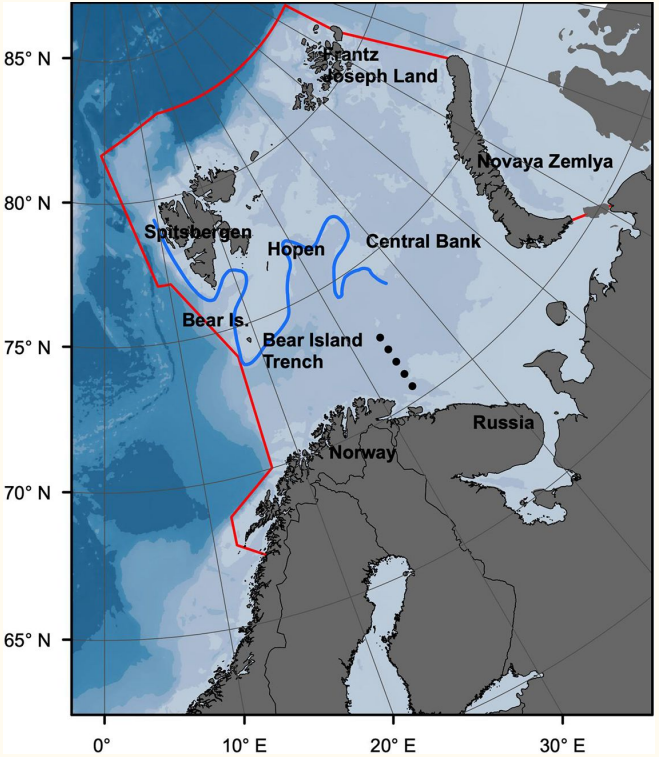
DS 5510 GMRI

https://norseaco.no/pollack/

Michael Massone, Josh Nougaret, Ned Hallahan

Context and Question

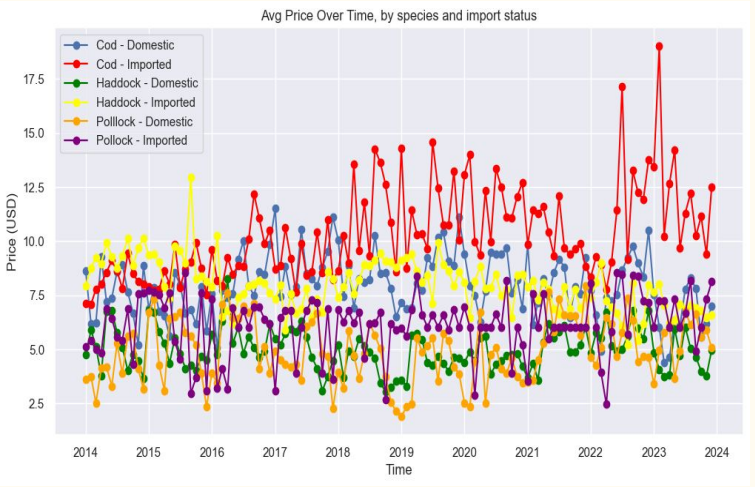
- Since 2022, the Barents Sea fisheries (Norway, Russia, Iceland) have reduced Cod and Haddock quotas by 20% each year in response to declining fish stocks.
- Our goal will be to see if recent changes in the Barents Sea region impact prices in Maine for Domestic landings of Pollock.
- By focusing on a specific import region and species we hope to provide insight to stakeholders and policymakers.

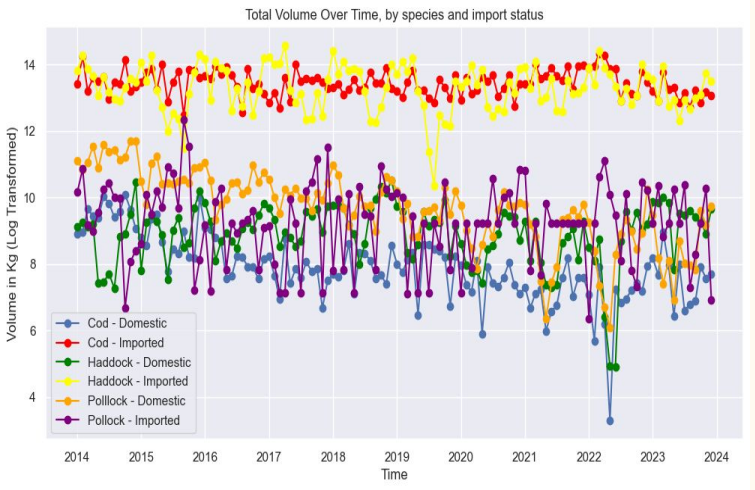


https://www.frontiersin.org/articles/10.3389/fmars.2021.732637

What we did

- Gathered data from the NOAA Fisheries Database, Portland Fish Exchange, and Consumer Price Index (CPI).
- Performed EDA and selected a timeframe from 2014 to 2024, following the collapse of the Northeast Groundfish Fishery.
- Preprocessing steps included data cleaning, price adjustment, visualization, and feature selection.
- We then applied various modeling techniques including univariate, multivariate, PCR, LASSO, and Ridge Regression.





Statmodel Summary report for the OLS Regression of our Features Matrix (X) on the Target Vector (y)

- This summary demonstrates that the Ordinary Least Square regression model explains about 52% of the total variance in the domestic Pollock value.
- The Adjusted R-squared value indicates that some of our variables are not contributing to the overall score, since R-squared tends to increase with more features.
- Imported Cod and Haddock prices do seem to have a significant effect on Pollock prices. The p-values of the F-Statistic for these features is quite low indicating that the regression model is likely producing a significant result.

OLS Regression Results

Dep. Variable:	Pollock_Domestic_USD	R-squared:	0.518
Model:	OLS	Adj. R-squared:	0.459
Method:	Least Squares	F-statistic:	8.769
Date:	Sun, 21 Apr 2024	Prob (F-statistic):	6.77e-12
Time:	14:21:53	Log-Likelihood:	-168.12
No. Observations:	120	AIC:	364.2
Df Residuals:	106	BIC:	403.3
Df Model:	13		
Covariance Type:	nonrobust		

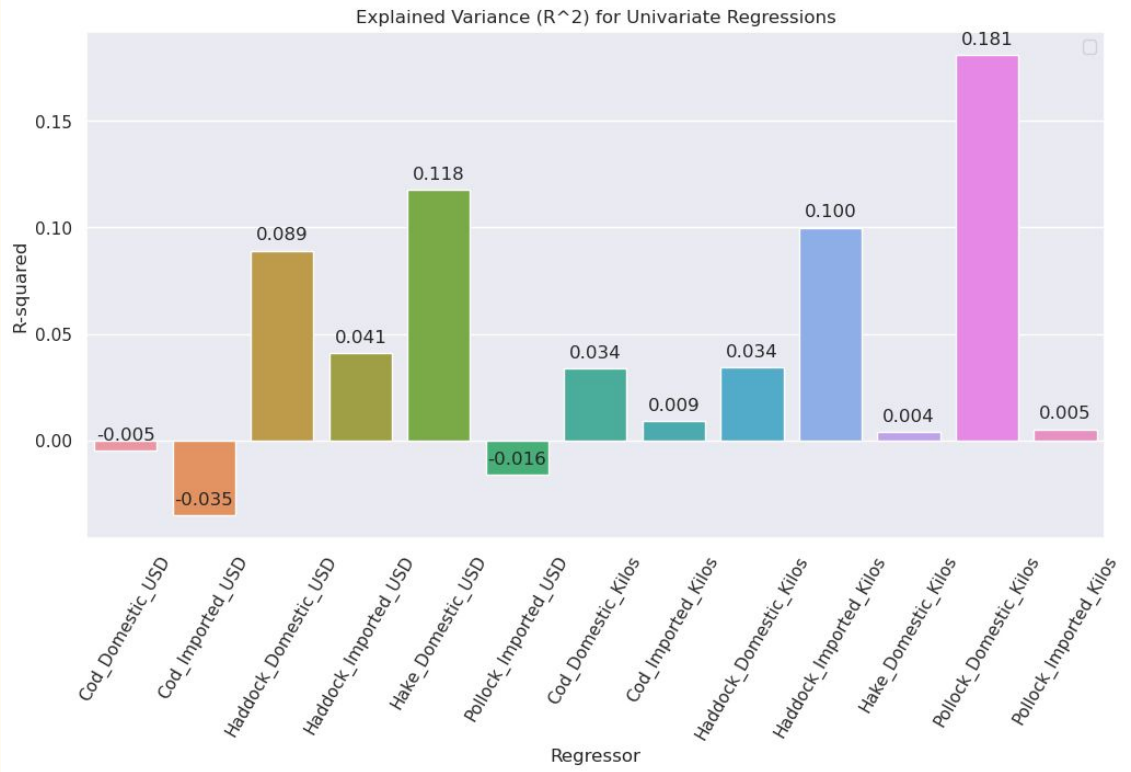
	coef	std err	t	P>|t|	[0.025	0.975]
const	4.9085	0.095	51.450	0.000	4.719	5.098
Cod_Domestic_USD	0.0087	0.114	0.077	0.939	-0.217	0.234
Cod_Imported_USD	-0.3615	0.127	-2.855	0.005	-0.612	-0.110
Haddock_Domestic_USD	0.2617	0.124	2.111	0.037	0.016	0.507
Haddock_Imported_USD	-0.2267	0.119	-1.910	0.059	-0.462	0.009
Hake_Domestic_USD	0.4876	0.121	4.030	0.000	0.248	0.728
Pollock_Imported_USD	0.1238	0.111	1.113	0.268	-0.097	0.344
Cod_Domestic_Kilos	0.1381	0.161	0.857	0.393	-0.181	0.458
Cod_Imported_Kilos	0.0979	0.110	0.893	0.374	-0.120	0.315
Haddock_Domestic_Kilos	0.0818	0.123	0.663	0.509	-0.163	0.326
Haddock_Imported_Kilos	-0.5231	0.128	-4.099	0.000	-0.776	-0.270
Hake_Domestic_Kilos	0.4379	0.142	3.091	0.003	0.157	0.719
Pollock_Domestic_Kilos	-0.8893	0.178	-5.002	0.000	-1.242	-0.537
Pollock_Imported_Kilos	0.1046	0.107	0.980	0.329	-0.107	0.316

Omnibus:	0.776	Durbin-Watson:	1.521
Prob(Omnibus):	0.678	Jarque-Bera (JB):	0.796
Skew:	-0.188	Prob(JB):	0.672
Kurtosis:	2.866	Cond. No.	4.07

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Univariate Regression

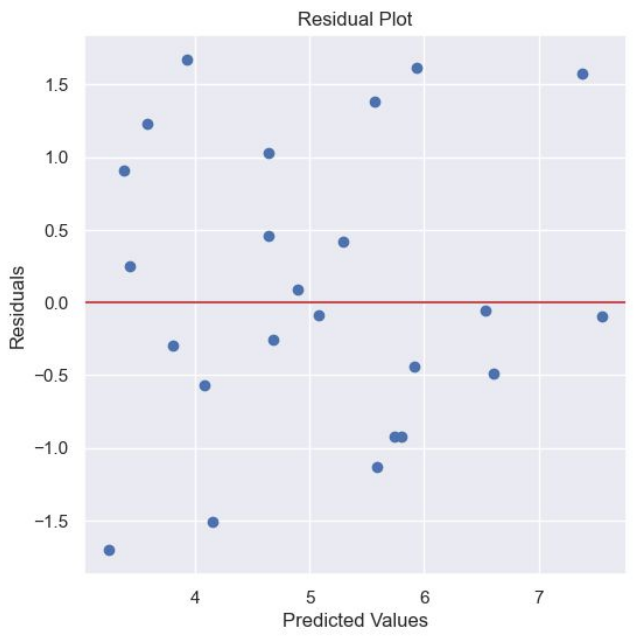
- Individual R^2 scores for each features.
- Features with the largest R^2 values match the features with significant (<0.01) p-values from OLS.
- Residual plots (not pictured) show the impact of covariance on univariate regressions.

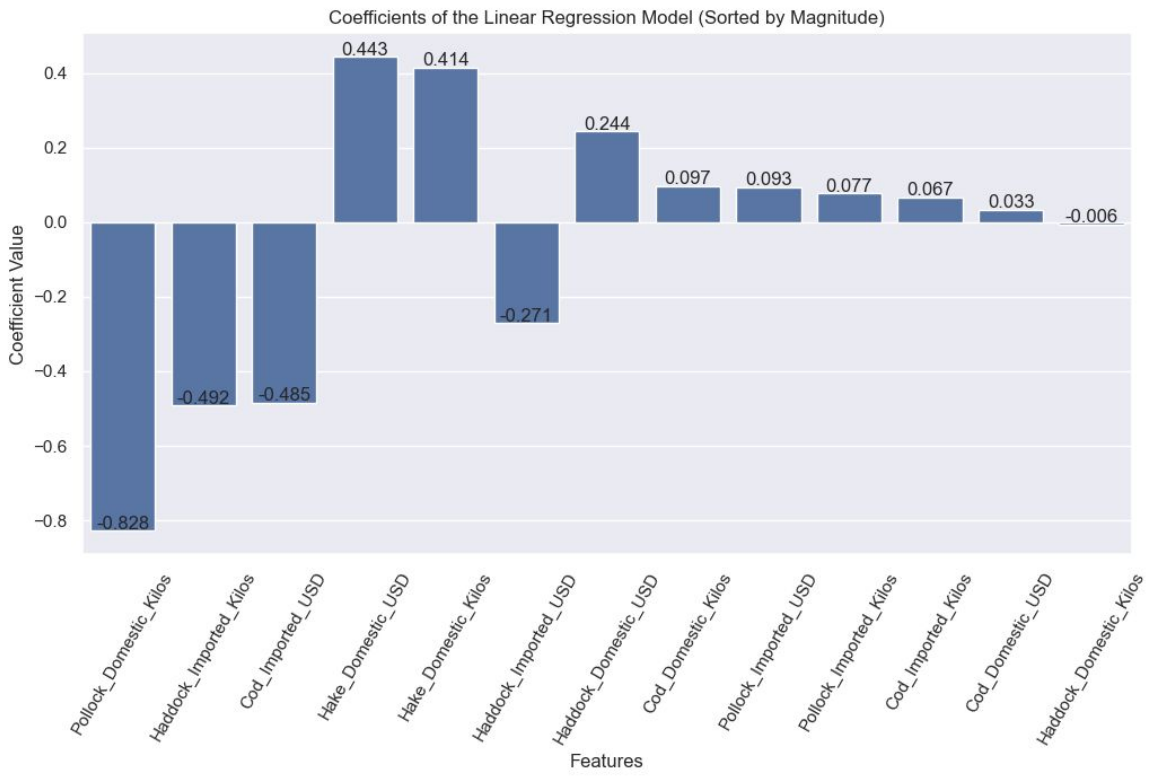


	coef	std err	t	P>|t|	[0.025	0.975]
Cod_*Imported*_USD	−0.3615	0.127	−2.855	0.005	−0.612	−0.110
Hake_*Domestic*_USD	0.4876	0.121	4.030	0.000	0.248	0.728
Haddock_*Imported*_Kilos	−0.5231	0.128	−4.099	0.000	−0.776	−0.270
Hake_*Domestic*_Kilos	0.4379	0.142	3.091	0.003	0.157	0.719
Pollock_*Domestic*_Kilos	−0.8893	0.178	−5.002	0.000	−1.242	−0.537

Multivariate Regression

R-squared: 0.5824
Mean Squared Error: 0.9472



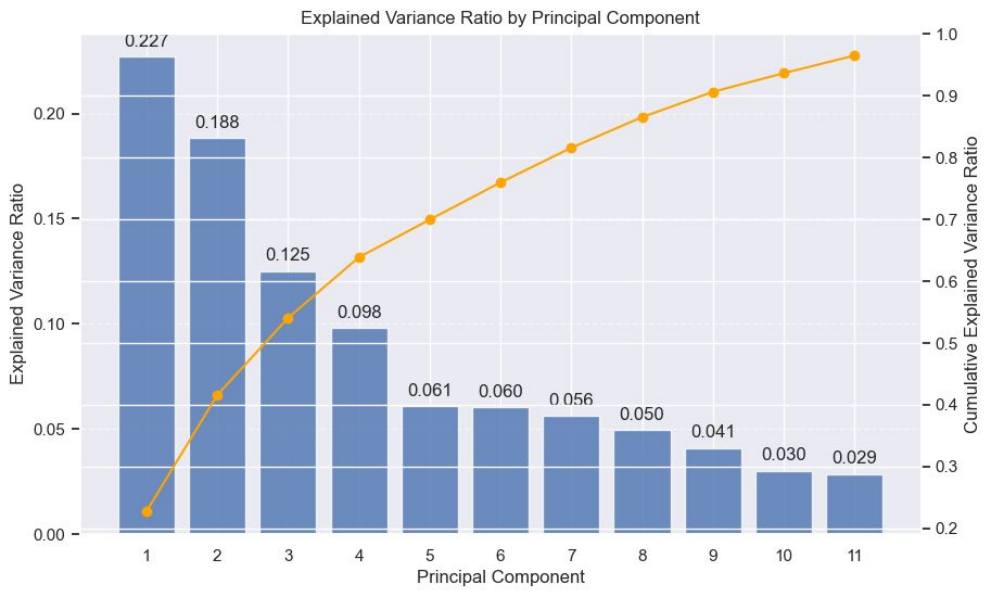


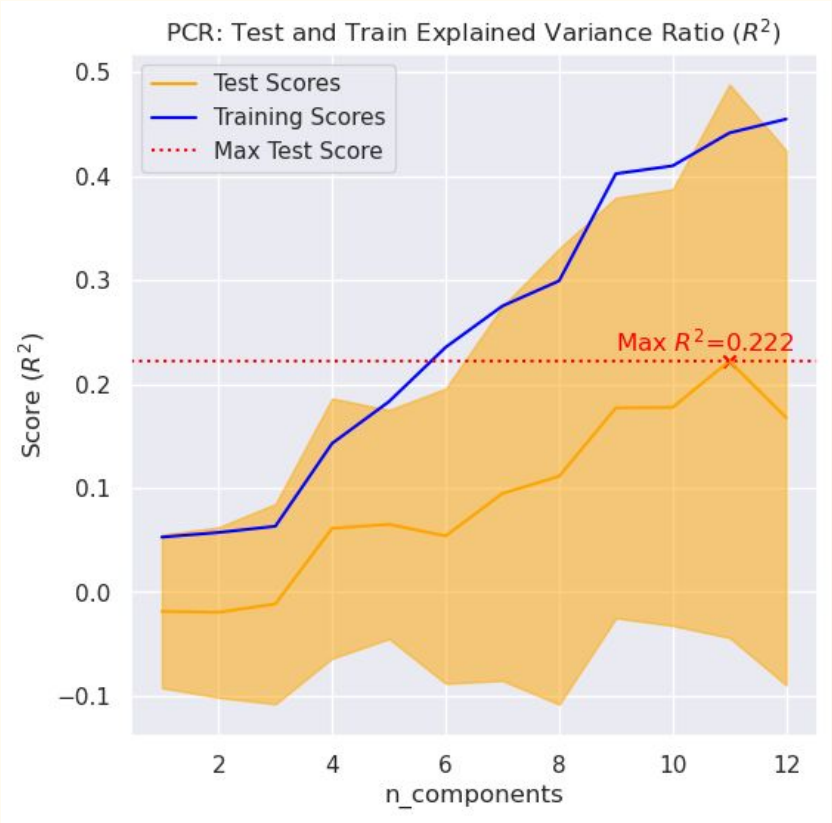
Principal Component Regression

R-squared: 0.5221

Mean Squared Error: 1.0840

Best Parameters: {'n_components': 11}



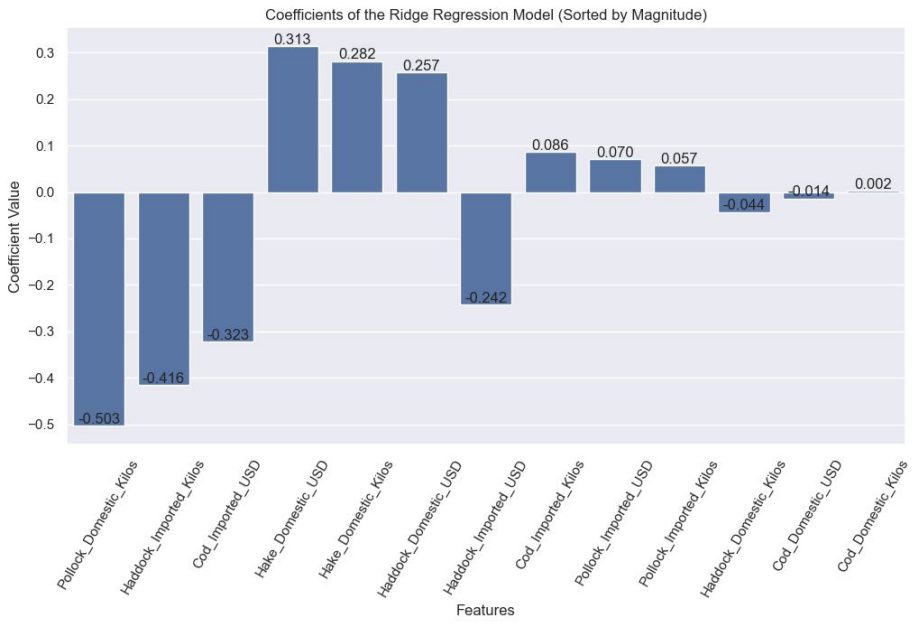


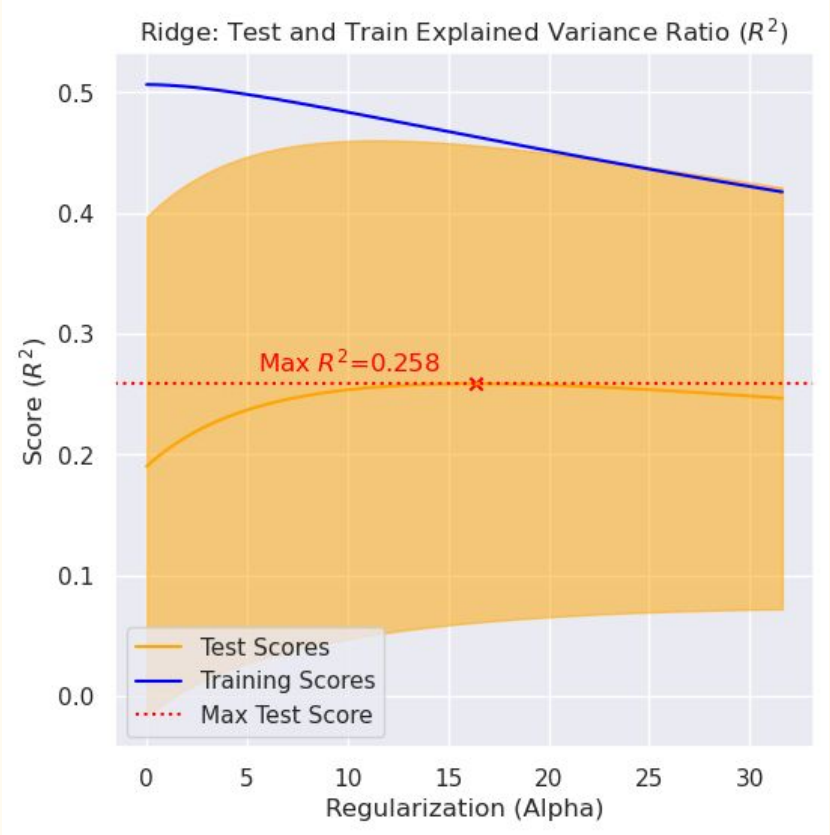
Ridge Regression

R-squared: 0.5473

Mean Squared Error: 1.0268

Best alpha: 16.3574



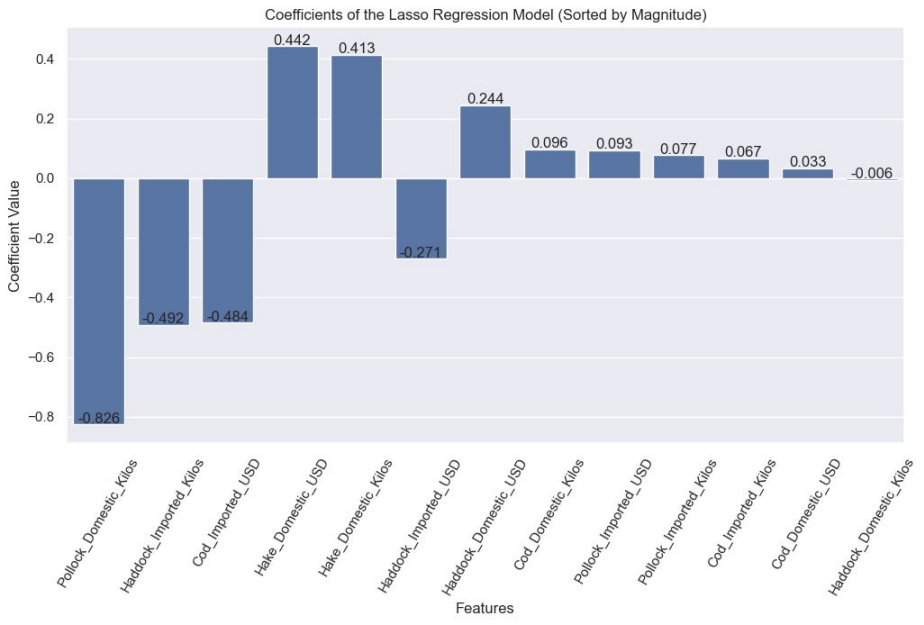


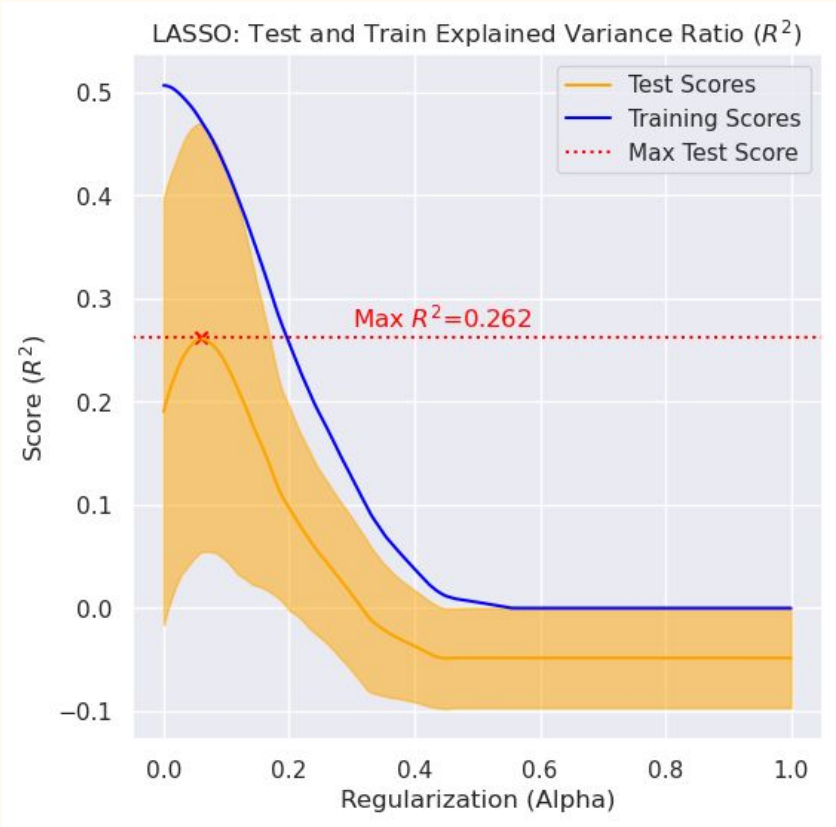
Lasso Regression

R-squared: 0.5824

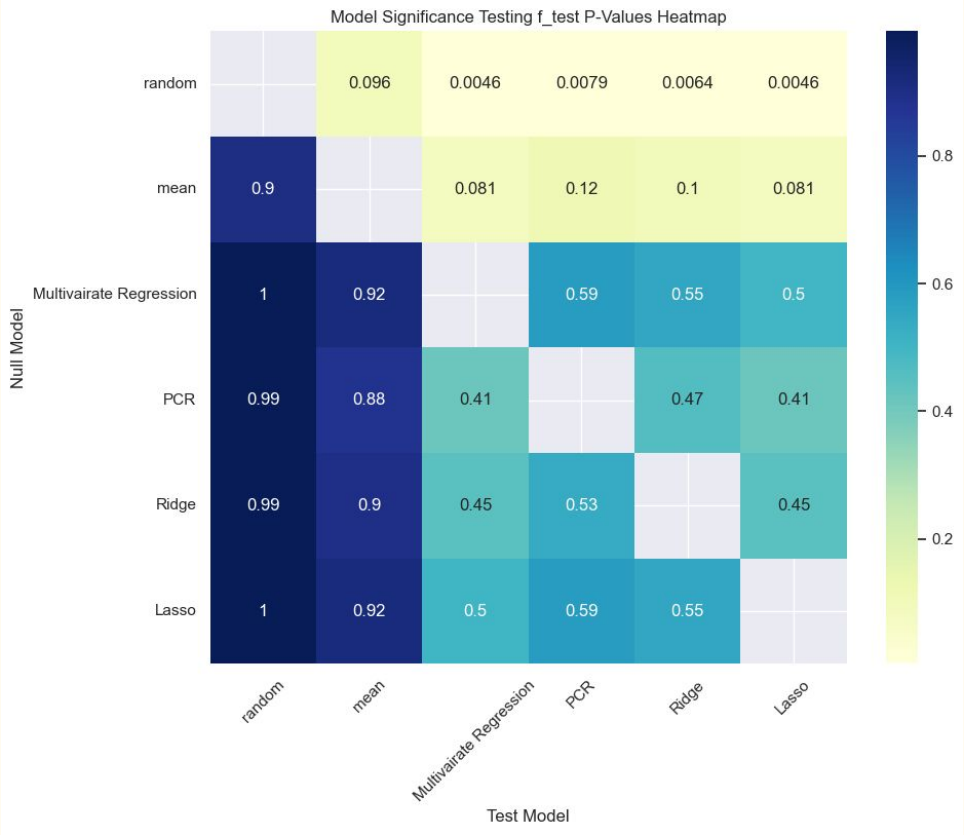
Mean Squared Error: 0.9471

Best alpha: 0.0590





Significance Testing



Results

- Our best model (Multivariate Model) returned an explained variance ratio of around ~.58.
- However, this wasn't significantly different than any of other models including a simple one that just assigned the mean of the target values to every target.
- Essentially, we cannot say that any of our models are significantly better than just guessing the average price
- We can't say the Barents Sea catch significantly impacts the Pollock Price in Maine but also given that our models did seem to show some improvement over the mean it isn't a closed question.

https://www.portlandoldport.com/listing/harbor-fish-market/

Reproducibility, Limitations and Next Steps

- Key limitations of our work include the limited scope of our selected data sample as well as the inherent complexity of price prediction.
- Unconsidered disruptors to the fishing industry, market forces and features not included in our analysis likely influence Domestic Pollock Price.
- We hope that our organized and methodical approach, thorough documentation, and implementation of the DRY principles provides a solid foundation for future analysis in this area.
- Next steps for the project may include seeking additional input from domain experts at GMRI; searching for more granular data or implementing further modeling approaches not yet explored.

Conclusion

- Overall, we don't have a 'smoking gun' conclusive result to provide today.
- Our group is reluctant to claim that ecological or policy changes in the Barents Sea have significantly impacted the price of Domestic Pollock in Maine to date.
- Accordingly, we believe that giving firm predictions regarding the impacts of decreased Cod and Haddock quotas in the Barents Sea would be irresponsible.
- Ultimately we hope that our work might provide insight and a jumping off point for future interested parties.

https://thebarentsobserver.com/en/industry-and-energy/2023/10/barents-sea-cod-quotas-lowered-20-third-year-row

Acknowledgments

1. Stakeholder meeting with Dr. Kanae Tokunaga, Senior Scientist at GMRI, March 12, 2024.
2. [NOAA 2020 Fisheries of the United States, May 2022](https://media.fisheries.noaa.gov/2022-05/Fisheries-of-the-United-States-2020-Report-FINAL.pdf)
3. [Barents Sea cod quota drops by 20 percent for third straight year](https://www.seafoodsource.com/news/supply-trade/barents-sea-cod-quota-drops-by-20-percent-for-third-straight-year#:~:text=Norway)
4. [Tight cod supplies, better for pollock | GLOBEFISH | Food and Agriculture Organization of the United Nations](https://fao.org/in-action/globefish/market-reports/resource-detail/en/c/1655476/)
5. [Groundfish Forum predicts wild-caught whitefish supplies will remain flat in 2024](https://www.seafoodsource.com/news/supply-trade/groundfish-forum-predicts-wild-caught-whitefish-supplies-remain-flat-in-2024)
6. [Groundfish: Supplies slightly down in 2023 | GLOBEFISH | Food and Agriculture Organization of the United Nations](https://www.fao.org/in-action/globefish/market-reports/resource-detail/en/c/1634023/)
7. [Supplies may become tighter | GLOBEFISH | Food and Agriculture Organization of the United Nations](https://www.fao.org/in-action/globefish/market-reports/resource-detail/en/c/1460139/)
8. [An Introduction to Statistical Learning with Applications in Python](https://www.statlearning.com/resources-python)
9. [Scikit Learn - Supervised Learning](https://scikit-learn.org/stable/supervised_learning.html)
10. * [Ordinary Least Squares](https://scikit-learn.org/stable/modules/linear_model.html#ordinary-least-squares)
11. * [Ridge Regression and classification](https://scikit-learn.org/stable/modules/linear_model.html#ridge-regression-and-classification)
12. * [Lasso](https://scikit-learn.org/stable/modules/linear_model.html#lasso)
13. [Statmodels - Regression and Linear Models](https://www.statsmodels.org/dev/user-guide.html#regression-and-linear-models)