

EVALUATING COMMUNITY HEALTH AWARENESS USING SOFA CHI-SQUARE AND DESCRIPTIVE ANALYSIS

1. Overview

Client:

A public health nonprofit in the United Kingdom focused on improving awareness of cardiovascular health in underserved communities

Objective:

To assess public awareness of heart disease risk factors and identify which demographic groups were most and least informed, using SOFA's statistical tools to analyze and visualize survey results.

2. Background

The organization conducted a community survey across four boroughs in Greater London. The aim was to measure understanding of common cardiovascular risks (diet, exercise, smoking, etc.) and identify populations for future targeted health campaigns. SOFA was selected for its user-friendly interface and visualization capabilities.

3. Data Summary

Dataset:

Responses from 520 individuals collected via street-intercept survey

Key Variables:

Variable	Type	Description
Aware_of_Risks	Binary	1 = Yes, 0 = No (correctly identified 3+ risk factors)
Age_Group	Categorical	18–29, 30–49, 50–69, 70+
Gender	Categorical	Male, Female
Education_Level	Categorical	High School, Vocational, University+
Ethnicity	Categorical	White, Black, Asian, Mixed, Other

Borough	Categorical	Camden, Lambeth, Hackney, Southwark
---------	-------------	-------------------------------------

4. Methodology

Software Used:

SOFA Statistics (v1.5.3)

Statistical Workflow:

1. Data Preparation:

- Imported CSV into SOFA
- Verified data types and renamed columns for readability
- Checked for missing values and filtered incomplete rows

2. Descriptive Analysis:

- Used SOFA's **Row % Tables** to summarize awareness rates by demographic
- Generated bar charts and pie charts for visualization
- Compared average awareness across boroughs

3. Chi-Square Tests:

- Cross-tabulated Aware_of_Risks with each categorical demographic variable
- Used **Chi-square test for independence** (via SOFA's GUI)
- Calculated standardized residuals to detect major contributors to associations

5. Key Results

Variable	χ^2 Value	df	p-value	Interpretation
Age_Group	18.7	3	<0.001	Older adults more likely to be aware
Education_Level	22.3	2	<0.001	Higher education associated with greater awareness
Gender	1.8	1	0.18	No significant difference
Ethnicity	7.6	4	0.10	Not statistically significant, but trends observed

Borough	12.9	3	0.004	Awareness varied significantly across boroughs
---------	------	---	-------	--

Additional Insight:

- Residents of Camden and Hackney had the **lowest awareness rates**
- Over **60% of university-educated respondents** correctly identified risk factors vs. only 35% among high school educated

6. Visual Outputs (via SOFA):

- Clustered bar chart: Awareness rate by education level
- Crosstab table: Awareness vs. borough
- Pie chart: Awareness levels (Yes/No) overall
- Mosaic plot: Awareness $\times$ Age_Group (color-coded by contribution to χ^2)

7. Deliverables

- Cleaned SOFA database file (.db) and exportable CSV
- Analytical report (12 pages) with:
 - Executive summary
 - Chi-square test output
 - Graphical dashboards and interpretation
- Summary slide deck (3 slides):
 - Key demographic findings
 - Borough-level intervention priorities
 - Recommendations for awareness campaigns

8. Application & Outcome

- Targeted awareness workshops were piloted in Camden and Southwark
- Printed materials were redesigned to include visuals that appealed to younger and less-educated audiences
- Client replicated the SOFA workflow internally for their next quarterly survey

9. Strategic Value Delivered

- Delivered **low-cost statistical insights** using open-source software
- Identified **education and geography** as key awareness drivers
- Provided **visual, actionable data** for community health outreach planning

Statssy