Describing project and dataset.

The project is to predict the future of a product by analyzing their data set. We have chosen to work with a wine dataset. It contains several factors that affect the wine quality. The wine quality directly affects the price of wine. We plan to predict the quality of future wine so that we can forecast the success of the wine.

Data Analysis.

While the dataset has a wine quality column, directly using this column to predict the quality of future wine is inaccurate as there are too many things that affect the wine quality. Instead, we are going to predict the wine features and use these predictions to predict the wine quality. For predicting wine quality using it's features, we have used random forest and SVM models.

Prediction

We will now predict the wine features of the future batches. For this, we will split the wine data into 12 batches (each representing a year) and train an ML model (still to be determined)on this. Once we have the predictions for 2 years, we will again use random forest and SVM models to predict wine quality.

Conclusion

Now that we have predicted the wine quality, we will communicate how the prediction is good enough to be actionable. We can also show our visualizations here. I also want to say how can the product be made better by using the correlation plot.