Fine Wine Price Prediction using Machine Learning

Michelle Yeo and Tristan Fletcher

Royal Institution 2016

invin.io

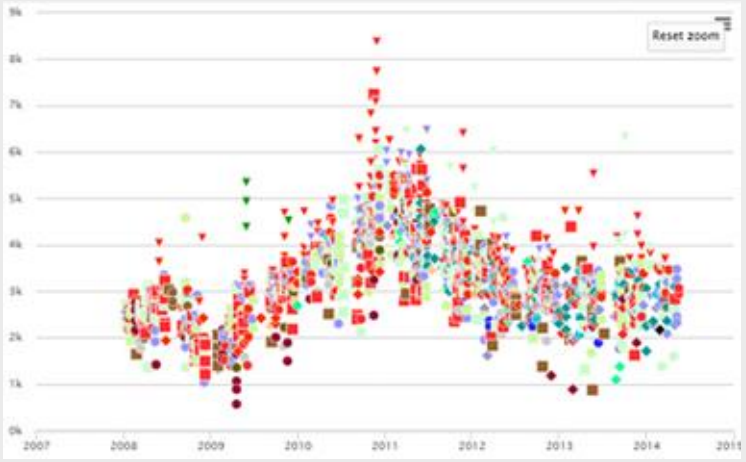
Introduction

1. Background to Fine Wine market
2. Invinio
3. Research

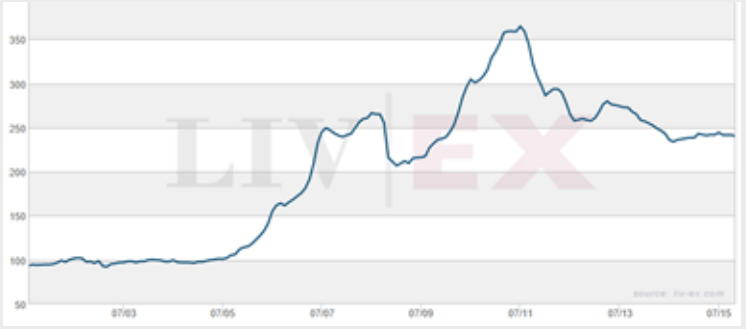
Background

- Many people invest in Treasure assets like classic cars and fine wines.
- Fine wines are inefficiently priced.
- They are more prone to price bubbles than other markets.

1982 Lafite Rothschild



Liv-ex 100



Invinio

- Invinio was set up to address these issues:
  - Intelligently price fine wines
  - Offer portfolio diversification
  - Facilitate trading
- Desire was to bring tools from hedge fund domain into this world.
  - Be completely numbers focused - be proud that we know nothing relating to the product from a consumptive perspective.
- Make it easier for outsiders to access and start investing in fine wine market.
  - Ended up with a "Nutmeg for Fine Wines".

Research

- Love using ML in new domains (e.g. medicine, supply chain, financial services).
- Great opportunity to use ML on the wine price data set.
- My question is always:
  - Is complexity justified by improved performance?
  - Michelle did great job - publication and media attention
  - Research had significant input from Prof. John Shawe-Taylor

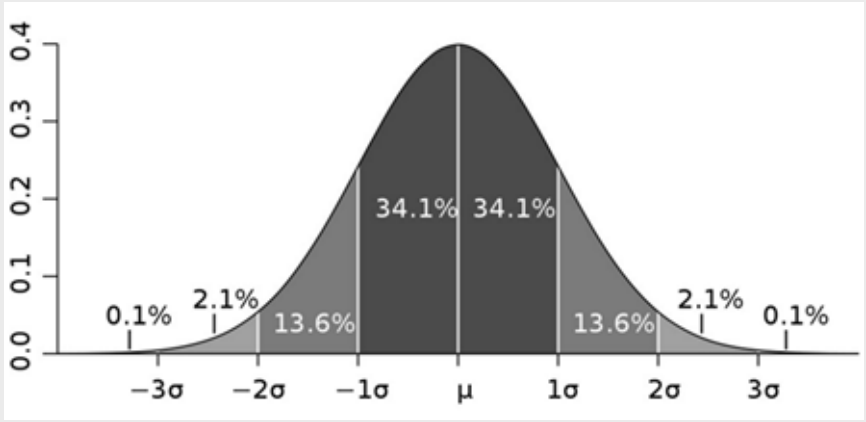
Explanation vs Prediction

- Explanation
  - Coming up with a hypothesis and testing
  - Simple linear regression
  - Common in wine literature
- Prediction
  - Focus purely on predictive accuracy
  - “Black box” algorithms like neural nets
  - Common in finance
- Idea: to use predictive machine learning algorithms on wine prices
- Questions:
  - How accurately we can predict the price of wine tomorrow based on today?
  - Is the increase in accuracy worth the effort?

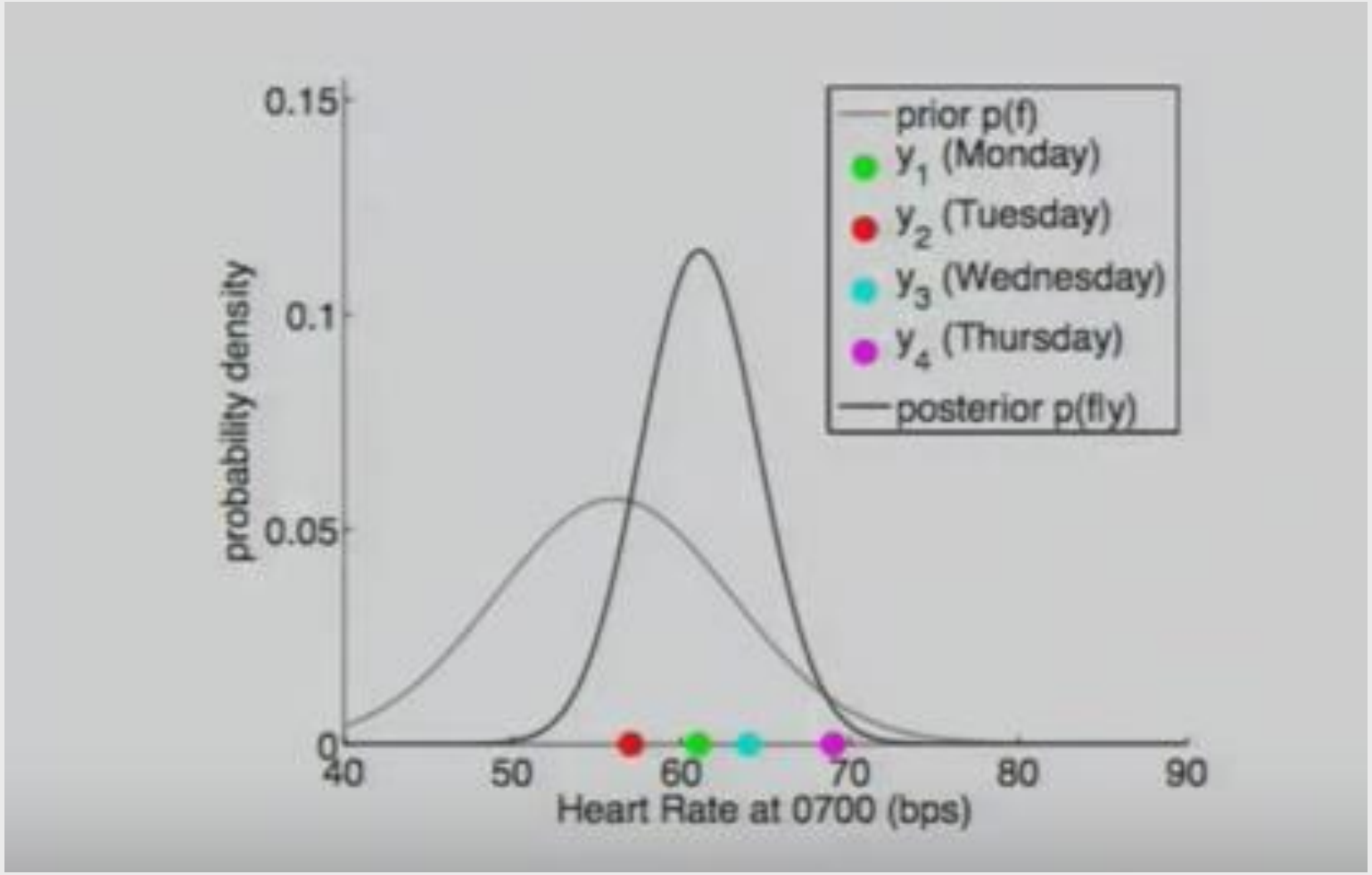
Predictive methods that purely focus on previous wine prices are unprecedented in wine price prediction.

What is a Gaussian?

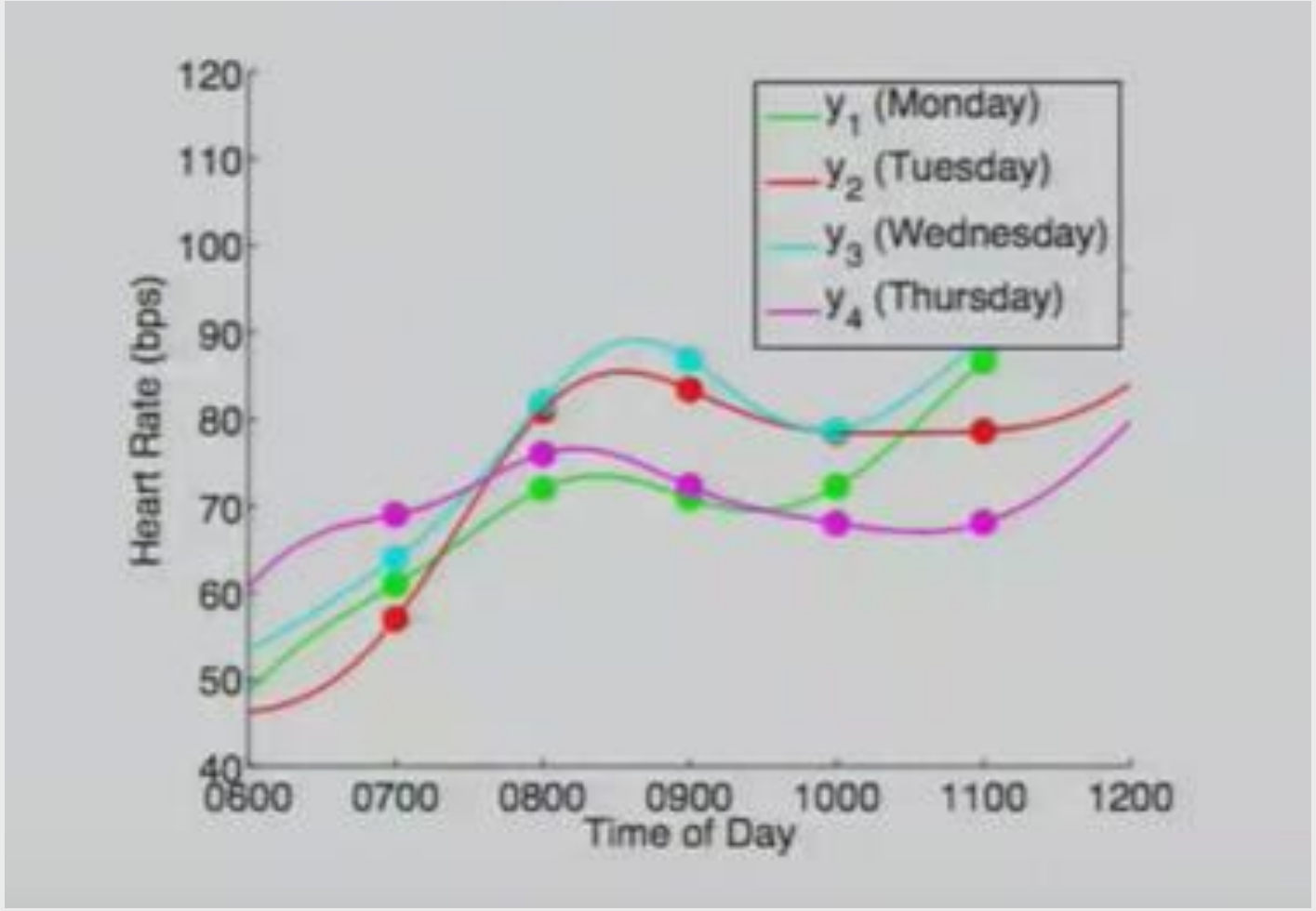
- Normal distribution with mean and variance
  - Gives the probability of an event within an interval happening
- Bayesian Inference
  - Prior: belief about a quantity before taking into account evidence
  - Evidence: observed data about the quantity
  - Posterior: updated belief in the quantity after considering evidence



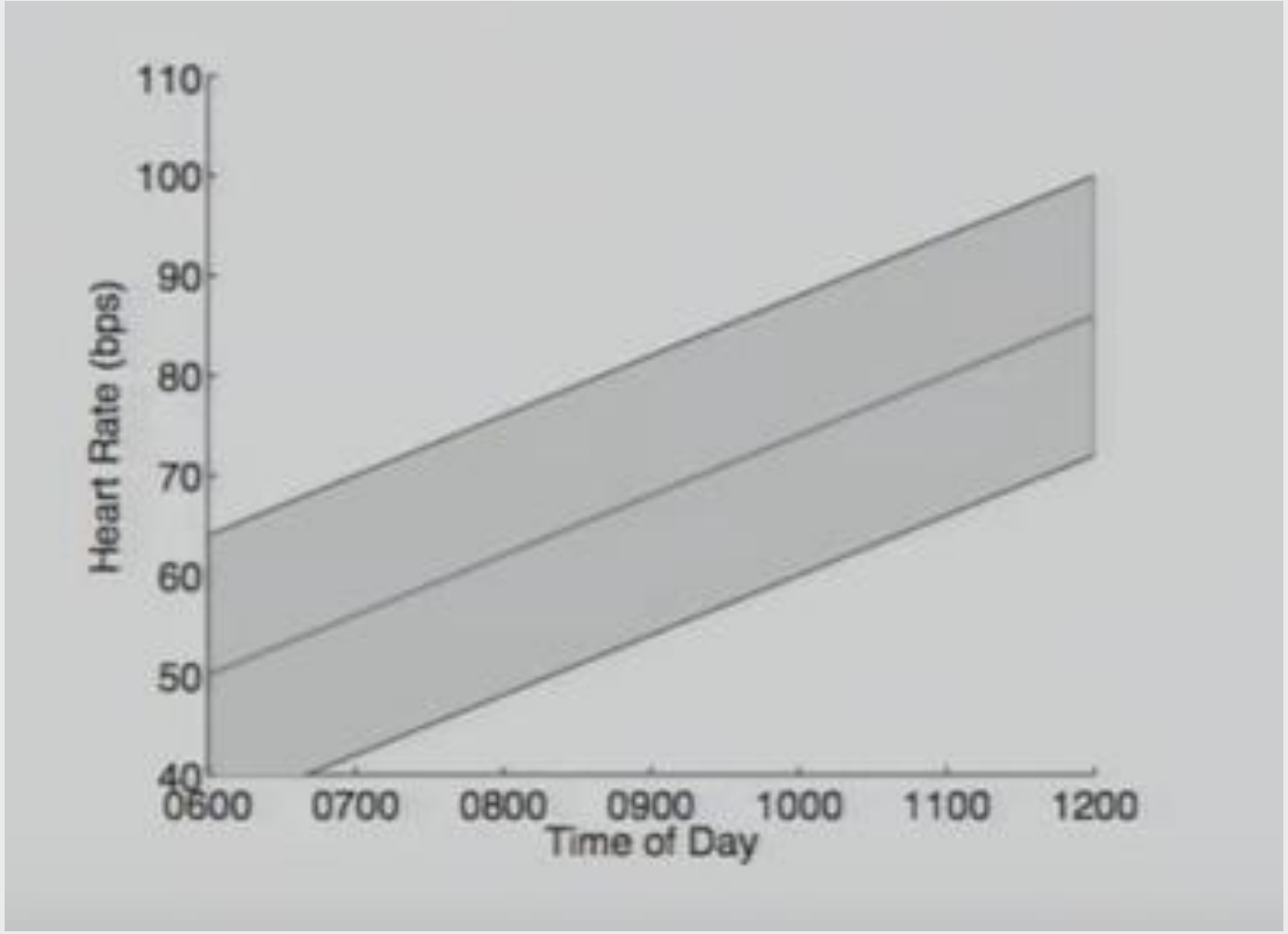
Bayesian Inference



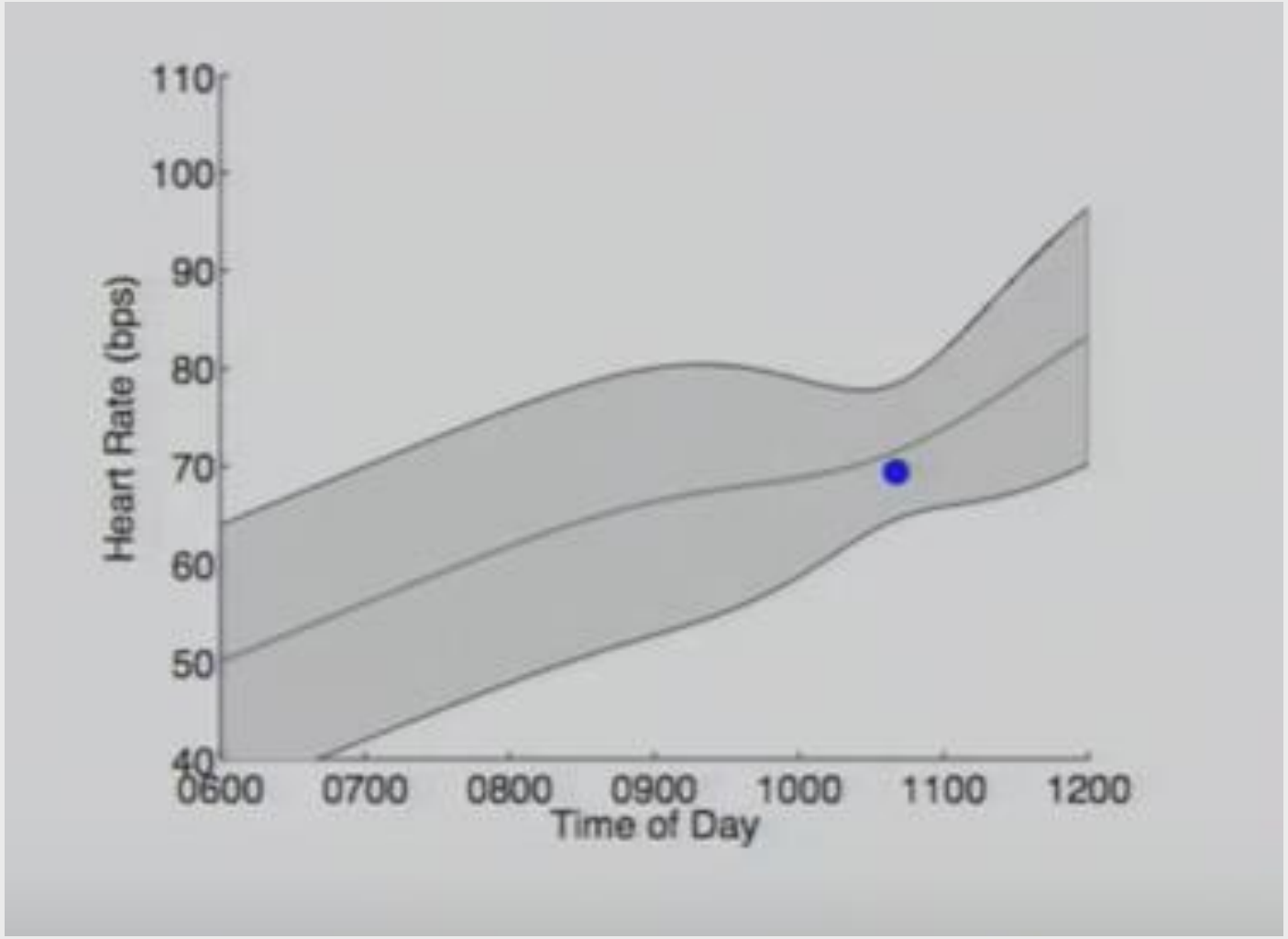
What if we are interested in a function over time?



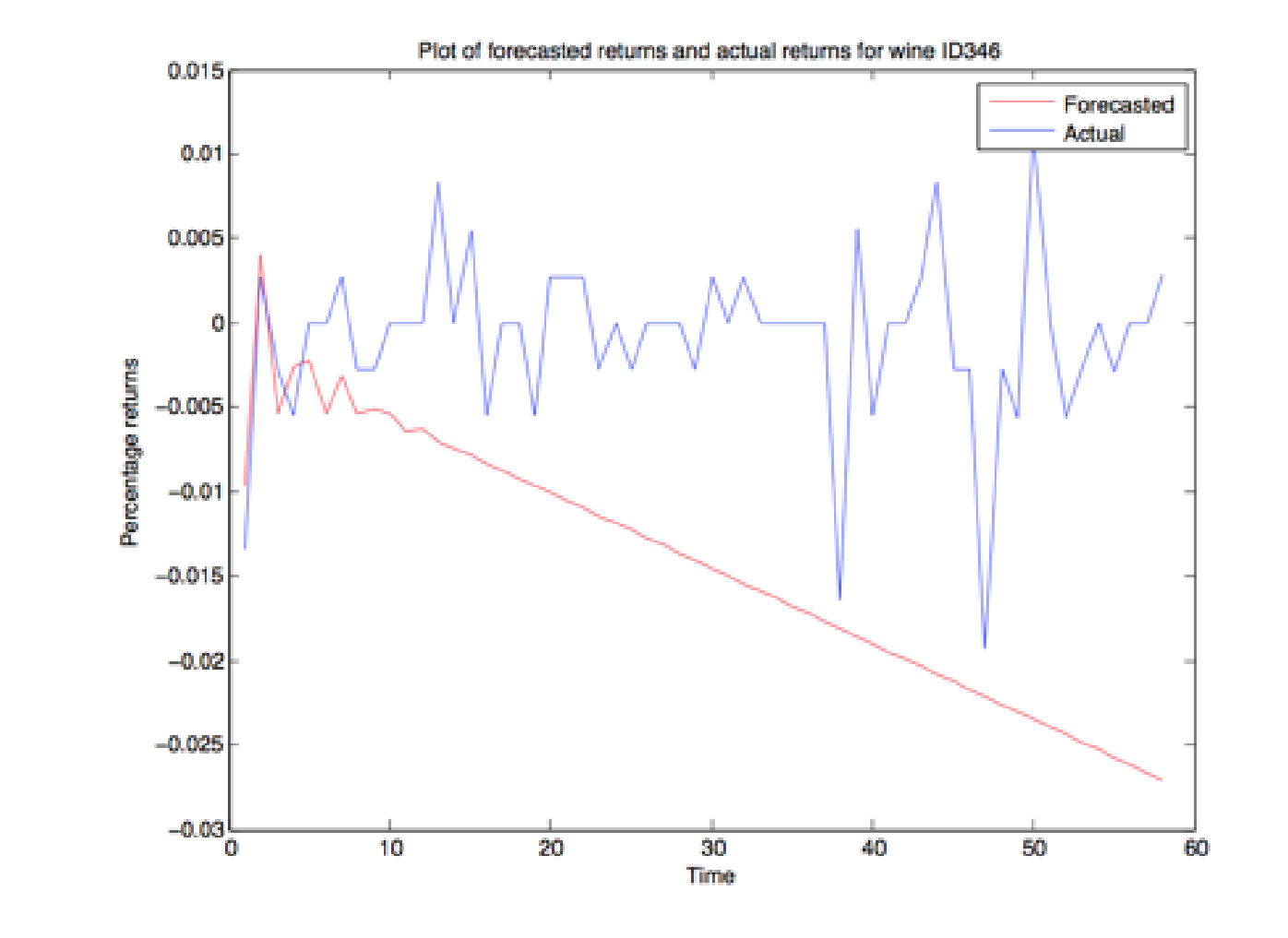
Bayesian Inference (again)



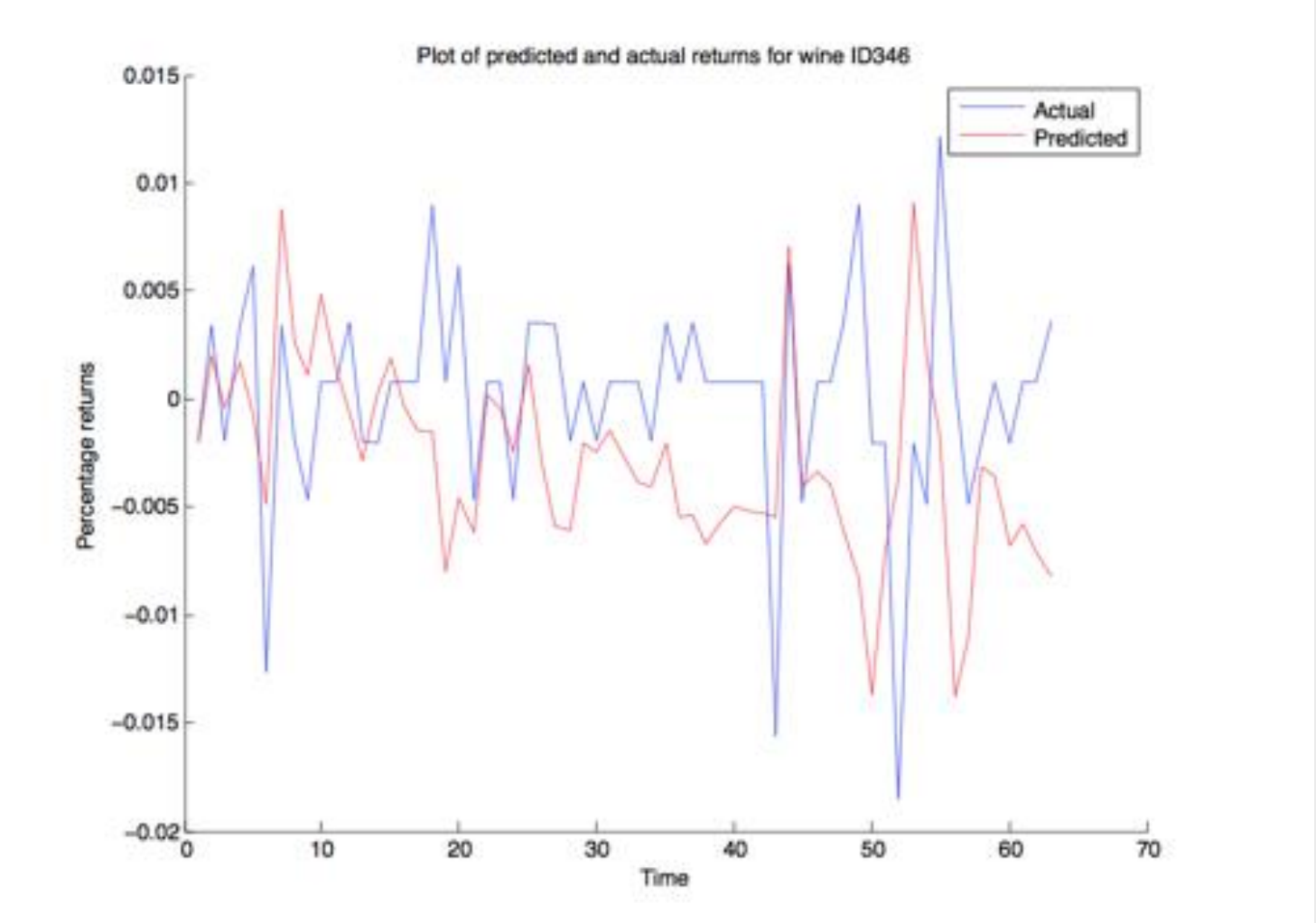
Bayesian Inference (again)



Simple time series forecasting



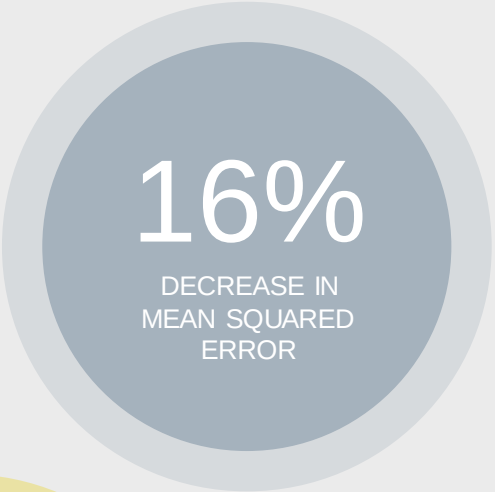
Gaussian Process Regression

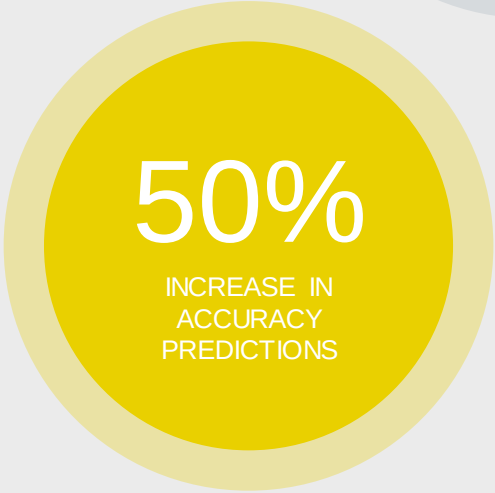


There is a fundamental link between GP regression and Bayesian linear regression

Results

- Liv-Ex 100 wines
- Trained on data from August 2013 - March 2014
- Tested on data from Apr 2014 - July 2014
- 16% decrease in mean squared error
- 50% increase in accuracy in determining positive/negative returns for the next time step





Explanation vs. Prediction (again)

- Higher predictive power with Gaussian Processes
- Explanatory too!
  - Weights in linear regression explains the observed data
  - Gaussian Process regression is essentially Bayesian linear regression with a prior on the weights
  - Gaussian Process regression lets us specify our belief on the relationship between predictors and response

Is all this worth it?

- Significantly improved predictive performance over simpler models
- Plots: learning more about the nature of wine prices
- Using Gaussian Processes we get both an explanatory & predictive model

Thank you for listening!

Any questions?