Introduction, in which you describe the question your project answers or the new data your project creates, and why it is important or useful;

Data, in which you describe the source and content of your data, present tables of key summary statistics, and visualizations of key variables or relationships;

Methods, in which you describe the statistical and machine learning methods you use, including how you choose any tuning parameters;

Results, in which you present in tables or graphs the results of your methods, including (if relevant) out-of-sample prediction accuracy;

Conclusion, in which you summarize the implications of your findings, discuss possible future extensions or uses of your results, or mention limitations

1. Introduction

The United States Department of Agriculture’s National Agricultural Statistics Service (NASS) performs an “Agricultural Census” to see the state of agricultural production and health in the United States economy (USDA, 2023). It provides critical information that cannot easily be found anywhere else to inform farmers, investors, and policymakers about the effects of weather, climate, and institutional decisions. However, this census is only collected once every five years, and an accompanying survey performed every year frequently fails to collect enough data to publish results for all counties. In fact, if not enough farmers in one county fail to respond, the NASS frequently withholds data from other counties in the state to prevent third parties from calculating the unresponsive county’s yield themselves. It’s not because there isn’t agricultural production there, either; which counties are not included changes year-to-year (Johanns & Thessen, 2020).

We plan to use environmental and economic data to generate a dataset filling in these gaps in non-census years for counties in which surveys produced insufficient responses.

1. Data

We obtained our data from:

* US Census

1. Methods

ML: doing PCA with weather features, predicting crop yields

1. Results
2. Conclusion