STUDY ON AVOCADO

STATS 131 – FALL 2018

ARIA WANG | ERIN YANG | IRIS YANG | JOANNA SHEN

WHY AVOCADO

➤ Avocado is gaining its popularity among the United States with its health benefits and unique taste

➤ The price of avocado is easily fluctuated by natural influences

➤ We are interested at analyzing the price changes of avocado in recent years to see whether people's purchasing behaviors adjust with prices
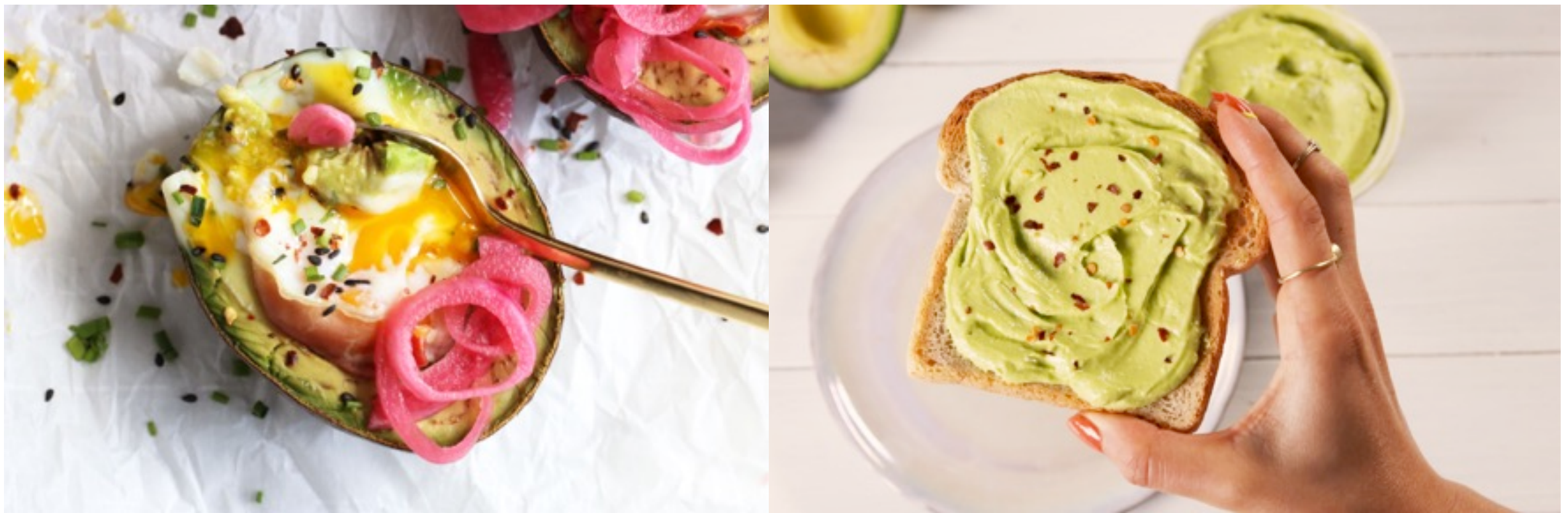
EXPLORING THE AVOCADO DATASET

	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags	Small Bags	Large Bags	XLarge Bags	type	year	region
0	2015-12-27	1.33	64236.62	1036.74	54454.85	48.16	8696.87	8603.62	93.25	0.0	conventional	2015	Albany
1	2015-12-20	1.35	54876.98	674.28	44638.81	58.33	9505.56	9408.07	97.49	0.0	conventional	2015	Albany
2	2015-12-13	0.93	118220.22	794.70	109149.67	130.50	8145.35	8042.21	103.14	0.0	conventional	2015	Albany
3	2015-12-06	1.08	78992.15	1132.00	71976.41	72.58	5811.16	5677.40	133.76	0.0	conventional	2015	Albany
4	2015-11-29	1.28	51039.60	941.48	43838.39	75.78	6183.95	5986.26	197.69	0.0	conventional	2015	Albany

➤ Total number of observations: 18,249

➤ Total number of variables: 13

   ➤ 2 Categorical variables: region, type (conventional, organic)

➤ Data collected from 2015 to 2018

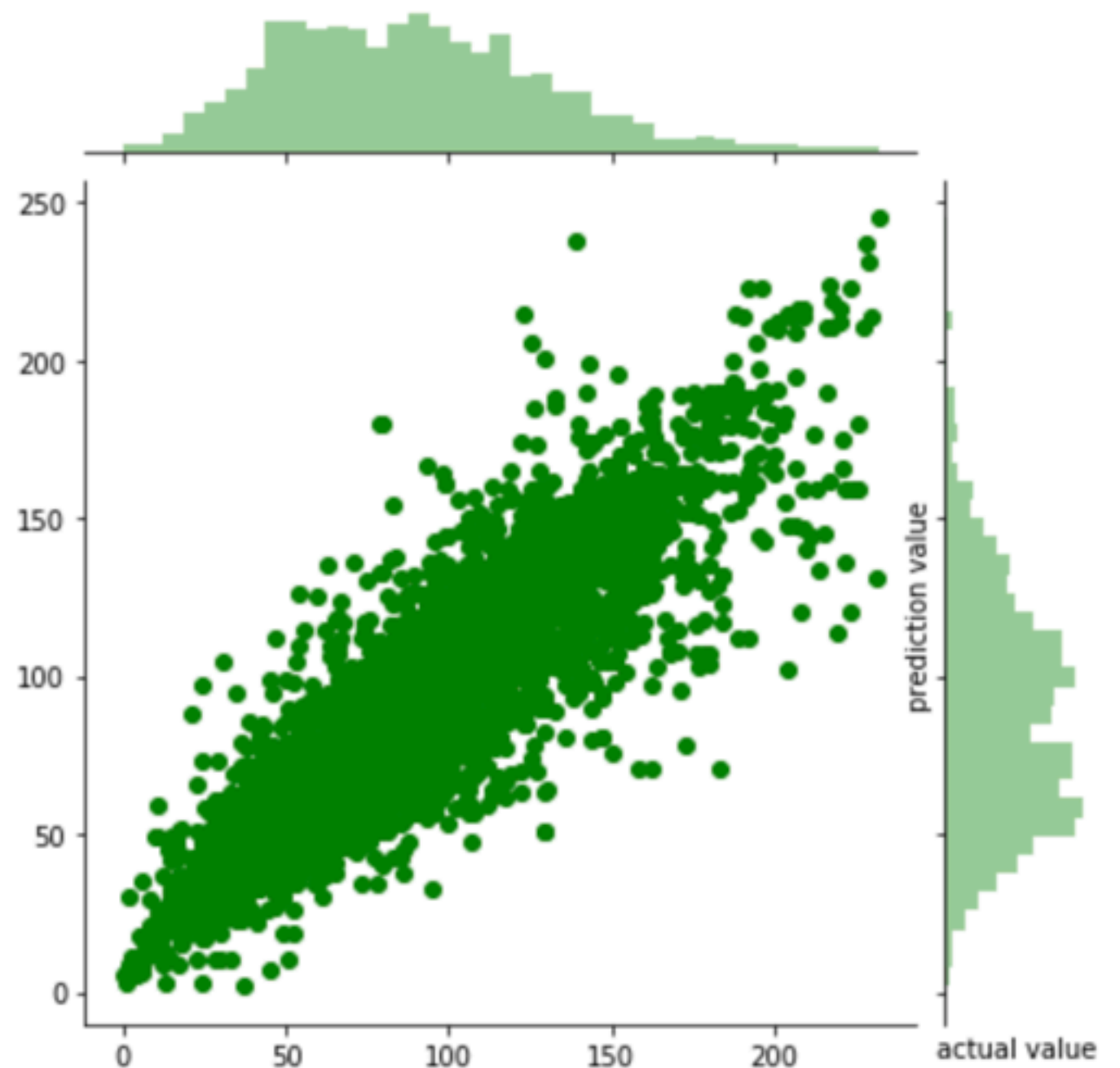# KEY FINDINGS FROM EXPLORATORY ANALYSIS | COMPARISON
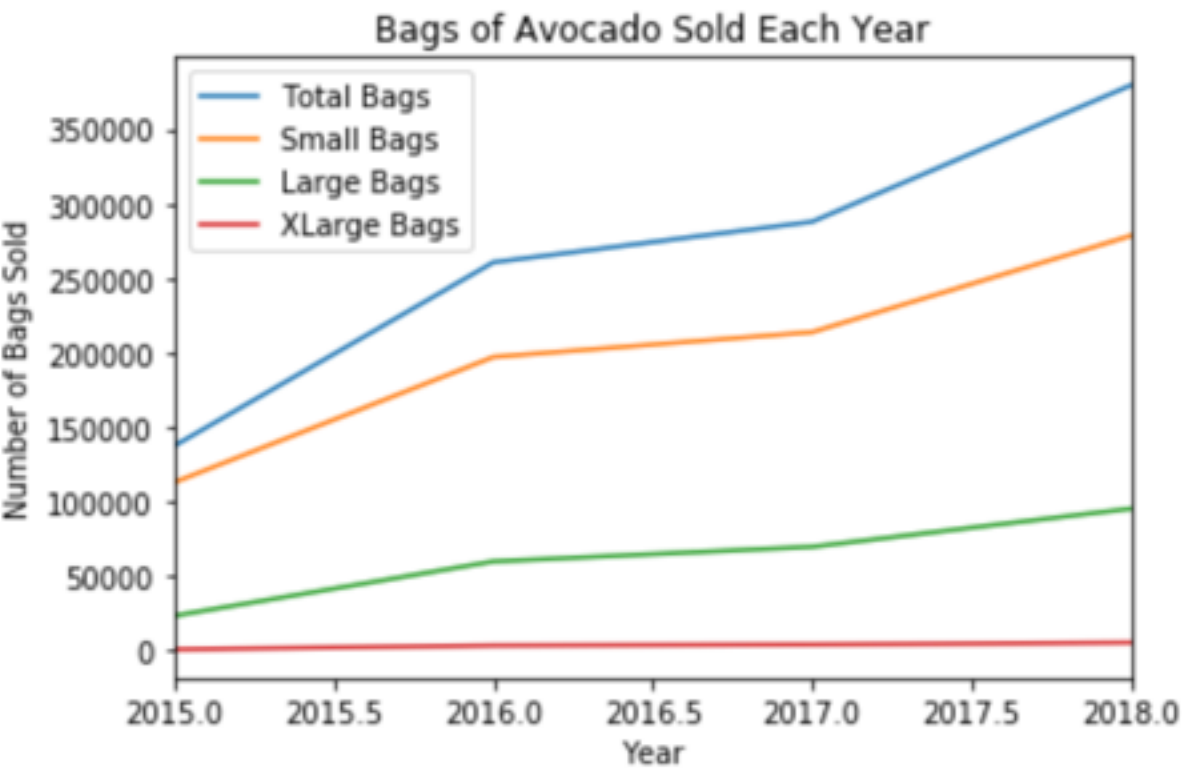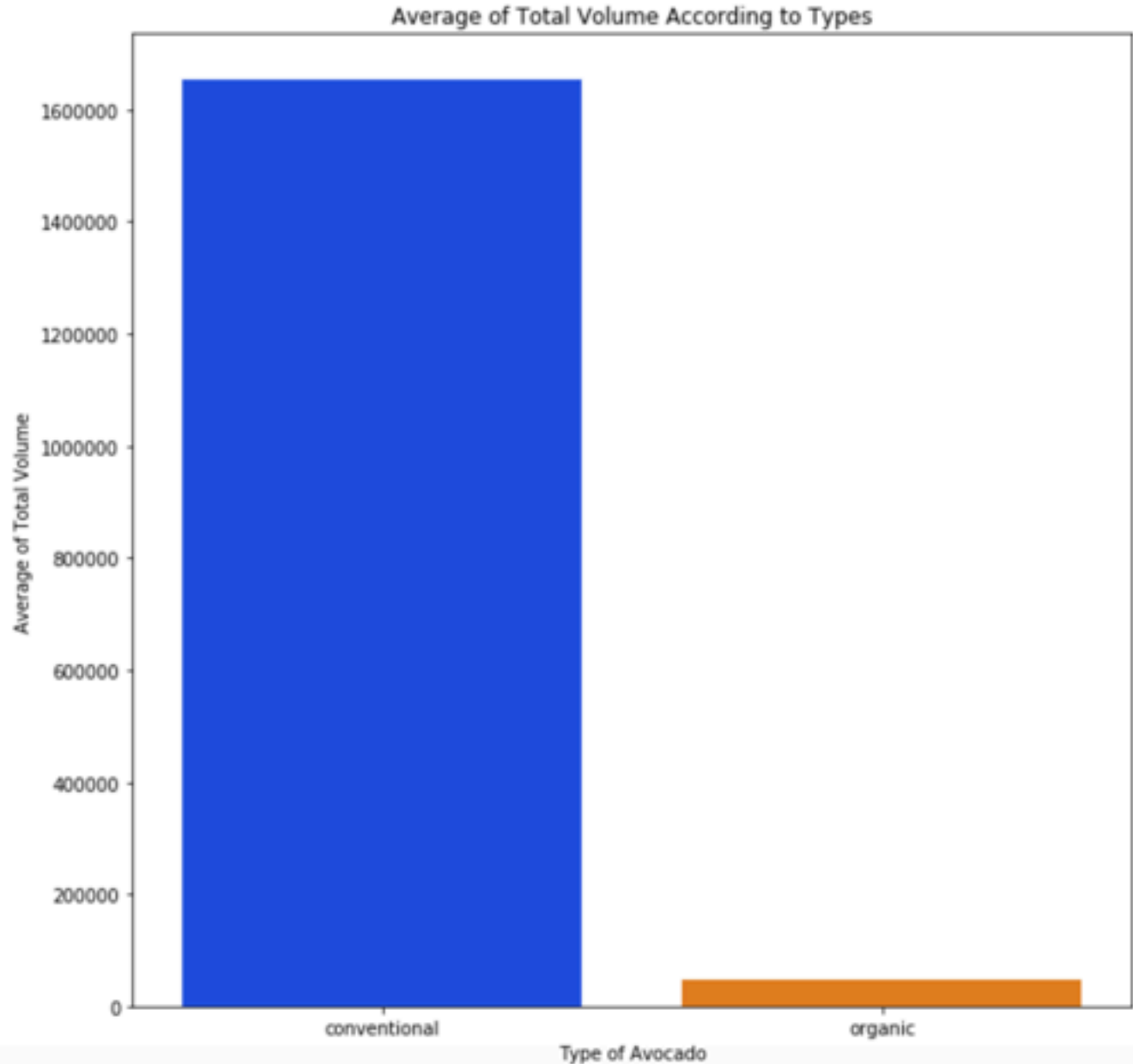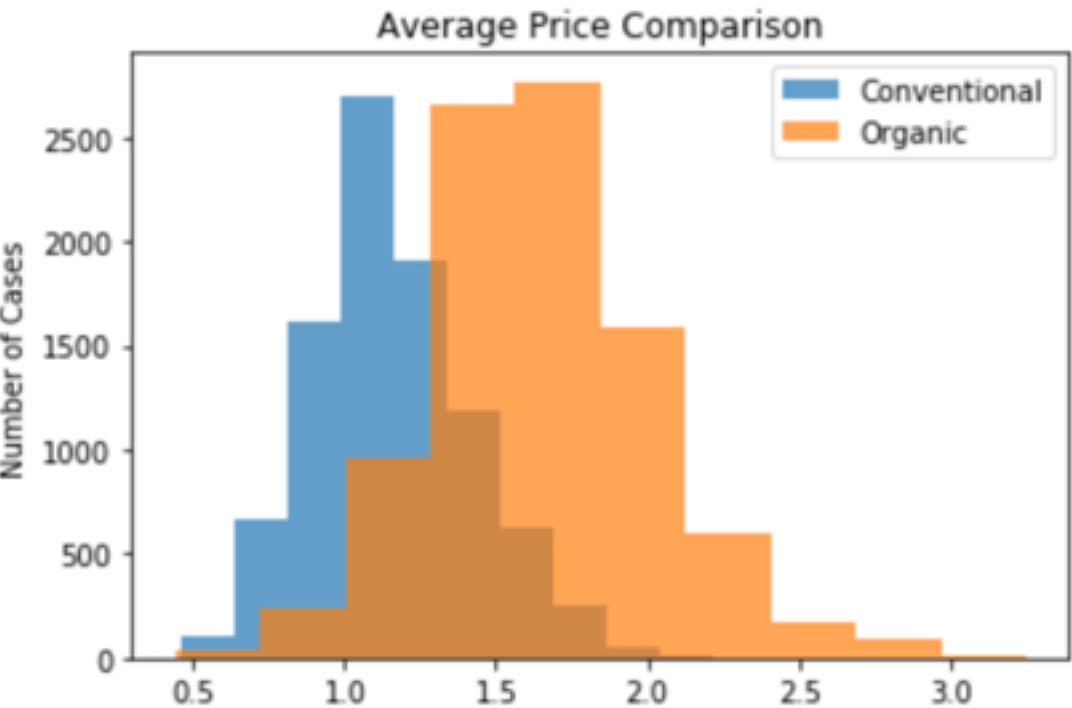
KEY FINDINGS FROM EXPLORATORY ANALYSIS | TIME PLOTS
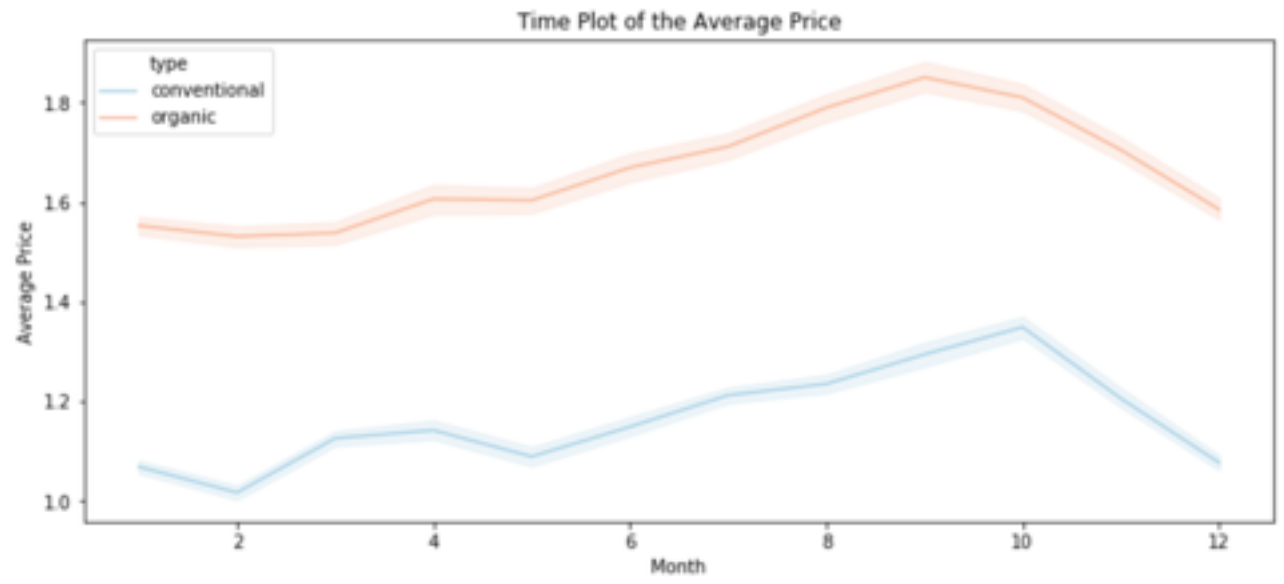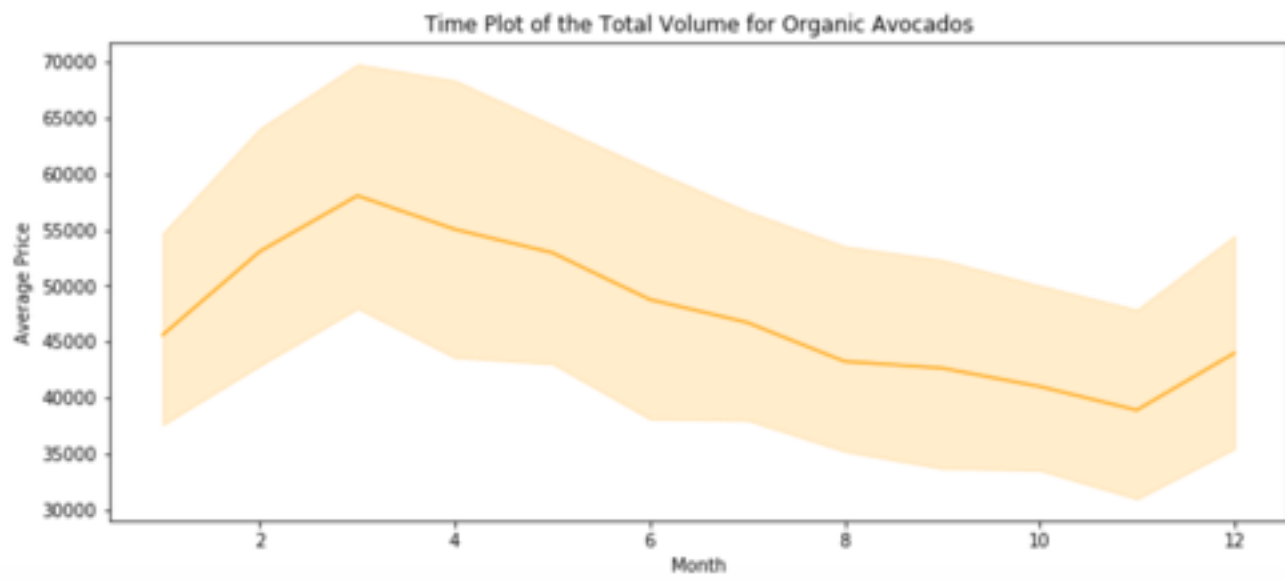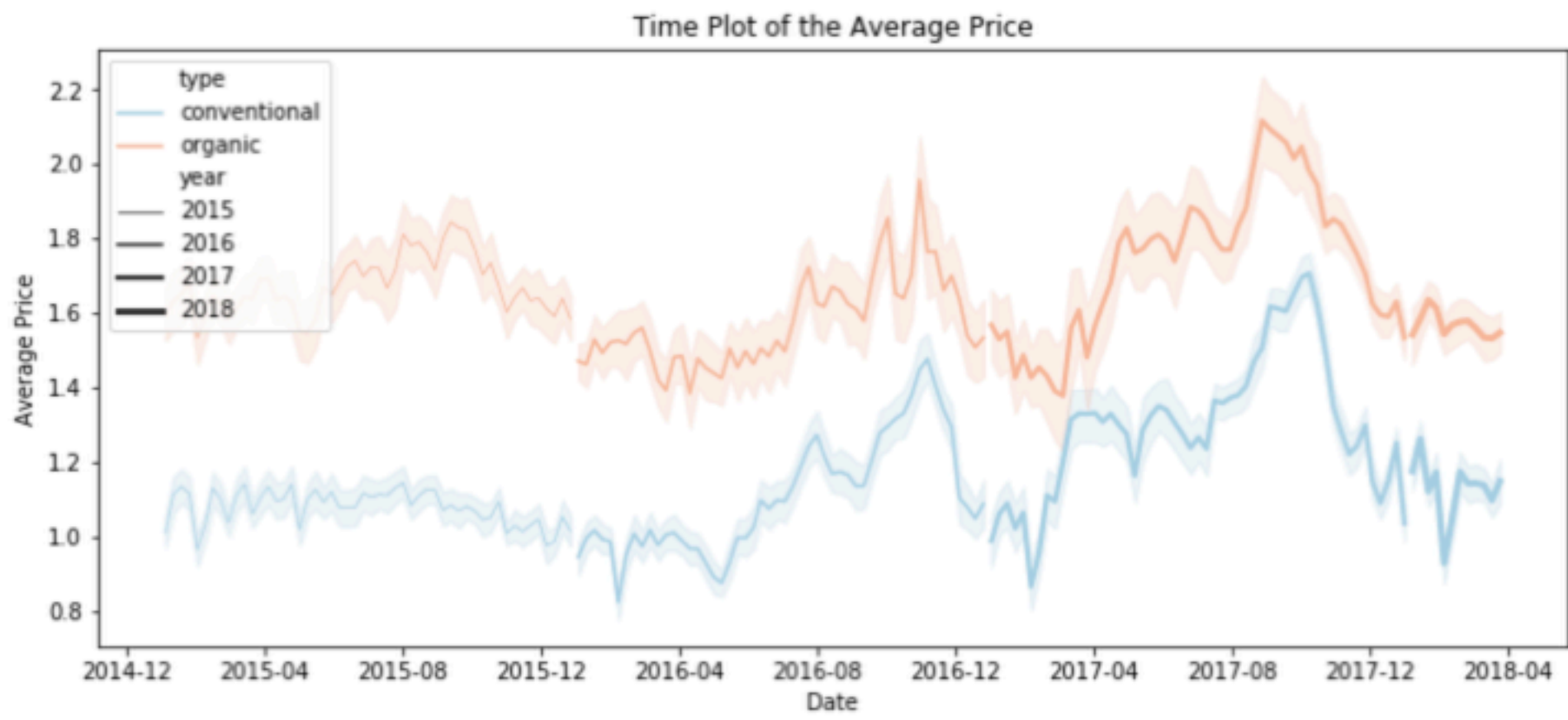
MODELING AND PREDICTION

➤ Methods attempted in predicting price: Linear Regression, Decision Tree, Random Forest, KNN

➤ Best prediction: KNN, with the lowest mean squared error.

THANK YOU !