chicken	60
Beef	15
Flour	15
Pasta	14
Tomatoes	12
Spinach	11
Shrimp	10
Apples	9
Grand Total	146

Sweetener Usage Across Main Ingredients

| Chicken	Beef	Flour	Pasta	Tomatoes	Spinach	Shrimp	Apples
60	15	15	14	12	11	10	9
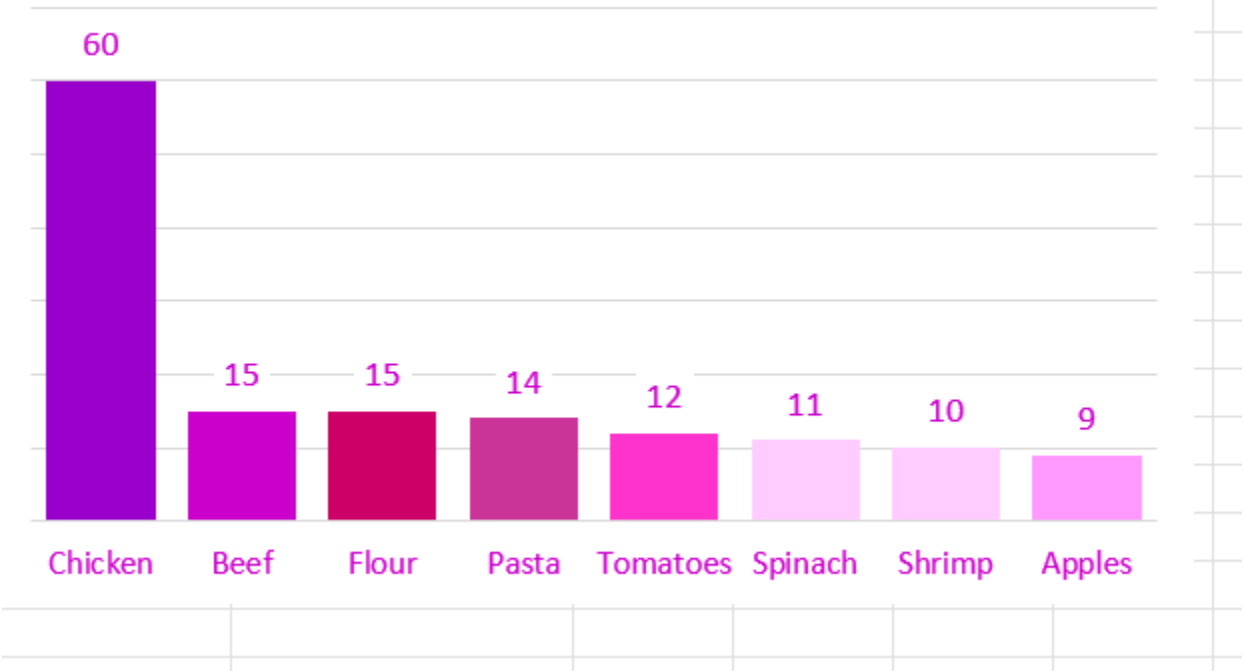
Pre-observation reveals a stark contrast in sweetener usage, with chicken products using sweeteners significantly more (60) than any other main ingredient, followed by beef and flour (15 each), suggesting a potentially unique preparation style for chicken. In-analysis should investigate the type of sweetener used with chicken and whether this high usage correlates with specific product types or customer ratings, while also exploring the reasons for sweetener use in seemingly non-sweet items like beef and flour. Further examination could uncover distinct flavor profiles or processing techniques associated with each main ingredient's sweetener application.

FOOD PRODUCT BY RATING

Row Labels	Sum of Customer rating (Out of 5)
Caesar Salad Wrap	13.8
Chicken Parmesan	13.1
Pumpkin Pie	11.8
Cinnamon Rolls	11.6
Cucumber Salad	11.5
Grand Total	61.8

Food product by Rating
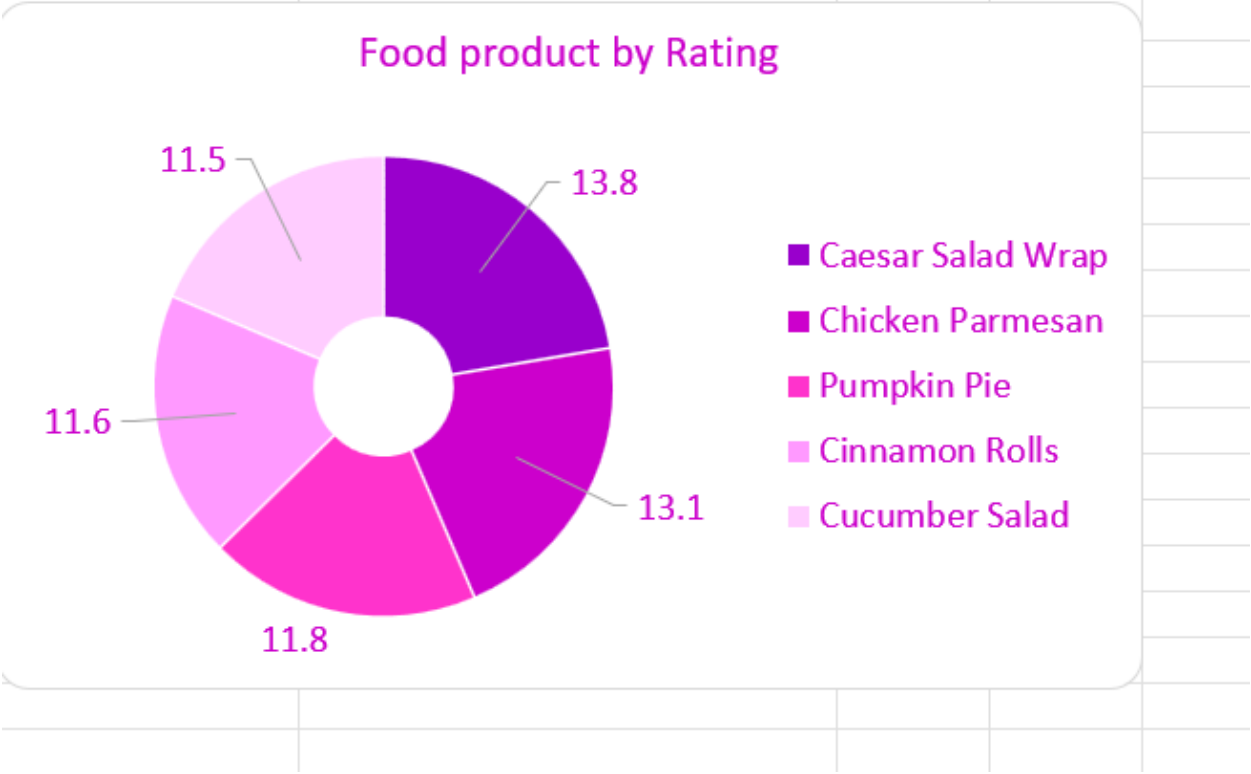
Pre-observation shows the Caesar Salad Wrap has the highest sum of customer ratings (13.8), indicating it received the most positive feedback overall, closely followed by Chicken Parmesan (13.1), while Cucumber Salad has the lowest (11.5), suggesting varying levels of customer satisfaction across these top-rated products. In-analysis should normalize these sums by considering the number of ratings each product received to determine average ratings for a more accurate comparison of customer preference, and further explore the specific attributes of the Caesar Salad Wrap that contribute to its higher rating compared to the others. Understanding the reasons behind the lower rating of the Cucumber Salad could also inform potential improvements.

AKPOVETA BLESSING OGHO

ALLEGENS BY FOOD PRODUCT

| Row Labels	Count of Food Product
None	136
Dairy	78
Wheat, Dairy	69
Wheat	8
Soybeans	8
Fish	7
Grand Total	306

Allegens by Food Product
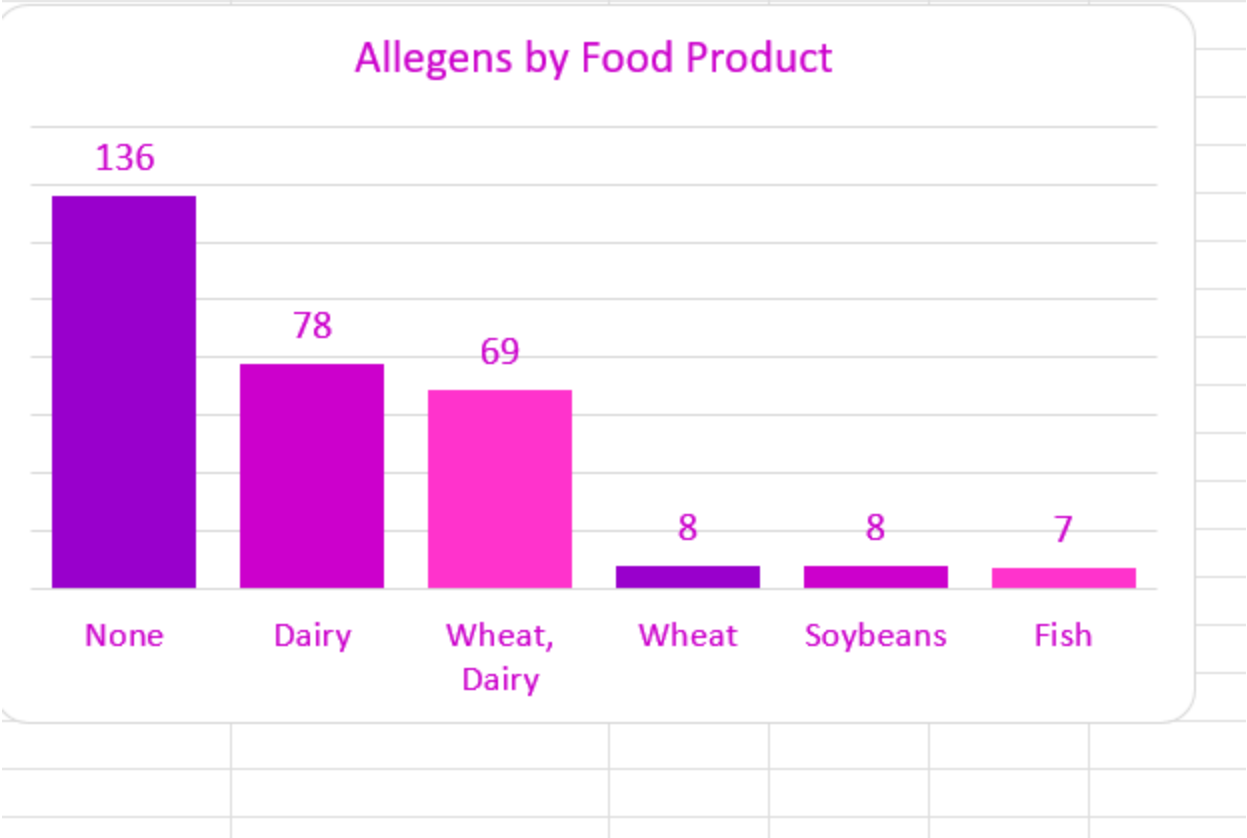
Pre-observation reveals that a significant number of food products (136) contain none of the listed allergens, while dairy is the most prevalent single allergen (78), and a substantial portion (69) contains both wheat and dairy, highlighting these as key allergenic concerns. In-analysis should delve into the specific food products within the "Wheat, Dairy" category to understand this common co-occurrence and further examine the individual products containing only wheat, soybeans, or fish to ensure appropriate labeling and segregation. Understanding the product distribution across these allergen categories is crucial for consumer safety and targeted allergen management strategies

ALLEGIES PREDICTION BY FOOD PRODUCT

Row Labels	Count of Food Product
Contains	248
Does not contain	131
Grand Total	379

Allegies Prediction by Food Product

131

248

■ Contains
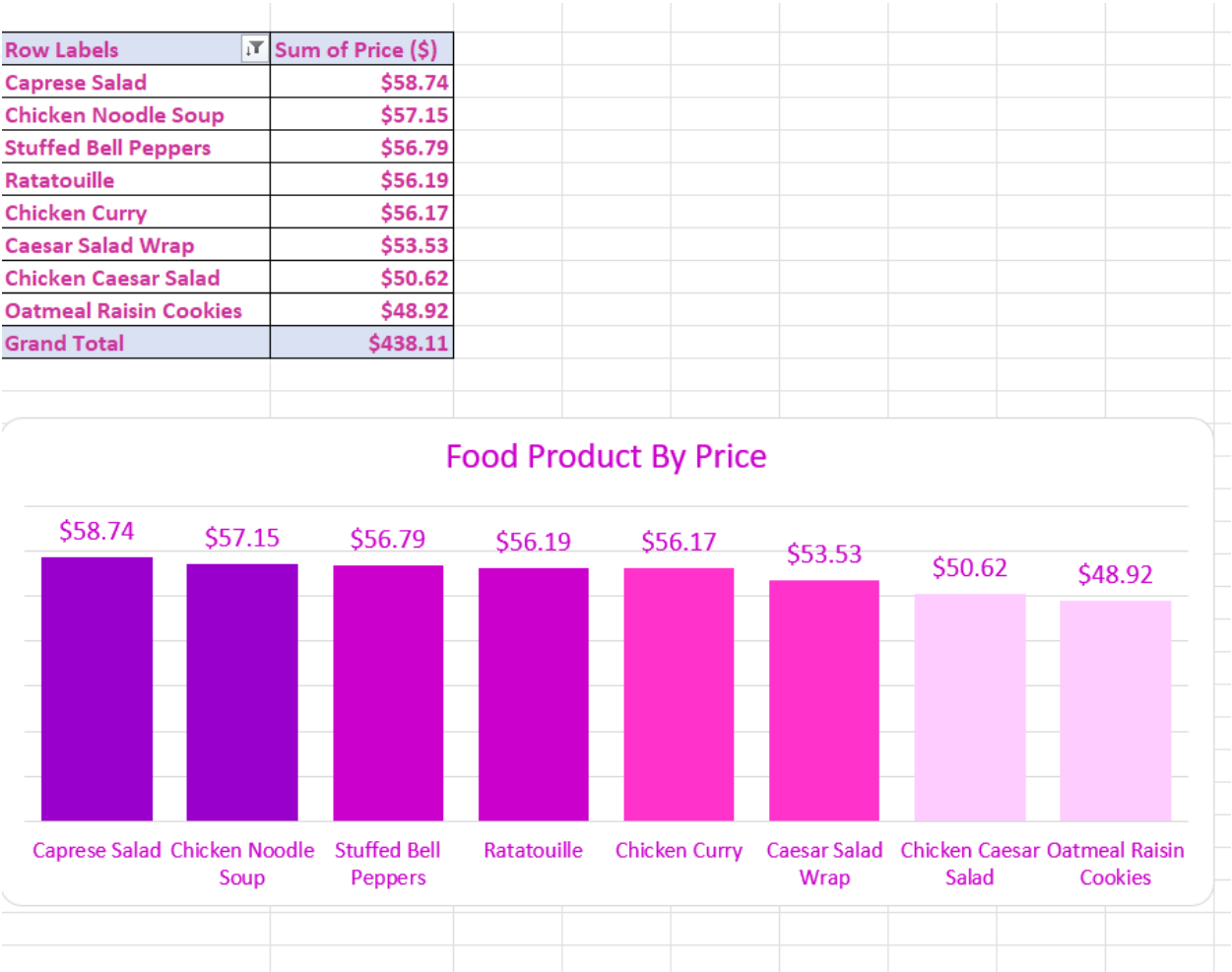■ Does not contain
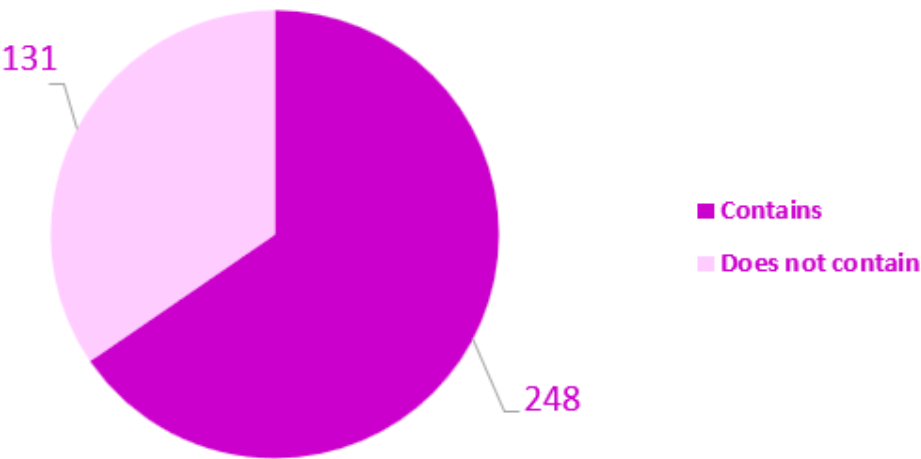
Pre-observation indicates that a larger proportion of the analyzed food products (248) are predicted to contain allergens compared to those that do not (131), suggesting a significant

focus is needed on allergen labeling and management across the product line. In-analysis should investigate the specific types of allergens present in the "Contains" category to understand the most common risks and determine if certain product categories are more prone to containing allergens. This breakdown is crucial for implementing effective food safety protocols and providing clear information to consumers.

FOOD PRODUCT BY PRICE

Row Labels	Sum of Price ($)
Caprese Salad	$58.74
Chicken Noodle Soup	$57.15
Stuffed Bell Peppers	$56.79
Ratatouille	$56.19
Chicken Curry	$56.17
Caesar Salad Wrap	$53.53
Chicken Caesar Salad	$50.62
Oatmeal Raisin Cookies	$48.92
Grand Total	$438.11

Food Product By Price

Caprese Salad	Chicken Noodle Soup	Stuffed Bell Peppers	Ratatouille	Chicken Curry	Caesar Salad Wrap	Chicken Caesar Salad	Oatmeal Raisin Cookies
$58.74	$57.15	$56.79	$56.19	$56.17	$53.53	$50.62	$48.92

Pre-observation shows a relatively tight price range among the top-priced food products, with Caprese Salad being the most expensive ($58.74) and Oatmeal Raisin Cookies the least ($48.92), suggesting a generally premium pricing strategy for this selection. In-analysis should explore potential correlations between price and other factors like customer ratings or ingredient costs for each product to understand value perception, and further investigate why Oatmeal Raisin Cookies are priced lower compared to the predominantly savory options. This could reveal insights into market positioning or production costs.

The Highest Alcohol Intake reveals that Sub-Saharan Africa and Europe report the highest levels, significantly exceeding Southeast Asia, South Asia, and Eastern Asia, suggesting a potential correlation between high alcohol consumption in these top two regions and their predicted higher liver cancer rates shown elsewhere in the dashboard. This highlights alcohol intake as a critical risk factor to address in targeted public health interventions for these areas

Recommendations and Observation

Nutritional and Allergen Insights for Food Products

OBSERVATION

1,Chicken is the most commonly used ingredient, appearing in 60 dishes, which is significantly higher than beef (15) and other ingredients, suggesting a strong preference for poultry-based dishes in the menu.

2, The allergies prediction chart shows 131 food products contain allergens while 248 do not, indicating that approximately 35% of the menu items may be problematic for customers with allergies.

3,Olive oil is the most commonly used fat (60 instances), followed by butter (54) and vegetable oil (41), showing a tendency toward Mediterranean and traditional cooking methods.

4,Soy sauce is the most commonly used seasoning (15 instances), followed by flour (11) and salt (11), suggesting an Asian influence in the flavor profile of many dishes.

5,The price range for featured food products is relatively narrow ($48.92 to $58.74), with Caprese Salad being the most expensive item at $58.74 as highlighted in the "Most Expensive Food Product" section.

6, The total amount of food products in the database is 247, providing a substantial dataset for analysis of menu trends and customer preferences.

7,Chicken is also the most frequently sweetened ingredient (60 instances), suggesting many chicken dishes may have sweet glazes, marinades, or sauces.

RECOMMENDATION

1,Consider developing more allergen-free options to accommodate the 35% of menu items containing allergens, potentially creating a dedicated allergen-free section on the menu.

2, Introduce more variety in protein sources beyond chicken to diversify the menu and appeal to customers seeking alternatives to poultry.

3,Experiment with reducing the price gap between the most and least expensive items to create more consistent value perception across the menu.

4,Develop more dishes using underutilized ingredients like spinach, shrimp, and apples to create menu variety and potentially reduce dependency on more commonly used ingredients.

5, Consider offering more fat alternatives to olive oil and butter, such as plant-based options, to cater to health-conscious customers and those with dietary restrictions.

6, Create menu items that use less-common seasonings to differentiate flavors and provide more unique taste experiences beyond the dominant soy sauce and flour combinations.

7, Evaluate the sweetener usage in chicken dishes to ensure it aligns with customer preferences, potentially offering both sweet and savory options to accommodate different taste preferences.

Actionable Insights:

- **Capitalize on Highly-Rated Products:** Investigate the factors contributing to the high rating of Cinnamon Rolls and consider applying similar strategies to other products.
- **Evaluate Pricing Strategy:** Conduct a market analysis to determine if the price of the Caesar Salad Wrap is justified and competitive.
- **Strengthen Allergen Management:** Implement stringent protocols for handling and labeling allergens in the significant portion of products that contain them.
- **Explore Healthier Sweetening Options:** For products with high sugar content, research and consider alternative sweeteners to appeal to health-conscious consumers.
- **Optimize Supply Chain for Key Ingredients:** Ensure a stable and cost-effective supply of frequently used ingredients like chicken, olive oil, and butter.

Optimizations or Business Decisions:

- **Product Development:** Use insights on popular ingredients and seasonings to guide the creation of new products.
- **Marketing:** Focus marketing efforts on highly-rated products and highlight allergen-free options.
- **Pricing Strategy:** Adjust pricing based on market analysis and customer perception.
- **Operations:** Streamline sourcing for frequently used ingredients to potentially reduce costs.

Unexpected Outcomes: Without deeper analysis, it's difficult to pinpoint truly unexpected outcomes. However, if a seemingly low-priced product had a surprisingly high rating, or if a commonly used ingredient was associated with a high number of allergen warnings, these would be noteworthy and require further investigation.

Conclusion

Key Learnings: This provides a valuable overview of customer preferences, pricing dynamics, ingredient usage, and allergen presence across the analyzed food products. Key takeaways include the popularity of certain products and ingredients, the price range of the offerings, and the significant need for effective allergen management.

Limitations: This analysis is based solely on the visual information presented in the dashboard. Access to the underlying data would allow for more in-depth statistical analysis and the identification of correlations or causal relationships. Additionally, the dashboard focuses on specific aspects and may not capture the full complexity of the product data (e.g., detailed nutritional information beyond fat and sweeteners).

Future Research:

- **Correlation Analysis:** Investigate the statistical relationships between customer ratings, prices, ingredients, and nutritional content.
- **Market Basket Analysis:** Explore which products are frequently purchased together.
- **Customer Segmentation:** Analyze if different customer segments have varying preferences for ingredients, price points, or allergen considerations.
- **Trend Analysis Over Time:** If historical data is available, analyze how these metrics have changed over time.

References & Appendices

References: The primary reference is the **"Nutrition and Allergen Analysis Insights in Food Products" dataset from kaggle itself.**