

BUS 5100
Introduction
To Business
Analytics
Spring 2021

Avocado Buying Trends In The United States Using SAC

*Velma Jones, Kendra Keyse, Alfredo Melgoza,
Karen Perez, Tammy Qamar, Jason Villalpando*

Overview

I. Introduction

II. Related Work

III. Specifications

IV. Data Analysis

**V. Data
Visualization**

VI. Key Findings



Introduction

- Hass Avocado Board Data From 2015-2020
- <https://www.kaggle.com/timmate/avocado-prices-2020>
- Dataset Size: 3.37 MB
- File Format: CSV
- Conventional Avocados vs. Organic Avocados
- Pricing Trends
- Geographical Prevalence
- GitHub Link: https://github.com/velmajones/Bus5100-93_Group_6

Related Work

- **George Washington University**
 - Focuses on the correlation between price and volume sold for conventional and organic avocados
 - Our study extends the trend analysis to the three common avocado purchased throughout the United States
- **Agronometrics In Charts**
 - The installment tracked the changes in price over 2020 and attributed price anomalies to events
 - Our analysis extends to the end of 2020 and includes a time series and regression analysis

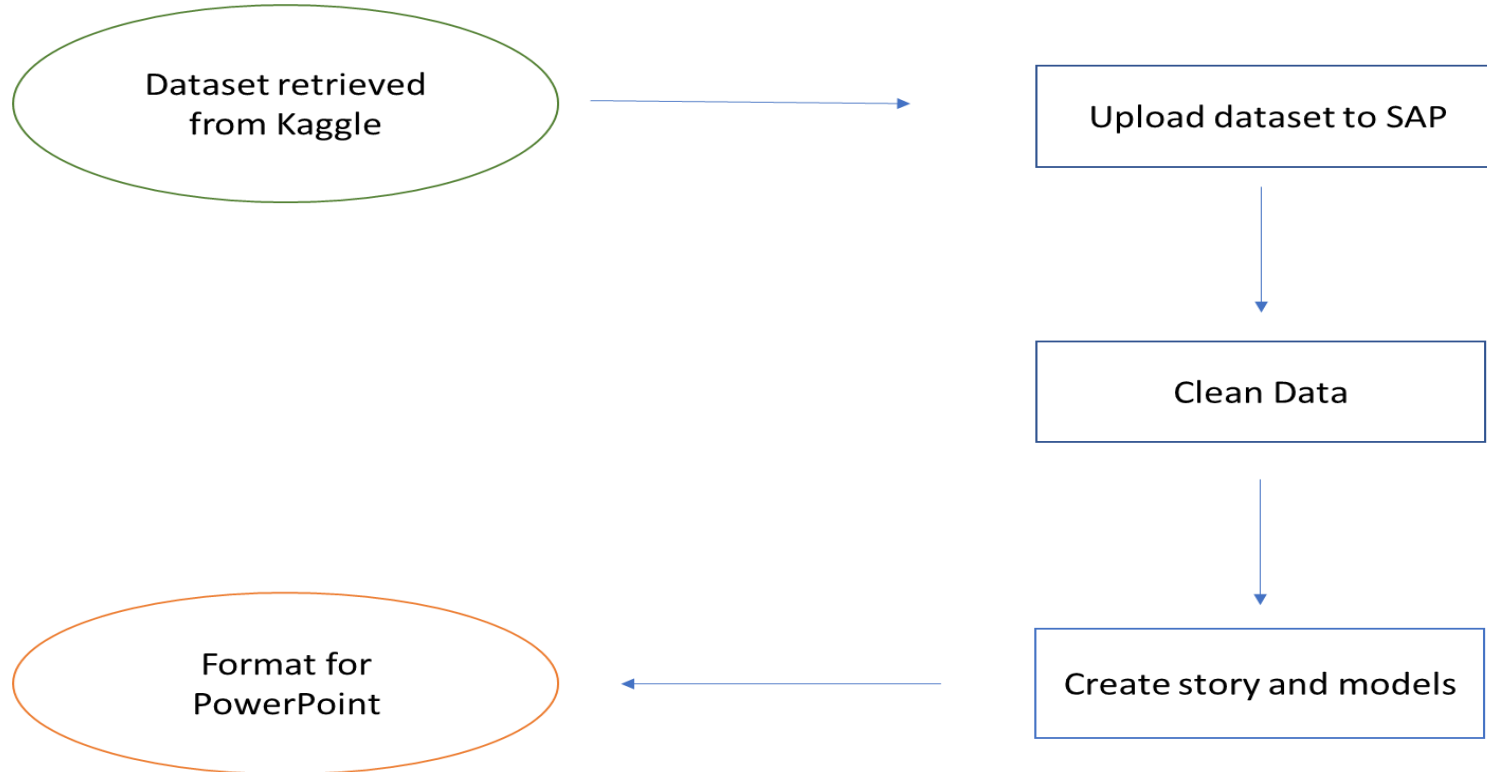
Specifications

Data Set - Avocado Size	Size (Total 3.37 MB)
4046 – Small/Medium Hass Avocado	1688 KB
4225 – Large Hass Avocado	1562 KB
4770 – Extra Large Hass Avocado	120 KB

- The data we used was retrieved from Kaggle
- The data has been updated on a consistent basis with data from Hass Avocado Board to be inclusive of 2015-2020
- The dataset consists of actual scan data from retailers' cash registers as well as multi-outlet reporting

Data Analysis

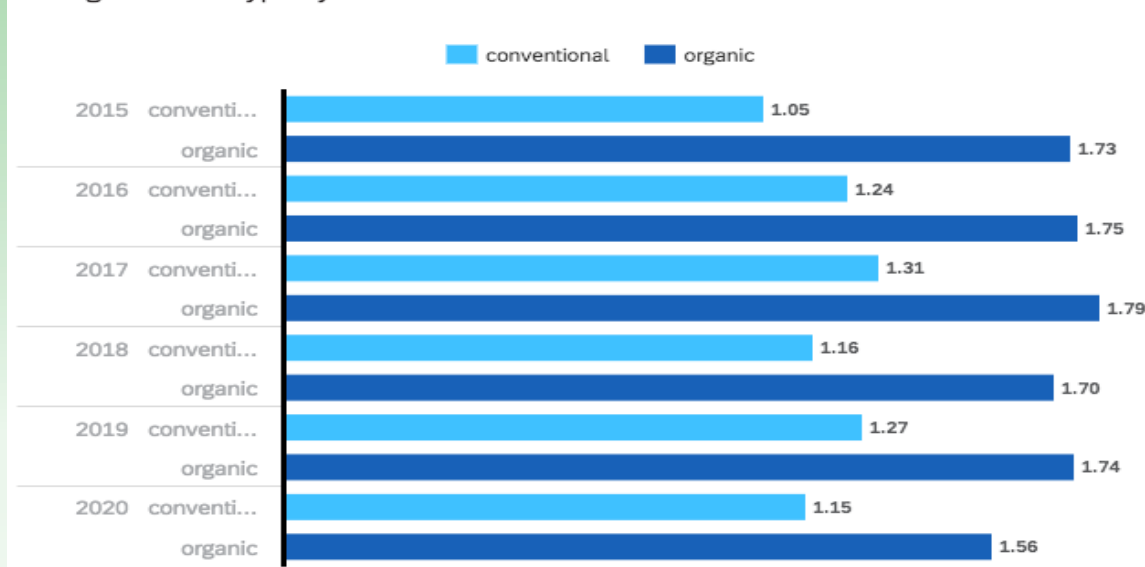
Implementation Flowchart



Data Visualization

Average Price Per Type & Year

Average Price of Type by Year



Data Visualization

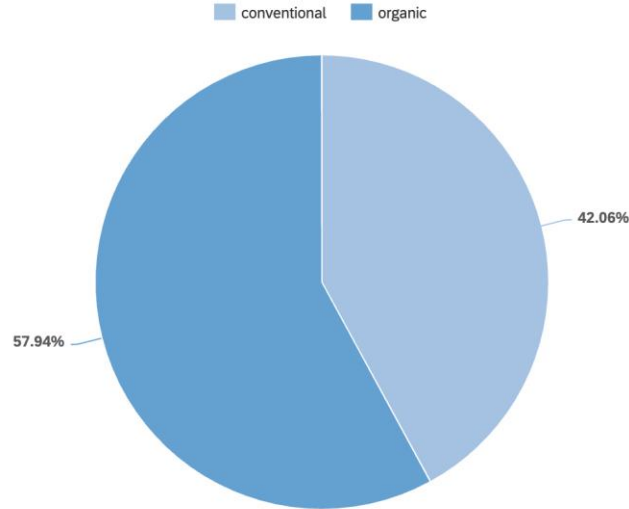
Average Price Per Type & Year

- Highlights difference in average price per type of avocado (conventional/organic) during the timeline of years 2015-2020
- Conventional increased slightly from \$1.05 in 2015 to \$1.15 in 2020
- Organic has trended downward from \$1.73 in 2015 to \$1.56 in 2020

Data Visualization

Average Price Per Type, Conventional & Organic

Average Price per type



Data Visualization

Average Price Per Type, Conventional & Organic

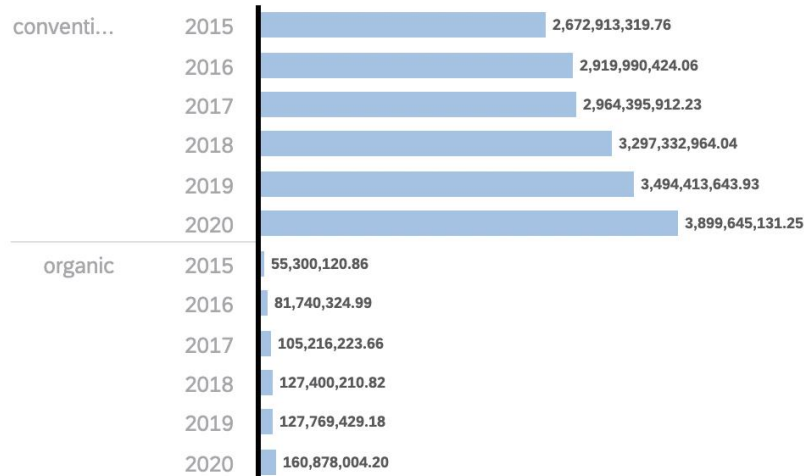
- Illustrates the average price per type of conventional and organic avocados
- Nearly 58% of organic avocado sales averaged \$1.80 per avocado
- Roughly 42% of conventional avocados averaged \$1.30 per avocado

Data Visualization

Total Volume Per Type & Year

Historical Data

total_volume per type, year



Data Visualization

Total Volume Per Type & Year

- The total volume output of conventional avocados far outnumbers the total volume output of organic avocados
- The total volume output of both conventional and organic avocados has grown steadily year after year confirming its popularity

Data Visualization

Time Series

Time Series Analysis and Forecast of Avocado Size and Total Volume Purchased



Data Visualization

Time Series

- The time series analysis and forecast of avocado size and total volume purchased over the course of one year
- 4046 represents small/medium Hass avocados, 4225 represents large Hass avocados, and 4770 represents extra-large Hass avocados
- The graph shows a gradual increase in total volume as the year progressed

Data Visualization

Total Volume Purchased By Geography

Total volume and average price

Location (total_volume)

Subtitle |  | 



Data Visualization

Total Volume Purchased By Geography

- The bubble color denotes the average price and the bubble size refers to total volume
- Syracuse, NY, and Boise, ID are the lowest two cities on the chart with approximately 24.1 million and 30.3 million, respectively
- As the highest-ranking city, Los Angeles contributes significantly with over 959 million purchases

Data Visualization

Predictive Price Model – Regression Analysis

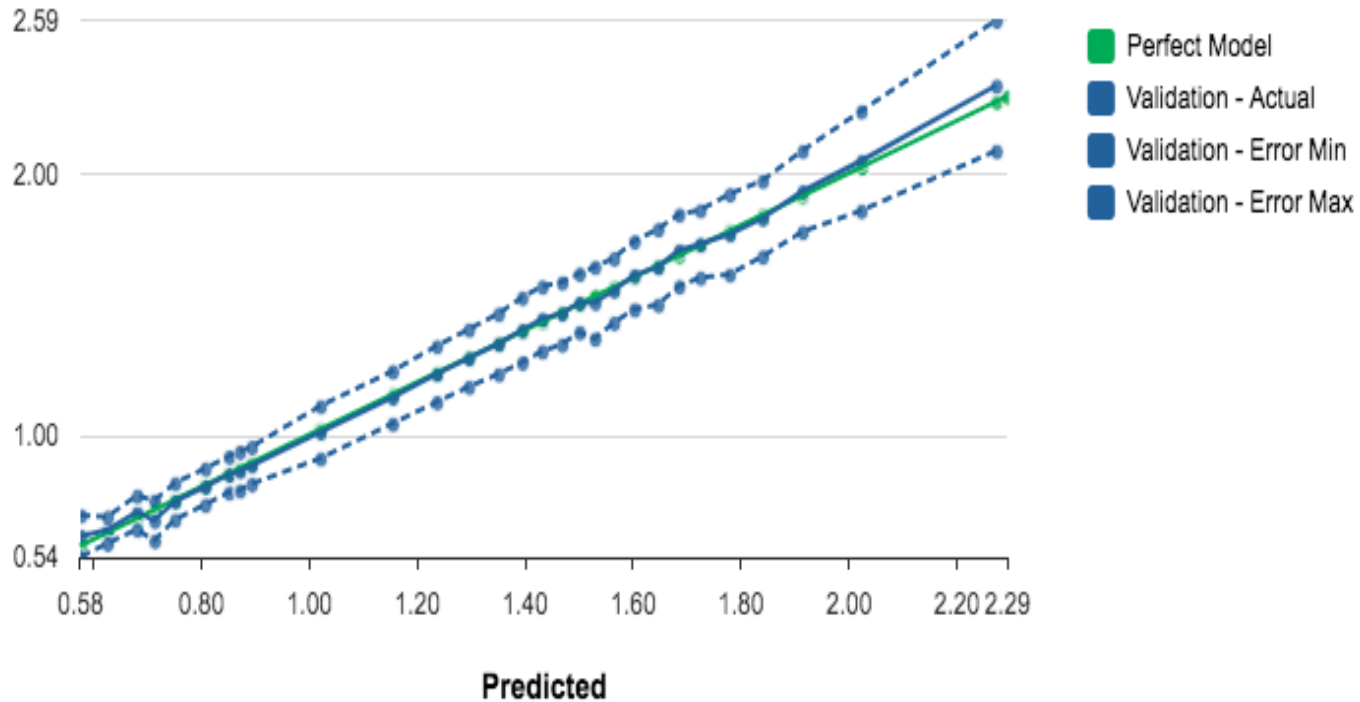
Global Performance Indicators

Root Mean Square Error (RMSE)	Prediction Confidence
0.12	98.29%

- Predictive Goal = Average Price of Avocados
- RMSE = .12 where the closer to 0, the better the model
- Prediction confidence = 98.29% which measures if the predictive model can do the predictions with the same reliability when new cases arrive [100% is ideal]

Data Visualization

Predicted vs. Actual



Data Visualization

- Validation – Actual is \$2.05 while the Predicted Value is \$2.02
- The regression analysis shows a Validation - Error Max of \$2.24 and a Validation - Error Min of \$1.86
- A Perfect Model would have shown \$2.03, so our model's \$2.02 prediction is very close

Measure: **Validation - Actual**

Predicted: **2.02**

Validation - Actual: **2.05**

Measure: **Validation - Error Max**

Predicted: **2.02**

Validation - Error Max: **2.24**

Measure: **Validation - Error Min**

Predicted: **2.02**

Validation - Error Min: **1.86**

Measure: **Perfect Model**

Predicted: **2.03**

Perfect Model: **2.03**

Key Findings

- The price of organic avocados is on average 35-40% higher than conventional avocados
- The sales volume of conventional avocados per year is on average 30 times bigger than that of the organic avocado sales
- Seasonality trends reveal that the highest point of sales take place in early February as well as the first week of May
- Geographical location plays a role in avocado buying behaviors

References

[1] Hass Avocado Board (March, 2021) Inside HAB.

Retrieved from <https://hassavocadoboard.com/inside-hab/>

[2] Holland, S. (July, 2019) Towards Data Science: Millennials' Favorite Fruit: Forecasting Avocado Prices with ARIMA Models. <https://towardsdatascience.com/millennials-favorite-fruit-forecasting-avocado-prices-with-arima-models-5b46e4e0e914>

[3] Colin, F. (June, 2020) Agronometrics in Charts: Avocado prices begin to recover as Mexico winds down. Retrieved from

<https://www.freshfruitportal.com/news/2020/06/02/agronometrics-in-charts-avocado-prices-begin-to-recover-as-mexico-winds-down/>

[4] Kaggle. <https://www.kaggle.com/timmmate/avocado-prices-2020>

[5] GitHub Link: https://github.com/velmajones/Bus5100-93_Group_6

[6] Avocado Story in SAC: [https://higher-](https://higher-education.us10.sapanalytics.cloud/sap/fpa/ui/tenants/81639/app.html#;view_id=story;storyId=E8A9304AA249809FEF2A1F903A0BFA9)

[education.us10.sapanalytics.cloud/sap/fpa/ui/tenants/81639/app.html#;view_id=story;storyId=E8A9304AA249809FEF2A1F903A0BFA9](https://higher-education.us10.sapanalytics.cloud/sap/fpa/ui/tenants/81639/app.html#;view_id=story;storyId=E8A9304AA249809FEF2A1F903A0BFA9)

[7] Predictive Model Regression Analysis in SAC: [https://higher-](https://higher-education.us10.sapanalytics.cloud/sap/fpa/ui/app.html#;view_id=smartPredict;scenarioId=2787)

[education.us10.sapanalytics.cloud/sap/fpa/ui/app.html#;view_id=smartPredict;scenarioId=2787](https://higher-education.us10.sapanalytics.cloud/sap/fpa/ui/app.html#;view_id=smartPredict;scenarioId=2787)

[8] Total Volume Purchased By Geography Geo-Map: [https://higher-](https://higher-education.us10.sapanalytics.cloud/sap/fpa/ui/app.html#;view_id=story;storyId=B6F2EB03635BEAD4220F09DA080EFF32;forceOpenView=true)

[education.us10.sapanalytics.cloud/sap/fpa/ui/app.html#;view_id=story;storyId=B6F2EB03635BEAD4220F09DA080EFF32;forceOpenView=true](https://higher-education.us10.sapanalytics.cloud/sap/fpa/ui/app.html#;view_id=story;storyId=B6F2EB03635BEAD4220F09DA080EFF32;forceOpenView=true)

THANK YOU

