

Key Insights from Principal Component Analysis (PCA) of Female Farmers' Health Study

Primary Insights

1. Four Key Dimensions of Health and Work Patterns

The PCA analysis revealed four principal components that together explain 63% of the variance in the numerical data:

- **PC1: Age and Experience Factor (23.6% of variance)**
 - Strongly associated with age, years in agriculture (Ancienneté agricole), and systolic blood pressure (TAS)
 - Represents the cumulative effects of aging and agricultural experience on health
- **PC2: Family Structure Factor (16.1% of variance)**
 - Negatively associated with number of children, number of dependents, and height
 - Represents family composition and physical characteristics
- **PC3: Cardiovascular Health Factor (13.5% of variance)**
 - Contrasts global arterial pressure (GAD, positive) with diastolic blood pressure and weight (negative)
 - Represents a complex physiological health dimension
- **PC4: Work Intensity Factor (9.9% of variance)**
 - Contrasts systolic blood pressure (positive) with work days per week and hours per day (negative)

- Represents the relationship between work intensity and cardiovascular stress

2. Four Distinct Worker Profiles Identified

Enhanced cluster analysis based on PCA dimensions identified four distinct profiles of female agricultural workers:

- **Enhanced Cluster 0 (36.2% of sample): Younger Workers with Better Protection**
 - Younger age (38.5 years, -18.4% vs overall)
 - Less agricultural experience (8.6 years, -50.7% vs overall)
 - Higher education levels (58.6% secondary education, +13.6% vs overall)
 - Higher protection scores (+8.5% vs overall)
 - Lower exposure to traditional practices (55.2% Tabouna exposure, -18.6% vs overall)
- **Enhanced Cluster 1 (30.0% of sample): Older High-Intensity Workers**
 - Older age (59.3 years, +25.8% vs overall)
 - Extensive agricultural experience (27.8 years, +58.8% vs overall)
 - Higher work intensity (7.3 hours/day, +18.2%; 6.9 days/week, +6.2%)
 - Higher blood pressure (TAS: 138.3, +11.0%; TAD: 79.2, +7.3%)
 - Lower protection scores (-2.2% vs overall)
- **Enhanced Cluster 2 (2.5% of sample): Elderly High-Risk Workers**
 - Eldest age group (67 years, +42.2% vs overall)
 - Most agricultural experience (42.5 years, +143.2%)
 - Abnormal cardiovascular indicators (GAD: 2.02, +91.4% vs overall)
 - Lowest protection scores (-20.0% vs overall)
 - All members illiterate (100% analphabète)
- **Enhanced Cluster 3 (31.2% of sample): Mid-Age Workers with Higher Family Burden**

- Middle-aged workers (43.9 years, -6.8% vs overall)
- Higher numbers of children and dependents (+31.4% and +49.2% vs overall)
- Higher weight (85.8 kg, +17.4% vs overall)
- Strong marriage rate (96% married, +17.2% vs overall)
- Lower protection scores (-6.1% vs overall)

3. Age-Experience-Protection Paradox

The PCA analysis revealed an unexpected negative relationship between agricultural experience and protective equipment usage:

- Older, more experienced workers show **lower** protection scores despite higher health risks
- PC1 (Age/Experience) shows a slight negative correlation (-0.025) with protection scores
- PC2 (Family Structure) shows a stronger negative correlation (-0.171) with protection
- The oldest cluster (Enhanced Cluster 2) shows significantly lower protection (-20% vs overall)

4. Socioeconomic-Physiological Risk Correlation

The PCA uncovered important relationships between socioeconomic factors and physiological health:

- Workers with higher cardiovascular risk indicators (Enhanced Cluster 3) have higher socioeconomic status (24% "bon" status, +7.8%)
- Workers with abnormal GAD values (Enhanced Cluster 2) have the lowest socioeconomic status (50% "bas", +31.2%)
- Education level correlates positively with protection behaviors but negatively with agricultural experience

Integration with MCA Findings

1. Enhanced Understanding of Worker Profiles

When combining PCA numerical clusters with MCA categorical patterns:

- **Young-Educated Profile:** Enhanced Cluster 0 from PCA corresponds closely with MCA Clusters 6 and 9, showing higher protection despite less experience
- **Older-Experienced Profile:** Enhanced Cluster 1 from PCA overlaps with MCA Clusters 1 and 7, highlighting experienced workers with poor protection habits
- **High-Risk Profile:** Enhanced Cluster 2 from PCA corresponds with the most vulnerable subgroup in MCA Cluster 2, facing compound health risks
- **Family-Burdened Profile:** Enhanced Cluster 3 from PCA matches with portions of MCA Clusters 3 and 5, showing how family responsibilities affect protection behavior

2. Regional Variations Explained by Numerical Factors

The integration of PCA with MCA regional findings reveals:

- **Monastir Workers:** Higher presence in Enhanced Cluster 0 (younger, better educated) with better protection despite less experience
- **Sfax Workers:** Concentrated in Enhanced Cluster 1 (older, high-intensity) with minimal protection and higher health risks
- **Mahdia Workers:** Predominantly in Enhanced Cluster 3 (family-burdened) with specific protection patterns shaped by economic needs

3. Work Pattern and Protection Relationship

Combined PCA-MCA analysis provides deeper insights on work patterns:

- Higher work intensity (Enhanced Cluster 1) correlates with specific MCA protection patterns (prioritizing head protection over respiratory protection)
- The relationship between work hours and protection shows a negative correlation (-0.17 between PC4 and protection scores)
- Seasonal workers (found more in Enhanced Cluster 0) show different protection patterns than permanent workers (more common in Enhanced Clusters 1 and 3)

4. Cardiovascular Health Risk Patterns

The PCA cardiovascular dimension offers crucial context for MCA protection findings:

- Workers with highest cardiovascular risk (Enhanced Cluster 2) engage least in protective behaviors
- Traditional practices (Tabouna smoke exposure) correlate with higher blood pressure values in Enhanced Clusters 1 and 3
- The combination of high TAS values (PC1) and minimal respiratory protection (from MCA) creates a compound health risk

Hidden Insights

1. Education-Protection-Age Triangle

The PCA-MCA integration reveals a complex relationship:

- Education level strongly predicts protection behavior (+8.5% protection in Enhanced Cluster 0 with highest education)
- Education effect is moderated by age (younger workers benefit more from education)
- A key insight is that education has a stronger effect on protection behavior than years of agricultural experience

2. Family Structure Impact on Protection

A surprising pattern emerges when integrating family structure from PCA with protection patterns from MCA:

- Higher dependents (Enhanced Cluster 3, +49.2%) correlates with lower protection (-6.1%)
- Highest number of children (Enhanced Cluster 2, +152.6%) correlates with lowest protection (-20.0%)
- This suggests current caretaking responsibilities may reduce capacity for self-protection

3. Health Status and Protection Behavior Alignment

The PCA health dimensions show an alarming pattern when aligned with MCA protection clusters:

- Workers with the worst health indicators (Enhanced Cluster 2: GAD +91.4%) have the worst protection behaviors
- This creates a negative feedback loop where poor health may reduce ability/willingness to use protection, further worsening health

4. Work Intensity-Protection Trade-off

PCA work intensity measures (PC4) provide critical context for MCA protection patterns:

- Higher work hours and days (Enhanced Cluster 1: +18.2% hours, +6.2% days) correlate with specific protection gaps
- This suggests time pressure and productivity demands may compete with safety practices

Visualizations to Include in Presentation

1. **enhanced_clusters_pca.png** - Shows the distribution of worker clusters in the PCA space
2. **biplot.png** - Reveals relationships between numerical variables
3. **pca_scatter_protection.png** - Illustrates how protection scores distribute across PCA dimensions
4. **loadings_heatmap.png** - Displays the strength of variable associations with principal components
5. **pc_variable_correlation_heatmap.png** - Shows correlations between principal components and original variables
6. **scree_plot.png** - Visualizes the explained variance by each principal component

Alignment with Project Objectives

These PCA findings directly address the study's core objectives:

1. **Identifying key health determinants:** The PCA dimensions precisely quantify how age, experience, family structure, and work intensity influence health outcomes.
2. **Analyzing correlations:** The PCA-MCA integration reveals complex relationships between quantitative factors (age, work hours) and categorical variables (protection equipment usage, regional differences).
3. **Developing predictive insights:** The enhanced cluster profiles provide a robust framework for predicting which workers face highest risks based on their numerical characteristics.
4. **Enhancing prevention:** The findings highlight specific intervention opportunities targeting each enhanced cluster's unique risk profile.

Strategic Recommendations Based on PCA Findings

1. **Age-Appropriate Interventions:** Develop protection education specifically tailored to older workers (Enhanced Clusters 1 and 2) that addresses ingrained habits and physical limitations.
2. **Family-Centered Support:** Create support systems for workers with high family burdens (Enhanced Cluster 3) to enable better self-protection.
3. **Cardiovascular Risk Management:** Implement health monitoring for high-risk groups (Enhanced Cluster 2) alongside protection improvement.
4. **Work Pattern Optimization:** Develop protection systems compatible with high-intensity work patterns (Enhanced Cluster 1) that don't compromise productivity.
5. **Education-Experience Integration:** Leverage the higher education in younger workers (Enhanced Cluster 0) to develop peer-training programs for older workers.

These integrated PCA-MCA insights establish a comprehensive framework for developing effective, targeted prevention solutions that account for the quantitative dimensions of health risk among female agricultural workers.