Predicting Canadian Food Prices

Patrick Walter

Background: Canada's Food Price Report

- This year the 8th edition of Canada's Food Price Report was published as a collaborative effort between Dalhousie University and University of Guelph.
- The report looks at factors that affect the future prices seen by consumers for food over the next 12-month period.
- Includes the key drivers of the upcoming year's food prices such as Climate, Energy Costs, Inflation, Policy Context, Food Processing Industry, Consumer Debt and Deleveraging and many more.
- The report divides the CPI basket into food categories including Dairy and Eggs, Fruit and Nuts, Meats, Vegetables, and Fish and Seafood.

Background: Consumer Price Index

- An indicator of the changes in prices experienced by Canadians, obtained by comparing the rising or falling cost of a fixed basket of good and services.
- Fixed basket means that products and services are of the same quality and quantity and therefore these changes in cost are a true reflection of the pure price changes.
- Data is released monthly and open to the public on Statistics Canada's website as part of the CANSIM database.

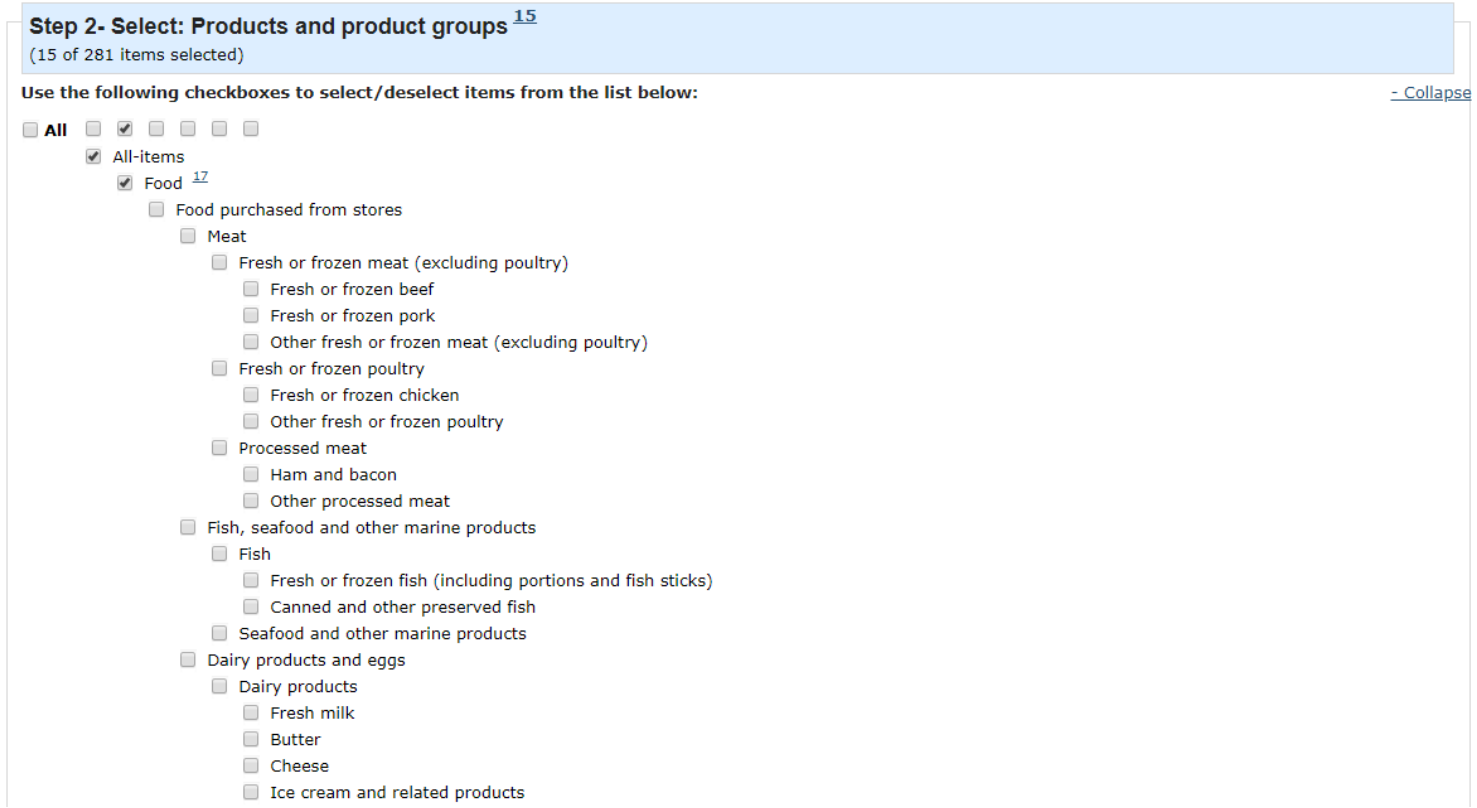
Background: Consumer Price Index

Consumer Price Index, by province (monthly) (Canada)

	February 2017	January 2018	February 2018	January 2018 to February 2018	February 2017 to February 2018
	2002=100			% change	
Canada					
All-items	**129.7**	**131.7**	**132.5**	**0.6**	**2.2**
Food	141.7	144.7	**144.7**	0.0	2.1
Shelter	137.6	139.7	**140.0**	0.2	1.7
Household operations and furnishings	121.7	122.4	**123.4**	0.8	1.4
Clothing and footwear	92.9	91.7	**93.3**	1.7	0.4
Transportation	131.9	137.2	**137.7**	0.4	4.4
Health and personal care	123.3	125.1	**125.6**	0.4	1.9
Recreation, education and reading	113.2	111.8	**114.1**	2.1	0.8
Alcoholic beverages and tobacco products	159.2	163.2	**164.1**	0.6	3.1
Special aggregates					
All items excluding food	127.3	129.2	**130.1**	0.7	2.2
All items excluding energy	127.8	129.5	**130.3**	0.6	2.0
Energy	151.1	159.5	**159.1**	-0.3	5.3

Source: Statistics Canada, CANSIM, table 326-0020 and Catalogue nos. 62-001-X and 62-010-X.
Last modified: 2018-03-23.

Background: Consumer Price Index



Background: Consumer Price Index

Selected items [Add/Remove data]

Geography = Nova Scotia
Products and product groups [15] = Coffee and tea

2017											2018	
February	March	April	May	June	July	August	September	October	November	December	January	February
Percentage Change (period-to-period)												
	3.5	0.4	-1.0	0.7	-2.1	0.8	2.1	-1.0	-1.5	4.5	-1.3	1.4

Background: Consumer Price Index

Selected items [Add/Remove data]

Geography = Nova Scotia
Products and product groups [15] = Coffee and tea

2017											2018	
February	March	April	May	June	July	August	September	October	November	December	January	February
Percentage Change (year-to-year)												
												6.5

Research Objectives

- Train and test predictive models for predicting the average national food price as reported by the CPI for a twelve month period for 5 years, from 2012 to 2016.
- Train and test predictive models for predicting 21 other targets in the CPI basket for the years 2016 and 2017.
- Gain insight on what years were difficult to predict, and what products were difficult to predict.

Dataset

- The original dataset was built by Jay Harris for his work on the 2017 and 2018 Canada Food Price Report and his thesis research titled *A Machine Learning Approach to Forecasting Consumer Food Price*
- It was constructed using a wide range of financial and econometric data from public and private institutions totaling 280 attributes
- The data set began at January 1985 with monthly records concluding at August 2017 giving 291 records.
- 22 targets from the CPI basket, mostly food products.

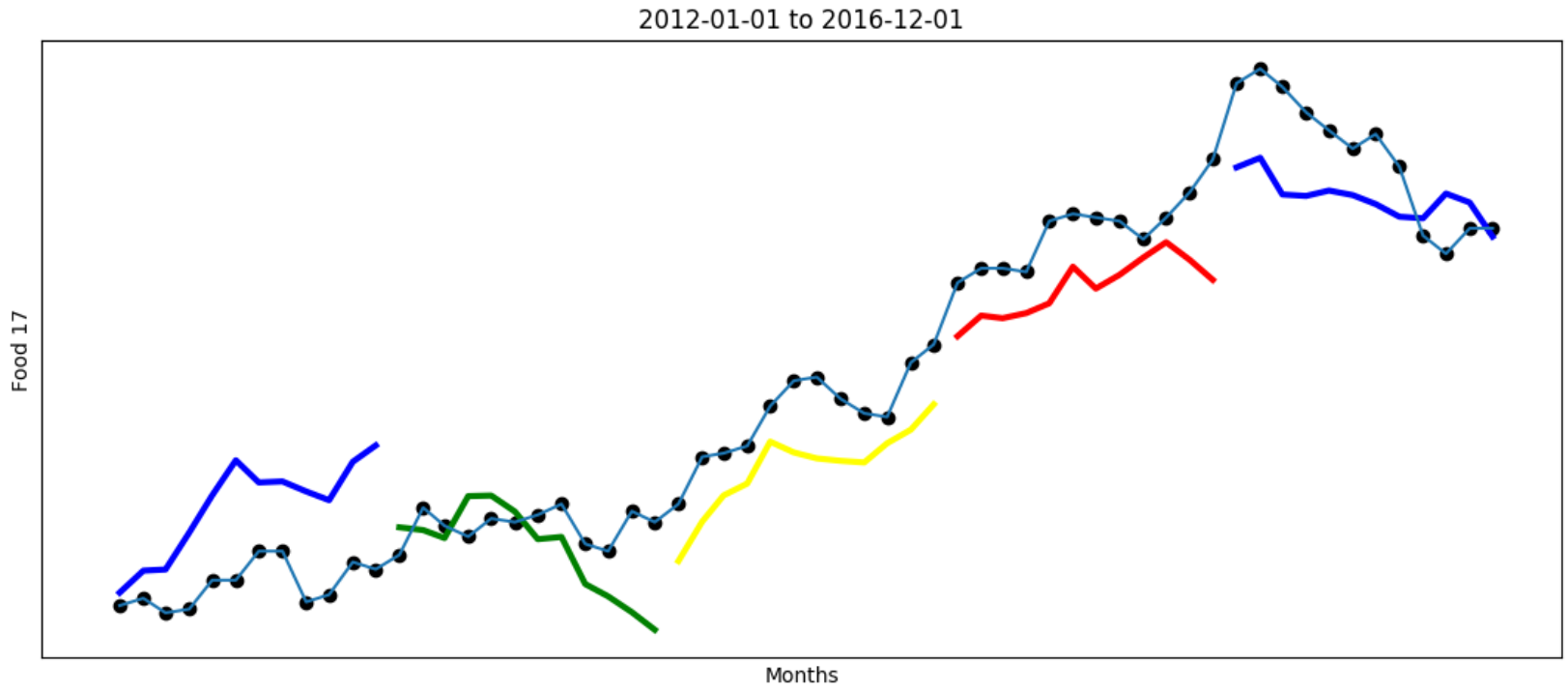
Data Preprocessing

- Pandas *dataframes* was used to manipulate the data.
- *Dataframe* is a tabular data structure that has labeled axes.
- Dividing the dataset in targets and attributes
- Capping the dataset at 1999
- Removing any columns that were incomplete
- Creating train and test *dataframes* for each of the five years
- Creating *dataframes* for each of the individual targets

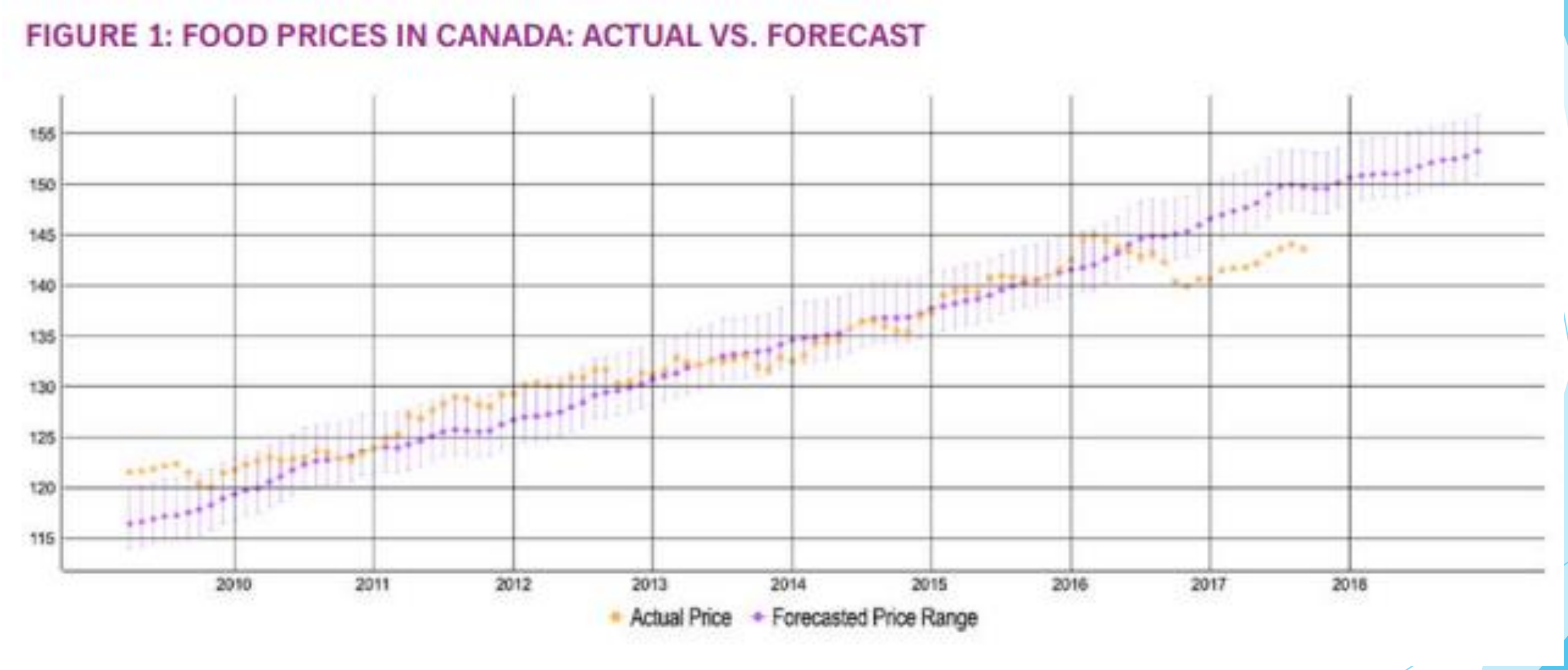
Tools: Python Scikit Learn

- A machine learning library for Python with a wide variety of implementations for popular algorithms.
- This research used the ordinary least squares linear regression implementation. Easy to use to train and test linear regression models.
- Feature selection, using mutual information and a selector to select the highest scoring features.

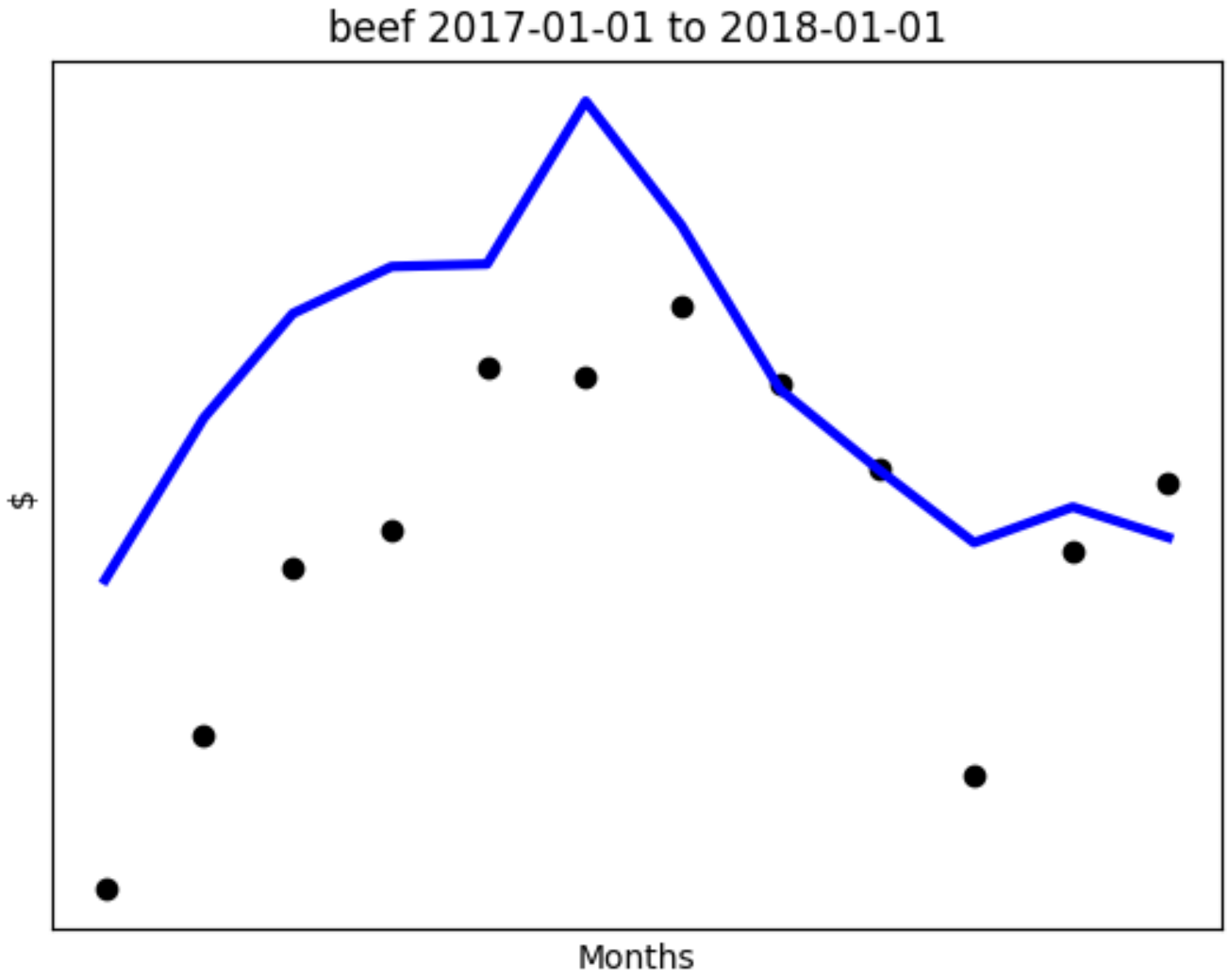
Capturing downward trends, k=25



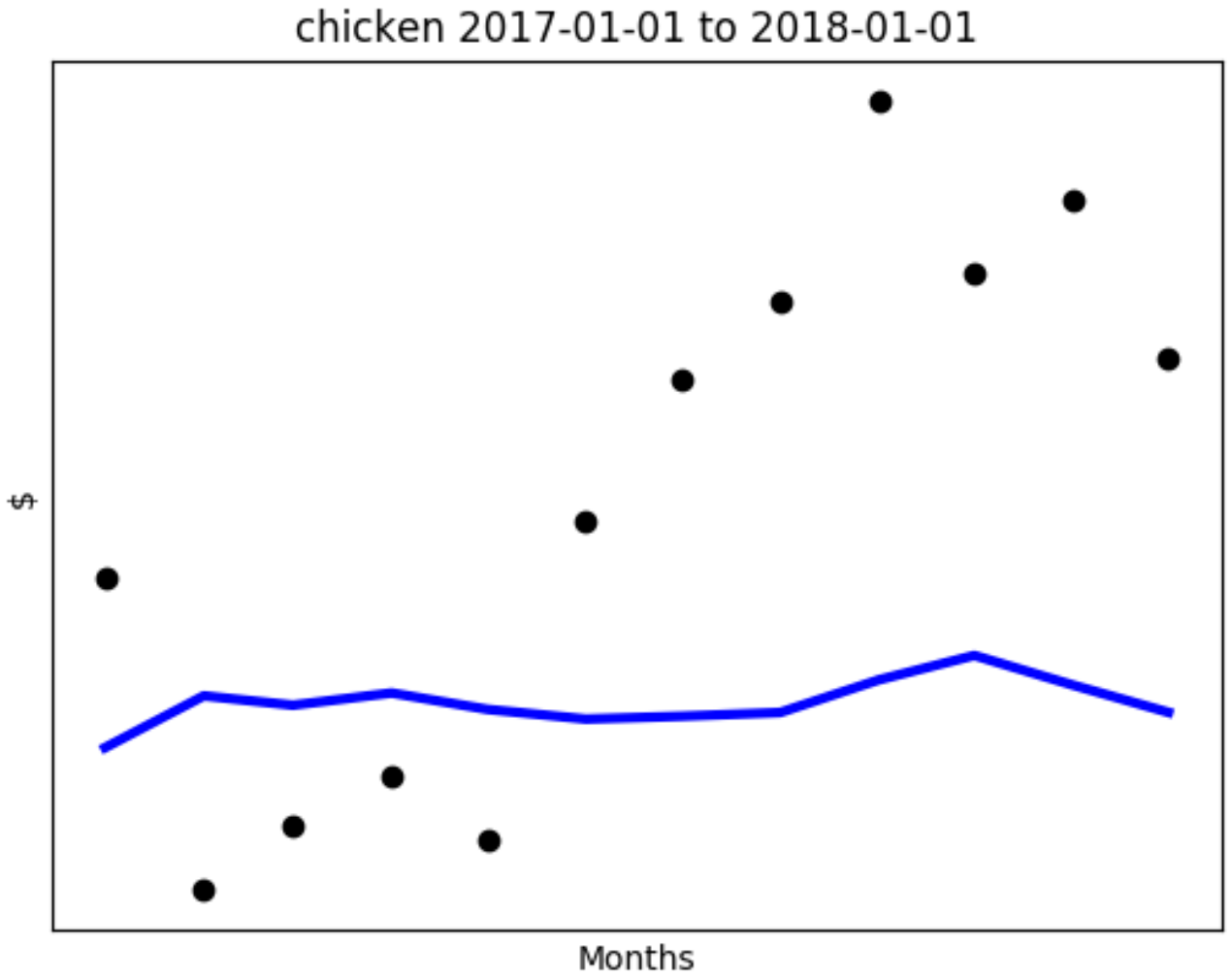
Food Price Report 2018



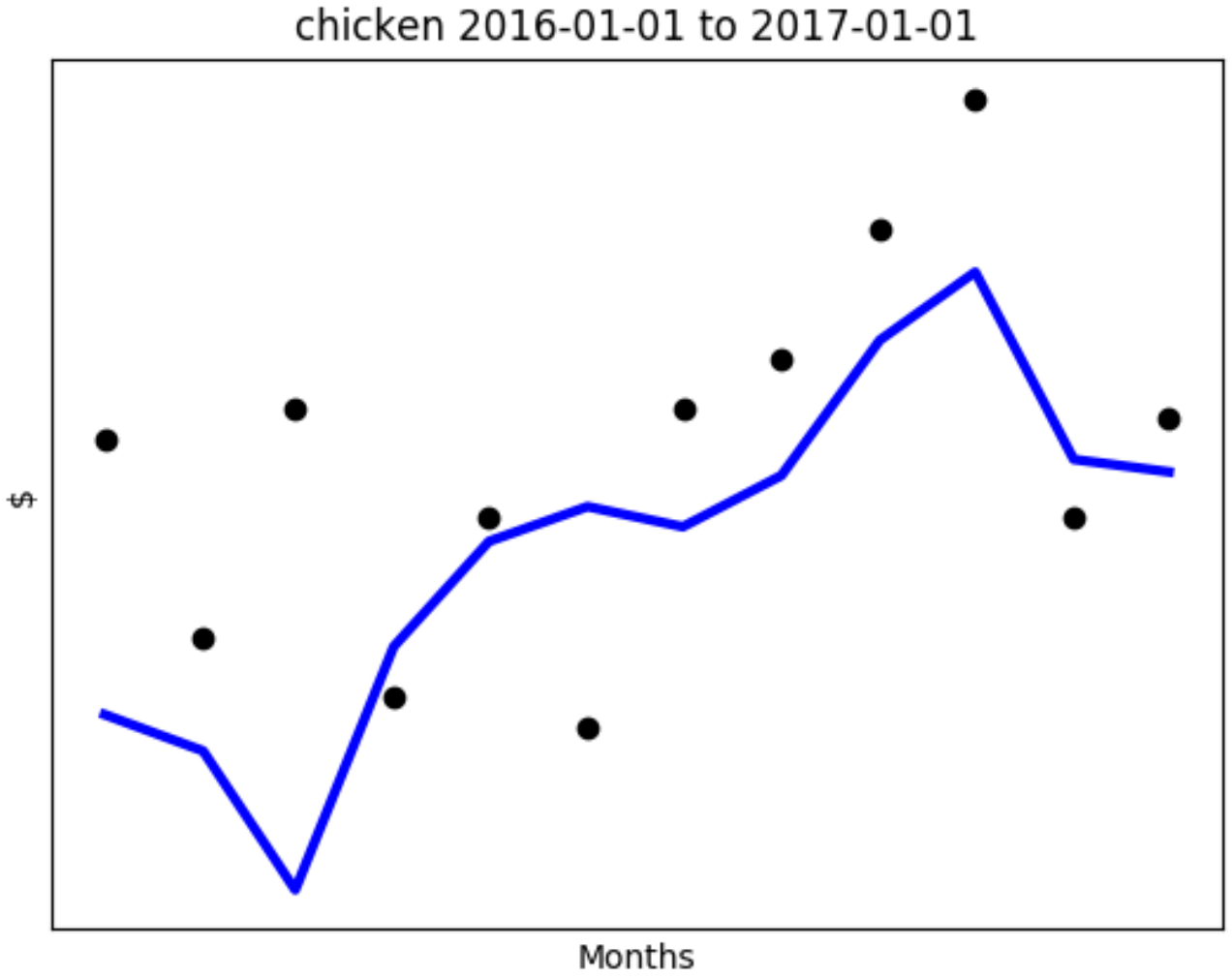
Individual Foods



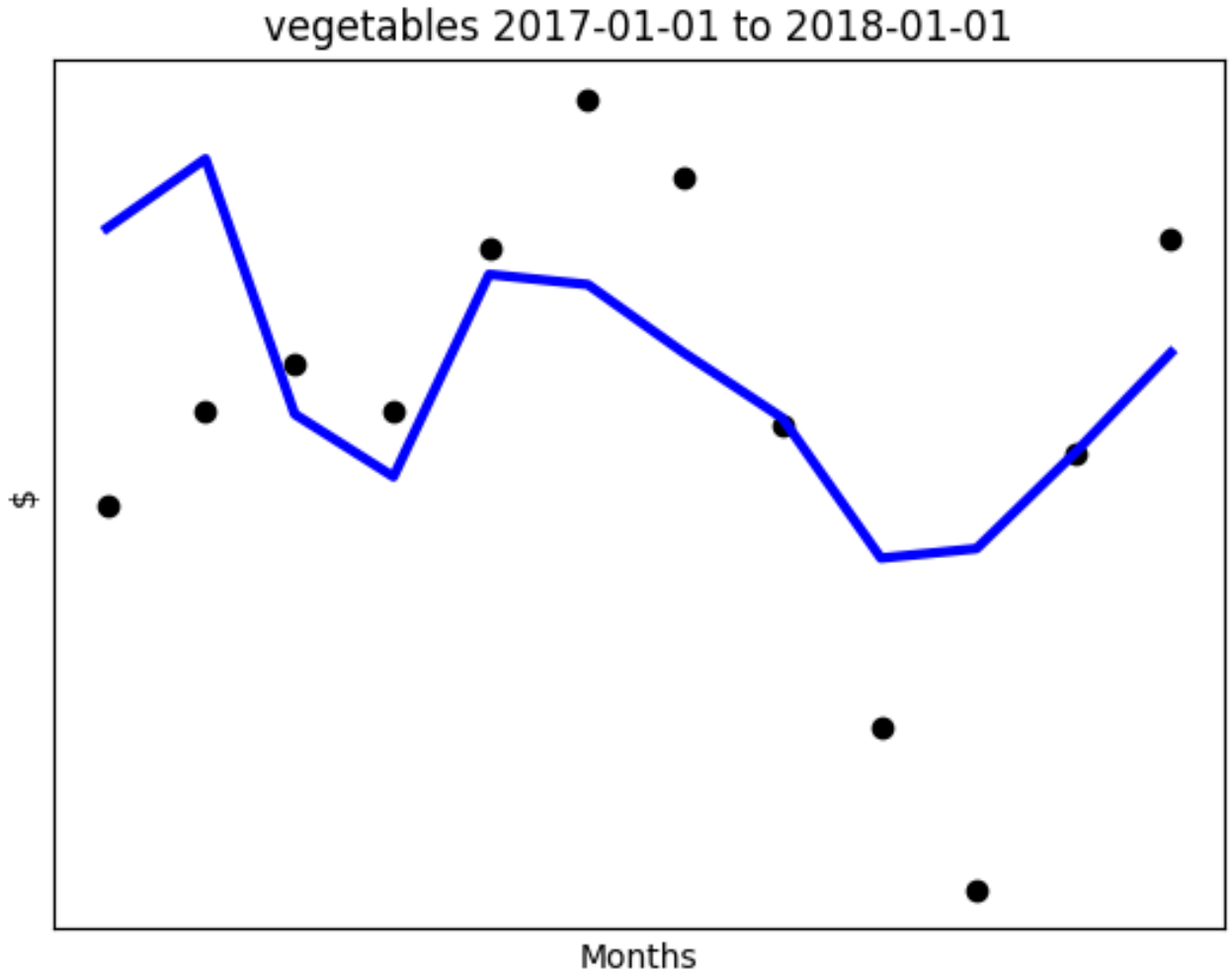
Individual Foods



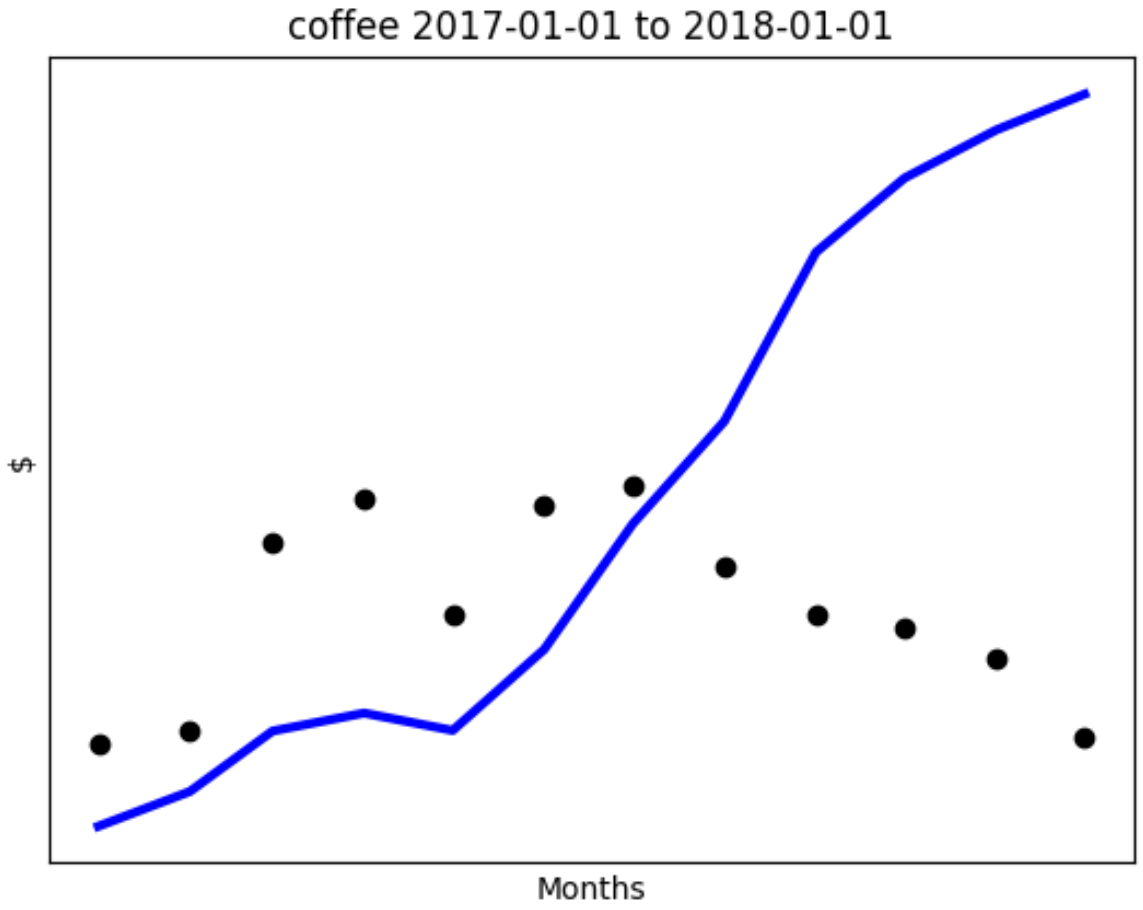
Individual Foods



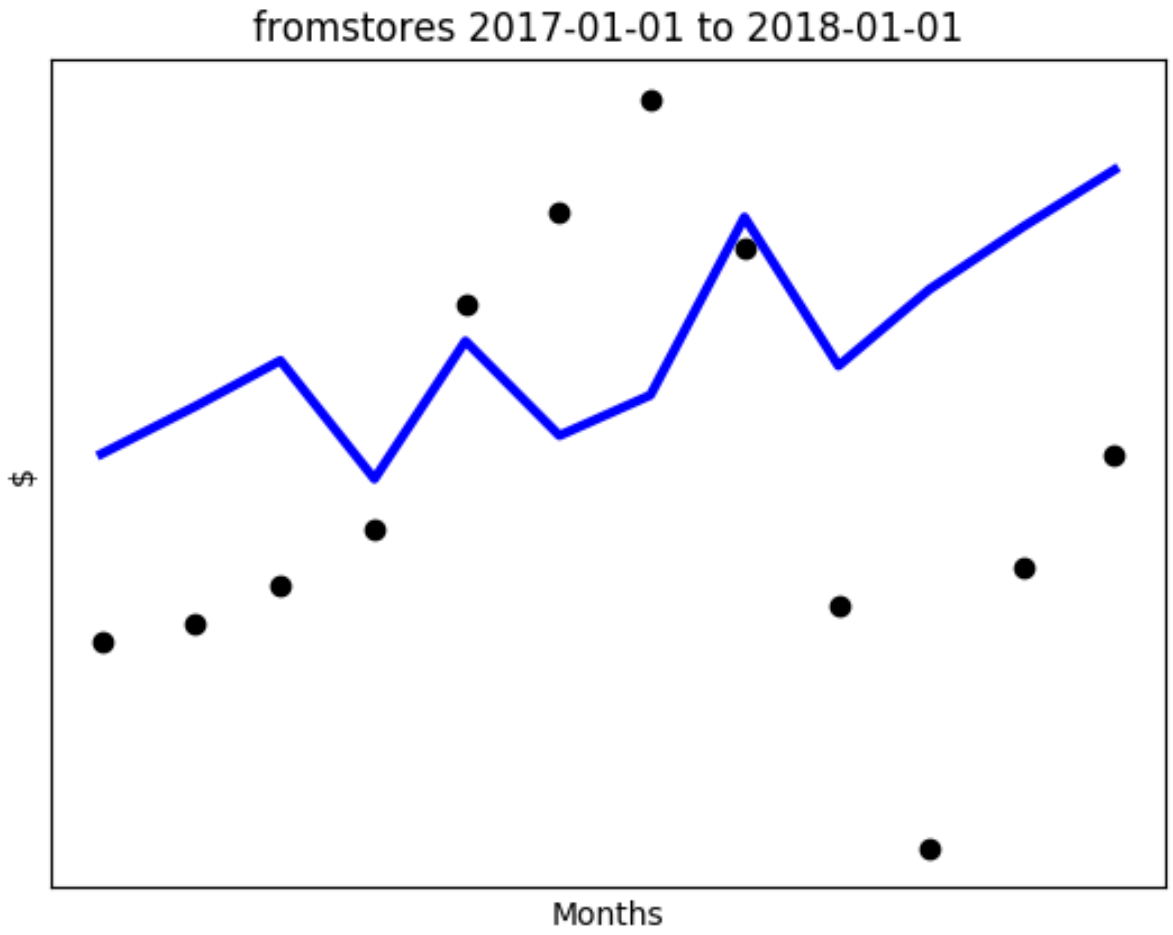
Individual Foods



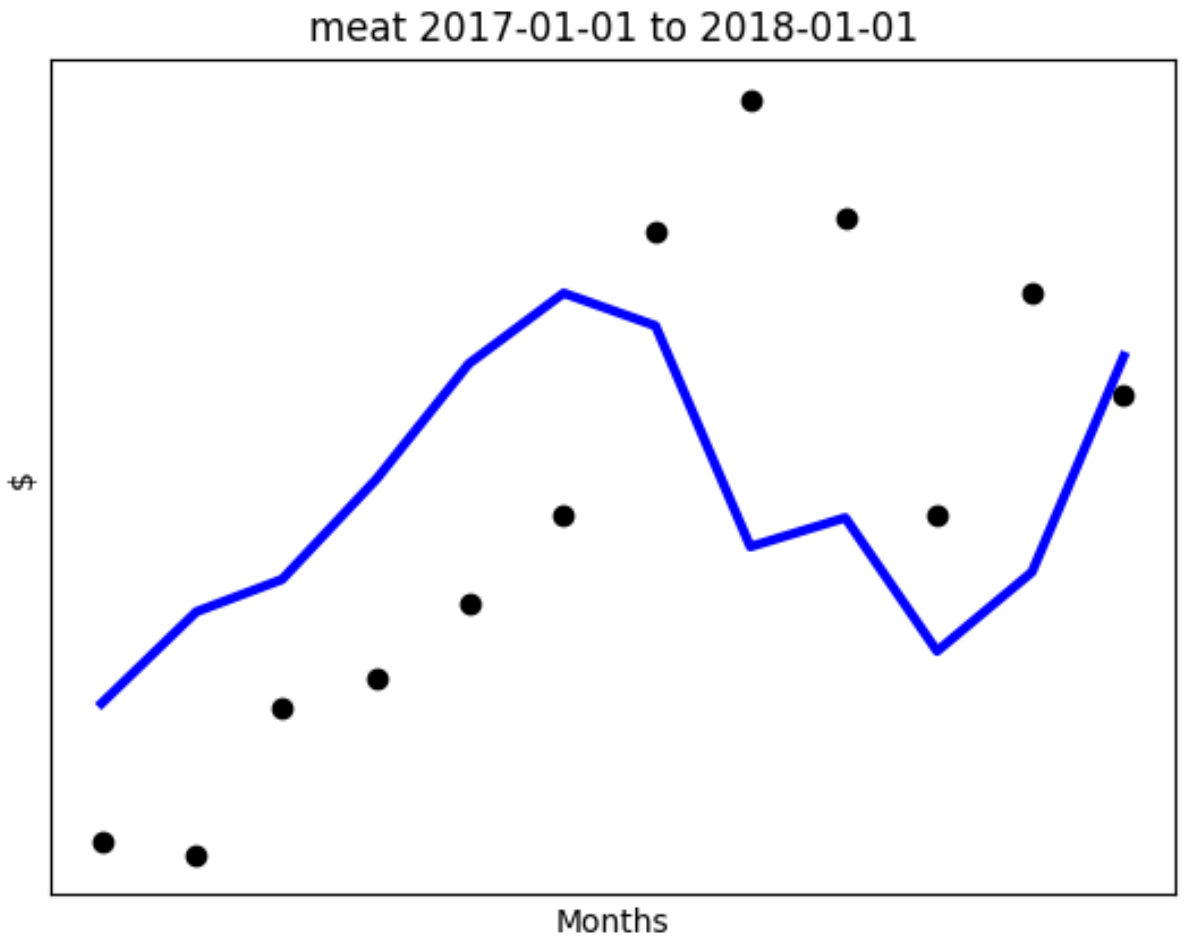
Individual Foods



Individual Foods



Individual Foods



References

Charlebois, S., Jabez, H., Tyedmers, P., Bailey, M., Keselj, V., Conrad, C., . . . Chamberlain, S. (2017). Canada's Food Price Report. Halifax: Dalhousie University.

Harris, J. (2017, August). A Machine Learning Approach to Forecasting Consumer Food Prices. Halifax, Nova Scotia, Canada.

Hunter, J., Dale, D., Firing, E., & Droettboom, M. (2012). Matplotlib. Retrieved from Matplotlib: https://matplotlib.org/

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating Mutual Information. Physical Review, E(69).

Pandas. (2017, December). Python Data Analysis Library. Retrieved from Pandas: https://pandas.pydata.org/

Scikit Learn. (2017). Generalized Linear Models. Retrieved from Scikit Learn: http://scikitlearn.org/stable/modules/linear_model.html

Scikit Learn. (2017). sklearn.feature_selection.mutual_info_regression. Retrieved from Scikit Learn: http://scikitlearn.org/stable/modules/generated/sklearn.feature_selection.mutual_info_regression.html#sklearn.feature_selection.mutual_info_regression

Statistics Canada. (2018, March). Consumer Price Index (CPI). Retrieved from Statistics Cnada: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2301