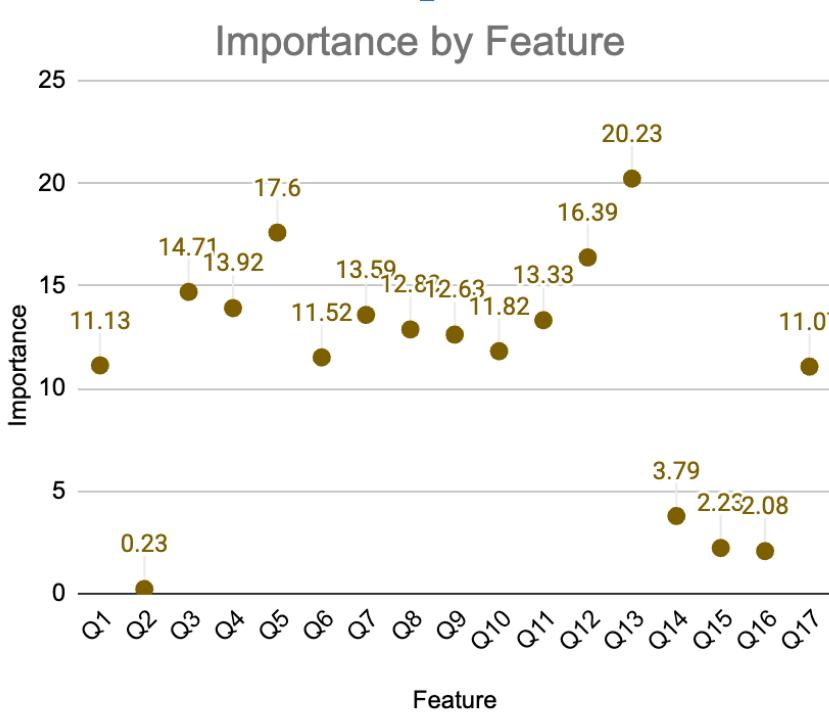


# Predictive Analysis Classification - GSB 544

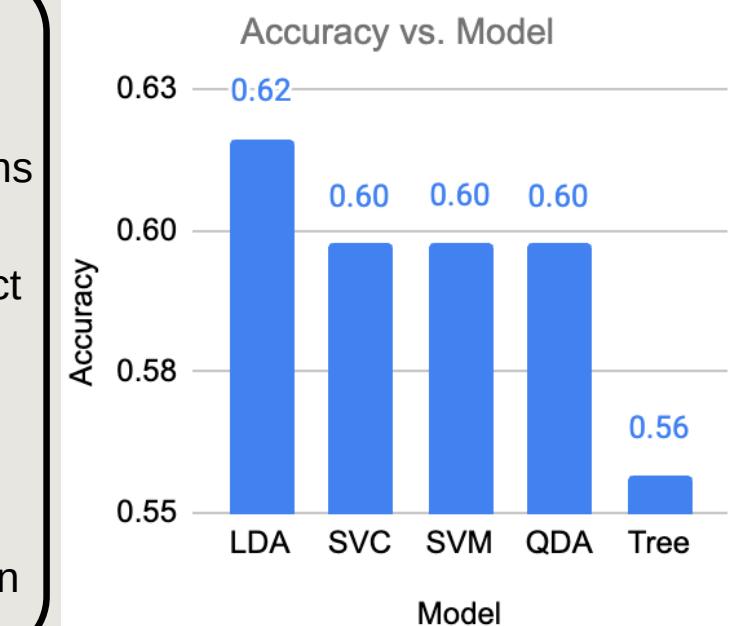
## Step 1 - Data Cleaning

- Check for missing values
- Check for outliers
- Dummify variable



## Step 2 - Feature Selection

- Top Features:
1. How Do Abortion Laws Effect Casual Sex Opinions
  2. Race
  3. How Does Smoking Effect Hook Up Culture
  4. Education Level
  5. Political Views
  6. Marajuana Opinion
  7. Casual Hook Ups Opinion



## Step 3 - Model Evaluation

The LDA (Linear Discriminant Analysis) model had the highest accuracy with a Cross Valudated accuracy score of 0.62

## Step 4 - Parameter Tuning and Cross Validation

```
params_lda = {
    "lda_solver": ['lsqr'],
    "lda_shrinkage": ['auto',
        None, 0.2, 0.4, 0.6, 0.8, 1]
grid_search = GridSearchCV(lda_model,
    params_lda, cv=5, scoring='accuracy')
grid_search.fit(X, y)
```

## Step 5 - Conclusion and Test Data

- Model Selection Based On:
- Different models
  - Different features
  - Different parameters

Accuracy on Test Data: 0.634

## Step 6 - Interpretation

- Benefits of building a political predictive model:
- Trend Recognition
  - Targeted Outreach
  - Resource Allocation (money and time)
  - Oppositional Positioning