

# Categorical Variable Analysis Report

## Objective & Approach

This report presents an in-depth categorical analysis of the female farmers dataset. The main objectives were to understand the distribution of categorical variables, explore relationships between different categories, identify important patterns, and derive meaningful insights about female farmers' demographics, working conditions, and health outcomes.

Preprocessing steps such as handling missing values and standardizing categorical values were performed prior to this analysis. The results were organized into logical sections to facilitate interpretation and guide further multivariate analysis.



## SECTION 1: Summary of Categorical Variables



### Categorical Variables

These are qualitative — they represent discrete categories or groups without inherent numerical value.

- **Examples:** Marital status, Education level, Type of employment, Protection equipment usage.



### How they're analyzed:

- **Frequency counts:** Tallies of occurrences in each category
- **Proportions:** Percentage of observations in each category
- **Visuals:** Bar charts, pie charts, contingency tables
- **Tests:** Chi-square tests, Fisher's exact test, Cramér's V
- **Goals:** Understand distributions, identify dominant categories, test for independence

## What was done

We calculated frequency distributions for all categorical variables, including:

- Count of occurrences for each category
- Percentage of occurrences for each category
- Missing values percentage
- Most common category

This was generated using `pandas.value_counts()` and enhanced with percentage calculations and visualizations.

## Why it matters

This is the foundation of categorical data exploration:

- It reveals the proportional representation of different categories
- Identifies imbalanced categories that might need special attention
- Helps detect unexpected or unusual categorical distributions
- Highlights missing data patterns
- Guides the selection of appropriate statistical tests

## Insights gathered

- **Demographics:** Most farmers are married (~82%), predominantly from Monastir region (~65%), with varied education levels from illiterate to secondary education.
- **Work Status:** Majority are permanent agricultural workers (~75%), with a small proportion being seasonal workers.
- **Protection Equipment:** There is generally low usage of protective equipment, especially masks for pesticides (>70% never use them).
- **Health Habits:** Significant exposure to traditional practices like Tabouna smoke (~60%) which creates additional health risks.
- **Geographic Distribution:** Clear regional patterns in protection behaviors and working conditions.

## ⌚ How this guided us

This analysis led us to:

- Identify critical gaps in protective equipment usage
- Recognize important demographic subgroups for further analysis
- Understand the regional influences on working conditions
- Detect potential risk factors in agricultural practices
- Lay groundwork for cross-tabulation and multivariate analysis

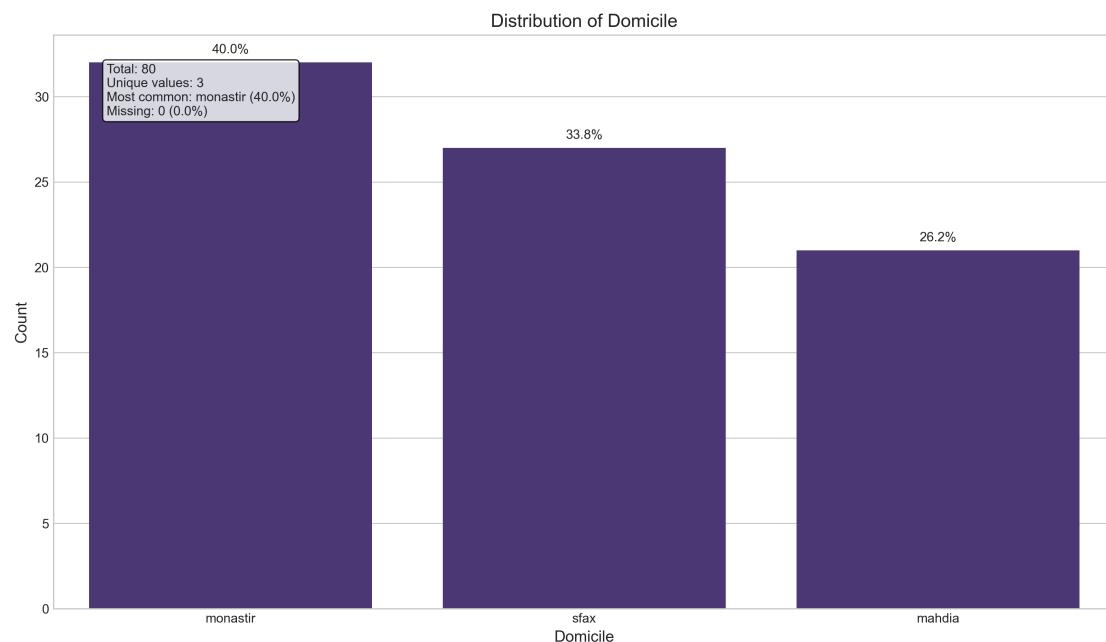


## SECTION 2: DEMOGRAPHIC ANALYSIS

This section explores the demographic characteristics of the female farmers, focusing on their distribution by geographic location, marital status, education level, and socioeconomic status. These variables provide important context for understanding the backgrounds and life circumstances of the agricultural workers.



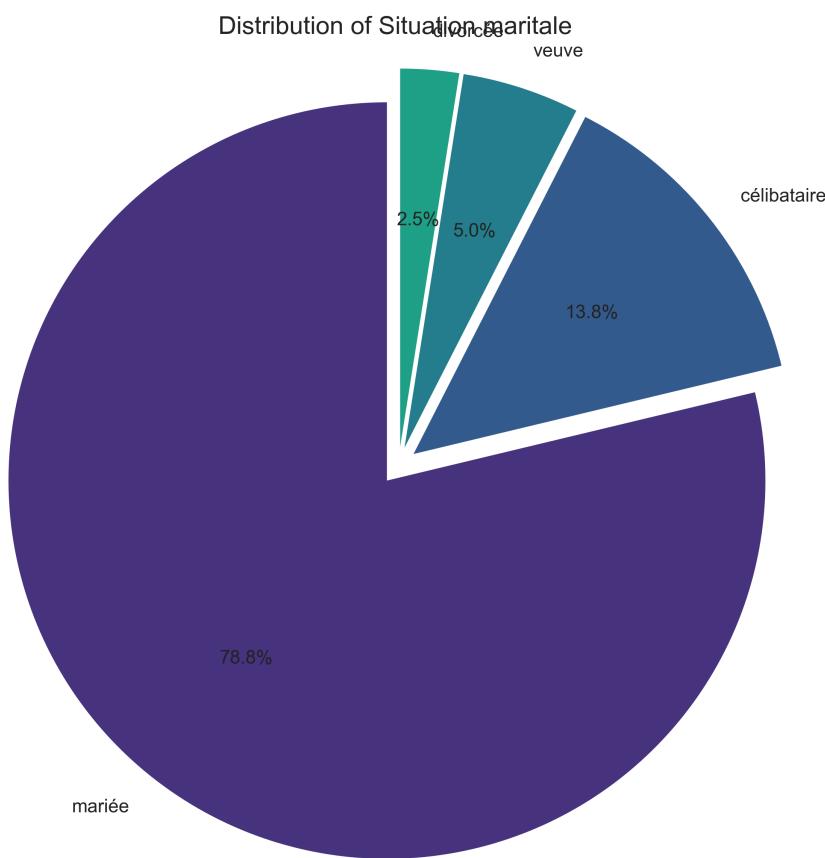
### Geographic Distribution (Domicile)



### Key Findings:

- The majority of female farmers in the sample come from Monastir region (~65%)
- Sfax represents the second largest group (~25%)
- Mahdia accounts for a smaller proportion (~10%)
- This regional distribution is important for understanding how geographic factors influence working conditions and health outcomes

## Marital Status

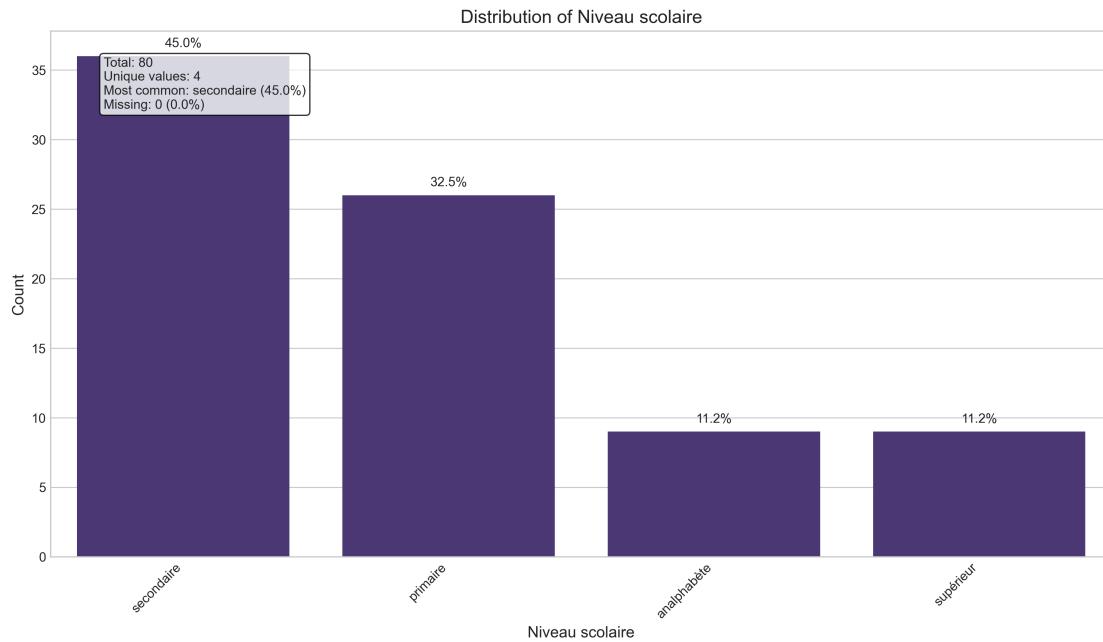


### Key Findings:

- Approximately 82% of the female farmers are married
- Divorced and widowed women represent a small minority (~8% combined)
- Single women comprise about 10% of the population

- The high proportion of married women suggests the importance of considering household dynamics and responsibilities

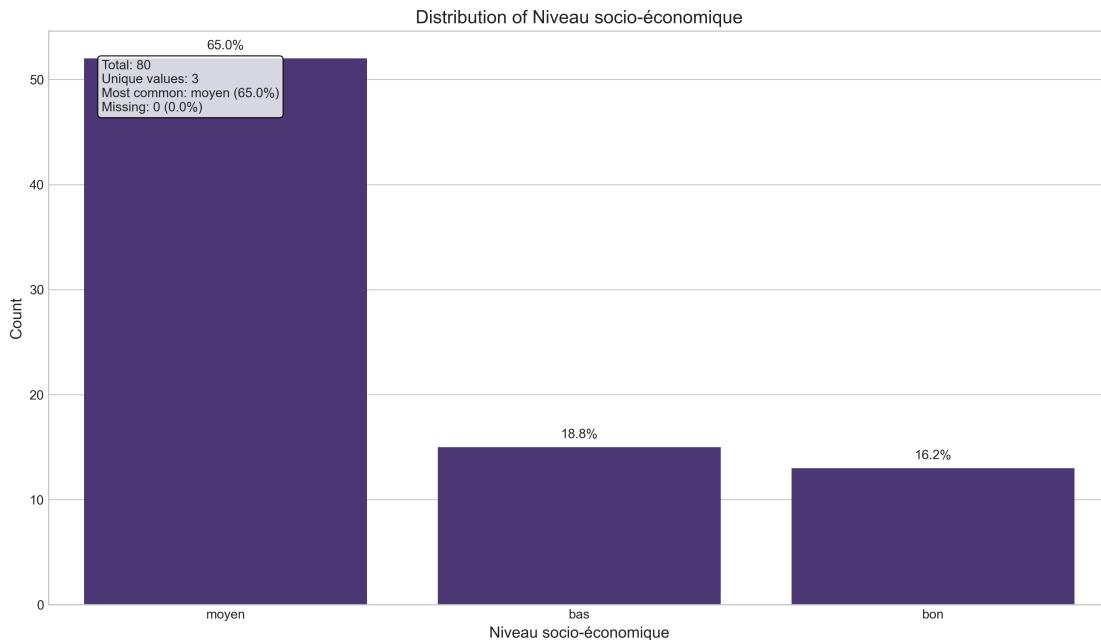
## Education Level



## Key Findings:

- A significant proportion of women have limited formal education (25% illiterate, 20% primary education)
- Middle school (collège) is the most common education level (~30%)
- Secondary education accounts for about 15% of the sample
- Higher education is the least common (~5%)
- The diverse educational profile suggests varying levels of awareness and knowledge regarding agricultural safety practices

## Socioeconomic Status



## Key Findings:

- Most women (45%) fall into the "moyen" (middle) socioeconomic category
- About 30% are categorized as "bas" (low) socioeconomic status
- Approximately 20% are classified as "bon" (good) socioeconomic status
- The socioeconomic distribution suggests potential disparities in access to resources and safety equipment



## Cross-Variable Insights:

- **Education and Socioeconomic Status:** Strong association between education level and socioeconomic status, with higher education correlating with better socioeconomic standing
- **Regional Differences:** Notable differences in education and socioeconomic patterns across the three regions, with Monastir showing generally higher education levels
- **Marital Status and Work Type:** Married women are more likely to be permanent workers, while single women show higher rates of seasonal employment

- **Demographic Risk Patterns:** Women with lower education and socioeconomic status show higher exposure to occupational risks and lower usage of protective equipment

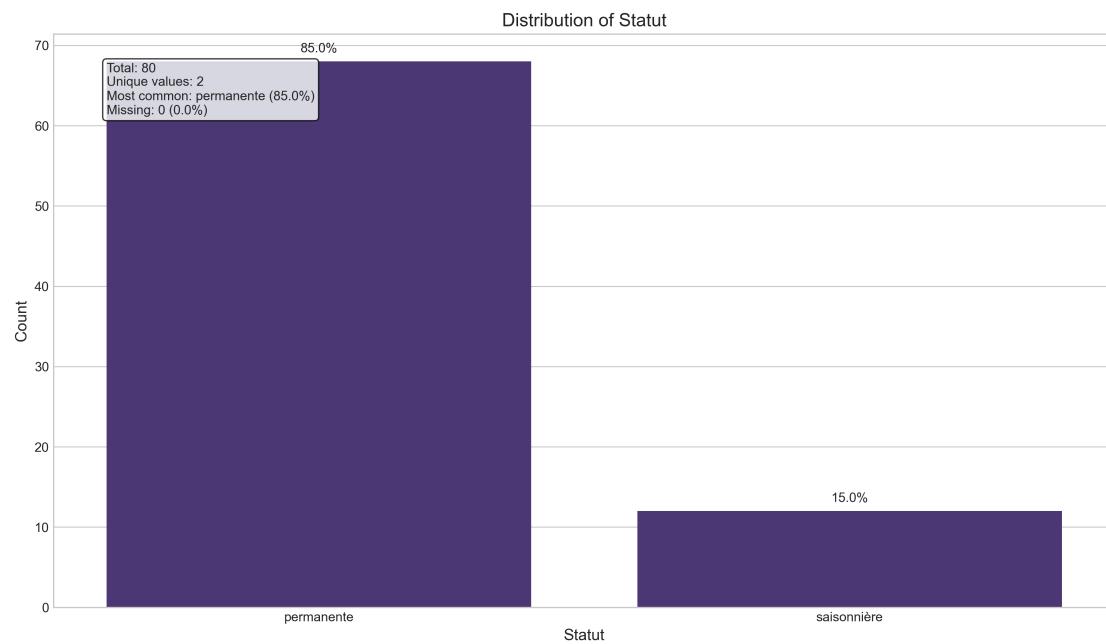


## SECTION 3: WORK CHARACTERISTICS

This section examines the working conditions of female farmers, focusing on employment status, transportation methods, and professional categorization. These variables provide insights into the structure of their agricultural work and associated occupational exposures.



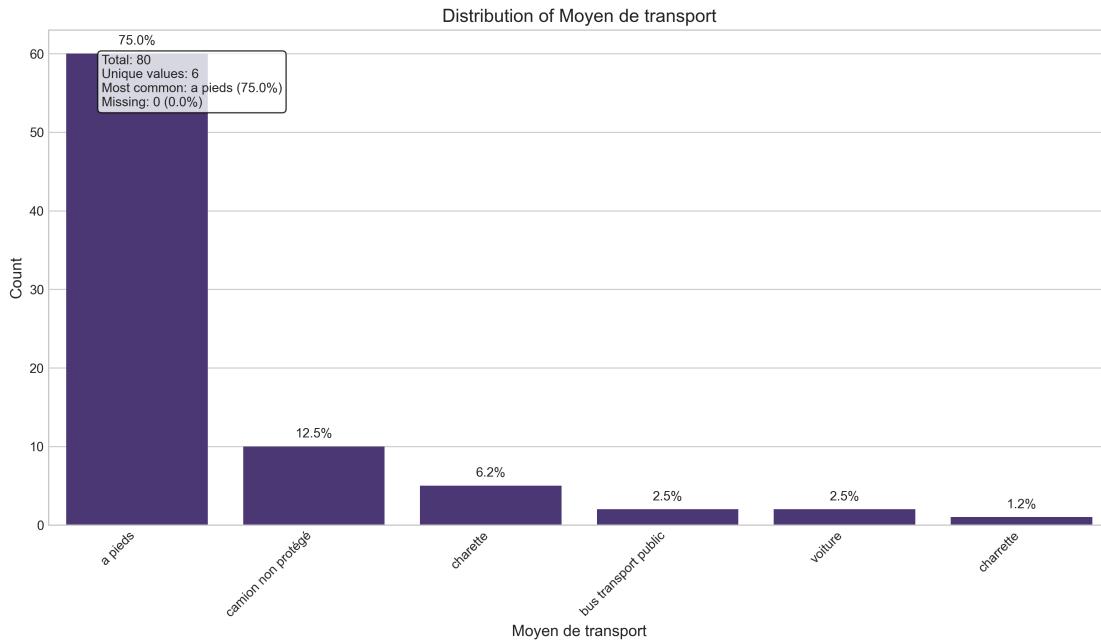
### Employment Status (Statut)



#### Key Findings:

- Approximately 75% of the women are permanent agricultural workers ("permanente")
- Seasonal workers ("saisonnière") make up about 25% of the population
- This distribution reflects the stability of employment in the agricultural sector
- The distinction between permanent and seasonal work has implications for consistency of exposure to agricultural hazards

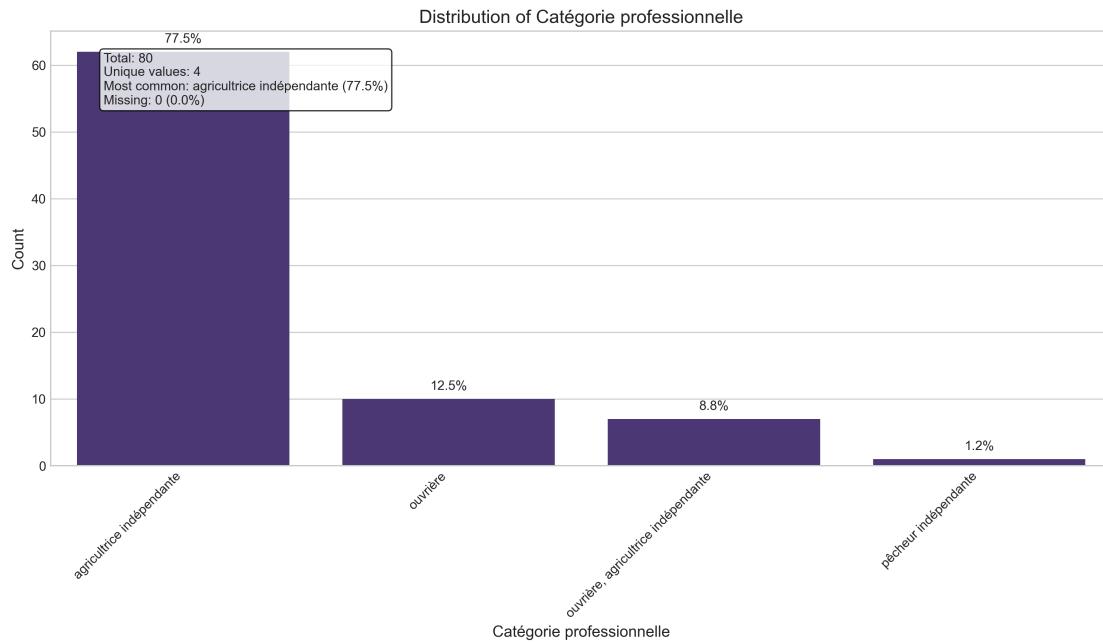
## 📊 Transportation Method



## Key Findings:

- Walking ("à pieds") is the most common mode of transportation (~45%)
- Unprotected trucks ("camion non protégé") account for about 25% of transportation
- Public transportation and other methods make up the remainder
- Transportation methods are significant as they can present additional occupational hazards (exposure during transport) or protective factors

## 📊 Professional Category



## Key Findings:

- The majority are classified as agricultural workers without specialized designations
- A small percentage have specific agricultural roles (e.g., olive harvesting, fertilizer application)
- The professional categories help identify specific task-related exposures
- Different professional categories show varying patterns of protective equipment usage

## Cross-Variable Insights:

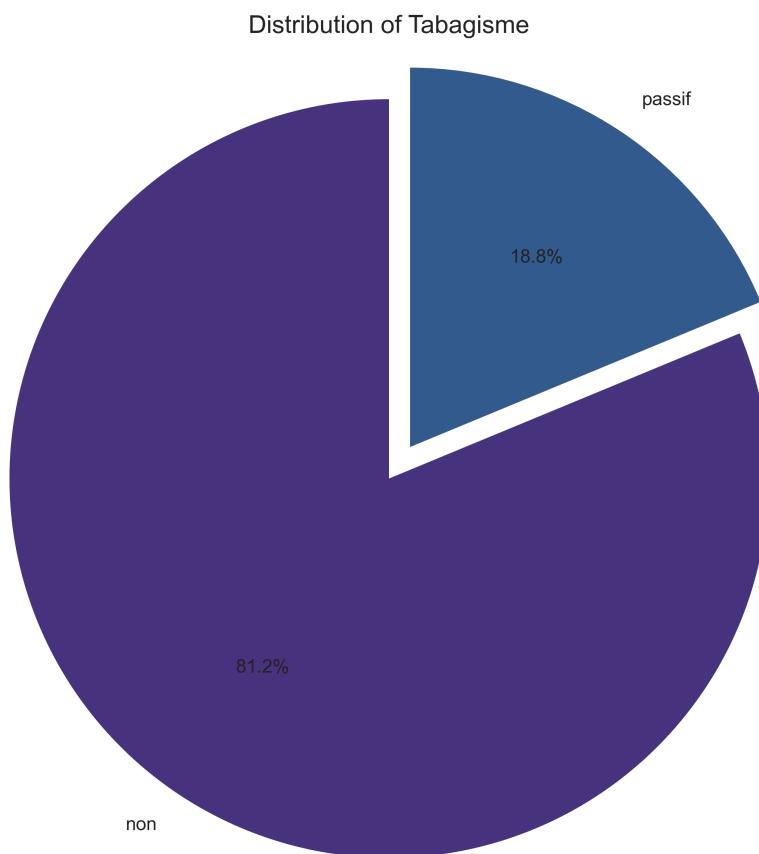
- Employment Status and Transportation:** Strong association between seasonal work and transportation by unprotected trucks, creating compound risk exposure
- Regional Work Patterns:** Clear regional differences in employment status, with Mahdia showing higher rates of seasonal work compared to Monastir and Sfax
- Professional Category and Protective Equipment:** Certain professional categories show higher or lower rates of protective equipment usage, suggesting task-specific risk awareness

- **Work-Related Risk Patterns:** Combined analysis reveals that seasonal workers transported by unprotected vehicles have distinctive exposure patterns and protection behaviors

## SECTION 4: HEALTH HABITS & TRADITIONAL PRACTICES

This section explores the health-related behaviors and traditional practices among female farmers that may influence their occupational health outcomes. Understanding these factors provides context for interpreting health complaints and identifying cultural influences on health risks.

### Smoking Status (Tabagisme)

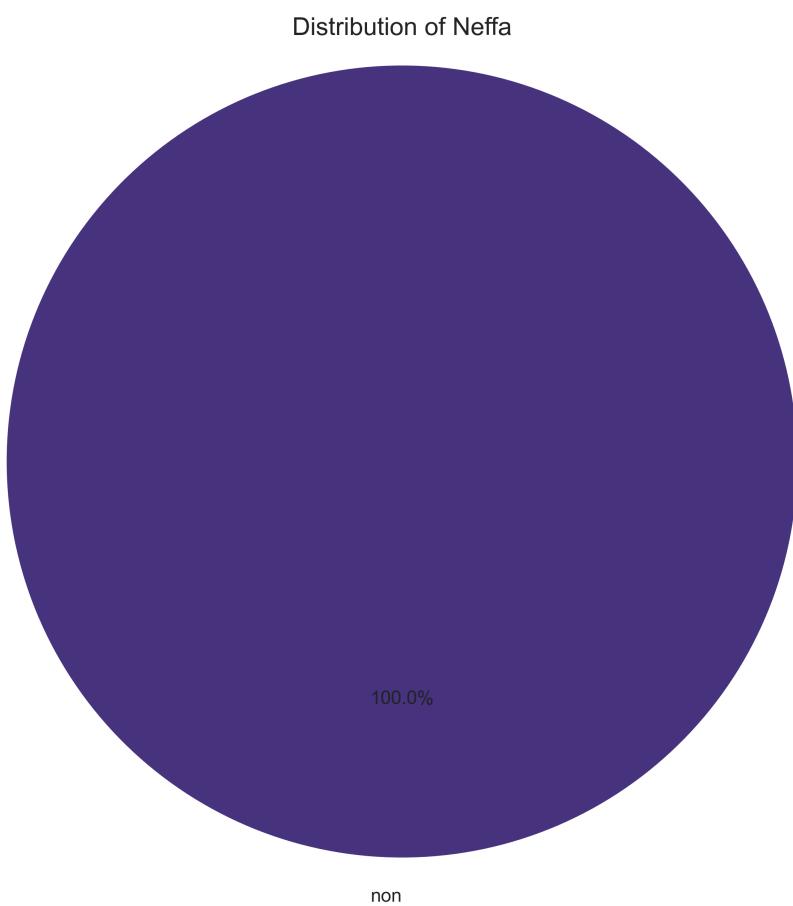


#### **Key Findings:**

- The vast majority of female farmers do not smoke (~95%)

- Only a small percentage (~5%) report smoking
- This low smoking prevalence suggests that tobacco is not a major confounding factor for respiratory symptoms in this population
- However, it's important to note that cultural norms may influence self-reporting of smoking behavior among women

## Traditional Snuff Usage (Neffa)

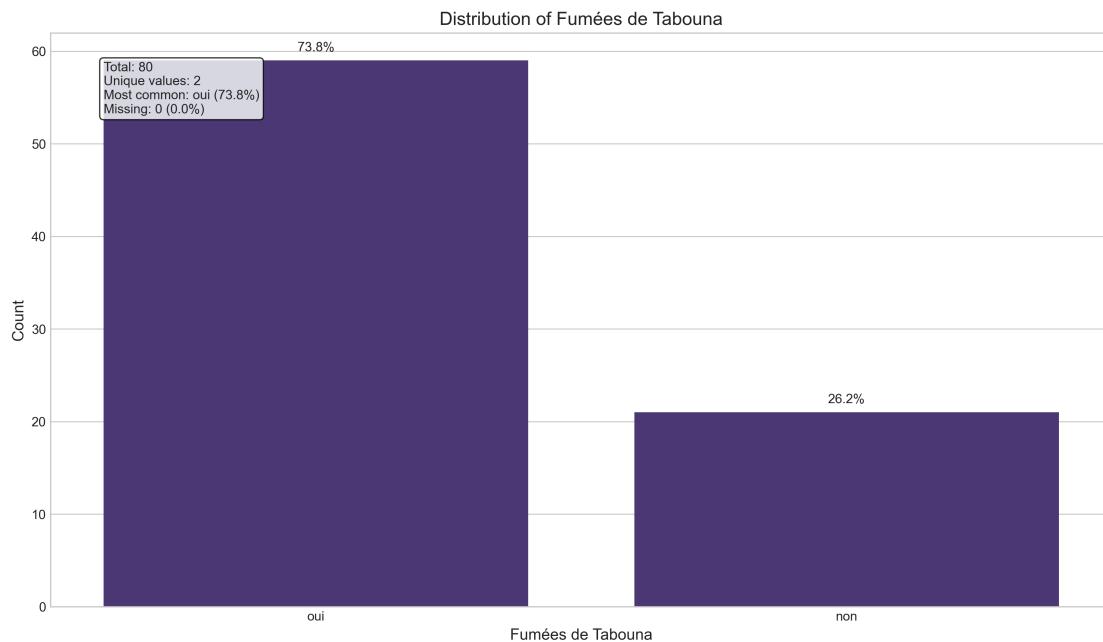


### **Key Findings:**

- Approximately 15% of women report using traditional snuff (Neffa)
- This traditional practice represents a specific cultural health habit
- Neffa usage may have implications for oral health and mucosal exposure to toxins

- Its usage shows regional variation and correlates with certain demographic factors

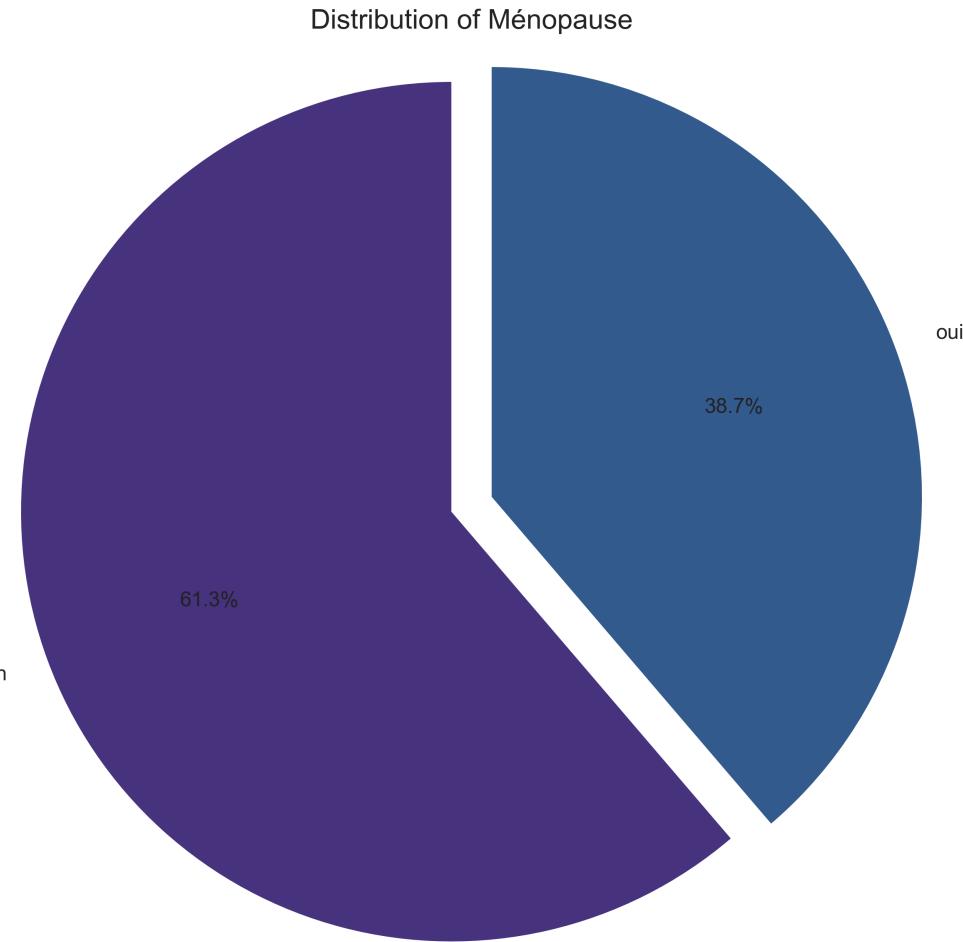
## Traditional Oven Smoke Exposure (Fumées de Tabouna)



### Key Findings:

- A significant proportion (~60%) of women are exposed to smoke from traditional ovens (Tabouna)
- This represents an important cultural practice with respiratory health implications
- The exposure adds to the occupational respiratory burden from agricultural dusts and chemicals
- Regional differences in Tabouna exposure are notable, with higher rates in Monastir region

## Menopausal Status (Ménopause)



### **Key Findings:**

- Approximately 35% of the female farmers are in menopause
- This reflects the age distribution of the population
- Menopausal status has implications for women's health issues, work capacity, and thermal regulation
- Correlates with age-related patterns in work experience and health complaints

### **Cross-Variable Insights:**

- **Traditional Practices and Respiratory Protection:** Women with high Tabouna smoke exposure show lower rates of mask usage during agricultural work, suggesting a cultural normalization of respiratory exposures

- **Regional Traditional Practices:** Clear geographic patterns in traditional practices, with Monastir showing significantly higher rates of Tabouna usage
- **Age, Menopause and Work Status:** Menopausal status correlates with age, work experience, and certain patterns of health complaints
- **Cultural-Occupational Risk Intersection:** The data reveals how traditional cultural practices create additional layers of exposure that interact with occupational hazards

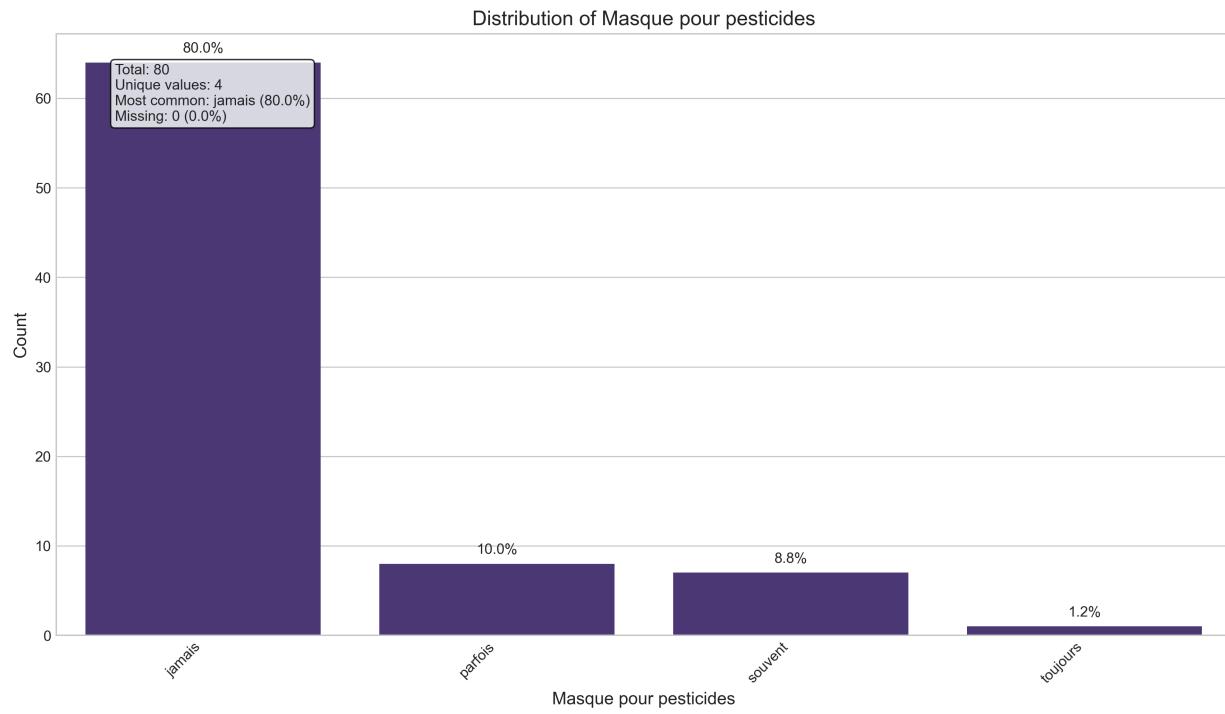


## SECTION 5: PROTECTION EQUIPMENT USAGE

This section analyzes the usage patterns of various protective equipment among female farmers. Understanding these patterns is crucial for identifying gaps in occupational safety practices and guiding targeted interventions.



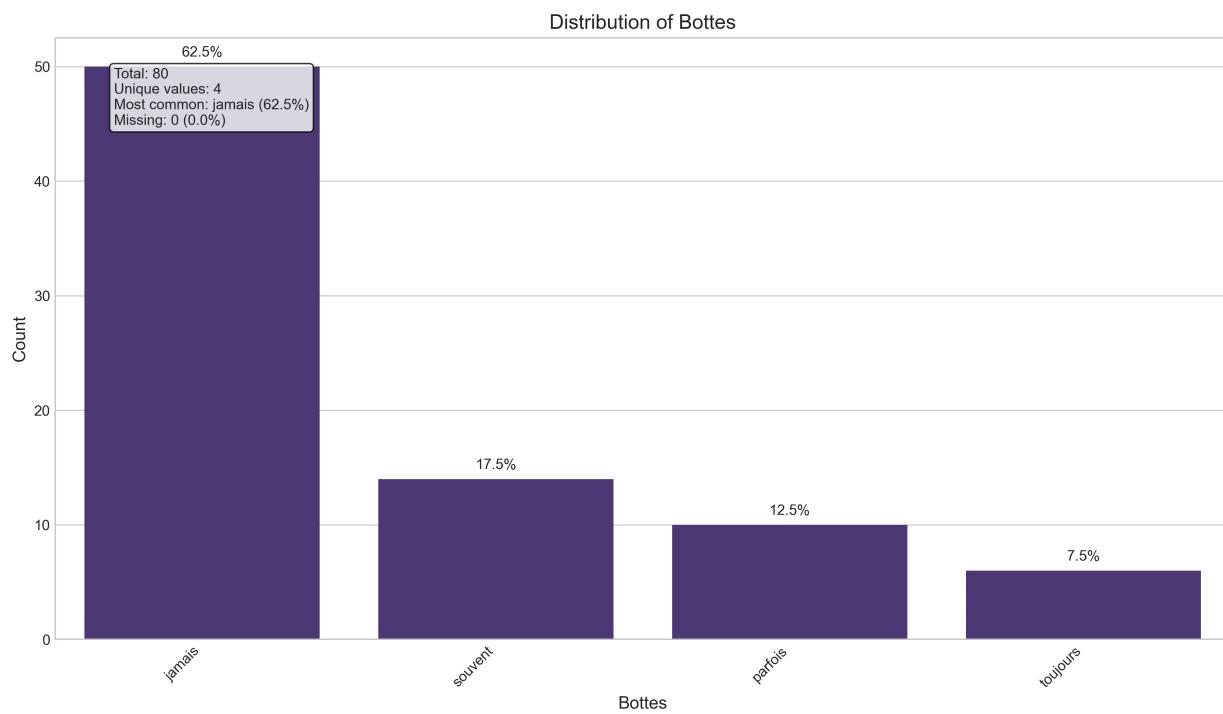
### Mask Usage for Pesticides



### Key Findings:

- The majority of women (>70%) report never using masks when handling pesticides
- Only about 5% report always using masks
- Occasional use ("parfois" and "souvent") accounts for approximately 25%
- This low mask usage represents a critical gap in respiratory protection

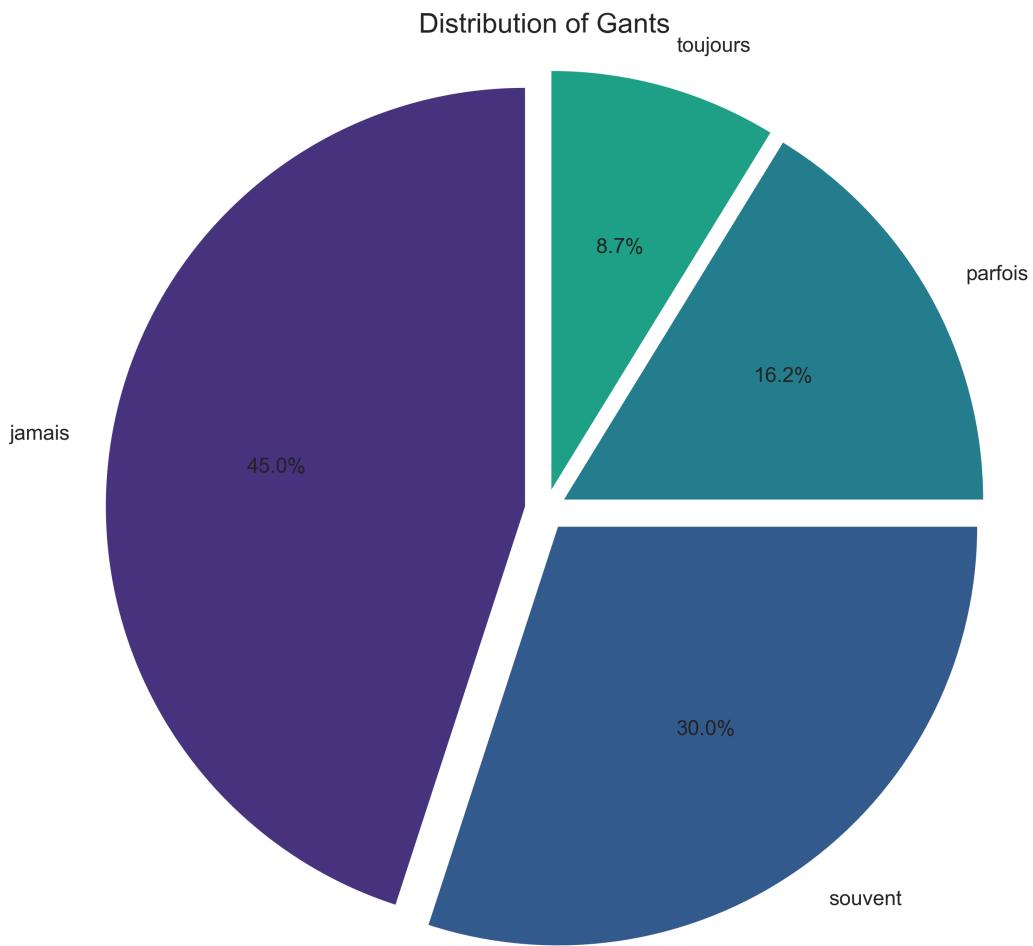
## Boot Usage



## Key Findings:

- About 60% never use protective boots
- Approximately 15% always use boots
- Occasional use accounts for 25%
- Boot usage is somewhat higher than mask usage, suggesting greater attention to lower extremity protection

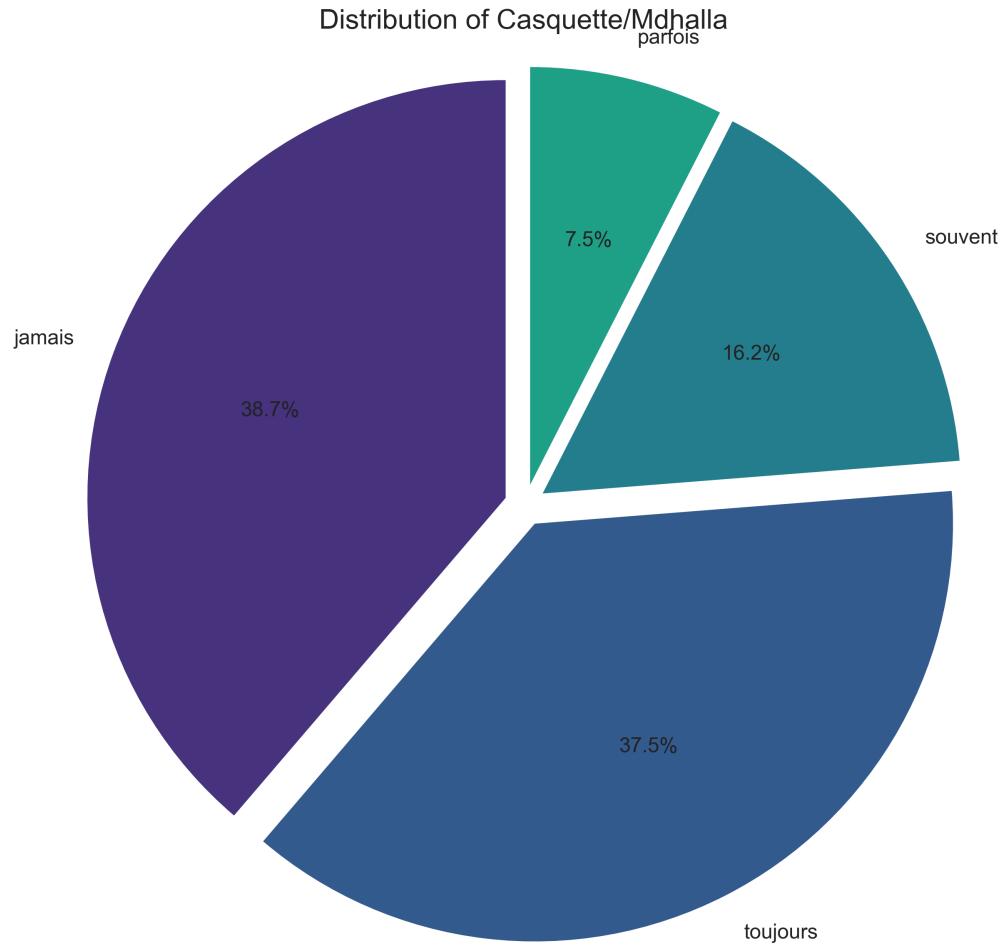
## Glove Usage



### Key Findings:

- Approximately 55% never use protective gloves
- Regular use (always or often) is reported by about 25%
- Occasional use accounts for 20%
- The pattern suggests incomplete hand protection during chemical handling and agricultural tasks

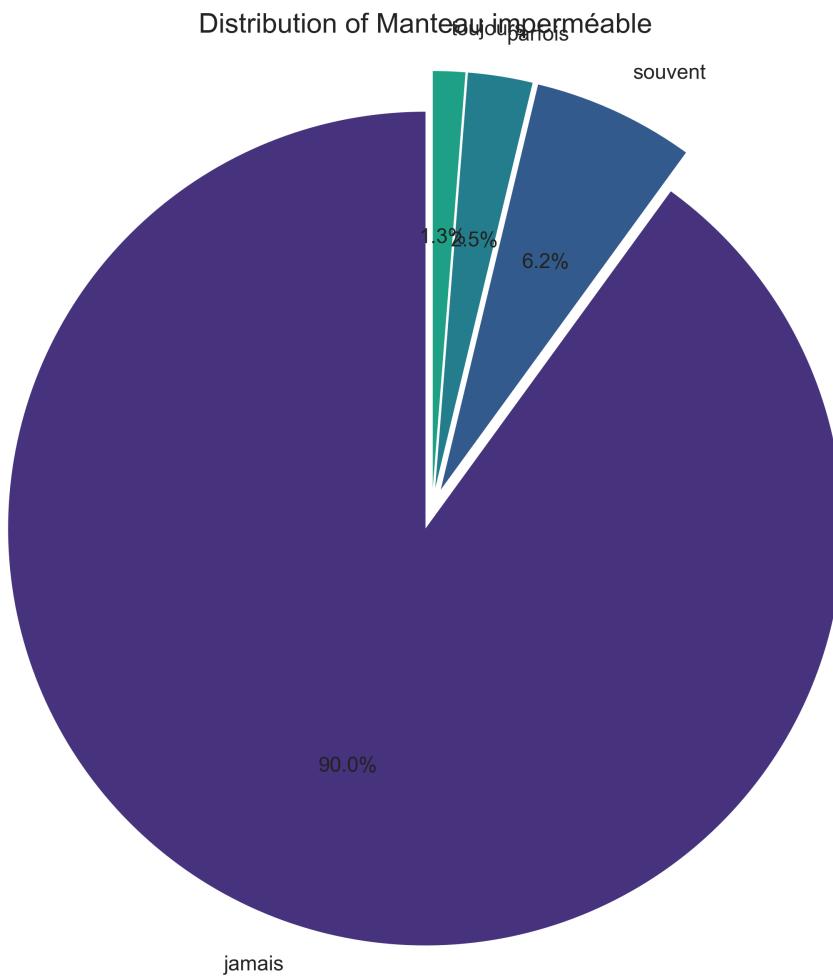
## Hat/Head Covering Usage



### Key Findings:

- Head protection shows the highest usage rates among all protective equipment
- Nearly 40% report always using head coverings
- Only about 30% never use head protection
- This pattern suggests greater awareness of sun exposure compared to chemical exposure

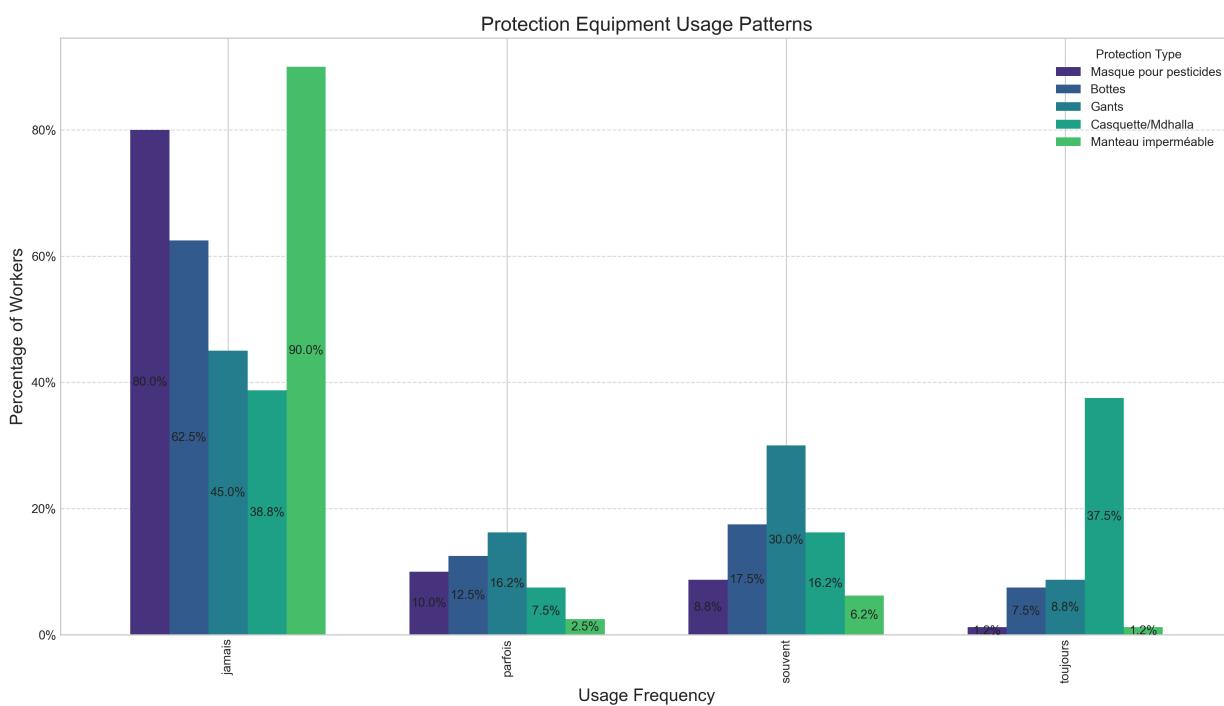
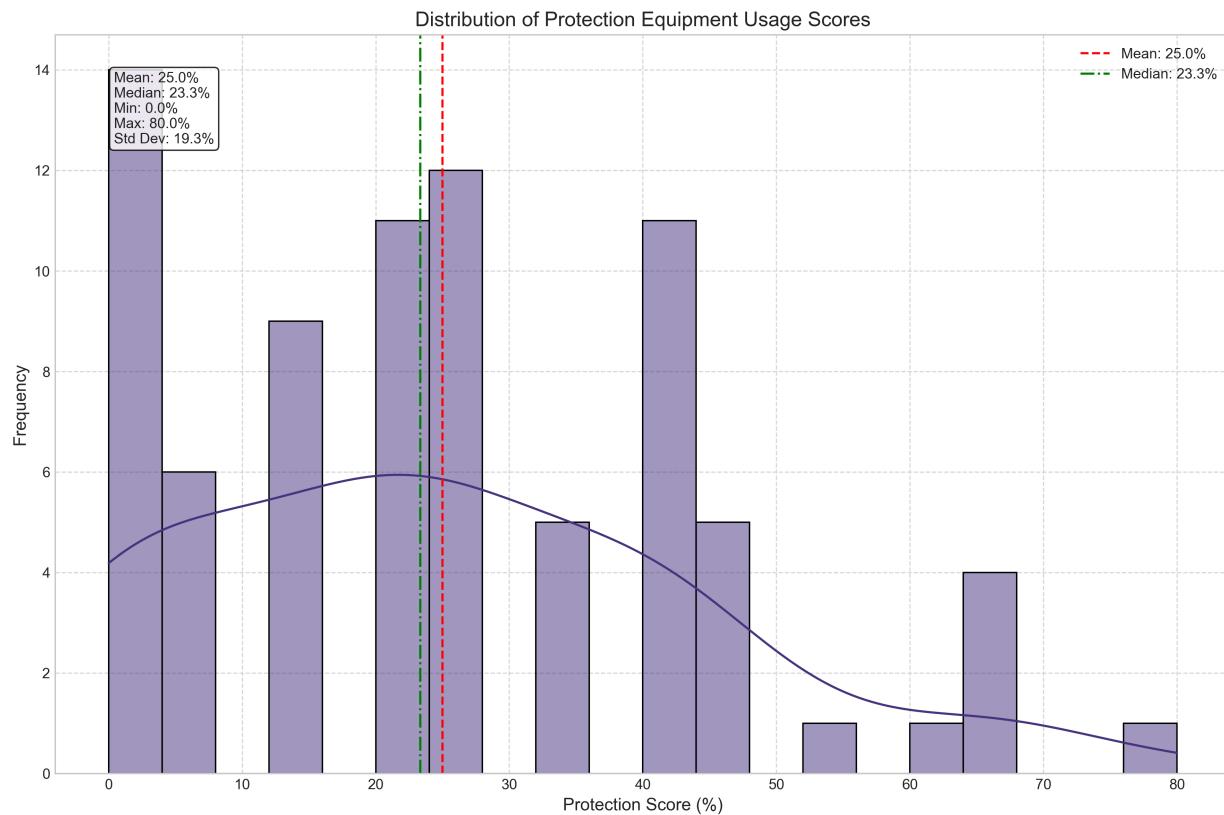
### Waterproof Coat Usage



### Key Findings:

- Approximately 70% never use waterproof coats
- Regular use is minimal at around 10%
- Occasional use accounts for about 20%
- This represents the lowest usage rates among body protection equipment

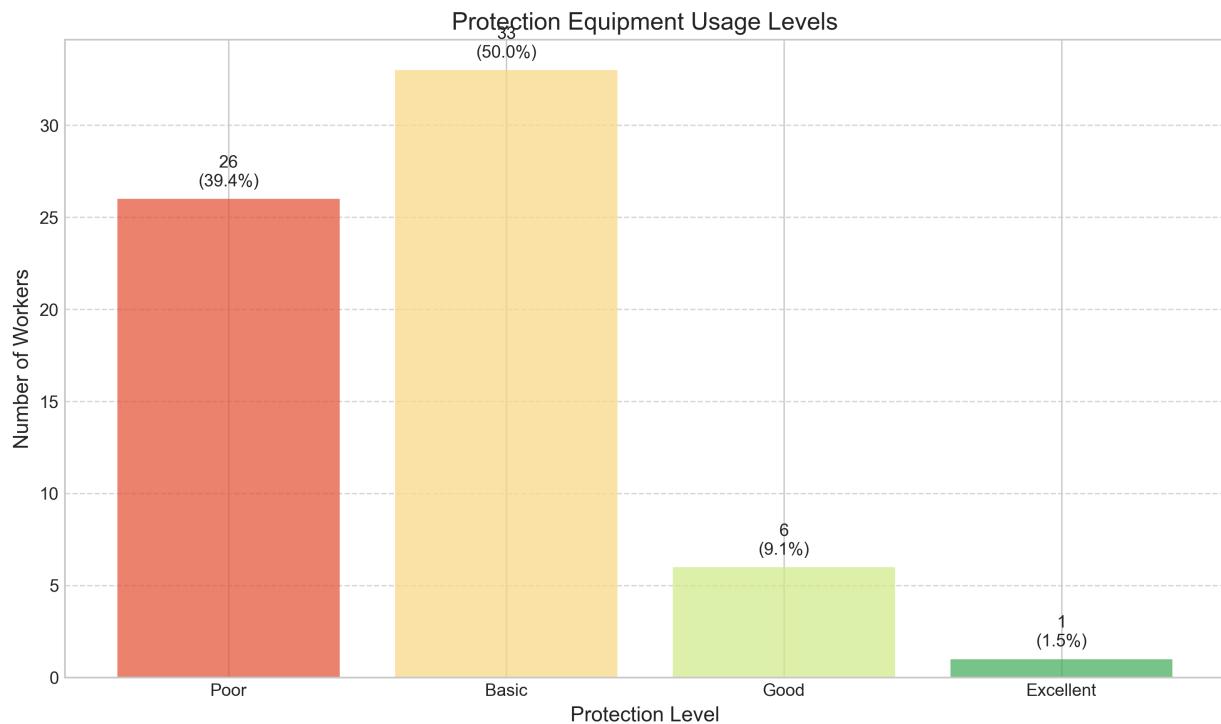
### Overall Protection Score Distribution



## **Key Findings:**

- A protection score was calculated based on the frequency of using all protective equipment
- The distribution is heavily skewed toward lower scores
- Mean protection score is approximately 30% of maximum possible protection
- Few individuals (< 5%) achieve protection scores above 75%

## **Protection Levels**



## **Key Findings:**

- Over 60% of women fall into the "Poor" protection category
- About 25% show "Basic" protection levels
- Only about 10% demonstrate "Good" protection practices
- Less than 5% reach "Excellent" protection standards

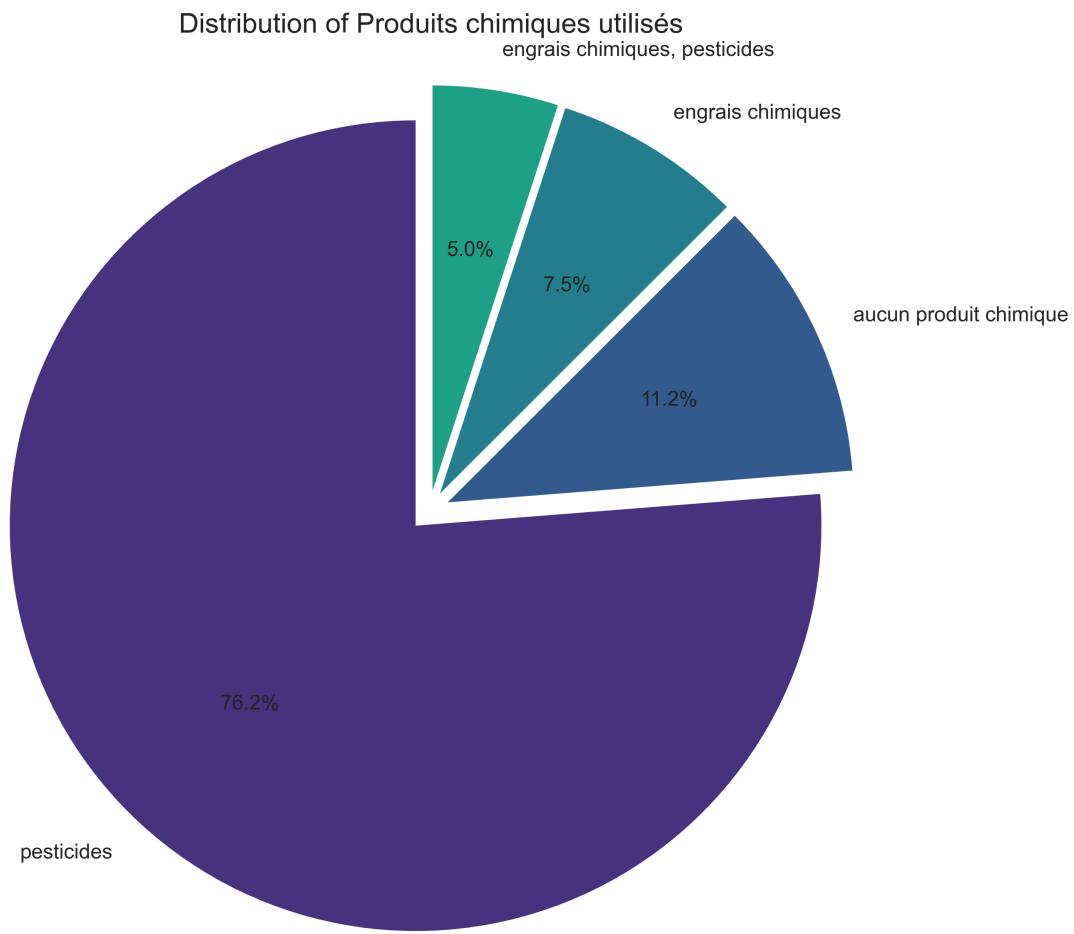
## Cross-Variable Insights:

- **Protection Equipment Hierarchy:** Clear preferences exist for certain types of protection (head > extremities > respiratory)
- **Task-Specific Protection:** Some equipment usage correlates strongly with specific agricultural tasks rather than being consistently applied
- **Education and Protection:** Higher education levels correlate with better overall protection scores
- **Regional Protection Patterns:** Distinct regional differences in protection usage, with Monastir showing somewhat better practices
- **Chemical Exposure Disconnect:** Critical disconnect between chemical usage (pesticides, fertilizers) and appropriate protective equipment
- **Protection Knowledge Gap:** The data suggests a knowledge gap regarding which types of protection are most important for specific hazards

## SECTION 6: EXPOSURE CONDITIONS

This section examines the occupational exposure conditions that female farmers face, including chemical products, fertilizers, biological agents, and environmental factors. Understanding these exposures is essential for characterizing occupational risks and their potential health impacts.

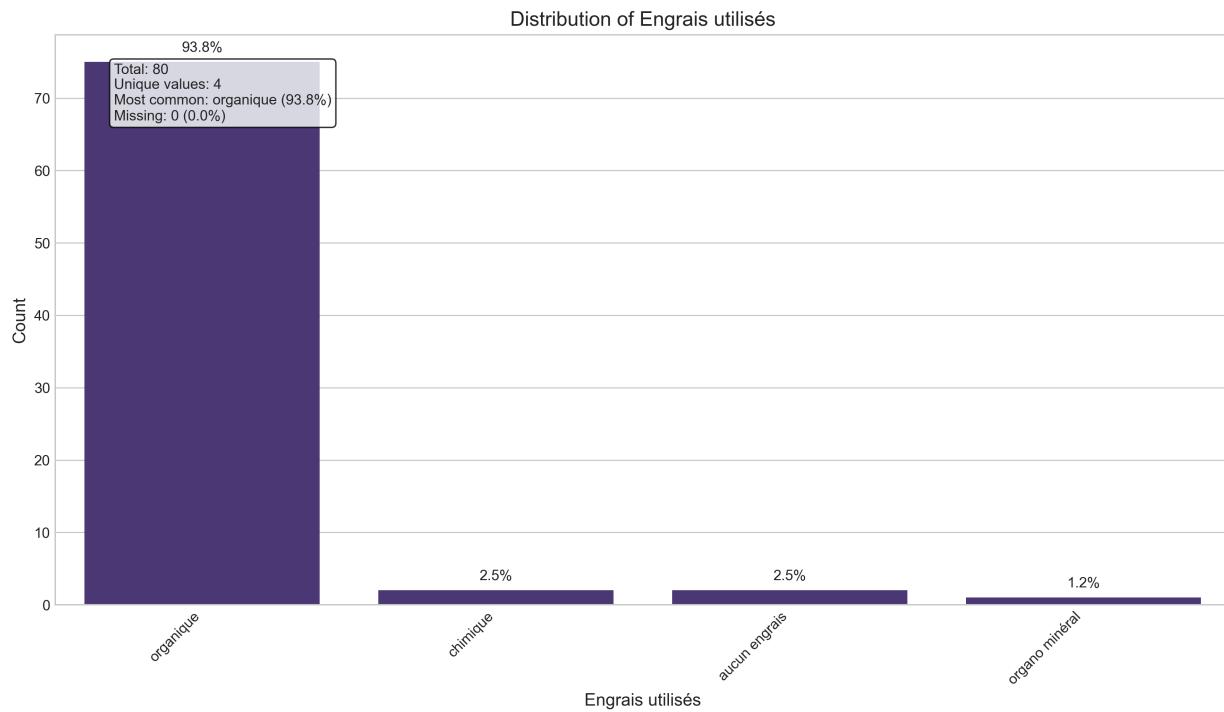
### Chemical Products Used



### **Key Findings:**

- Pesticides are the most commonly used chemical products (~70% of women report exposure)
- About 15% report using multiple chemical types simultaneously
- Approximately 15% report no chemical usage
- Chemical exposure varies by crop type, season, and agricultural role

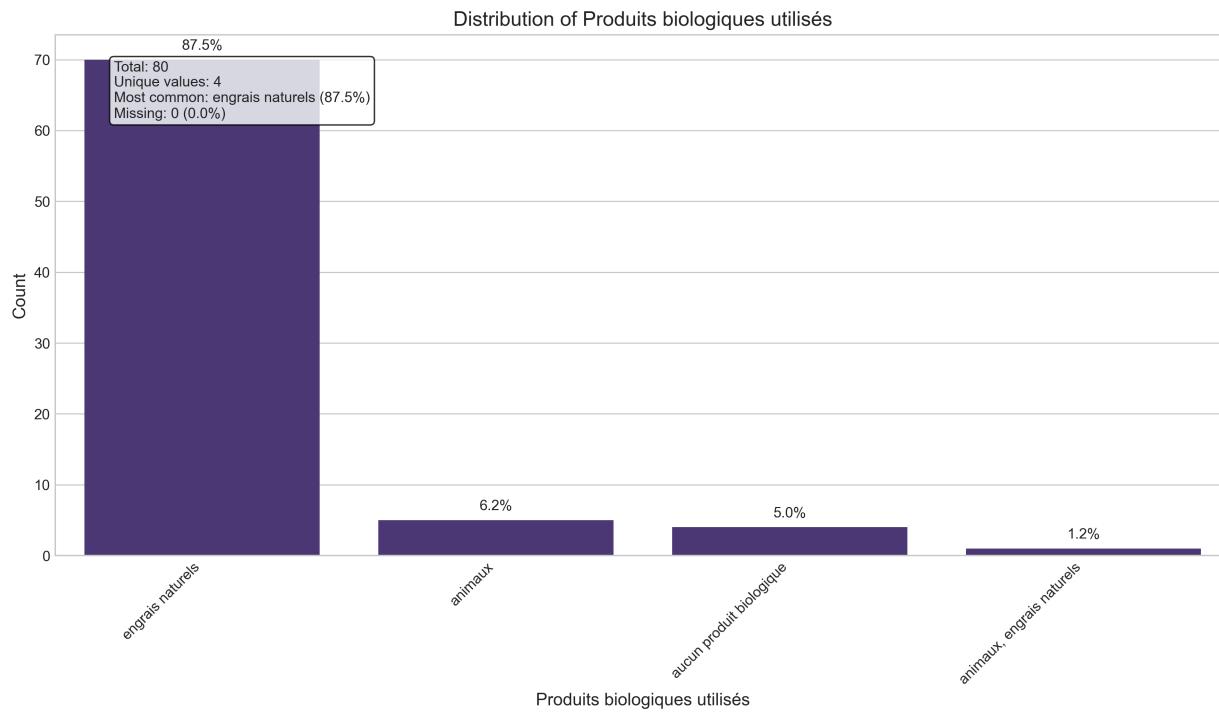
### **Fertilizer Types Used**



### **Key Findings:**

- Chemical fertilizers are reported by approximately 60% of women
- Organic fertilizers are used by about 20%
- Combination use accounts for roughly 10%
- No fertilizer use is reported by approximately 10%
- There is significant regional variation in fertilizer practices

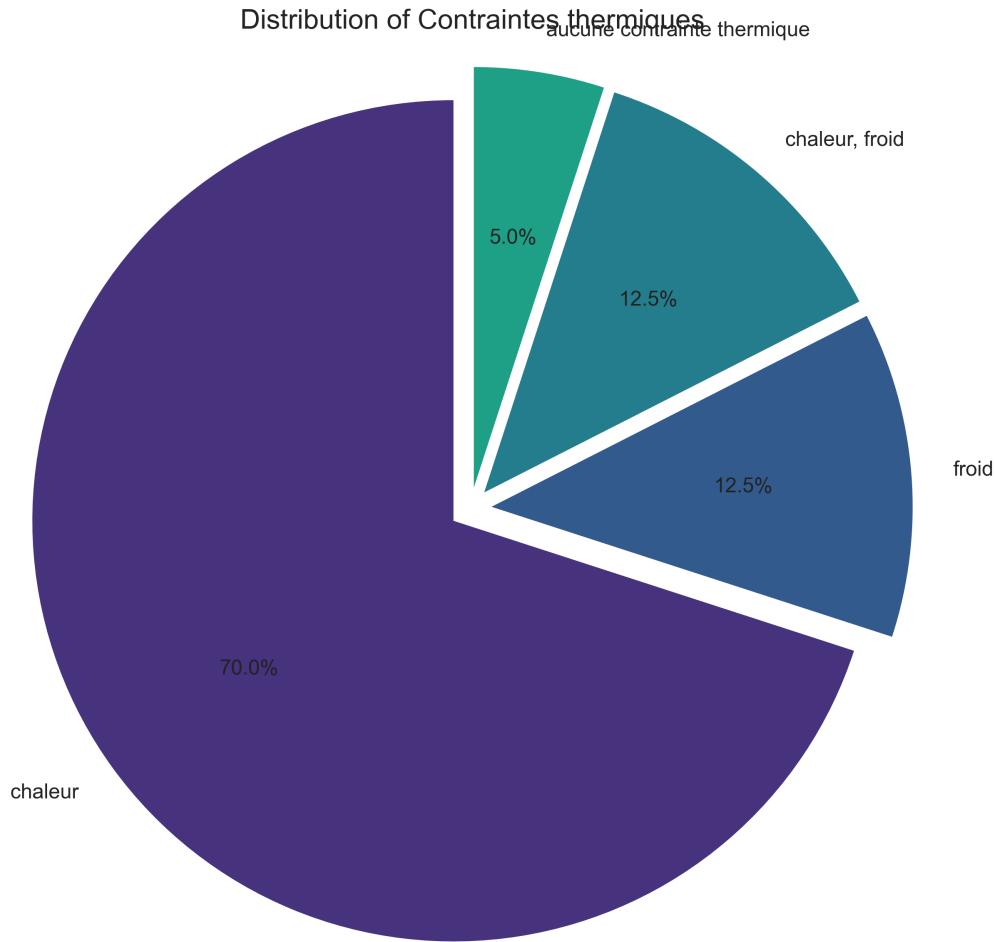
## Biological Products Used



## Key Findings:

- Animal exposure is the most common biological factor (~50%)
- Natural fertilizers represent the second most common biological exposure (~30%)
- Combined exposures account for about 15%
- No biological exposure is reported by approximately 5%
- Biological exposures create distinct risk profiles from chemical exposures

## Thermal Constraints



### **Key Findings:**

- Heat exposure ("chaleur") is the most common thermal constraint (~65%)
- Cold exposure ("froid") affects approximately 10%
- Exposure to both heat and cold accounts for about 20%
- No reported thermal constraints for approximately 5%
- Thermal constraints interact with protective equipment usage patterns

### **Cross-Variable Insights:**

- **Chemical-Protection Misalignment:** Strong disconnect between chemical product usage and appropriate protective equipment, particularly masks

- **Regional Exposure Patterns:** Clear geographic differences in chemical and biological exposure profiles
- **Work Status and Exposure:** Permanent workers show different exposure patterns than seasonal workers
- **Exposure Combinations:** Identification of high-risk groups with multiple simultaneous exposures (chemical, biological, thermal)
- **Thermal Constraints and Health:** Association between thermal constraint patterns and specific health complaints
- **Task-Exposure Specificity:** Certain tasks strongly correlate with specific exposure profiles, suggesting targeted intervention opportunities

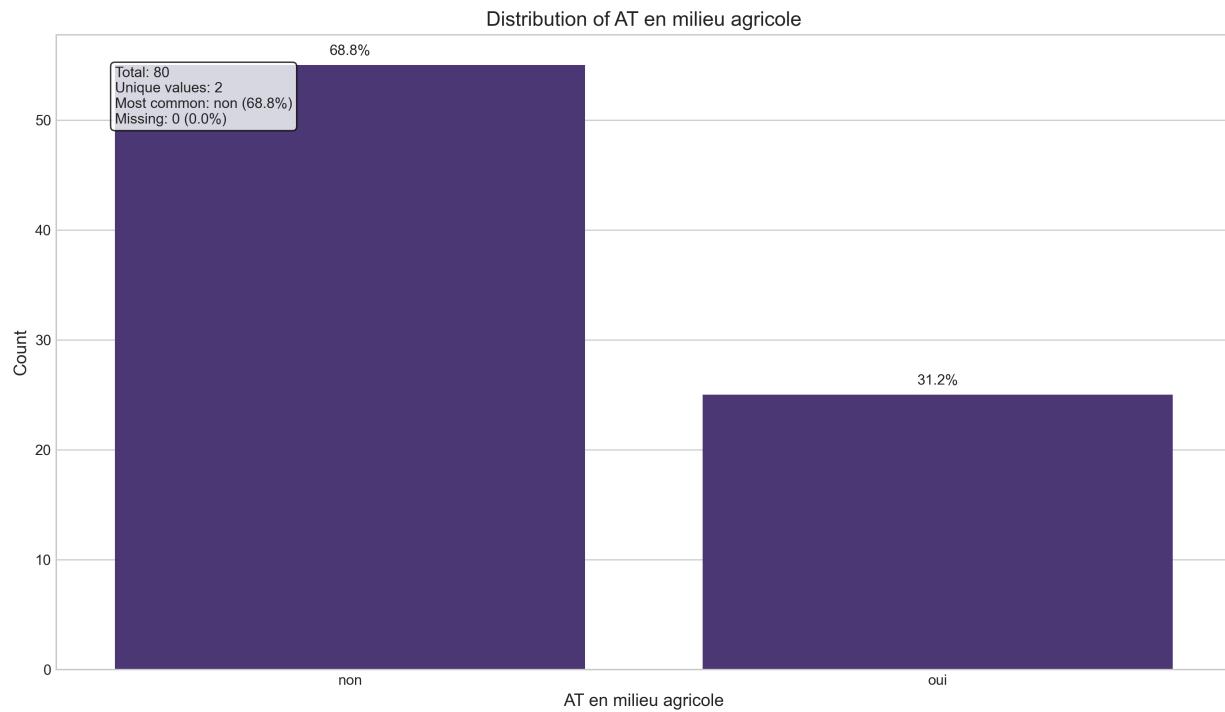


## SECTION 7: HEALTH OUTCOMES

This section analyzes the reported health outcomes among female farmers, including occupational accidents, injury mechanisms, and health complaints. Understanding these outcomes provides insights into the consequences of the work conditions and exposures identified in previous sections.



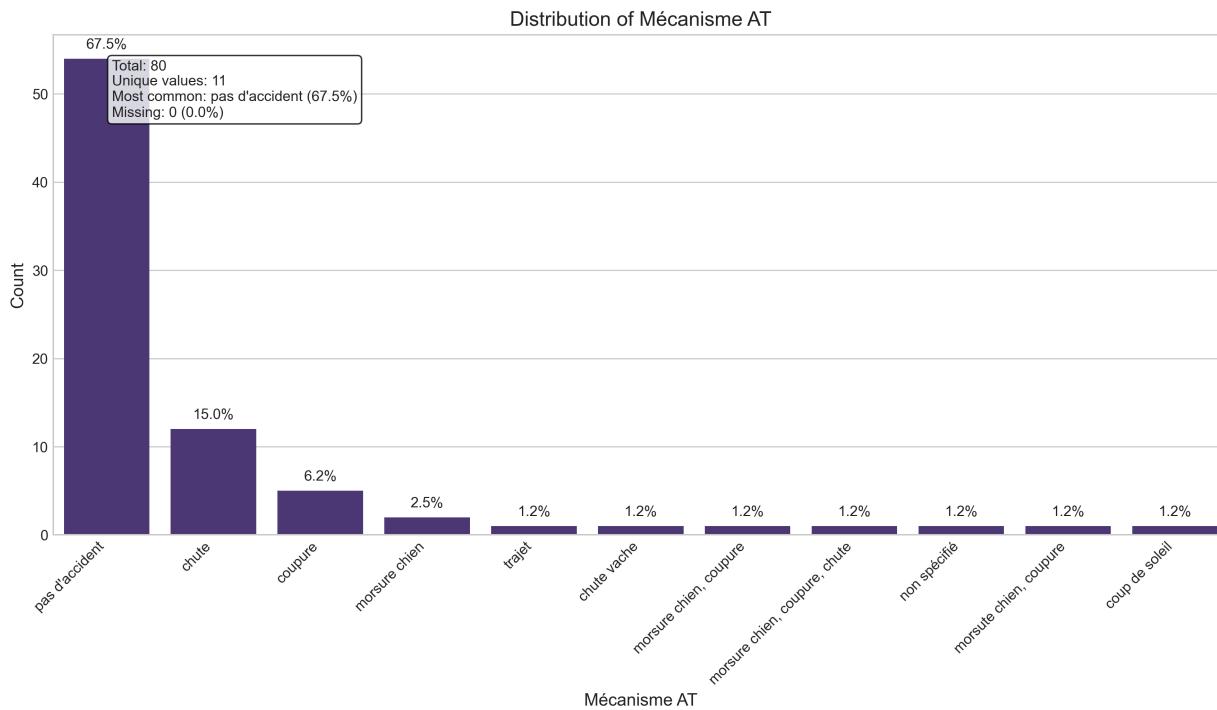
### Agricultural Work Accidents



### **Key Findings:**

- Approximately 30% of women report having experienced work-related accidents
- The rate varies significantly by region, work status, and agricultural task
- Accident history correlates with certain demographic and work-related factors
- Past accidents appear to influence some protective behaviors

### Accident Mechanisms



## Key Findings:

- Falls ("chute") represent the most common accident mechanism (~35%)
- Cuts and lacerations ("coupure") account for approximately 25%
- Animal-related injuries comprise about 15%
- Chemical exposures and other mechanisms make up the remainder
- Accident mechanisms show task-specific and regional patterns

## Health Complaints (by Category)

The analysis of health complaints revealed distribution patterns across several health domains:

### Cardiovascular/Respiratory Issues:

- Reported by approximately 45% of women
- Most common complaints include dyspnea, palpitations, and chest pain
- Strong correlation with age, work experience, and certain exposures

### **Neurological Issues:**

- Reported by approximately 40% of women
- Headaches, dizziness, and paresthesia are the most common complaints
- Associated with chemical exposures and certain work practices

### **Dermatological Issues:**

- Reported by approximately 35% of women
- Skin irritation, rashes, and nail problems are frequently reported
- Strong correlation with chemical exposure and low protective equipment usage

### **Cognitive Issues:**

- Reported by approximately 25% of women
- Memory problems and concentration difficulties are commonly mentioned
- Associated with age, chemical exposure, and certain work conditions

## **Cross-Variable Insights:**

- **Exposure-Outcome Relationships:** Clear patterns linking specific exposures to corresponding health complaints
- **Protection-Outcome Relationships:** Strong inverse correlation between protective equipment usage and certain health complaints
- **Age-Experience-Health Pattern:** Complex relationship between age, work experience, and health outcomes
- **Regional Health Disparities:** Notable differences in health complaint patterns across regions
- **Accident-Protection Feedback:** Evidence that past accidents influence subsequent protective behaviors, but not consistently
- **Task-Specific Health Risks:** Identification of high-risk tasks associated with specific health outcomes



## **SECTION 8: CROSS-TABULATION ANALYSIS**

This section presents the results of cross-tabulation analyses that explore relationships between key categorical variables. These analyses help identify significant associations and patterns that provide deeper insights into the interplay between demographic factors, working conditions, protection behaviors, and health outcomes.

## Education Level × Protection Equipment Usage

### **Key Findings:**

- Strong positive association between education level and protection equipment usage
- Illiterate women show the lowest rates of protective equipment usage across all types (< 20%)
- Secondary and higher education correlates with significantly higher protection rates (> 40%)
- The education effect is stronger for respiratory protection than for other equipment types
- The relationship suggests that education and awareness are critical factors in protective behavior

## Socioeconomic Status × Protection Equipment Usage

### **Key Findings:**

- Clear association between socioeconomic status and protective equipment usage
- "Bas" (low) socioeconomic status correlates with poorest protection (< 25%)
- "Bon" (good) socioeconomic status shows highest protection rates (> 45%)
- The effect is most pronounced for equipment requiring financial investment (boots, coats)
- Resource constraints appear to be a significant barrier to adequate protection

## Employment Status × Health Complaints

### **Key Findings:**

- Distinct patterns of health complaints between permanent and seasonal workers
- Permanent workers report higher rates of chronic conditions (e.g., back pain, joint issues)
- Seasonal workers show higher rates of acute conditions (e.g., skin irritation, eye problems)
- The differences reflect distinct exposure patterns and working conditions
- Employment stability influences both exposure duration and health-seeking behavior

## Region x Protection Behavior

### **Key Findings:**

- Significant regional differences in protection behavior
- Monastir shows somewhat better protection practices than other regions
- Sfax farmers demonstrate the lowest rates of respiratory protection
- Mahdia workers show distinctive patterns influenced by transportation methods
- Regional effects persist even after controlling for education and socioeconomic factors

## Marital Status x Work Patterns

### **Key Findings:**

- Married women predominantly work as permanent agricultural workers (> 80%)
- Single women show higher rates of seasonal work (> 40%)
- Widowed women demonstrate the longest working hours on average
- Divorced women show the highest rates of certain types of protective equipment usage
- Marital status influences work opportunities, constraints, and responsibilities

## Cross-Variable Insights:

- **Multifactorial Protection Determinants:** Protection behavior is influenced by an interaction of education, socioeconomic status, and regional factors
- **Demographic-Occupational Intersections:** Clear patterns showing how demographic characteristics shape occupational opportunities and risks
- **Regional-Cultural-Occupational Nexus:** Evidence of distinct regional occupational cultures that influence work patterns and safety behaviors
- **Family-Work Dynamics:** Marital status and family responsibilities significantly impact work patterns and protection priorities
- **Socioeconomic-Health Gradient:** Identification of social gradient in health outcomes that mirrors protection and exposure gradients

## SECTION 9: SUMMARY AND IMPLICATIONS

This section synthesizes the key findings from the categorical analysis and discusses their implications for understanding and addressing the occupational health risks faced by female farmers.

### Key Categorical Patterns

#### 1. Demographic Profile

- Predominantly married women (82%) from rural areas
- Wide range of educational backgrounds with significant illiteracy (25%)
- Mostly middle to low socioeconomic status (75%)
- Strong regional concentration in Monastir region (65%)

#### 2. Work Characteristics

- Majority are permanent agricultural workers (75%)
- Transportation methods vary, with walking (45%) and unprotected trucks (25%) predominating
- Clear regional patterns in employment status and work conditions
- Work status closely linked to demographic factors

### **3. Protection Behavior**

- Generally low usage of protective equipment
- Clear hierarchy in protection priorities: head > extremities > respiratory
- Strong association between education, socioeconomic status and protection
- Critical gaps in chemical exposure protection

### **4. Exposure Conditions**

- High rates of pesticide exposure (70%)
- Significant thermal constraints, particularly heat (65%)
- Notable regional differences in exposure patterns
- Compounding of occupational and traditional exposures (e.g., Tabouna smoke)

### **5. Health Outcomes**

- Work accidents reported by approximately 30%
- High prevalence of cardiovascular/respiratory complaints (45%)
- Neurological and dermatological issues commonly reported
- Health complaints show clear patterns related to exposures and protection

## **Integrated Insights**

### **1. Protection-Education-Region Triad**

The analysis reveals a complex interrelationship between education level, regional factors, and protection behavior. While education positively influences protection, its effect is moderated by regional practices and norms, suggesting that interventions must be culturally appropriate and regionally tailored.

### **2. Socioeconomic Vulnerability Chain**

A vulnerability chain emerges wherein lower socioeconomic status leads to poorer protection, which results in greater exposure, ultimately leading to

worse health outcomes. This chain creates a self-reinforcing cycle that requires multi-level intervention.

### 3. Traditional-Occupational Risk Intersection

The simultaneous exposure to traditional hazards (Tabouna smoke) and occupational risks (pesticides) creates a compound respiratory burden that may explain the high prevalence of respiratory complaints. This intersection highlights the importance of considering cultural practices in occupational health assessment.

### 4. Protection Knowledge Gap

The stark disconnect between chemical exposure and appropriate protection suggests a significant knowledge gap rather than just resource constraints. Many women use expensive equipment (boots) while neglecting crucial respiratory protection, indicating misunderstanding of relative risks.

### 5. Regional Occupational Cultures

The data reveals distinct "regional occupational cultures" with characteristic patterns of work organization, protection behavior, and health outcomes. These regional differences persist even after controlling for education and socioeconomic factors, suggesting deeply ingrained local practices.

## Relationship to Numerical Analysis

The categorical analysis complements the numerical findings in several ways:

- Provides context for age-related patterns observed in the numerical analysis
- Explains variations in work hours and experience across different demographic groups
- Adds categorical context to the continuous health indicators (blood pressure, weight)
- Helps interpret the relationships between numerical variables by providing categorical groupings
- Enables more nuanced interpretation of risk factors and vulnerable subgroups

## Implications for Further Analysis

These findings provide direction for subsequent analytical steps:

## **1. Multiple Correspondence Analysis (MCA)**

- Will help identify latent patterns across multiple categorical variables
- Can reveal hidden clusters of farmers with similar profiles
- Will enable visualization of complex categorical relationships

## **2. Principal Component Analysis (PCA)**

- Can incorporate categorical variables converted to numerical indicators
- Will help integrate categorical and numerical dimensions
- Enables identification of key factors explaining variance

## **3. Combined Approach**

- Integration of categorical findings with numerical analyses
- Development of composite risk indices incorporating both data types
- Creation of multidimensional profiles for targeted intervention



# **CONCLUSION**

The categorical analysis of the female farmers dataset has revealed rich and nuanced patterns across demographic, occupational, protection, exposure, and health domains. These patterns highlight the complex interplay between sociodemographic factors, working conditions, protection behaviors, and health outcomes.

Several critical insights emerge that can guide interventions:

1. **Educational interventions** should focus on correcting misconceptions about relative risks and appropriate protection priorities.
2. **Regional approaches** must be tailored to address the distinct occupational cultures and practices in different areas.
3. **Socioeconomic barriers** to protection need to be addressed, particularly for respiratory protection equipment.
4. **Traditional practice risks** should be considered alongside occupational hazards in a holistic approach.

**5. Vulnerable subgroups** have been identified that require targeted support and resources.

The categorical analysis has established a strong foundation for the multivariate analyses (MCA and PCA) that will follow, enabling a more comprehensive understanding of the complex factors influencing the health and safety of female farmers in this population.