FOOD_ADULERATION_DATA_REPORT

1. BACKGROUND
2. EXPLORATORY DATA ANALYSIS (EDA)
    1. DATA SOURCING
    2. DATA MIGRATION
    3. DATA TRANSFORMATION
    4. DEFINITION OF OBJECTIVES
    5. INSIGHTS GENERATION
    6. RECOMMENDATION

3.0 CONCLUSION

BACKGROUND

The food industry is a critical sector that affects public health, economic stability, and consumer trust. Ensuring the safety and quality of food products is a top priority for manufacturers, regulatory agencies, and consumers. However, food adulteration remains a significant challenge, with potentially severe consequences for human health and the economy. This dataset provides valuable insights into food adulteration cases, detection methods, and responses, offering a foundation for analysis and improvement.

The dataset contains information on food adulteration cases, including:

1. **Product categories**: Various food products, such as meat, dairy, beverages, and bakery items, are represented in the dataset.
2. **Adulterants**: A range of substances, including melamine, artificial sweeteners, coloring agents, and chalk, have been detected in food products.
3. **Detection methods**: Different methods, such as sensory evaluation, chemical analysis, and spectroscopy, have been employed to detect adulteration.
4. **Severity and health risks**: Adulteration cases have been classified according to their severity and potential health risks, ranging from minor to severe and low to high, respectively.
5. **Actions taken**: Various responses, including product recall, warning issued, investigation launched, and fine imposed, have been taken to address adulteration cases.

CONT'S OF BACKGROUND

Motivation for Analysis

The analysis of this dataset is motivated by the need to:

1. **Understand the scope and nature of food adulteration**: By examining the types and frequencies of adulterants, detection methods, and responses, we can gain a deeper understanding of the issue.
2. **Identify patterns and trends**: Analyzing the dataset may reveal patterns and trends in adulteration cases, detection methods, and responses, which can inform strategies for improvement.
3. **Inform policy and decision-making**: The findings from this analysis can provide valuable insights for policymakers, regulatory agencies, and manufacturers, enabling them to develop more effective strategies for ensuring food safety and quality.

Research Questions

The analysis of this dataset aims to answer the following research questions:

1. What are the most common adulterants detected in food products, and what are their associated health risks?
2. Which detection methods are most effective in identifying adulterated products, and what actions are typically taken in response to adulteration cases?
3. Are there patterns or trends in adulteration cases by product category, brand, or region?

Expected Outcomes

The analysis of this dataset is expected to provide insights into the nature and extent of food adulteration, the effectiveness of detection methods, and the responses to adulteration cases. The findings can inform strategies to:

1. **Improve food safety**: By understanding the types and frequencies of adulterants, detection methods, and responses, we can develop more effective strategies for ensuring food safety.
2. **Enhance compliance with regulations**: The analysis can identify areas for improvement in compliance with food safety regulations and standards.
3. **Promote best practices**: The findings can inform best practices for manufacturers, regulatory agencies, and consumers, enabling them to take proactive steps to prevent and respond to food adulteration cases.

DATA SOURCE

National Agency for Food and Drug Administration and Control (NAFDAC) in Nigeria

DATA MIGRATION

Steps: 1. Open MySQL Work Bench.

2. Click on ‘Local Host’.

3. (a) Click on ‘Create a new Schema’ icon.

(b) Rename the Schema in the dialogue box that pops up.

4. (a) Click on ‘Apply’. (b) Click ‘Apply’ again.

5. Click on Finish.

6. From the upper left-hand side of the Command line Interface, click on ‘Refresh’.

7. (a) Go to the new Schema created, and right click on the ‘Tables’.

(b) From the option box that pops up, click on ‘Table Data Import Wizard’.

(c) From the Table Data Import box, go to ‘File Path’ to select the <u>prepared</u> appropriate Microsoft Excel CSV File.

(d) Click on ‘Next’

(e) Check ‘Drop table, if exists field.

(f) Click on ‘Next’

(g) Click on ‘Next’ again and again

(h) Click on ‘Finish’

(i) Click on ‘Refresh’

Data migration to MYSQL workbench

```sql
select*
from `food_adulteration_data`;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

product_name	brand	category	adulterant	detection_date	detection_method	severity	health_risk	action_taken
Butter	BrandB	Meat	Artificial sweeteners	5/11/2024	Microbiological Analysis	Moderate	Low	Product Recall
Chicken	BrandC	Dairy	Coloring agents	5/23/2024	Sensory Evaluation	Severe	Medium	Warning Issued
Yogurt	BrandC	Meat	Artificial sweeteners	2/17/2024	Sensory Evaluation	Severe	High	Investigation Launched
Wine	BrandB	Beverages	Coloring agents	5/16/2024	Spectroscopy	Minor	Medium	Product Recall
Bread	BrandD	Dairy	Water	6/6/2024	Chemical Analysis	Severe	Medium	Warning Issued
Wine	BrandD	Dairy	Melamine	5/12/2024	Microbiological Analysis	Severe	Medium	Investigation Launched
Beef	BrandC	Meat	Melamine	4/1/2024	Chemical Analysis	Minor	Medium	Investigation Launched
Juice	BrandD	Beverages	Melamine	3/18/2024	Chemical Analysis	Moderate	High	Fine Imposed

food_adulteration_data 2

DATA TRANSFORMATION

This involves the removal of unwanted columns and their corresponding values. The data cleaning includes the various steps:

1. I clicked on the columns to be removed.
2. I then right-clicked.
3. From the appeared box, I selected the Delete option.
4. Changing the data type of date_detection.

Data set before transformation

adulteratio	product_n	brand	category	adulterant	detection_	detection_	severity	health_risk	action_taken
1	Butter	BrandB	Meat	Artificial sv	########	Microbiolo	Moderate	Low	Product Recall
2	Chicken	BrandC	Dairy	Coloring ag	5/23/2024	Sensory Ev	Severe	Medium	Warning Issued
3	Yogurt	BrandC	Meat	Artificial sv	2/17/2024	Sensory Ev	Severe	High	Investigation Launched
4	Wine	BrandB	Beverages	Coloring ag	5/16/2024	Spectrosco	Minor	Medium	Product Recall
5	Bread	BrandD	Dairy	Water	########	Chemical A	Severe	Medium	Warning Issued
6	Wine	BrandD	Dairy	Melamine	########	Microbiolo	Severe	Medium	Investigation Launched
7	Beef	BrandC	Meat	Melamine	########	Chemical A	Minor	Medium	Investigation Launched
8	Juice	BrandD	Beverages	Melamine	3/18/2024	Chemical A	Moderate	High	Fine Imposed
9	Chicken	BrandC	Meat	Artificial sv	6/14/2024	Microbiolo	Minor	Low	Investigation Launched
10	Milk	BrandA	Bakery	Melamine	4/23/2024	Spectrosco	Moderate	Medium	Fine Imposed
11	Cheese	BrandA	Beverages	Coloring ag	2/25/2024	Sensory Ev	Moderate	Low	Investigation Launched
12	Milk	BrandA	Condiment	Melamine	########	Sensory Ev	Moderate	Low	Warning Issued
13	Bread	BrandC	Dairy	Chalk	########	Chemical A	Minor	High	Product Recall

Data set after transformation

	product_name	brand	detection_date	category	adulterant	detection_method	severity	health_risk	action_taken
2	Butter	BrandB	05/11/2024	Meat	Artificial sweeteners	Microbiological Analysis	Moderate	Low	Product Recall
3	Chicken	BrandC	5/23/2024	Dairy	Coloring agents	Sensory Evaluation	Severe	Medium	Warning Issued
4	Yogurt	BrandC	2/17/2024	Meat	Artificial sweeteners	Sensory Evaluation	Severe	High	Investigation Launched
5	Wine	BrandB	5/16/2024	Beverages	Coloring agents	Spectroscopy	Minor	Medium	Product Recall
6	Bread	BrandD	06/06/2024	Dairy	Water	Chemical Analysis	Severe	Medium	Warning Issued
7	Wine	BrandD	05/12/2024	Dairy	Melamine	Microbiological Analysis	Severe	Medium	Investigation Launched
8	Beef	BrandC	04/01/2024	Meat	Melamine	Chemical Analysis	Minor	Medium	Investigation Launched
9	Juice	BrandD	3/18/2024	Beverages	Melamine	Chemical Analysis	Moderate	High	Fine Imposed
10	Chicken	BrandC	6/14/2024	Meat	Artificial sweeteners	Microbiological Analysis	Minor	Low	Investigation Launched
11	Milk	BrandA	4/23/2024	Bakery	Melamine	Spectroscopy	Moderate	Medium	Fine Imposed
12	Cheese	BrandA	2/25/2024	Beverages	Coloring agents	Sensory Evaluation	Moderate	Low	Investigation Launched
13	Milk	BrandA	03/07/2024	Condiments	Melamine	Sensory Evaluation	Moderate	Low	Warning Issued
14	Beef	BrandA	6/13/2024	Meat	Chalk	Chemical Analysis	Moderate	Medium	Warning Issued

OBJECTIVES

The potential objectives for analysis based on the provided dataset:

- **Identify the most common adulterants and their associated health risks**: Analyze the dataset to determine which adulterants (e.g., melamine, artificial sweeteners, coloring agents) are most frequently detected across different product categories and brands, and assess the corresponding health risks.
- **Evaluate the effectiveness of detection methods and actions taken**: Investigate the relationship between detection methods (e.g., microbiological analysis, sensory evaluation, spectroscopy) and the severity of adulteration, as well as the actions taken (e.g., product recall, warning issued, investigation launched) in response to adulteration cases.
- **Determine brand-specific and product category-specific trends in adulteration**: Examine the dataset to identify patterns and trends in adulteration cases by brand and product category, including the types of adulterants detected, severity, and health risks, to inform targeted quality control and regulatory measures.

INSIGHT GENERATION

1. Identify the most common adulterants and their associated health risks

```
-- To identify the most common adulterants and their associated health ris
SELECT
 adulterant,
 `health_risk`,
 COUNT(*) as frequency
ROM
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

adulterant	health_risk	frequency
Coloring agents	Low	86
Water	Low	74
Coloring agents	High	72
Chalk	High	71
Artificial sweeteners	Low	69
Chalk	Low	68
Melamine	High	67

Insights from the Analysis

Based on the query outcome and the given dataset, here are two insights that can be generated as a data analyst:

Insight 1: Melamine and Coloring Agents are the Most Prevalent Adulterants with Varying Health Risks

The query outcome shows that Melamine and Coloring Agents are among the top adulterants found in the dataset, with varying health risks associated with them. For instance, Melamine is associated with High, Medium, and Low health risks, with frequencies of 67, 66, and 61, respectively. Similarly, Coloring Agents are associated with Low, High, and medium health risks, with frequencies of 86, 72, and 57, respectively. This suggests that these adulterants are widespread and can have significant health implications for consumers. As a data analyst, it would be essential to investigate the sources of these adulterants and the reasons behind their prevalence in the food products.

Insight 2: Water and Chalk are Also Significant Adulterants with Notable Health Risks

The query outcome also reveals that Water and Chalk are notable adulterants in the dataset, with significant health risks associated with them. For example, Water is associated with Low, Medium, and High health risks, with frequencies of 74, 64, and 55, respectively. Chalk is associated with High, Low, and medium health risks, with frequencies of 71, 68, and 63, respectively. This suggests that these adulterants are not only prevalent but also pose significant health risks to consumers. As a data analyst, it would be crucial to investigate the reasons behind the presence of these adulterants in food products and recommend measures to mitigate their impact on public health.

These insights highlight the importance of monitoring and regulating food adulteration to ensure public health safety. By analyzing the data and identifying patterns and trends, data analysts can provide valuable insights that can inform policy decisions and interventions aimed at reducing food adulteration and promoting public health.

2. Evaluate the effectiveness of detection methods and actions taken

The query groups the data by adulterant, detection method, action taken, and health risk, and calculates the frequency of each combination. By analyzing the results of this query, we can identify patterns and trends in the effectiveness of detection methods and actions taken for specific adulterants and health risks.

```sql
-- To Evaluate the effectiveness of detection methods and actions taken
SELECT
adulterant,
`detection_method`,
`action_taken`,
`health_risk`,
COUNT(*) as frequency
from
`food_adulteration_data`
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

adulterant	detection_method	action_taken	health_risk	frequency
Artificial sweeteners	Microbiological Analysis	Investigation Launched	Medium	10
Coloring agents	Microbiological Analysis	Fine Imposed	Low	10
Chalk	Sensory Evaluation	Investigation Launched	Medium	9
Chalk	Chemical Analysis	Product Recall	Low	9
Water	Microbiological Analysis	Product Recall	High	9
Melamine	Chemical Analysis	Warning Issued	High	8
Chalk	Spectroscopy	Product Recall	Medium	8

Insights from the Analysis

Based on the query outcome, here are two detailed insights that can be generated as a data analyst**:**

Insight 1: Microbiological Analysis is Effective for Detecting Artificial Sweeteners and Coloring Agents, but Action Taken Varies

The query outcome shows that Microbiological Analysis is a commonly used detection method for Artificial Sweeteners and Coloring Agents. For Artificial Sweeteners, Investigation Launched is the most frequent action taken, with a frequency of 10 for medium health risk and 7 for Low health risk. In contrast, Fine Imposed is also a common action taken for Artificial Sweeteners, with a frequency of 8 for medium health risk. For Coloring Agents, Fine Imposed is the most frequent action taken, with a frequency of 10 for Low health risk and 8 for Low health risk when combined with Sensory Evaluation. This suggests that while Microbiological Analysis is effective in detecting these adulterants, the action taken may depend on the specific circumstances of the case, such as the level of health risk and the severity of the adulteration.

Insight 2: Different Detection Methods and Actions Taken are Used for Different Adulterants and Health Risks

The query outcome also shows that different detection methods and actions taken are used for different adulterants and health risks. For example, Chemical Analysis is commonly used for detecting Melamine and Chalk, while Spectroscopy is used for detecting Chalk and Melamine. The action taken also varies depending on the adulterant and health risk. For instance, Product Recall is a common action taken for Chalk and Water, while Warning Issued is used for Melamine. Investigation Launched is also a common action taken for several adulterants, including Artificial Sweeteners, Chalk, and Melamine. This suggests that the choice of detection method and action taken depends on the specific characteristics of the adulterant and the level of health risk associated with it. As a data analyst, it would be essential to further investigate the reasons behind these differences and identify best practices for detecting and mitigating food adulteration.

3. Determine brand-specific and product category-specific trends in adulteration:
The query can help us identify specific brands or product categories that are more prone to adulteration and take targeted measures to mitigate the issue.

```sql
-- Brand-specific trends
SELECT
  brand,
  adulterant,
  COUNT(*) as frequency,
  AVG(CASE WHEN `health_risk` = 'Low' THEN 1 WHEN `health_risk` = 'Medium' THEN 2 ELSE 3 END) as `avg_health_risk
FROM
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

brand	adulterant	frequency	avg_health_risk
BrandC	Coloring agents	54	2.0370
BrandD	Water	46	1.8696
BrandC	Artificial sweeteners	45	2.0000
BrandE	Artificial sweeteners	45	1.8667
BrandA	Chalk	44	1.9545
BrandC	Water	44	1.9545
BrandD	Melamine	43	2.0930

```sql
-- Product category-specific trends
SELECT
  category,
  adulterant,
  COUNT(*) as frequency,
  AVG(CASE WHEN `health_risk` = 'Low' THEN 1 WHEN `health_risk` = 'Medium' THEN 2 ELSE 3 END) as `avg_health_risk`
FROM
  `food_adulteration_data`
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

category	adulterant	frequency	avg_health_risk
Dairy	Water	54	1.9074
Meat	Melamine	48	2.0833
Meat	Chalk	48	2.2083
Meat	Coloring agents	46	1.8478
Dairy	Coloring agents	45	2.0889
Bakery	Chalk	45	1.8889
Beverages	Coloring agents	44	1.9318

Insights from the Analysis

Based on the query results, here are some insights that can be generated as a data analyst:

Insight 1: **BrandC and BrandD have High Frequencies of Adulteration**

The results show that BrandC has high frequencies of adulteration with Coloring agents (54), Artificial sweeteners (45), and Water (44). BrandD also has a high frequency of adulteration with Water (46). This suggests that these brands may have issues with their quality control processes or supply chains, leading to a higher incidence of adulteration.

Insight 2: **Meat and Dairy Products are More Prone to Adulteration**

The results also show that Meat and Dairy products are more prone to adulteration, with high frequencies of Melamine (48) and Chalk (48) in Meat products, and Coloring agents (45) and Melamine (43) in Dairy products. This suggests that these product categories may require additional scrutiny and monitoring to prevent adulteration.

Insight 3: **Coloring Agents and Melamine are Common Adulterants Across Multiple Brands and Product Categories**

The results show that Coloring agents and Melamine are common adulterants across multiple brands and product categories, with high frequencies and average health risks. This suggests that these adulterants may be more difficult to detect or may be more widely used in the food industry, highlighting the need for increased vigilance and monitoring. By analyzing these results, data analysts can identify trends and patterns in adulteration and provide insights that can inform quality control processes, supply chain management, and regulatory policies.

RECOMMENDATIONS:

Based on the insights generated from the data analysis, the following recommendations are proposed for a detailed analysis of the data:

1. **Investigate Brand-Specific Issues:** Conduct a detailed analysis of BrandC and BrandD to identify the root causes of their high frequencies of adulteration. This may involve reviewing their quality control processes, supply chain management, and production procedures.
2. **Focus on High-Risk Product Categories:** Prioritize Meat and Dairy products for additional scrutiny and monitoring due to their high propensity for adulteration. This may involve increasing the frequency of inspections, testing, and audits for these product categories.
3. **Develop Targeted Detection Methods:** Develop and implement targeted detection methods for common adulterants like Coloring Agents and Melamine, which are prevalent across multiple brands and product categories.
4. **Standardize Action Protocols**: Establish standardized action protocols for different adulterants and health risks to ensure consistency in response and minimize the impact on public health.
5. **Monitor and Evaluate Detection Methods:** Continuously monitor and evaluate the effectiveness of different detection methods, such as Microbiological Analysis, Chemical Analysis, and Spectroscopy, to identify best practices and areas for improvement.
6. **Analyze Supply Chain and Production Data:** Analyze supply chain and production data to identify potential vulnerabilities and risks that may contribute to adulteration.
7. **Develop Predictive Models:** Develop predictive models to forecast potential adulteration risks and identify high-risk products, brands, and product categories.

By implementing these recommendations, data analysts can provide valuable insights that can inform quality control processes, supply chain management, and regulatory policies, ultimately reducing the incidence of food adulteration and promoting public health

CONCLUSION

Based on the analysis of the food adulteration data, it is clear that certain brands and product categories are more prone to adulteration than others. BrandC and BrandD have high frequencies of adulteration, and Meat and Dairy products are more susceptible to adulteration. Common adulterants like Coloring Agents and Melamine are prevalent across multiple brands and product categories.

To mitigate the impact of food adulteration on public health, it is essential to implement targeted detection methods, standardize action protocols, and continuously monitor and evaluate detection methods. Analyzing supply chain and production data can also help identify potential vulnerabilities and risks.

By developing predictive models and prioritizing high-risk product categories, data analysts can provide valuable insights that can inform quality control processes, supply chain management, and regulatory policies. Ultimately, the goal is to reduce the incidence of food adulteration and promote public health.

Key Takeaways

Certain brands (BrandC and BrandD) and product categories (Meat and Dairy) are more prone to adulteration.

Common adulterants like Coloring Agents and Melamine are prevalent across multiple brands and product categories.

Targeted detection methods, standardized action protocols, and continuous monitoring and evaluation are essential to mitigate the impact of food adulteration.

Analyzing supply chain and production data can help identify potential vulnerabilities and risks.

Predictive models can forecast potential adulteration risks and identify high-risk products, brands, and product categories.