EXAMINING VOTER PARTICIPATION TRENDS USING PSPP CHI-SQUARE ANALYSIS

1. Overview

Client:

A nonpartisan civic engagement nonprofit in the United States focused on improving voter turnout among underrepresented groups

Objective:

To identify which demographic factors are significantly associated with voter participation in recent national elections, using PSPP to perform statistical testing on survey data. The goal was to refine outreach strategies for future voter awareness initiatives.

2. Background

Low voter turnout remains a persistent issue in many U.S. communities. The client collected survey data during the last election cycle and wanted to understand how factors such as age, education, race, and income related to voting behavior. Given budget constraints, PSPP was selected as a no-cost alternative to SPSS for analysis.

3. Data Summary

Dataset:

Post-election survey of 3,200 registered voters from six U.S. states

Key Variables:

Variable	Type	Description
Voted_Last_Election	Categorical	Yes / No
Age_Group	Categorical	18–29, 30–44, 45–64, 65+
Education_Level	Categorical	High School, Some College, College+, Postgrad
Race_Ethnicity	Categorical	White, Black, Hispanic, Asian, Other
Household_Income	Categorical	<$30K, $30K–$60K, $60K–$100K, $100K+
Region	Categorical	Northeast, Midwest, South, West

4. Methodology

Software Used:

PSPP (v1.6.2)

Workflow Steps:

1. **Data Cleaning and Preparation:**
    - Imported CSV file using PSPP's GUI
    - Recoded and labeled categorical variables
    - Verified variable distributions using **Frequencies** and **Crosstabs**
2. **Chi-Square Analysis:**
    - Tested for independence between Voted_Last_Election and each demographic variable
    - Used **Crosstabs → Statistics → Chi-Square** in PSPP
    - Examined expected counts and standardized residuals to interpret results
3. **Assumption Checks:**
    - All contingency tables checked to ensure >80% of cells had expected counts ≥5
    - Grouped sparse categories (e.g., income >$100K merged with $60K–$100K)

5. Key Results

Predictor	**χ^2 Value**	**df**	**p-value**	**Interpretation**
Age_Group	32.4	3	<0.001	Older age groups more likely to vote
Education_Level	24.6	3	<0.001	Higher education correlated with higher turnout
Household_Income	13.7	2	0.001	Higher income associated with increased participation
Race_Ethnicity	9.1	4	0.059	Not statistically significant (marginal)
Region	5.2	3	0.156	No significant geographic differences

6. Visual Outputs (Generated via PSPP & Supplemented with LibreOffice):

- Clustered bar charts: Voter turnout by age group
- Crosstab tables with row percentages
- Heat map matrix of chi-square contributions
- Summary infographic: Top 3 predictors of voting

7. Deliverables

- Cleaned PSPP .sav and exportable CSV files
- Final report (14 pages), including:
    - Executive summary of findings
    - Interpretation of significant chi-square results
    - Implications for outreach design
- Visual dashboard PDF (3 pages):
    - Turnout gaps by subgroup
    - Demographics ranked by impact
    - Suggestions for data-informed campaign messaging

8. Application & Outcome

- Client used findings to redesign email and SMS outreach by focusing on young and low-income voters
- Education-targeted messaging improved engagement in pilot campaigns
- PSPP analysis provided a replicable, cost-free approach for future survey cycles

9. Strategic Value Delivered

- Highlighted **age and education as primary turnout differentiators**
- Delivered **data-driven insights** with zero software cost
- Created **a scalable model** for volunteer-led survey analysis using open tools