Modeling U.S. Inflation

Presented by Benjamin Allen
Ben.allen@rice.edu
STAT410 Final Project

Agenda

- Introduction
- Project Overview
- Initial Insights
- Linear Model
- Time Series Models
- Conclusions
- Q&A

What's new with inflation?

- In recent months, Inflation has been everywhere in the news
- COVID-19 and subsequent lockdowns led to disinflation
- Government responded with immense economic aid
- Economic recovery was swift, at the cost of volatile inflation
- Inflation is a general increase in prices in the economy
- High inflation harms consumers and reduces standard of living

Project Overview

- For this project we will model the inflation rate for all goods
- Inflation is measured as the percent change of prices from a year ago versus now
- Main Goal: Create a model able to forecast inflation a year in advance

$$2021 \rightarrow 2022$$
$$\$50 \qquad \$60$$

Inflation Rate

$$r_{Inf} = \frac{60 - 50}{50} = 20\%$$

Project Data

Sourced from U.S. Bureau of Labor Statistics (bls.gov):

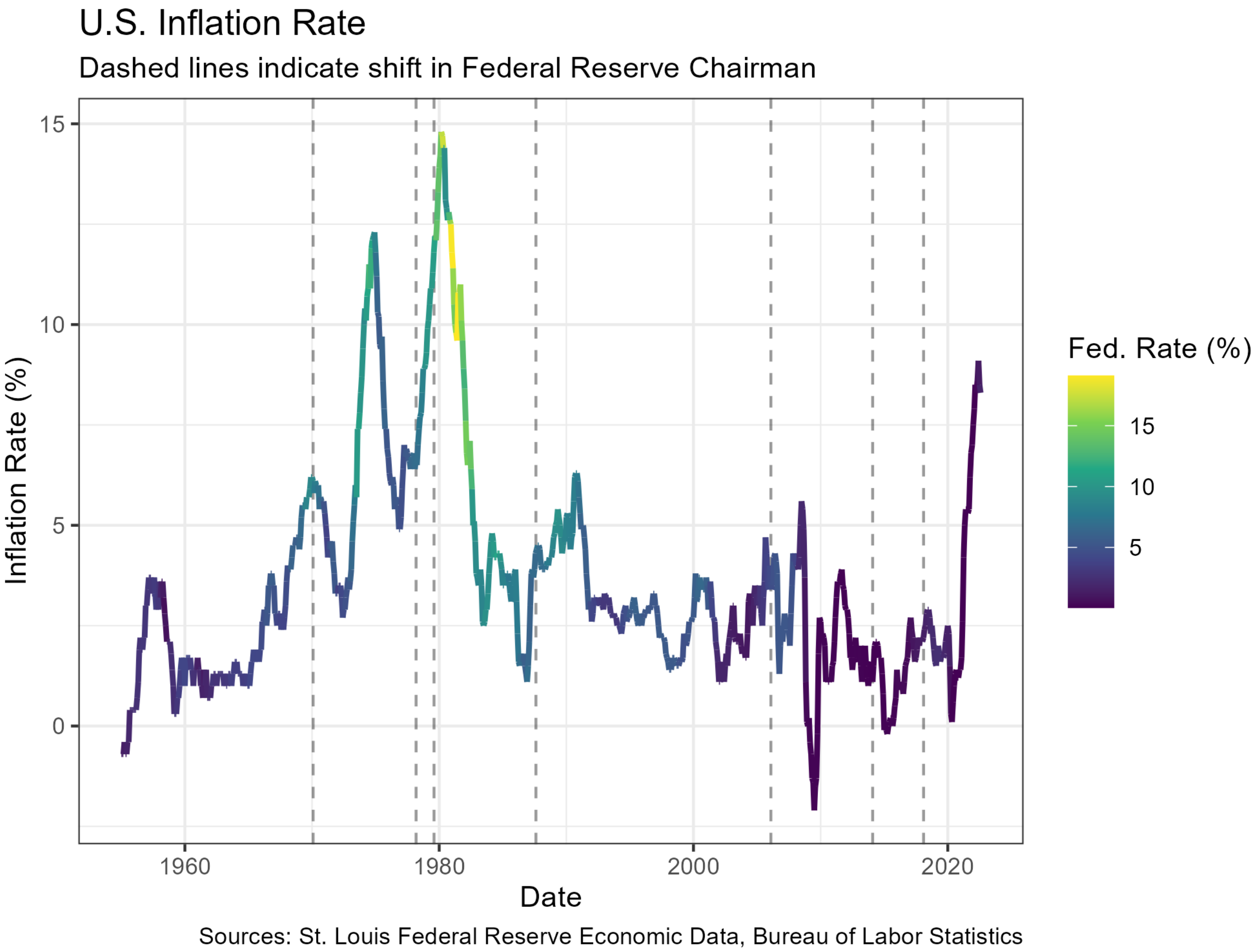
- Inflation Rate for All Goods
- Monthly Unemployment Rate
- WTI Crude Oil Price Index
- All Commodities Price Index

Sourced from St. Louis Federal Reserve (fred.stlouisfed.org)

- Federal Funds Rate
- M2 Monthly Money Supply
- Consumer Sentiment Index (University of Michigan)

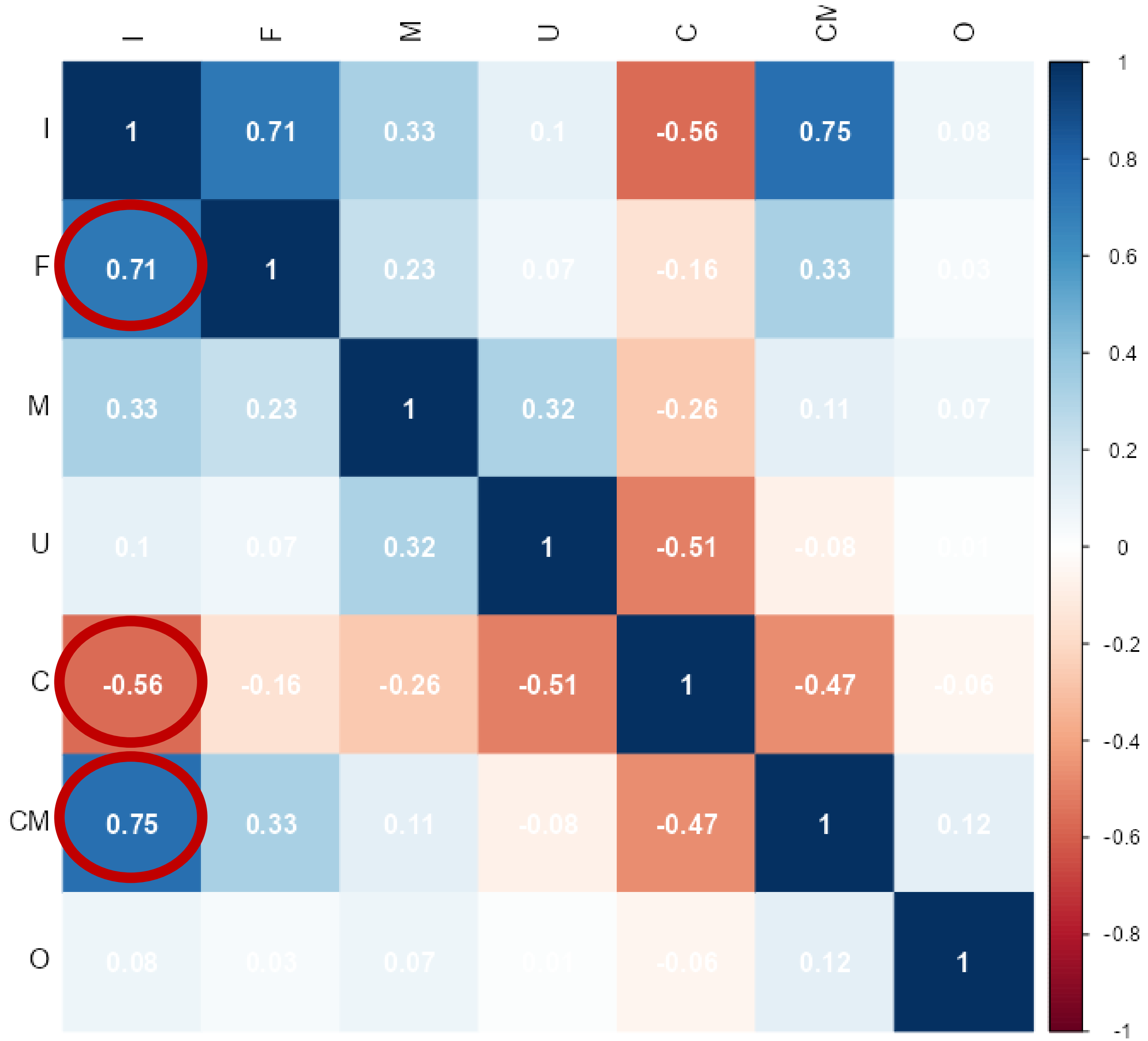
How has inflation changed over time?

- Stable inflation is a recent trend
- The 70s saw high inflation due to large increases in the money supply and an oil embargo
- Brief deflationary period during the Great Recession



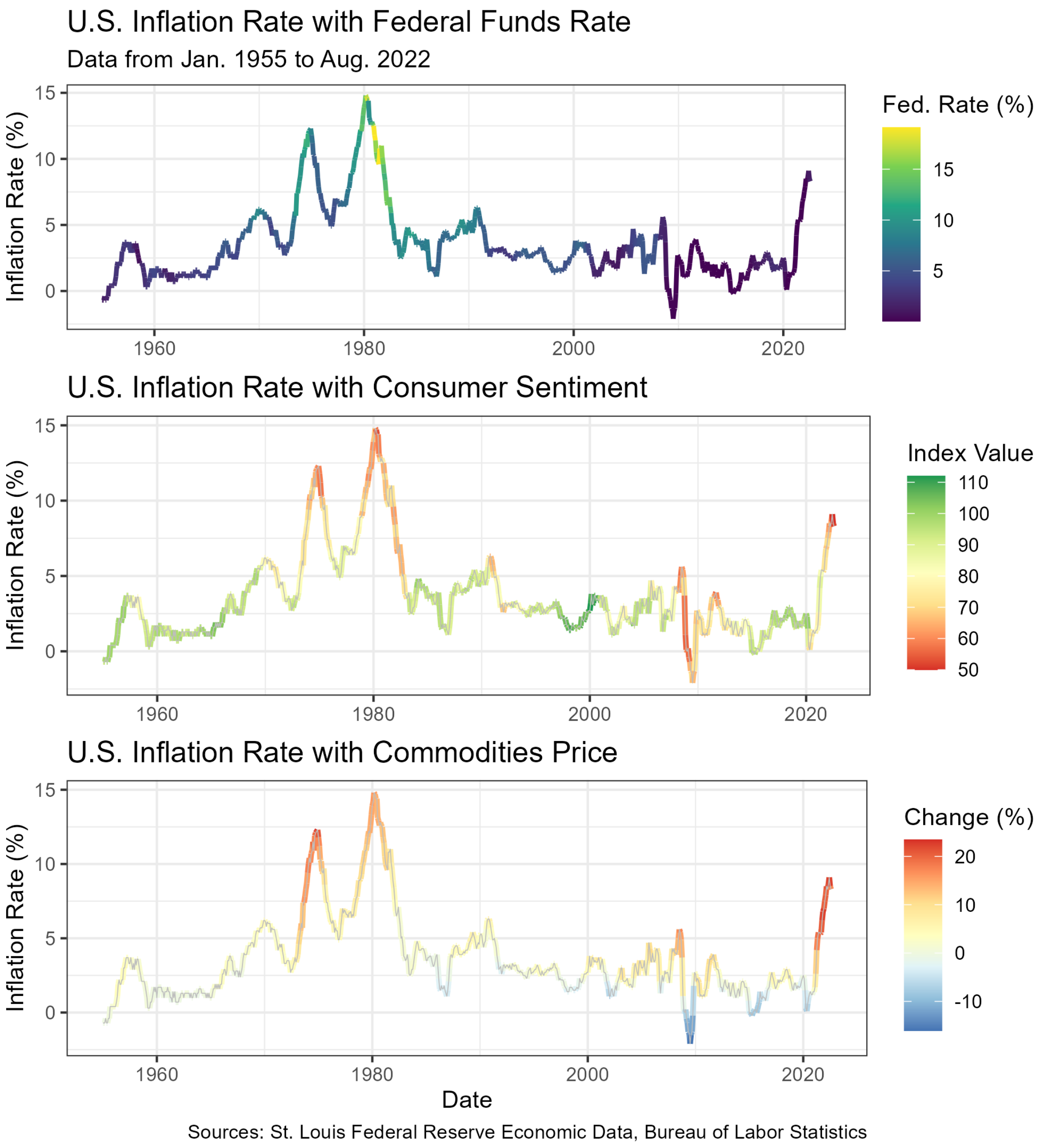
What is related to inflation?

- Federal funds rate and commodity prices are highly positively correlated
- Consumer sentiment is moderately negatively correlated
- Federal funds and commodity prices uniquely related to inflation



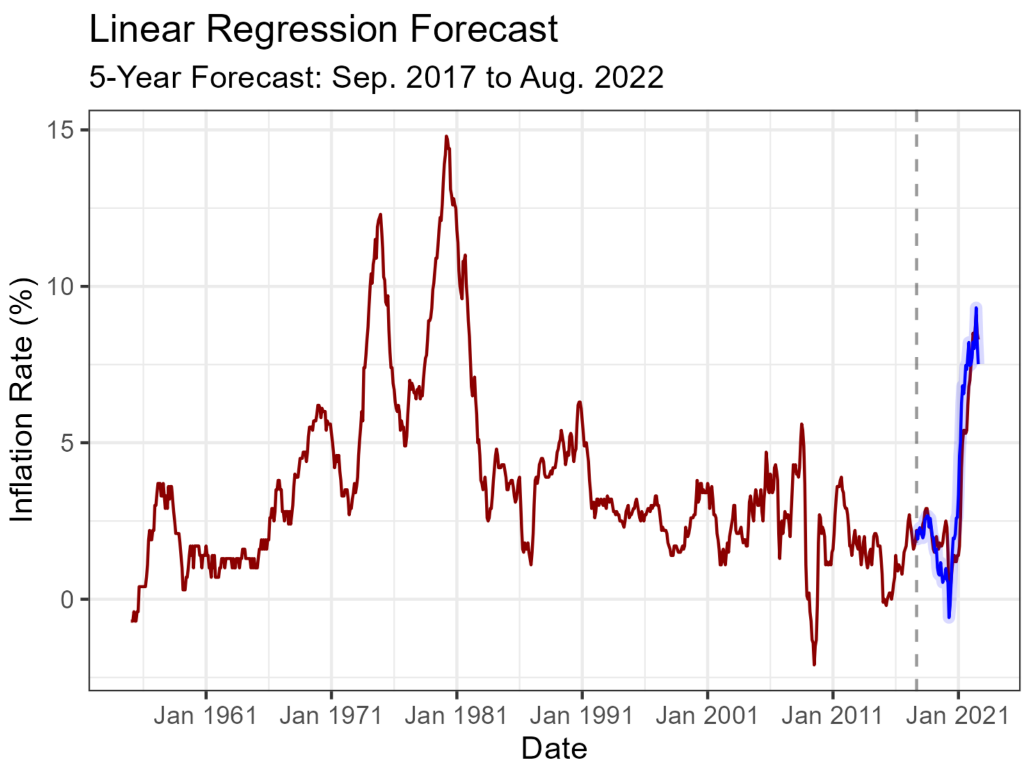
What is related to inflation?

- Federal funds is lower nowadays, but still rises with inflation
- Consumer sentiment seems to drop because of volatile inflation
- Commodity prices and inflation rise and drop together



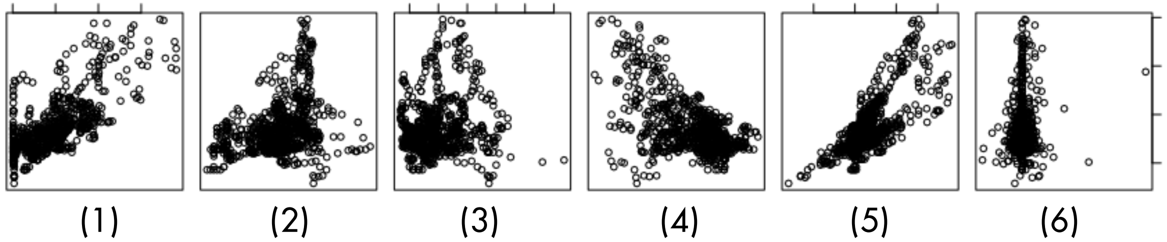
Setting up Linear Model

- Our first model is a multiple linear regression that uses all variables as regressors
- $Adj.\ R^2 = 0.8797$
- Oil price index was the only non-significant variable

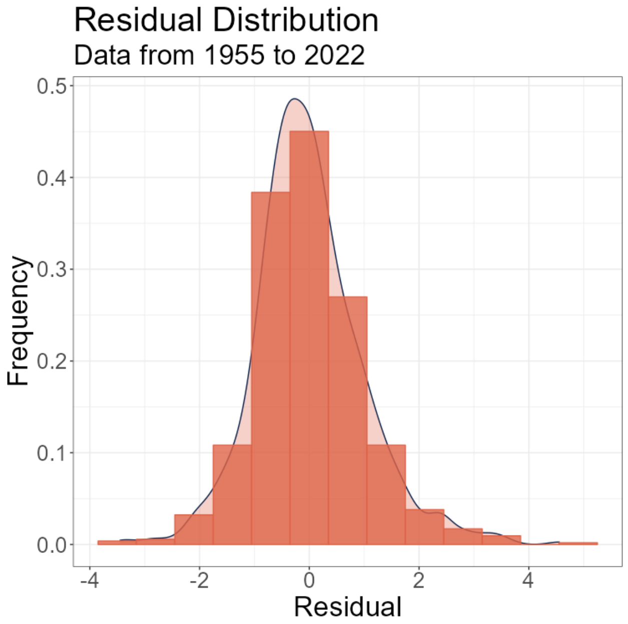


Evaluating the Linear Model

- Linearity?
  - Potentially true for federal funds rate (1), consumer sentiment (4), and commodity prices (5)
  - The rest do not appear linear



- Normal Errors?
  - Histogram shows slight positive skew in residuals
  - Shapiro and Kolmogorov-Smirnov test both rejected

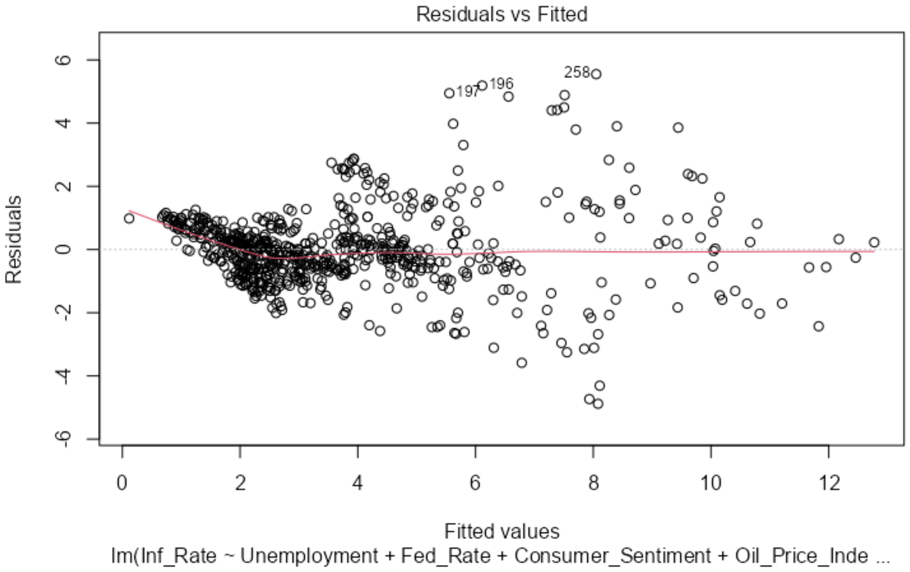



```
        Shapiro-Wilk normality test

data:  full_model$residuals
W = 0.97373, p-value = 2.24e-10
```

```
        One-sample Kolmogorov-Smirnov test

data:  full_model$residuals
D = 0.06745, p-value = 0.003427
alternative hypothesis: two-sided
```

Evaluating the Linear Model

- Homoscedasticity?
  - Looking at scatterplot, variance increases with larger fitted values
  - Corresponds to higher volatility



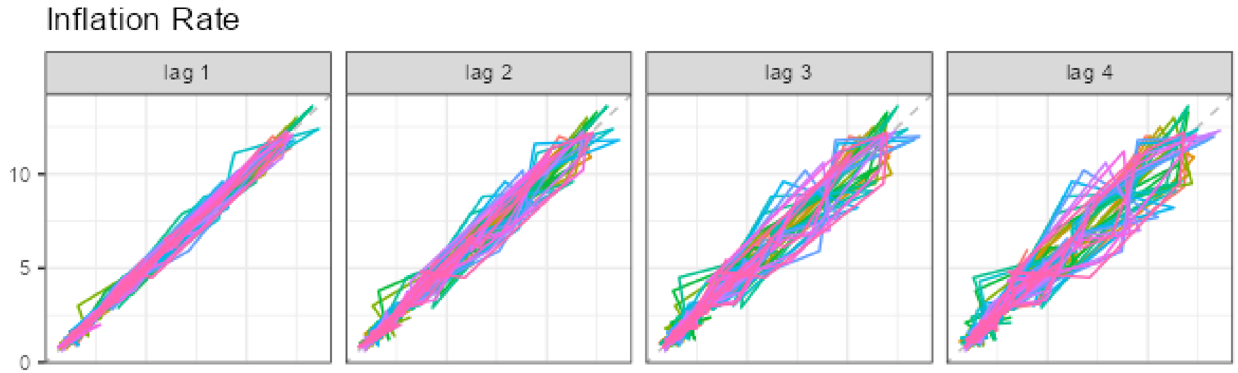
- Uncorrelated Errors?
  - Model has an autocorrelation of ~0.92
  - Durbin-Watson hypothesis firmly rejected for model

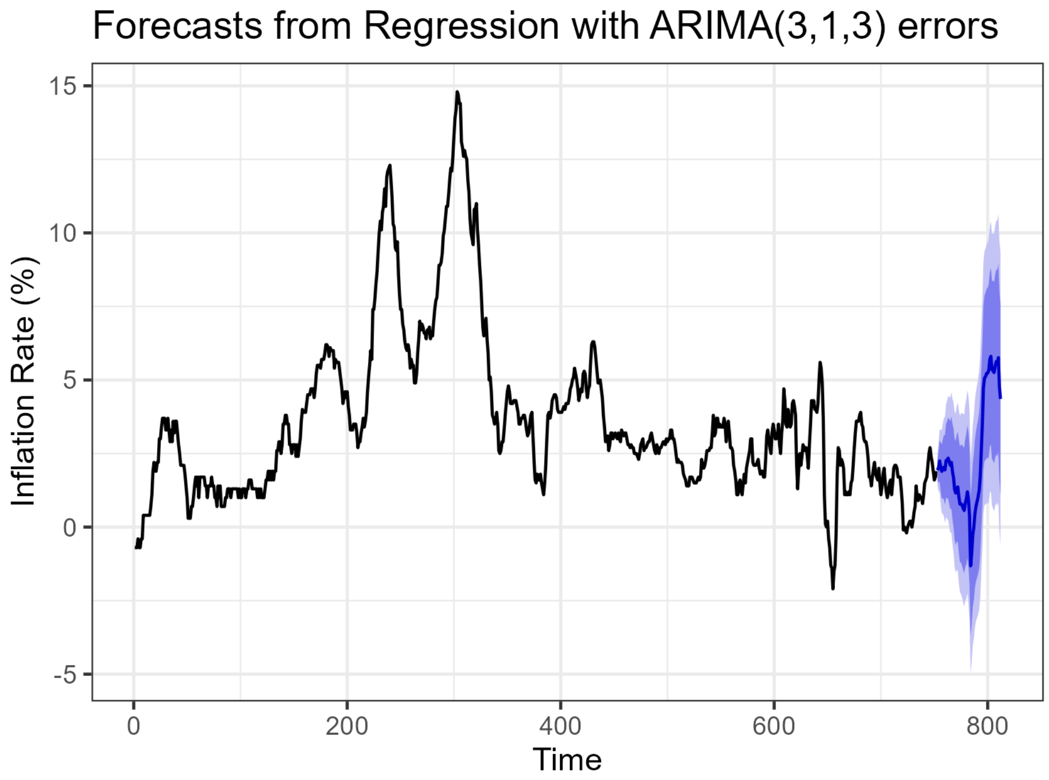
```
 lag Autocorrelation D-W Statistic p-value
   1       0.9221938     0.1545934       0
 Alternative hypothesis: rho != 0
```

Time Series Modeling

How can we model this data better?

- Include autocorrelation into model predictions
- ARIMA models are made for data with autocorrelations
- Tested a model with only autoregression and dynamic models including regressors
- Evaluate based on prediction error from 5-year forecast





Best Results

MAE: 0.948%

MedAE: 0.626%

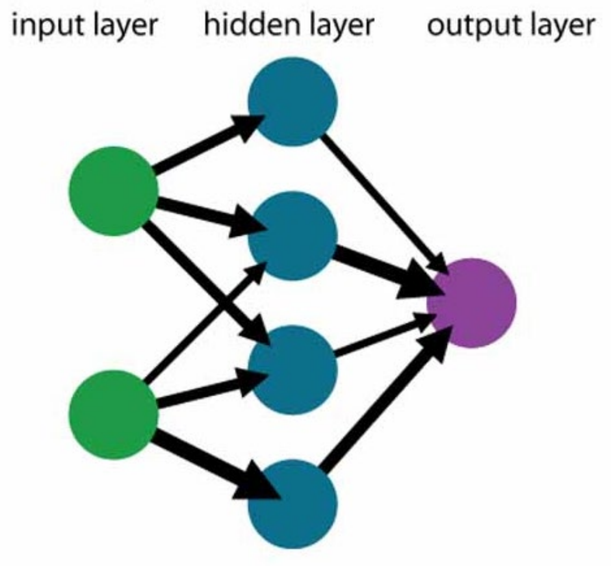
Max Error: 3.94%

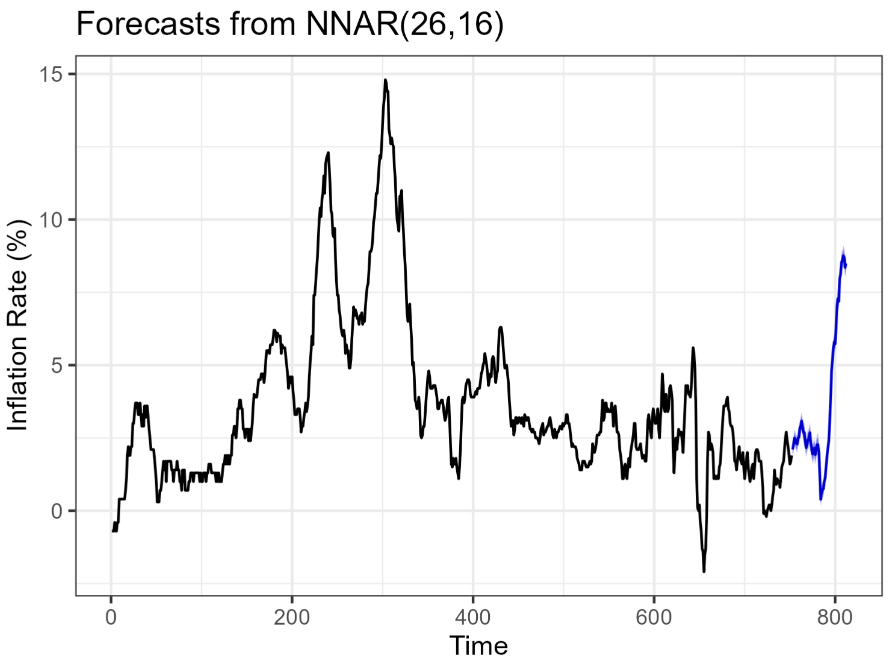
AIC: 212.99

BIC: 273.06

Another Approach: Neural Networks

- Neural networks are better at understanding complex non-linear relationships
- NNAR model incorporates autoregression
- Same testing and evaluation criteria as ARIMA





Best Results

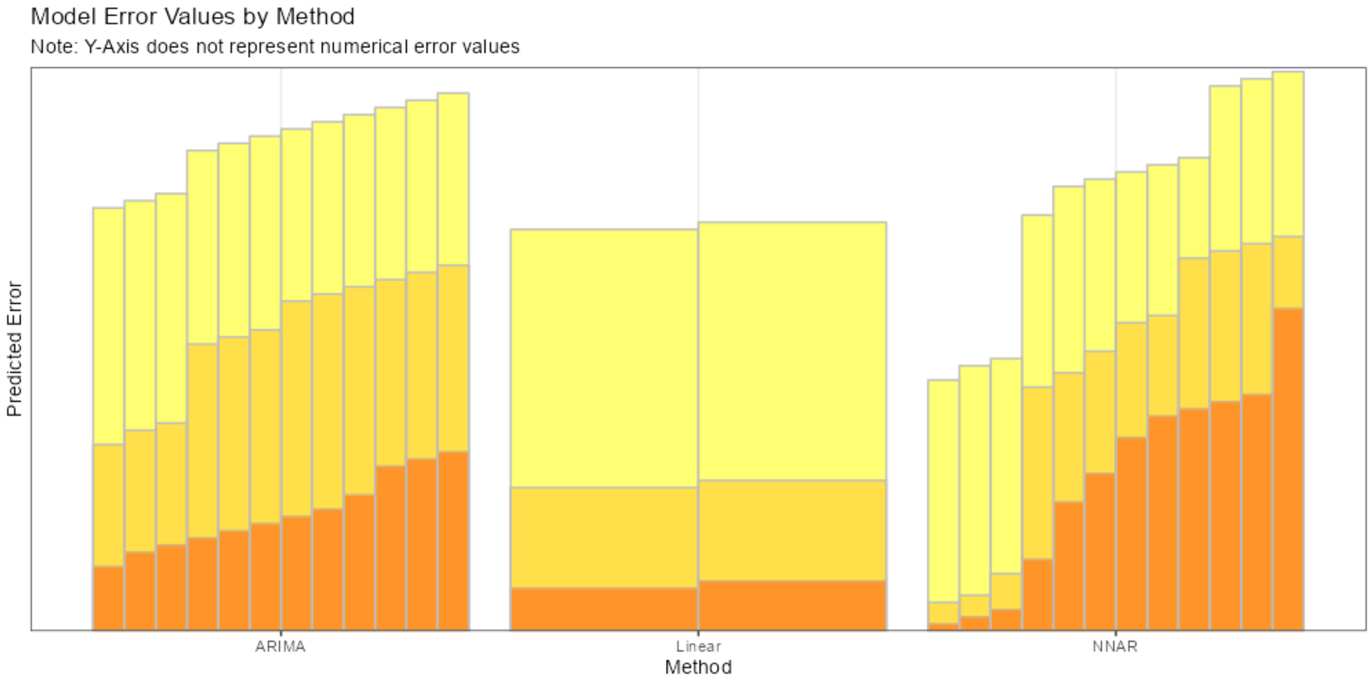
MAE: 0.406%

MedAE: 0.311%

Max Error: 1.33%

Time Series Model Evaluation

- Tested 26 models in total
  - 2 Linear Models
  - 12 ARIMA Models
  - 12 NNAR Models
- Evaluated based on
  - Predicted MAE
  - Predicted Median AE
  - Predicted Max Error
  - AIC (Linear and ARIMA)
  - BIC (Linear and ARIMA)



- ARIMA models had consistently decent performance
- Linear models had good MAEs and Median AEs, but high max errors
- NNAR models had varied performance but top models did the best overall

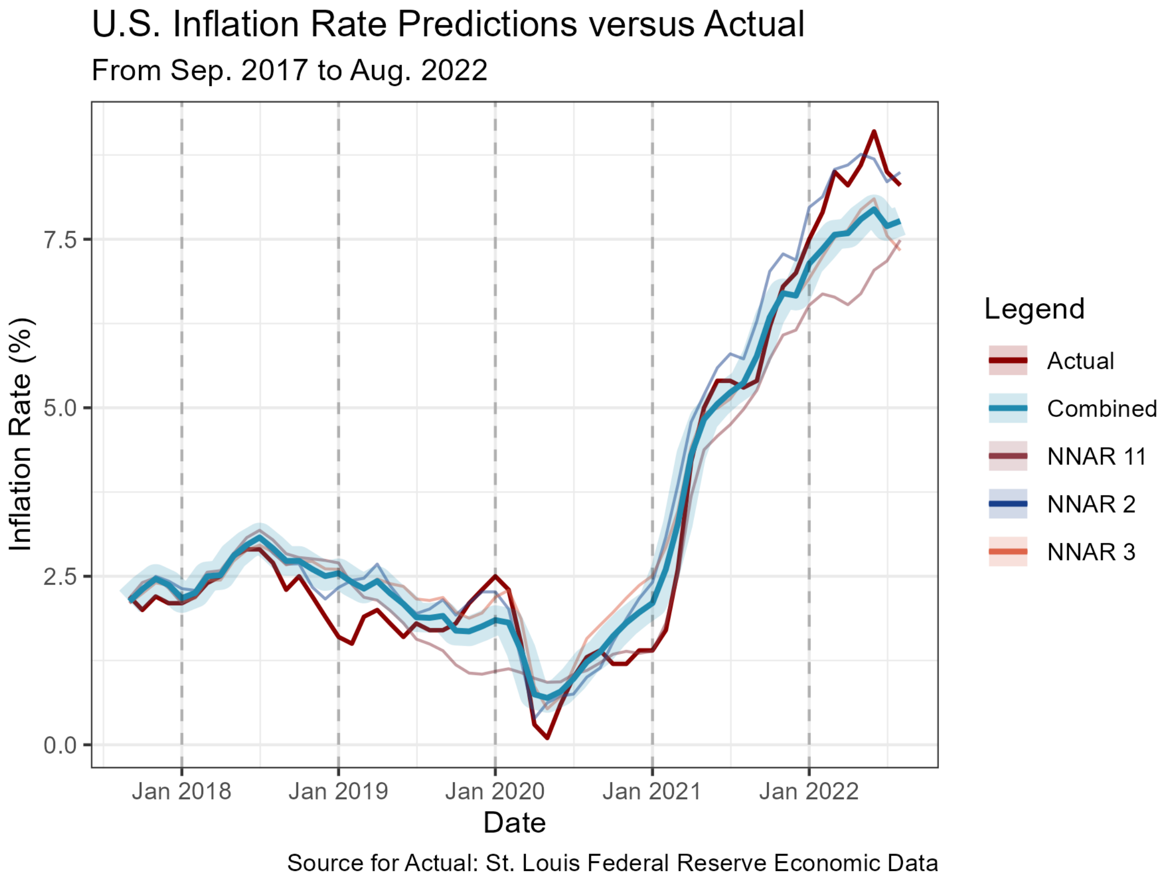
Final Model Selection

Selected Models:

- NNAR with all regressors (NNAR 2)
- NNAR with all except oil prices (NNAR 3)
- NNAR using only commodity prices (NNAR 11)

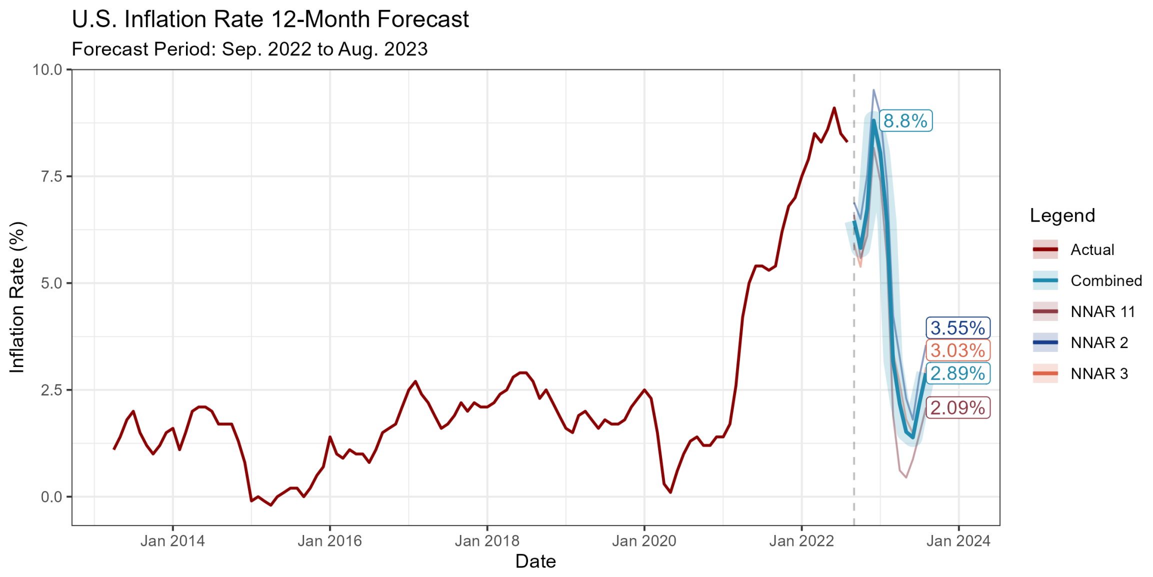
Combined model averages predictions of our three selected models

MAE: 0.349, Max Error: 0.944



12-Month Inflation Forecast

- Finally, we can forecast inflation into the future
- Requires forecasting regressors as well
- August 2023 Projections:
  - NNAR 2: 3.55%
  - NNAR 3: 3.03%
  - NNAR 11: 2.09%
  - Combined Model: 2.89%



All models expect inflation to drop significantly by March 2023

Conclusions

- Inflation is expected to drop substantially through 2023
- Commodity prices are especially useful for predicting inflation but tend to underestimate
- Consumer sentiment drops in times of volatile inflation

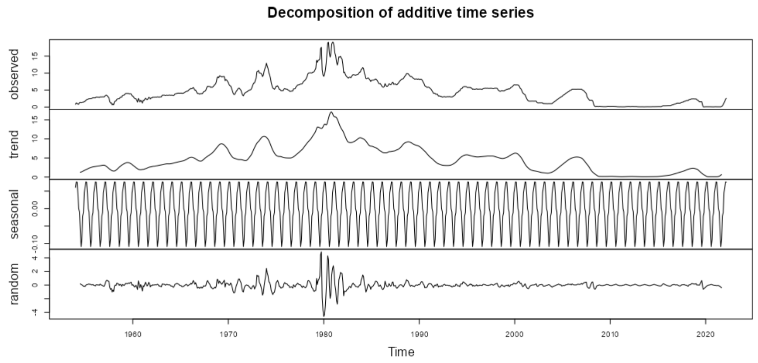
Questions

Appendix 1 – Data Cleaning Summary

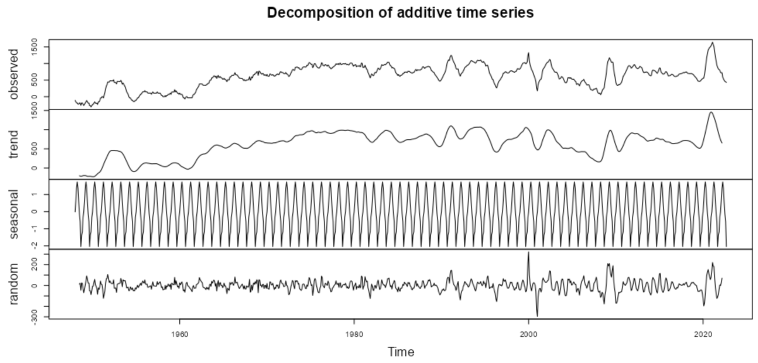
- Transformed most variables to be 12-month percent changes
- Some data was originally in a wide format, so it needed to be transformed to a long format
- Consumer Sentiment data featured a lot of missing data before 1978, so this was imputed using linear interpolation
- No other variables were missing data
- Dates separated into month and year columns
- Finally, data was combined into one data frame
- Log-transforming the data was tried, however it did not improve model performance

Appendix 2a – Time Series Decompositions

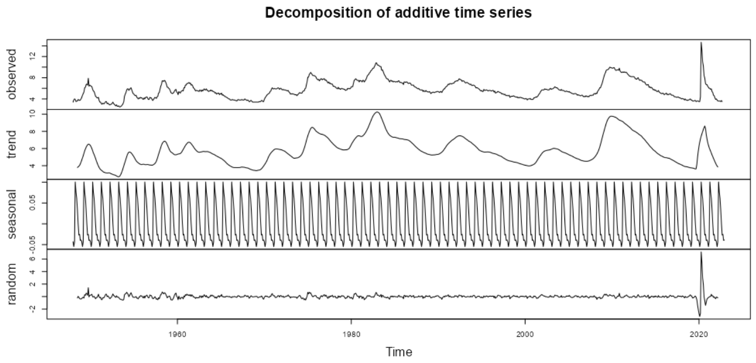
Federal Funds



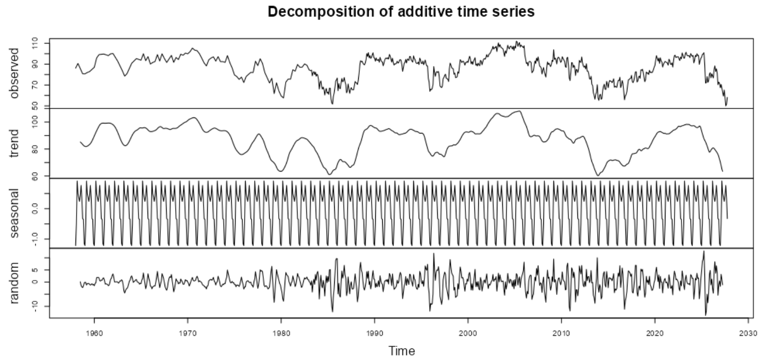
Money Supply



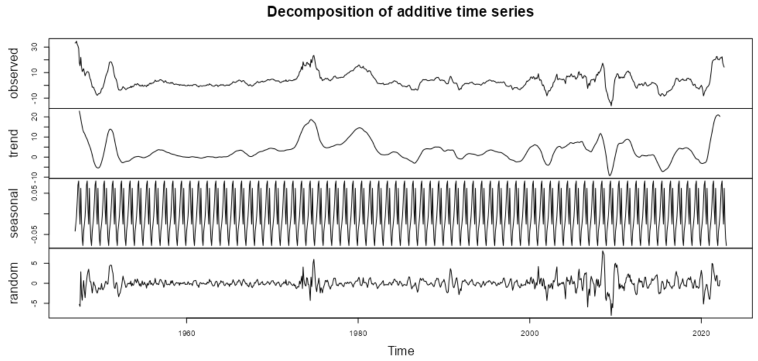
Unemployment



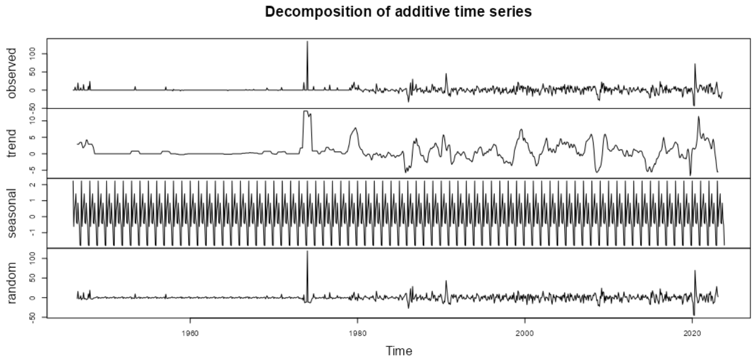
Consumer Sentiment



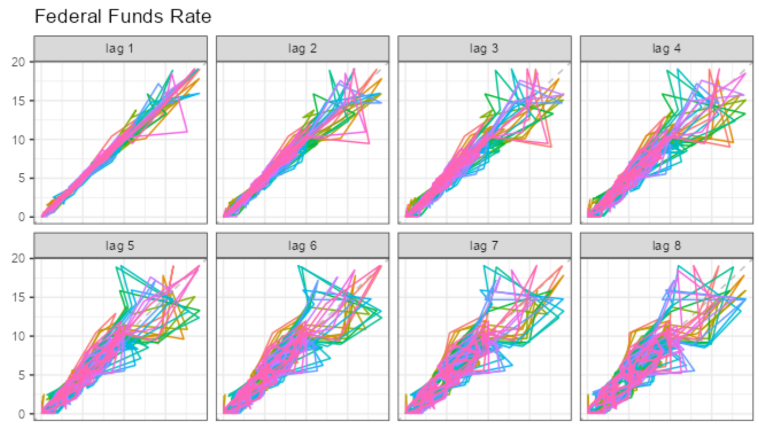
Commodities Price Index

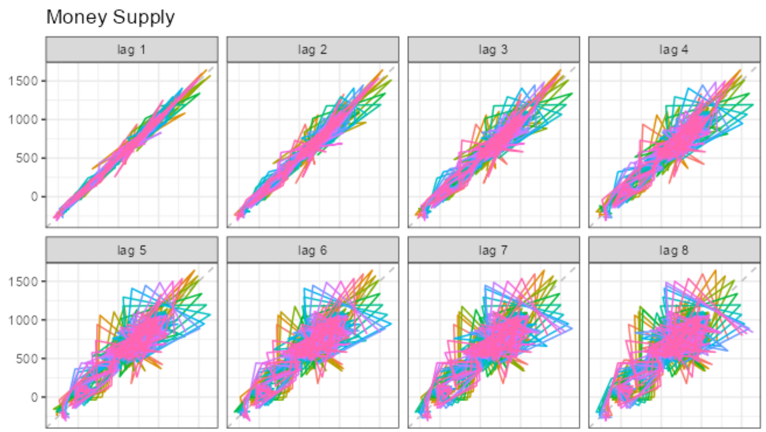


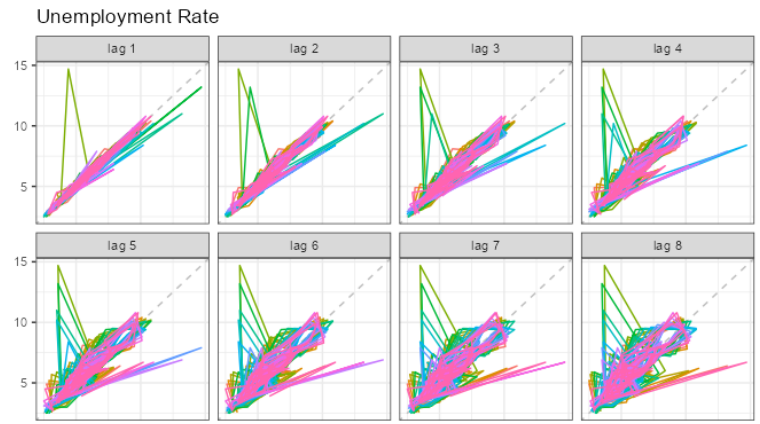
Oil Price Index

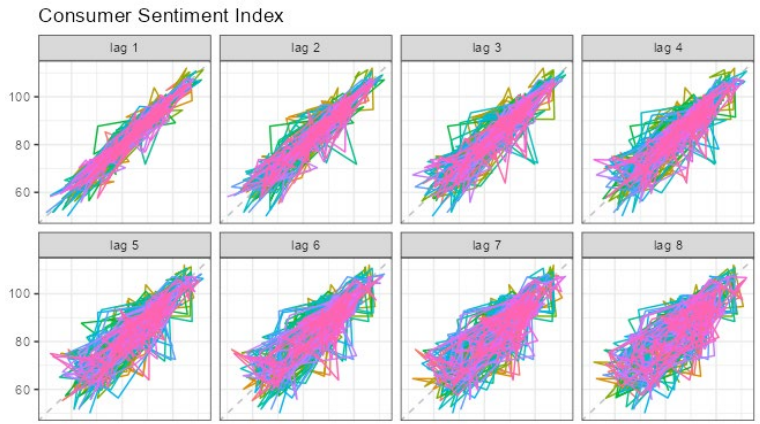


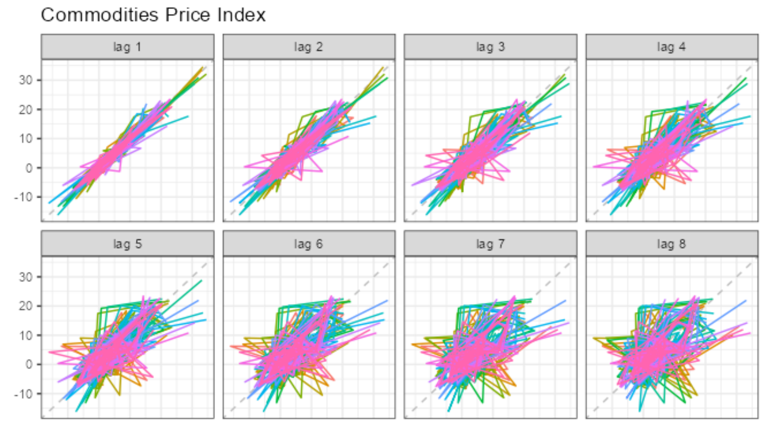
Appendix 2b – Lag Plots

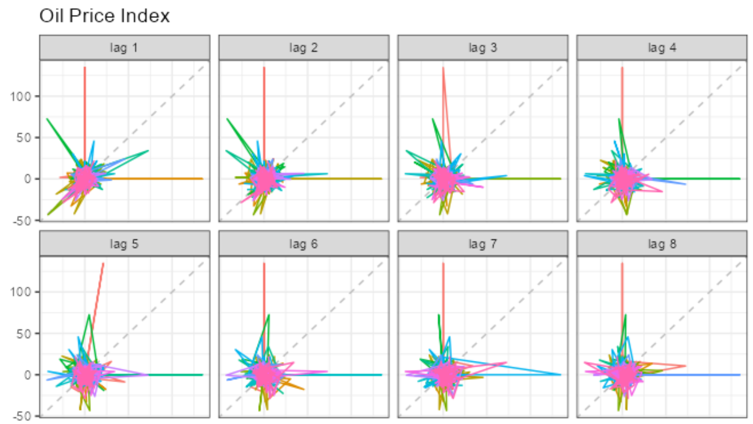




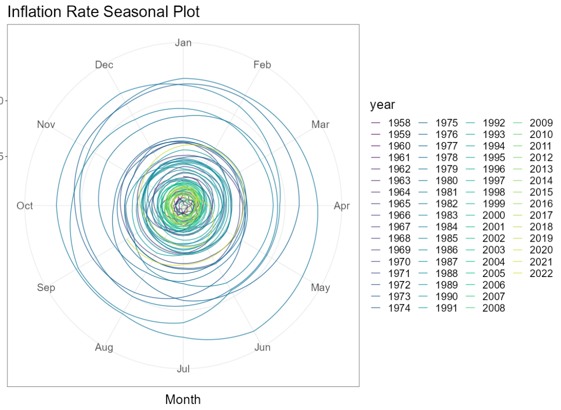


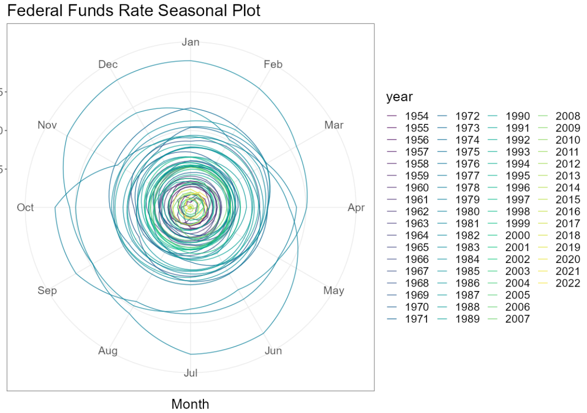


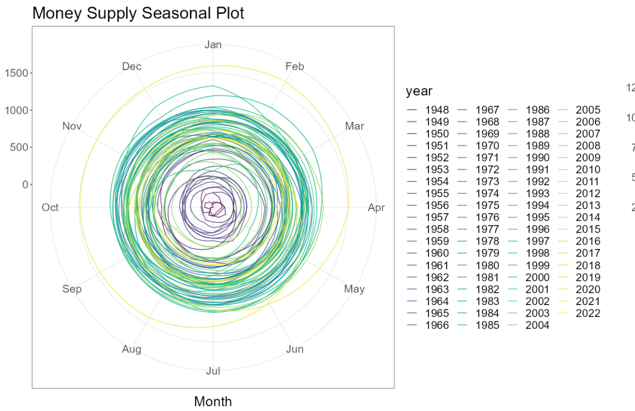


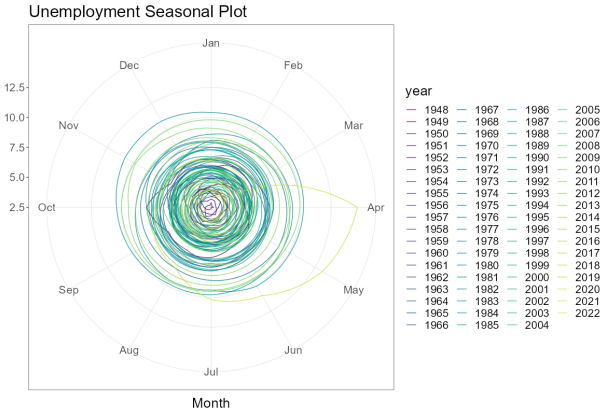


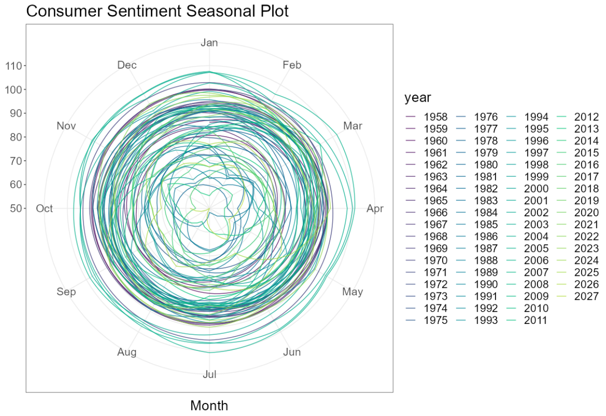
Appendix 2c – Season Plots

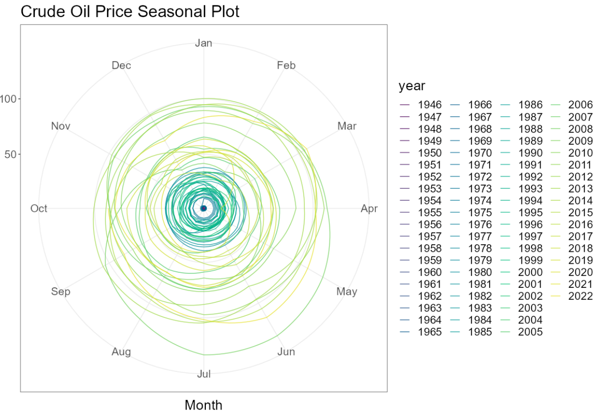


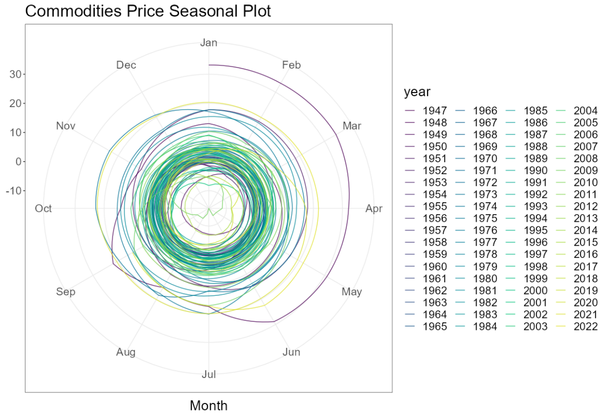




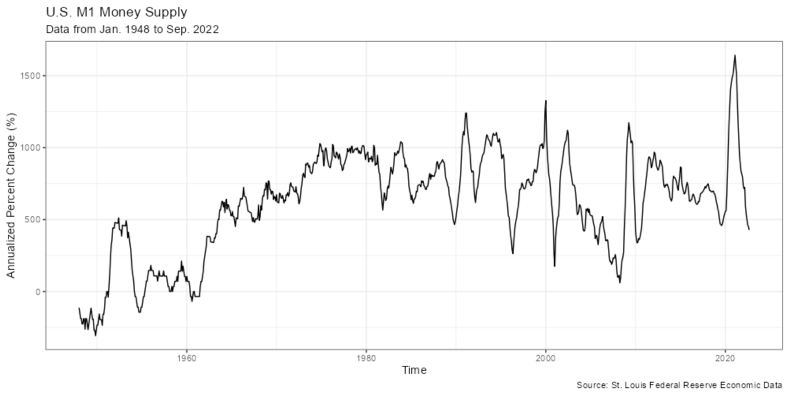


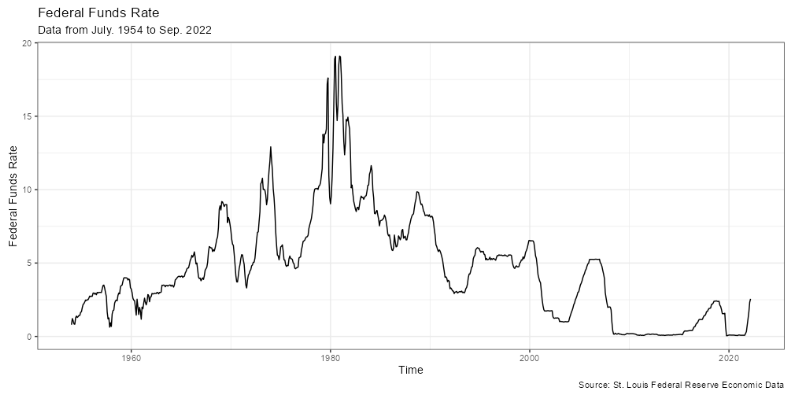


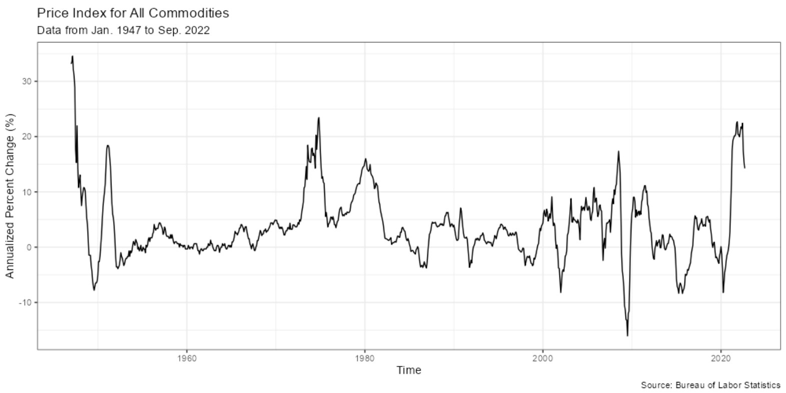


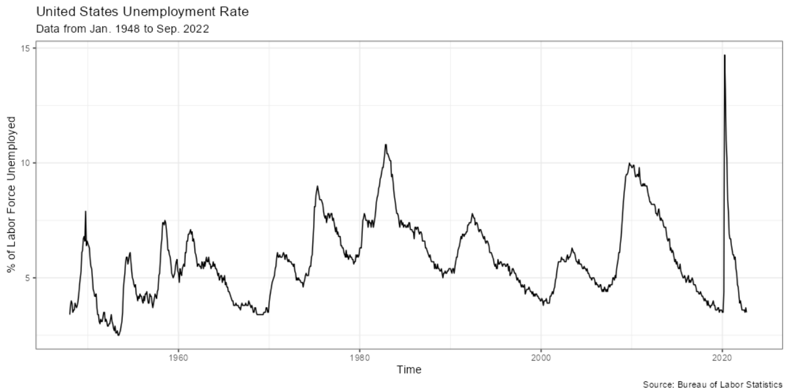


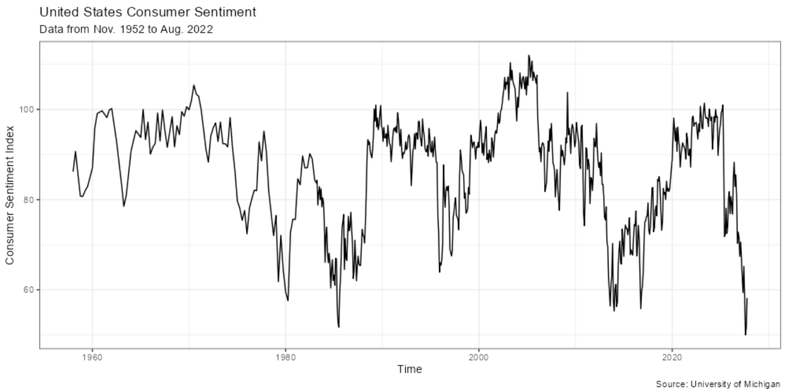
Appendix 3 – Line plots of variables

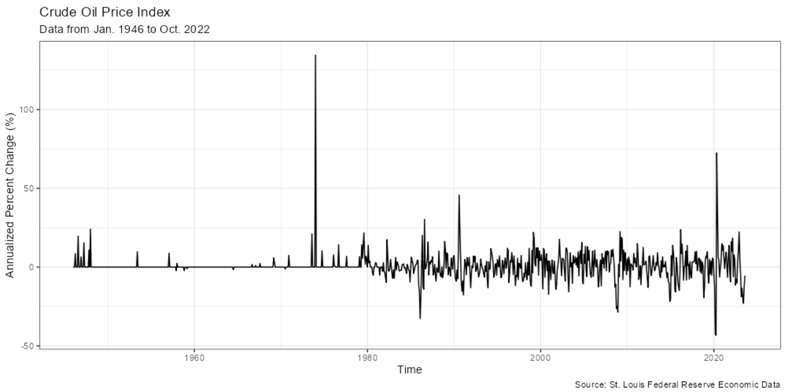






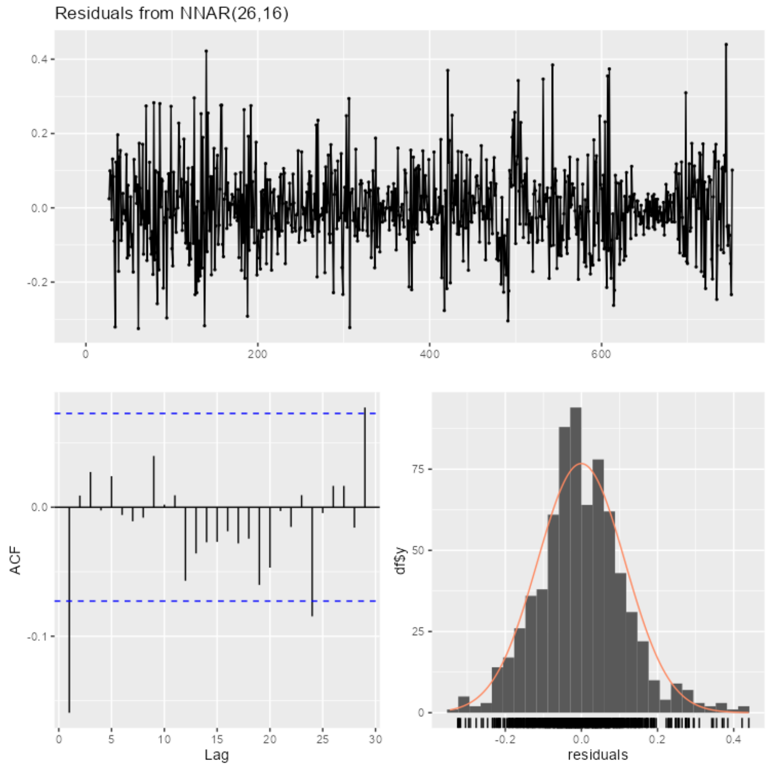




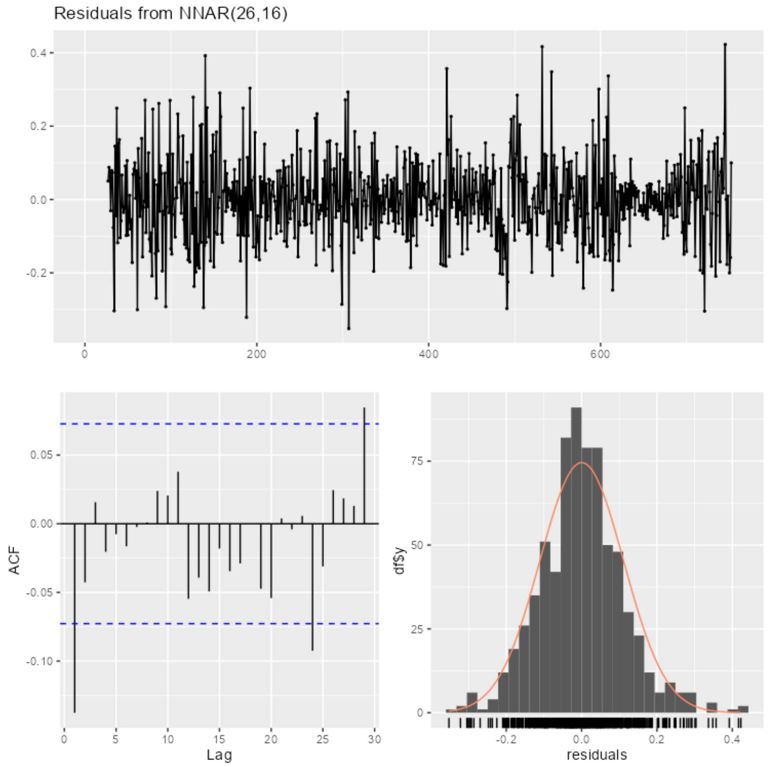


Appendix 4 – Residual Plots

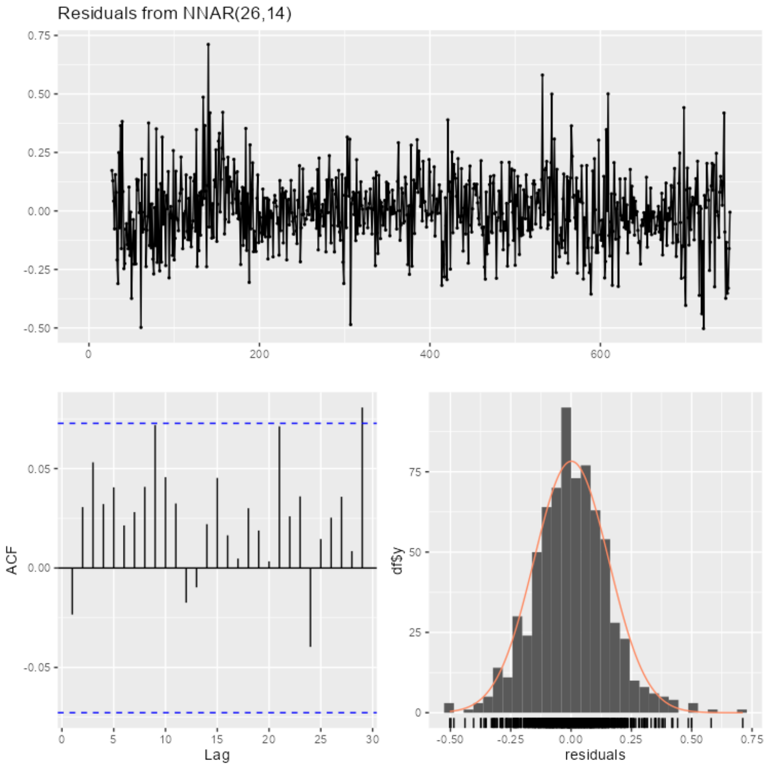
NNAR 2



NNAR 3

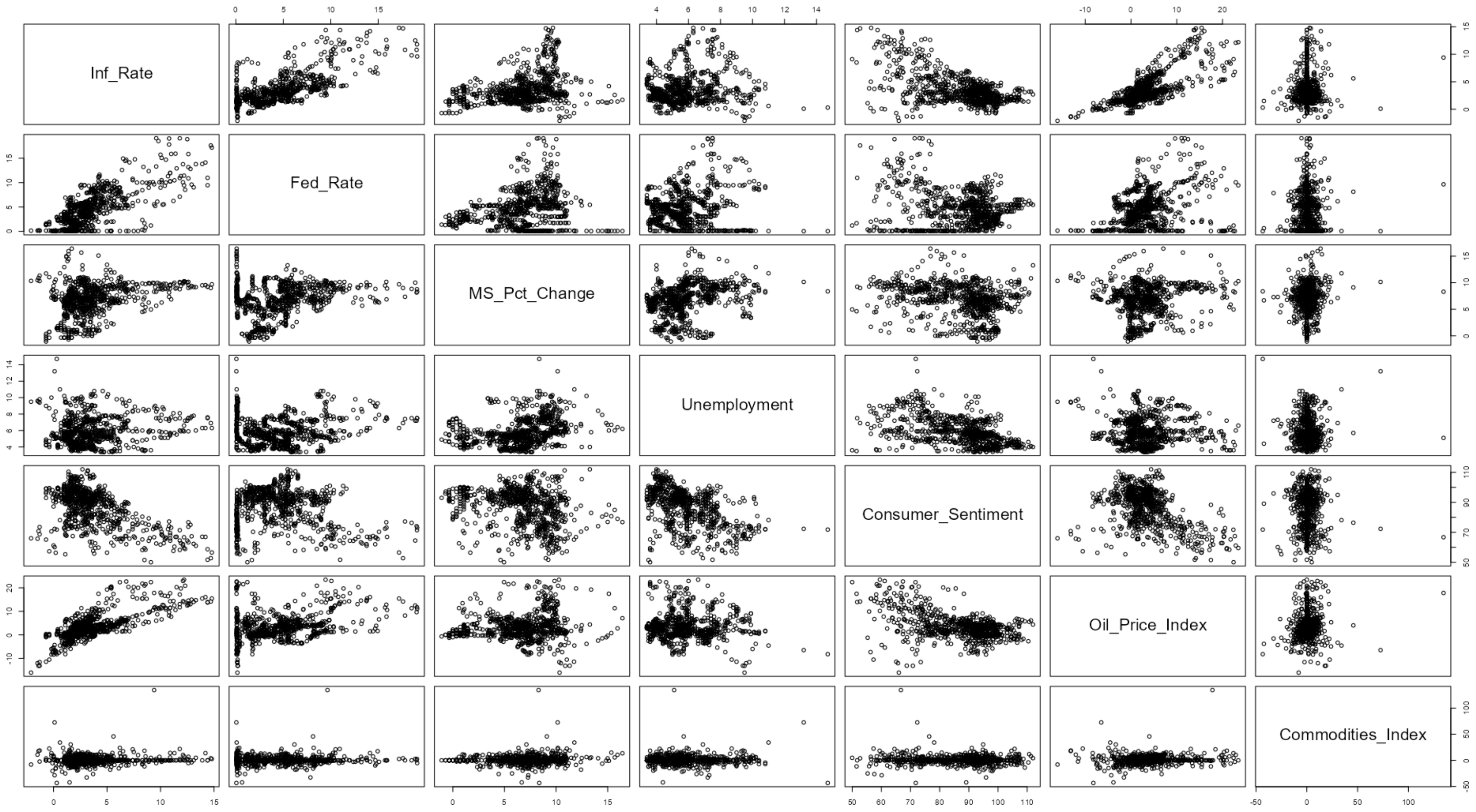


NNAR 11



Appendix 5 – Full Pair Plot

Note:
Oil price and commodities were mislabeled and should be switched



Appendix 6 – Modeling Methodology

- Linear model
  - Full model first tested
  - Stepwise regression used which removed commodities index, but performance did not dramatically improve
- ARIMA model
  - Model using only autoregression tested
  - Next, models using only one of the regressors were tested
  - Finally, a few combinations of regressors were tested
- NNAR model
  - Same combinations of regressors as specified for ARIMA models
- Combination model
  - Outside research has shown that averaging predictions from various models can improve predictions
  - 3 of the NNAR models were distinctly better than the rest, so these were averaged

Appendix 7 – Forecasting Methodology

- Given that our models used other variables to forecast, we need to forecast these variables themselves
- Quality of the inflation forecast is built on the quality of the other forecasts
- Federal funds rate is estimated to be 4.5% in 2023, so forecast was a linear interpolation
- Money supply, unemployment, and oil prices were forecasted similarly, but money supply and oil prices had 12-month percent changes calculated from linear interpolation
- Consumer sentiment and commodities were forecasted using a NNAR model based on autoregression

Appendix 8 – Full Inflation Forecasts

	date	nfit2	nfit3	nfit11	combination
1	2022-09-01	6.887261	5.902433	6.5980612	6.462585
2	2022-10-01	6.498829	5.380590	5.5827531	5.820724
3	2022-11-01	7.527604	6.559016	6.1075416	6.731387
4	2022-12-01	9.522352	8.723998	8.1684028	8.804918
5	2023-01-01	8.942776	7.824529	7.3885141	8.051940
6	2023-02-01	7.331693	6.358645	5.5461933	6.412177
7	2023-03-01	4.245423	3.388874	1.9045054	3.179601
8	2023-04-01	3.264035	2.637718	0.6157841	2.172512
9	2023-05-01	2.296421	1.811976	0.4470117	1.518470
10	2023-06-01	1.803327	1.481554	0.8748433	1.386575
11	2023-07-01	2.796937	2.269053	1.4537536	2.173248
12	2023-08-01	3.553174	3.033187	2.0868727	2.891078

Appendix 9 – Model Evaluation Results

	method	variables	predicted_MAE	predicted_MedAE	predicted_MaxError	AIC	BIC
27	combination	average	0.346296306520047	0.289183347415921	0.943672070814852	NA	NA
17	nnetar	FMUCO	0.406340385836921	0.311019459671546	1.32569265421826	NA	NA
16	nnetar	FMUCOR	0.408426042587719	0.327683381196413	1.30387030275022	NA	NA
25	nnetar	O	0.457865751852488	0.310473326137046	1.34827116720475	NA	NA
1	linear	FMUCOR	0.770997342582075	0.455922374680332	2.88746803555719	2128.08423284797	2165.06612343957
2	linear	FCOR	0.772374554057209	0.431818809414066	2.92510111442482	2126.70350816722	2159.06266243487
13	arima	O	0.936240166694307	0.61059340990341	4.02558005956096	231.237709106058	268.208954320172
4	arima	FMUCOR	0.9482209322149	0.626480452376875	3.94415264914671	212.985037528914	273.063311001847
5	arima	FMUCO	0.956261269100593	0.630671243840079	3.98551948507234	212.325974160765	267.782841981935