Change in Docstring Logit Difference when Ablating Edge **Activation Patching** Edge in circuit -2 Edge not in circuit Line of Best Fit: y = 0.531x + -0.049;R^2: 0.27 0 2