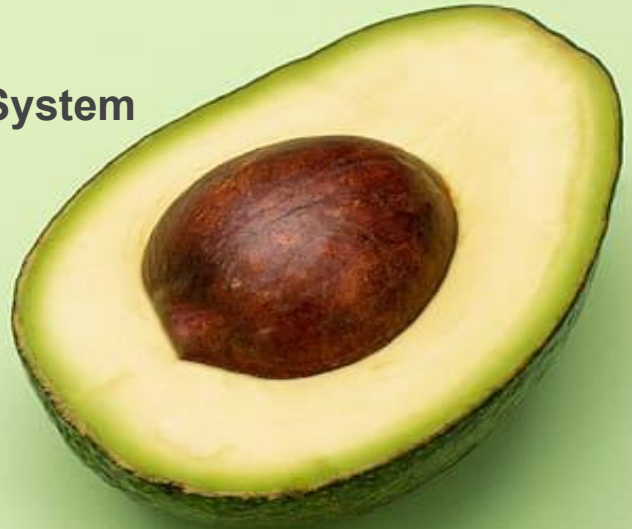
Final Presentation

Avocado Price and Sales Prediction System

Yunze Qiu yq2310

Yiran Lin yl4628

Stacy Lai sl4450



Motivation/Novelty

Surge in Market Value

Avocado industry increased at a steady rate from 2019 12.82 billion dollars to 14.33 billion dollars in 2021. Growth is projected to continue increasing at a steady rate until 2025 to 17.91 billion dollars

Heavy industry-funded research of marketing campaigns

Change the images of avocado as one of the superfoods that beneficial for your health.

Rise in consumption and sales revenue

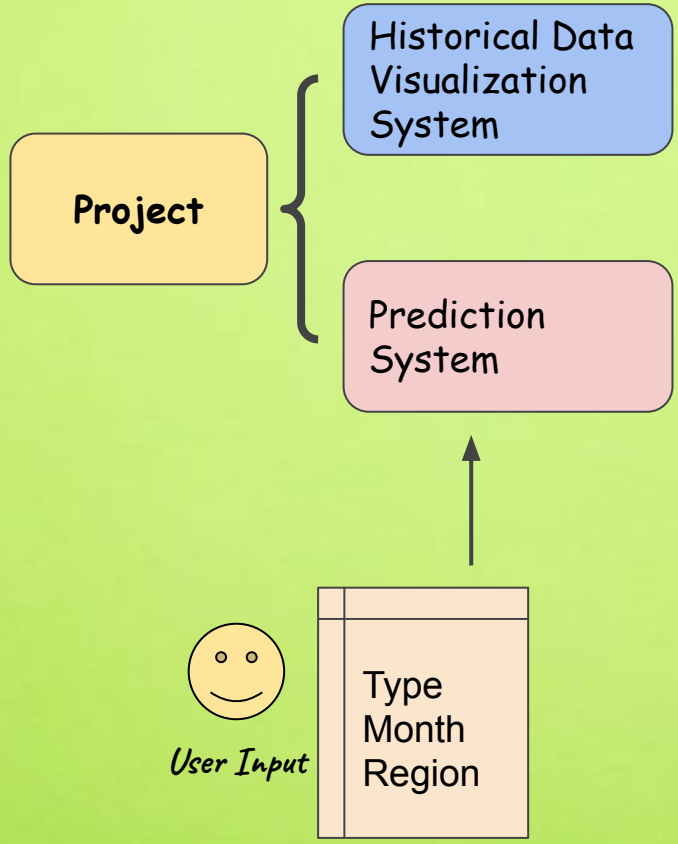
Goal: Build a system which can predict the price and consumption of avocados in certain markets as well as to generate a visual data report by analyzing the marketing data of avocados.

Business Value

Bridge the gap among previous works, where we can build a system that allows users:

- Estimate the price and consumption level of avocado
  - help sellers in the process of demand planning
  - estimate the sales revenue
- A visual data report analyzing the consumer purchasing behavior in the U.S.
  - present marketing insights of consumer buying behavior, ex: the most sought-after type of avocado and how preferences vary by region across the U.S.

System Overview



Data

Historical Data

Record

Three years Data with more than 18k records and 14 features

Date

Average Price

Total Volume

Region

Feature selection

Record

Year

Month

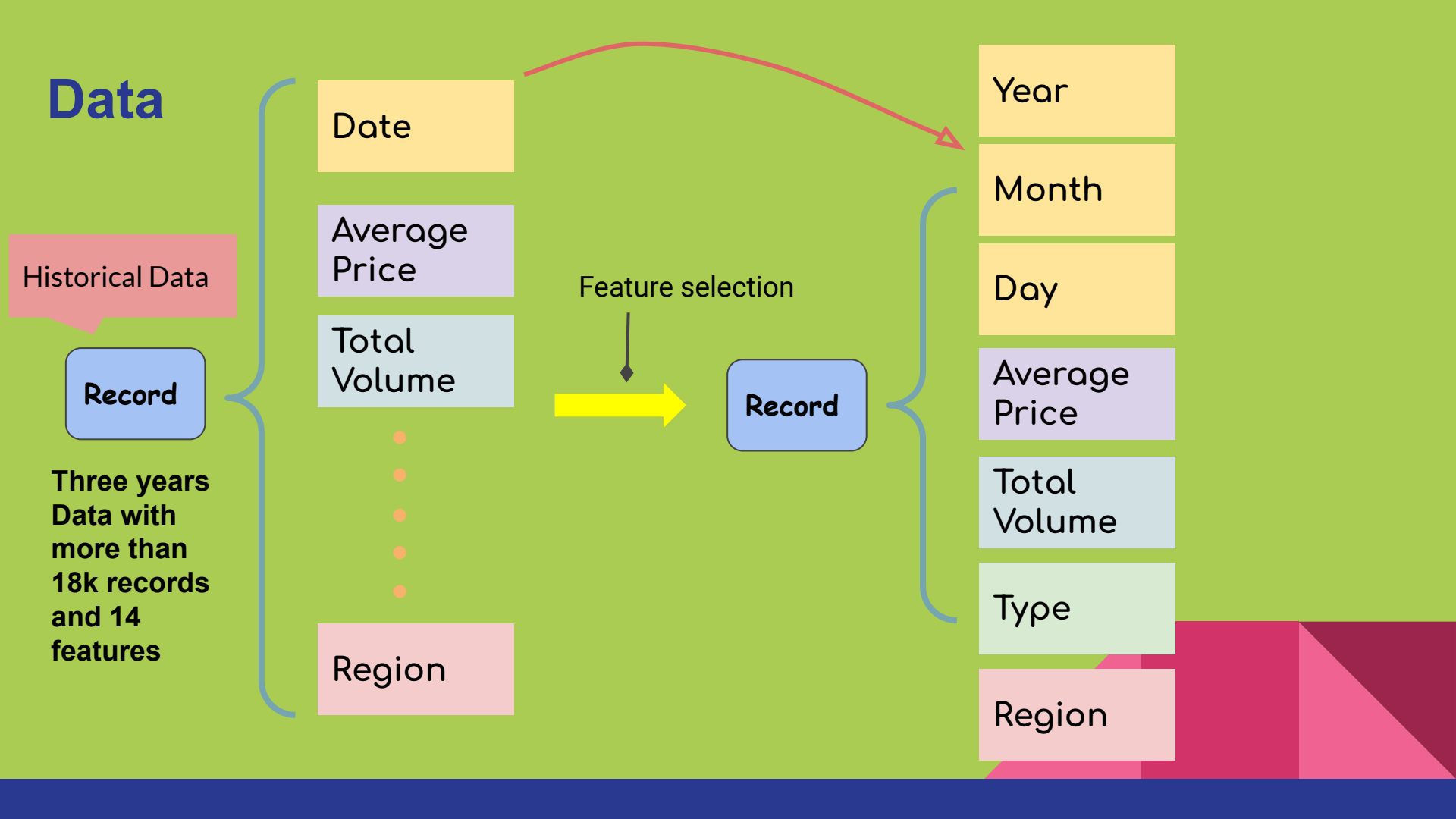
Day

Average Price

Total Volume

Type

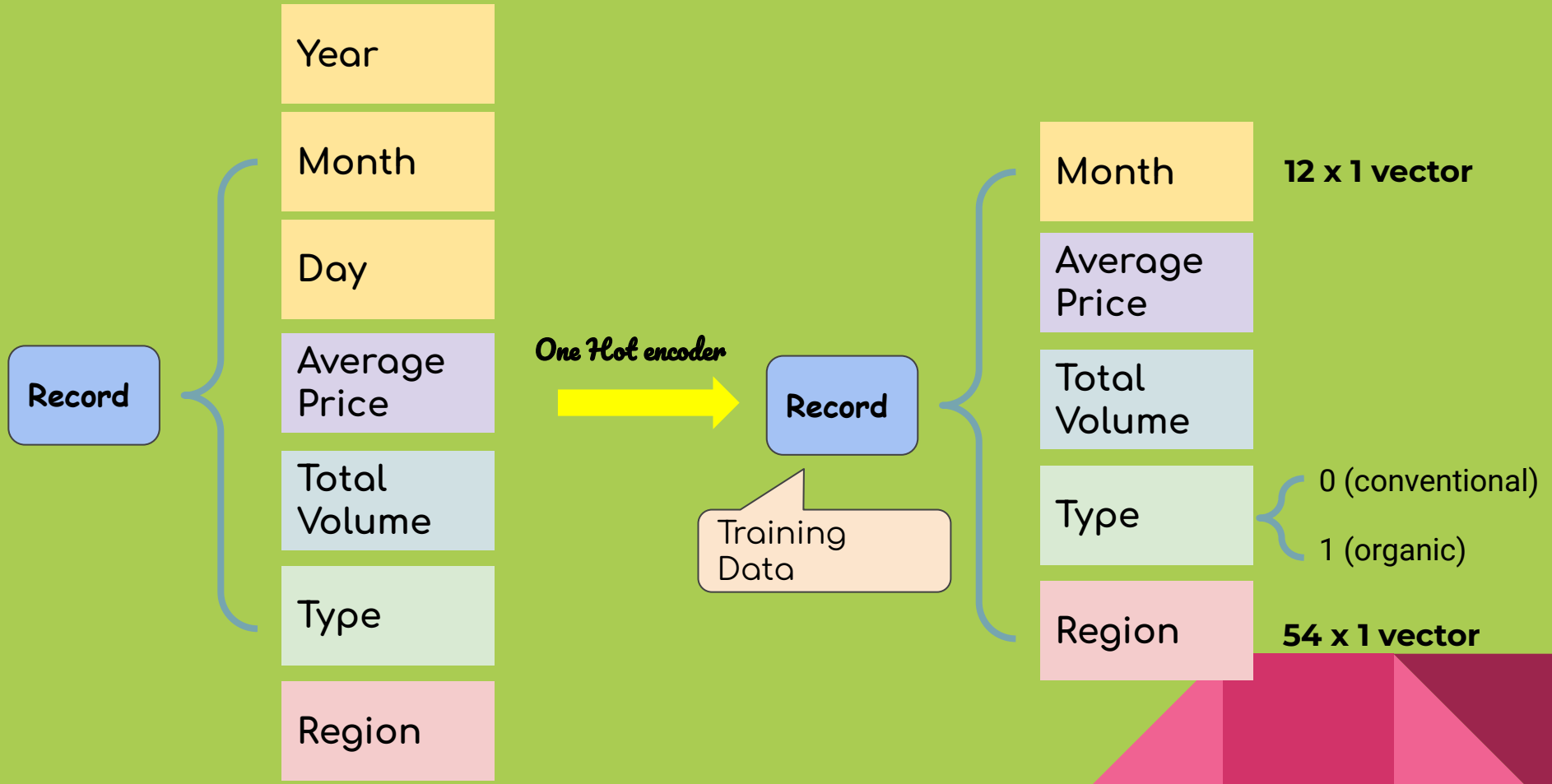
Region



Record

- Year
- Month
- Day
- Average Price
- Total Volume
- Type
- Region

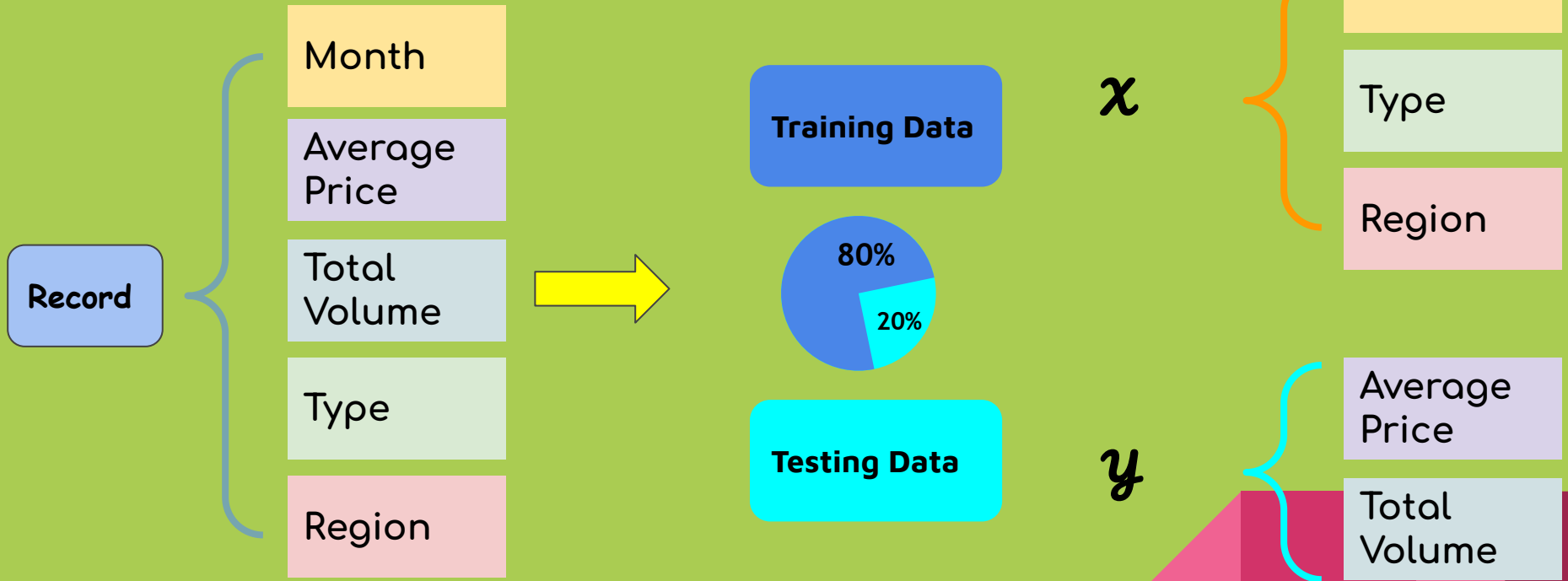
	AveragePrice	Total Volume	type	year	region	month
0	1.33	64236.62	conventional	2015	Albany	12
1	1.35	54876.98	conventional	2015	Albany	12
2	0.93	118220.22	conventional	2015	Albany	12
3	1.08	78992.15	conventional	2015	Albany	12
4	1.28	51039.60	conventional	2015	Albany	11
...	...	...	...	...	...	...
18244	1.63	17074.83	organic	2018	WestTexNewMexico	2
18245	1.71	13888.04	organic	2018	WestTexNewMexico	1
18246	1.87	13766.76	organic	2018	WestTexNewMexico	1
18247	1.93	16205.22	organic	2018	WestTexNewMexico	1
18248	1.62	17489.58	organic	2018	WestTexNewMexico	1


Record
Year
Month
Day
Average Price
Total Volume
Type
Region
One Hot encoder
Record
Training Data
Month
12 x 1 vector
Average Price
Total Volume
Type
0 (conventional)
1 (organic)
Region
54 x 1 vector

Model

- Linear Regression
- Bayes Ringe
- Ridge regression
- Random Forest
- Support Vector Machine
- Multilayer Perceptron

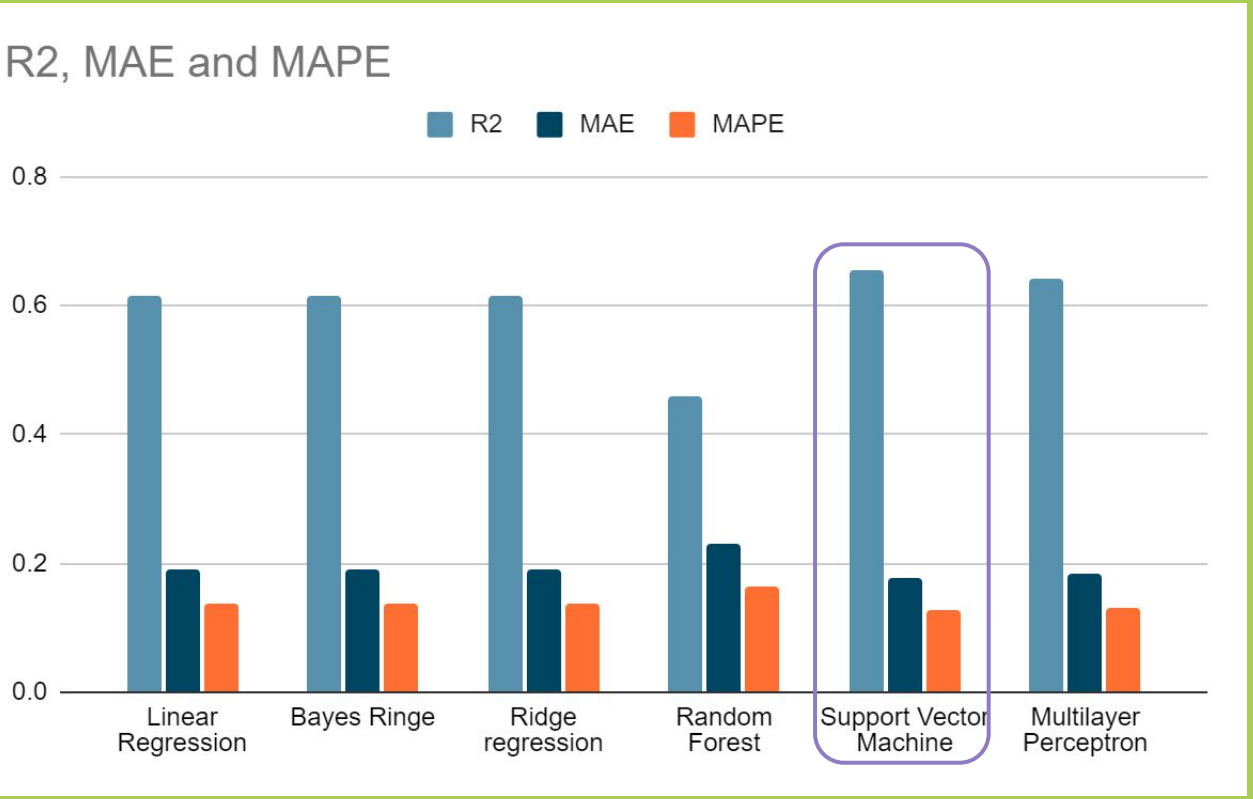
Training


Record
Month
Average Price
Total Volume
Type
Region
Training Data
80%
20%
Testing Data
x
y
Month
Type
Region
Average Price
Total Volume

Experiments

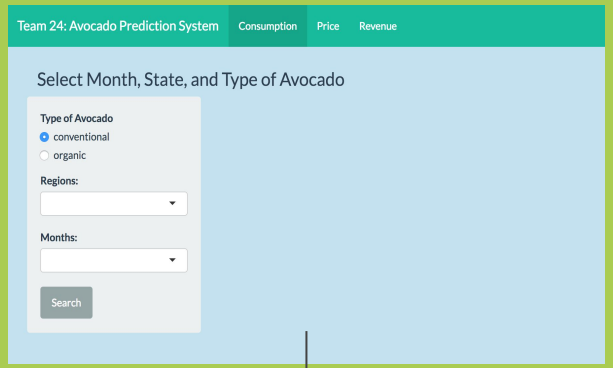
	Linear Regression	Bayes Ridge	Ridge regression	Random Forest	Support Vector Machine	Multilayer Perceptron
R2	0.6164	0.6162	0.6164	0.4587	0.6566	0.6417
MAE	0.1909	0.1909	0.1908	0.2293	0.1762	0.1851
MAPE	0.1362	0.1362	0.1362	0.1651	0.1278	0.1299

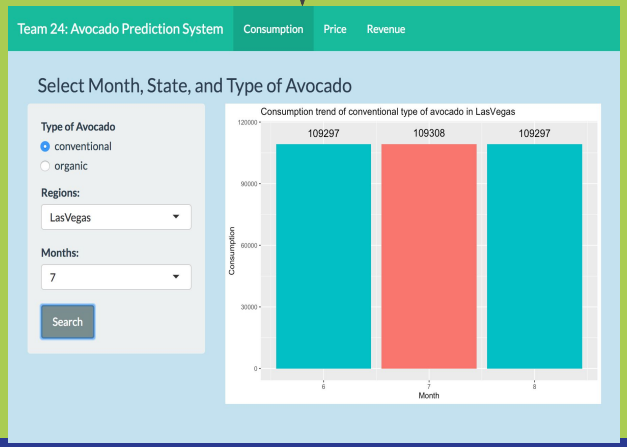
Experiments



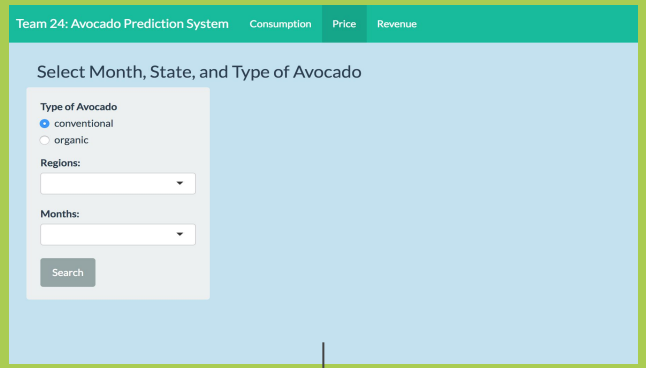
Prediction System Screenshot

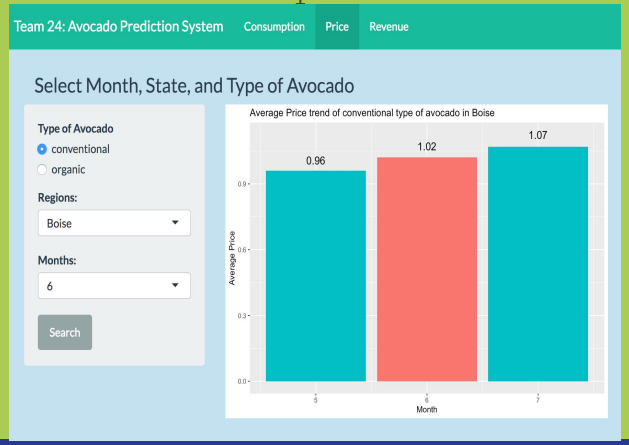
Consumption



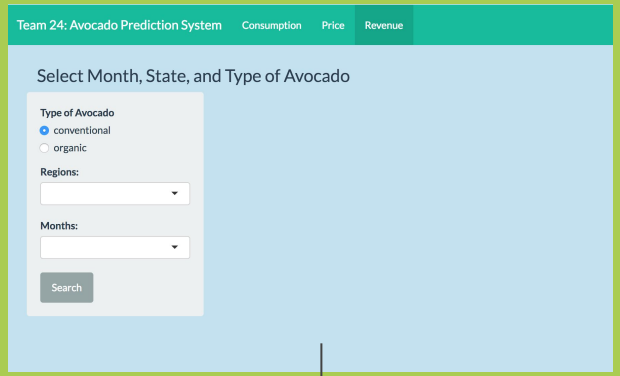


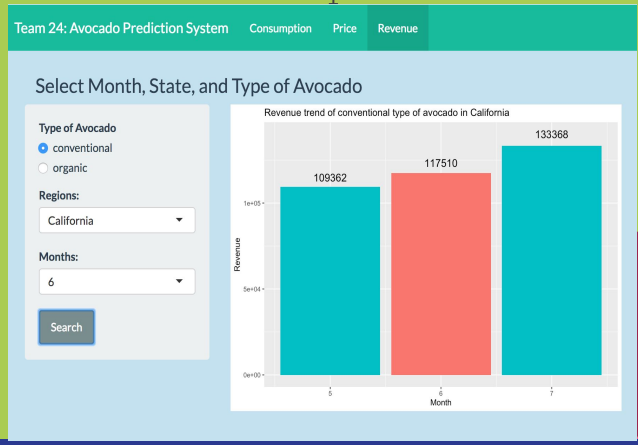
Price



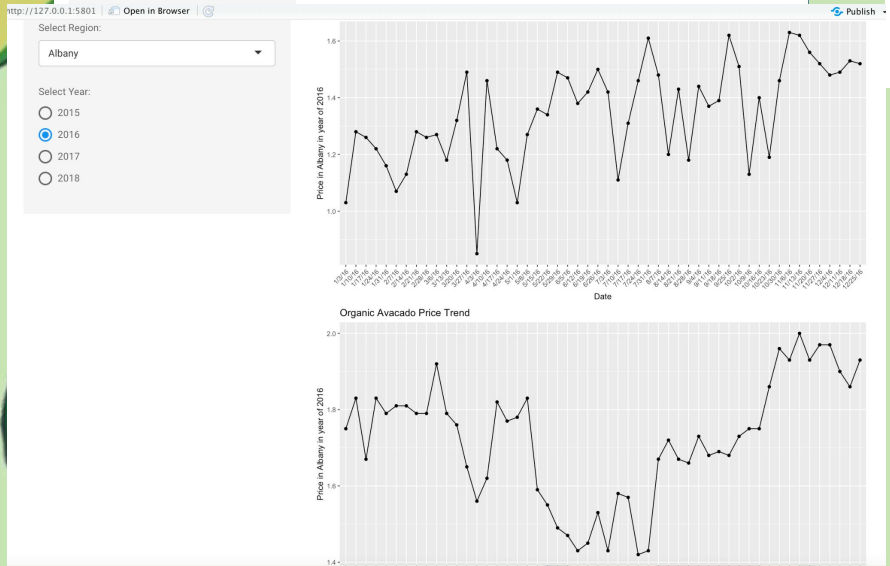


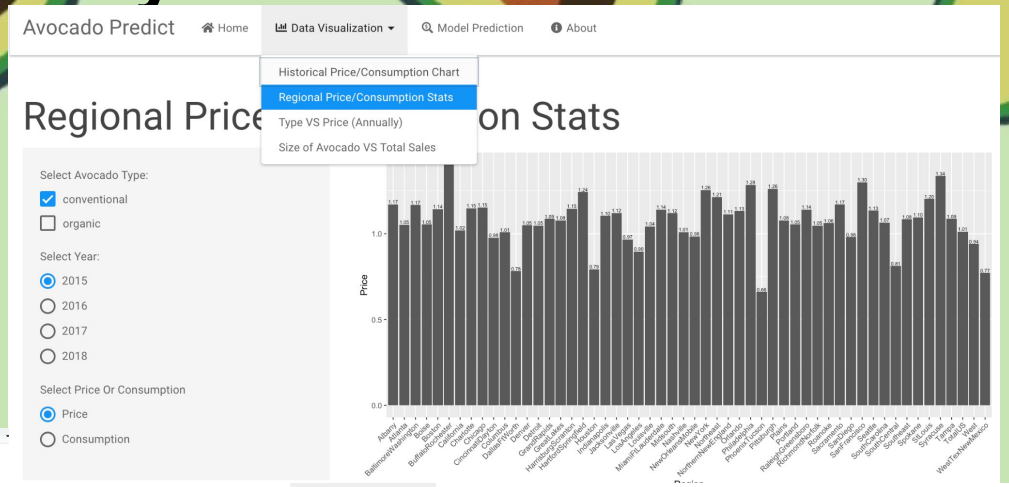
Revenue

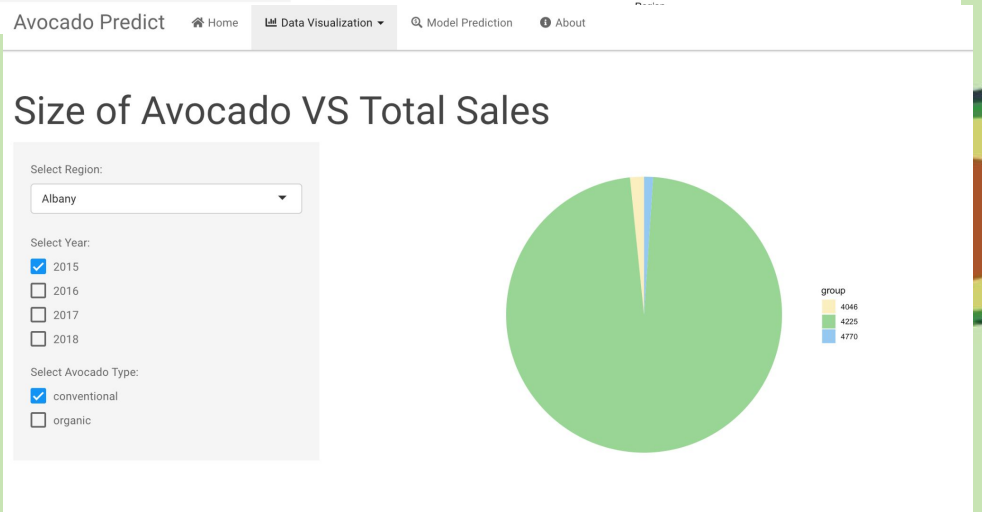




Data Visualization using Shiny







Reference/Related Works

[1] Rincon-Patino, Juan & Lasso, Emmanuel & Corrales, Juan. (2018). Estimating Avocado Sales Using Machine Learning Algorithms and Weather Data. Sustainability. 10. 3498. 10.3390/su10103498.

[2] Evans, E. A., & Ballen, F. H. (2015). An econometric demand model for Florida green-skin avocados. HortTechnology, 25(3), 405-411.

[3] Jones, M., Keyse, K., Melgoza, A., Perez, K., Qamar, T., Valladolid, J., & Woo, J. (2021). Avocado Buying Trends in the United States Using SAC. arXiv preprint arXiv:2104.04649.