The attached project titled "Lobbying and SOPA/PIPA" could fit well within the Data Analysis Project (DAP) expectations and criteria based on the following points:

**1. Research Problem and Relevance:**

* **Research Question**: The project aims to explore the correlation between lobbying contributions and Congressional voting behavior on SOPA/PIPA legislation. This aligns well with the DAP's expectation to analyze a real-world issue using data and statistical inference.
* **Relevance**: The study touches on an important policy issue related to lobbying and legislative decisions, which is highly relevant in political science and public policy contexts. It connects to broader themes such as political influence, policymaking, and the role of interest groups, fitting the DAP's goal of addressing a policy or public administration problem.

**2. Data Source and Data Description:**

* **Dataset**: The project uses secondary data available through ProPublica on lobbying contributions to U.S. Congress members regarding SOPA/PIPA. This is in line with DAP's requirements to use reliable publicly available datasets.
* **Dataset Description**: The data includes key variables such as member names, party affiliation, state, chamber, stance on the bill, and lobbying contributions from pro- and anti-SOPA/PIPA groups. The project provides summary statistics (e.g., mean, median, standard deviation) for the lobbying contributions, fulfilling the DAP’s expectation to clearly describe the dataset and its variables.

**3. Hypotheses and Statistical Analysis:**

* **Hypothesis Formulation**: The project tests whether a congressperson's stance on SOPA/PIPA correlates with the amount of lobbying money they receive from supporting or opposing industries. This fits the DAP's requirement to clearly define testable hypotheses related to the research question.
* **Statistical Tests**: The project uses a variety of statistical methods such as chi-square tests, two-sample hypothesis tests, and multivariate regression, which are appropriate for exploring the relationship between lobbying money and voting behavior. This meets the DAP's requirement for using statistical techniques to test hypotheses, including tests like t-tests and regression models.
* **Handling Confounders**: The project recognizes and addresses potential confounders, such as political party affiliation, by using a multivariate model. This is in line with the DAP’s expectation to account for variables that may affect the analysis.

**4. Results and Interpretation:**

* **Results Reporting**: The project reports key findings such as the statistically significant correlation between lobbying contributions and voting stance (with p-values) and provides detailed tables and visual aids (e.g., histograms, contingency tables) to support these findings. This matches DAP’s expectations to use tables and graphs to present results and to clearly connect them to the hypotheses.
* **Significance and Interpretation**: The findings are interpreted in relation to the research question, discussing how lobbying money is correlated with support or opposition to SOPA/PIPA, but not necessarily causal. This aligns with DAP’s expectations for a thorough discussion of how the results contribute to the understanding of the policy problem.

**5. Discussion and Limitations:**

* **Discussion**: The project discusses the implications of the findings, such as how lobbying money may influence legislative decisions and the limitations, like potential confounders (e.g., party affiliation). This satisfies DAP’s requirement to interpret results and discuss limitations.
* **Further Research**: The project suggests areas for further investigation, like considering members whose position differs from their party's stance. This aligns with DAP’s encouragement to explore broader implications and real-world impact.

**6. Presentation of Data and Variables:**

* The project includes summary tables for variables like lobbying contributions ("money pro" and "money con") and a detailed "Variables Dictionary" that categorizes the data (e.g., continuous or categorical variables), which meets the DAP’s data preparation and variable operationalization requirements.