

PSC 400

SYRACUSE UNIVERSITY

DATA ANALYTICS

FOR POLITICAL

SCIENCE

EXTENSIONS TO REGRESSION

EXPLORATION SURVEY

- <https://tinyurl.com/400topics>

CATEGORICAL VARIABLE

- **cces19.csv**
- **DV: Registered to vote (voters)**
 - 1 if registered, 0 if not
- **IV: Gender (female)**
 - 1 if female, 0 if male
 - categorical independent variable

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.94 - 0.035 * \text{female}$**
 - **What is the predicted probability that a woman is registered?**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.94 - 0.035 * \text{female}$**
 - **What is the predicted probability that a woman is registered?**
 - **$0.94 - 0.035 * 1 = 0.905$**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.94 - 0.035 * \text{female}$**
 - **What is the predicted probability that a man is registered?**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.94 - 0.035 * \text{female}$**
 - **What is the predicted probability that a woman is registered?**
 - **$0.94 - 0.035 * 0 = 0.94$**

CATEGORICAL VARIABLE

- **DV: Registered to vote or not**
- **IV: Partisanship**
 - **Democrat, Republican, Independent**
 - **Categorical variable**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.9425 - 0.0617 * \text{Independent} - 0.0004 * \text{Republican}$**
- **What is the predicted probability that a Republican is registered?**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.9425 - 0.0617 * \text{Independent} - 0.0004 * \text{Republican}$**
- **What is the predicted probability that a Republican is registered?**
- **$0.9425 - 0.0617 * 0 - 0.0004 * 1 = 0.9421$**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.9425 - 0.0617 * \text{Independent} - 0.0004 * \text{Republican}$**
- **What is the predicted probability that an Independent is registered?**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.9425 - 0.0617 * \text{Independent} - 0.0004 * \text{Republican}$**
- **What is the predicted probability that an Independent is registered?**
- **$0.9425 - 0.0617 * 1 - 0.0004 * 0 = 0.8808$**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.9425 - 0.0617 * \text{Independent} - 0.0004 * \text{Republican}$**
- **What is the predicted probability that a Democrat is registered?**

CATEGORICAL VARIABLE

- **$\text{Pr}(\text{Registered}) = 0.9425 - 0.0617 * \text{Independent} - 0.0004 * \text{Republican}$**
- **What is the predicted probability that a Democrat is registered?**
- **$0.9425 - 0.0617 * 0 - 0.0004 * 0 = 0.9425$**

CATEGORICAL VARIABLE

- **Key insight: If a categorical variable has x categories, the regression will estimate $(x-1)$ regression coefficients**
- **Category that is left out: baseline category**
 - **The other categories give the effect of being in a certain category relative to the baseline**
 - e.g. if baseline: male
 - then coefficient gives effect of being female vs. being male