

PROJECT 1: Support for Same-Day Voter Registration (Q19)

Overview

We want to see which factors (like trust in local officials or political affiliation) predict how people feel about **same-day voter registration** (Q19).

Variables

- **Target:** Q19 (Support for same-day registration)
- **Predictors:**
 - **Demographics** (Q1–Q8: Ward, Age, Gender, Marital Status, Ethnicity, Religion, Education, Income)
 - **Political Affiliation** (Q9)
 - **Confidence/Direction** (Q10, Q11, Q14–Q16)
 - **Presidential vote/excitement** (Q12, Q13)
 - **Information Sources** (Q17)
 - **Reasons for In-Person** (Q18, optional)

Steps

1. **Clean and Prepare Data**
 - Handle missing values (e.g., create “No Response” category if a question is blank).
 - [Encode data based on the mappings](#)
2. **Set Up Model**
 - Classification of support (binary or multi-class).
 - Model options: Logistic Regression or Random Forest
3. **Train/Test Model**
 - Split the dataset into training (e.g., 80%) and test (e.g., 20%).
 - [Perform k-fold cross-validation](#) (k = 5) on the training set to fine-tune model parameters.
 - Apply the final model to the test set and compare predicted vs. actual Q19 responses (accuracy, F1-score, etc.).
4. **Interpret Results**
 - Examine feature importance or coefficients to see which questions (Q10, Q11, Q14, etc.) drive support for same-day registration
 - Identify patterns (e.g., certain age groups, confidence levels, or political affiliations heavily favor or oppose)