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
- Health Prioritization Trends
- Pricing Trends
- Geographical Prevalence

# 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 and bag sizes 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

*Table 1. Avocado Size*

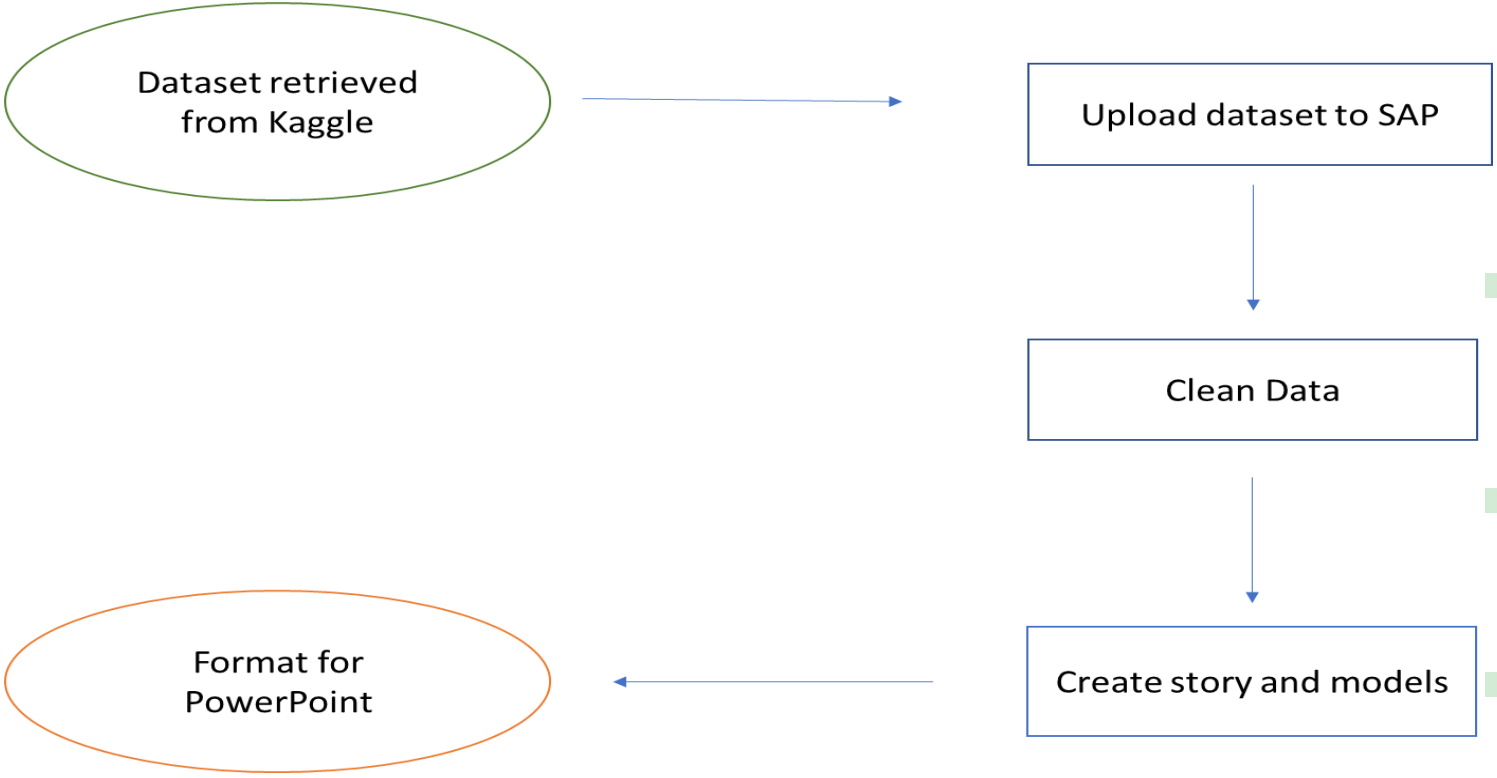
| **Data Set** | Size (Total 3.37 MB) |
|---|---|
| **4046** – Small/Medium Hass Avocado | 1688 KB |
| **4225** – Large Hass Avocado | 1562 KB |
| **4770** – Extra Large Hass Avocado | 120 KB |

*Table 2. Avocado Bag Size*

| **Data Set** | Size (Total 3.37 MB) |
|---|---|
| small_bags | 2312 KB |
| large_bags | 986 KB |
| xlarge_bags | 72 KB |

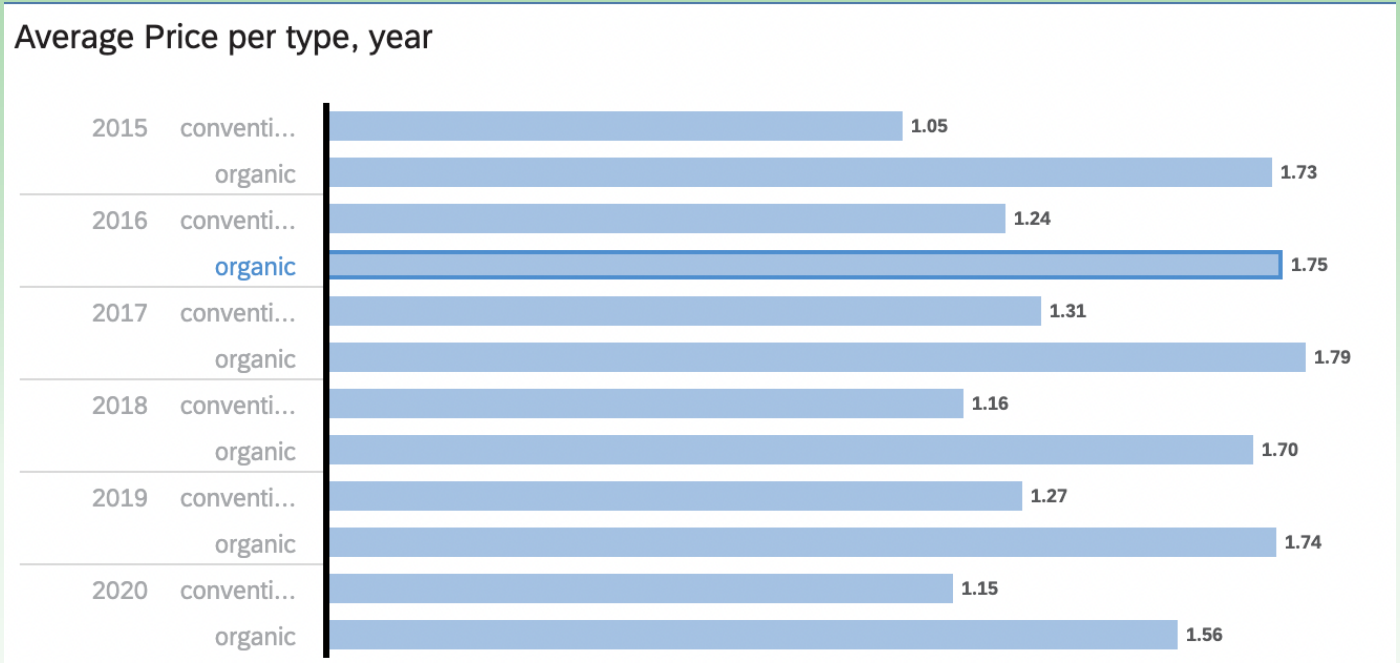
# Data Analysis

## Implementation Flowchart



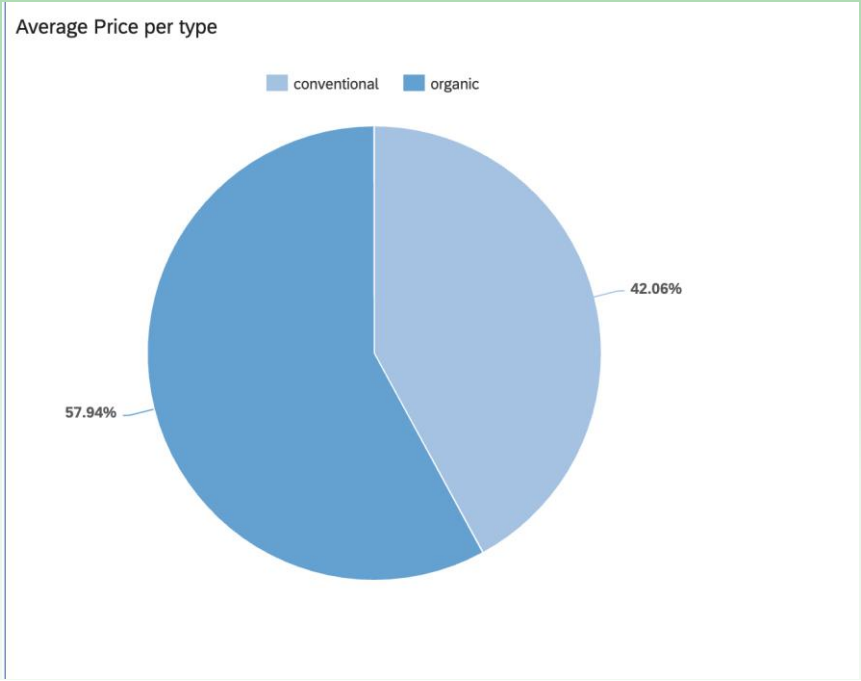
# Data Visualization

## Average Price Per Type & Year



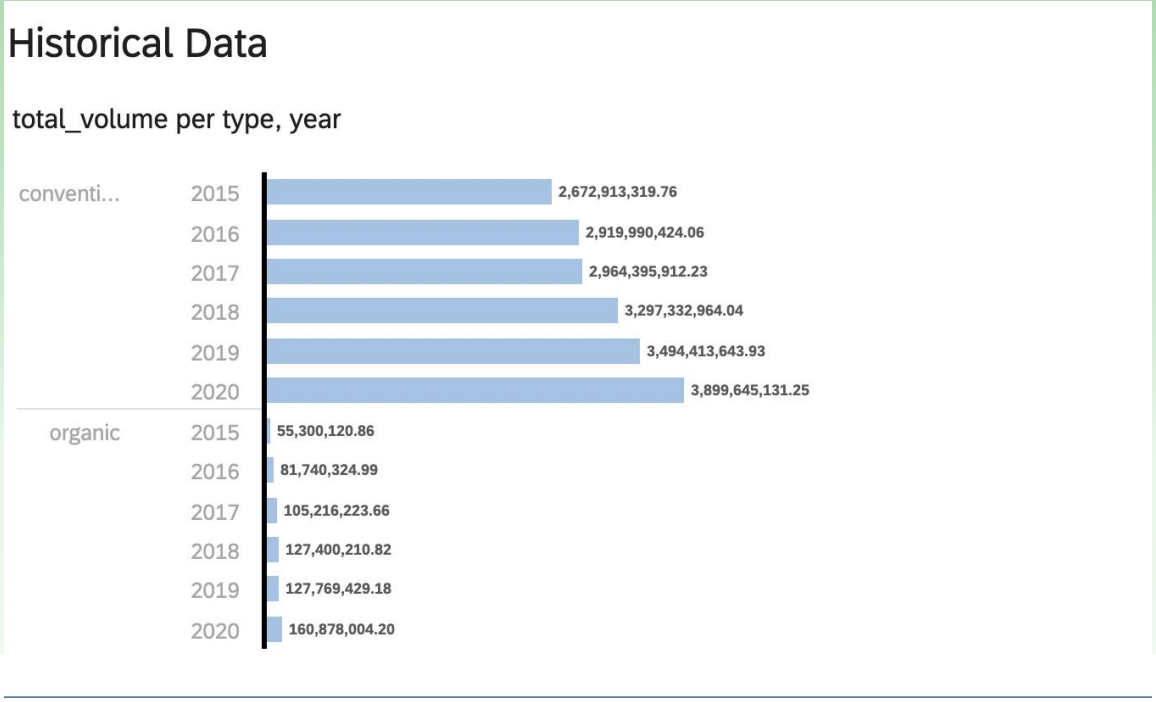
# Data Visualization

## Average Price Per Type, Conventional & Organic



# Data Visualization

## Total Volume Per Type & Year

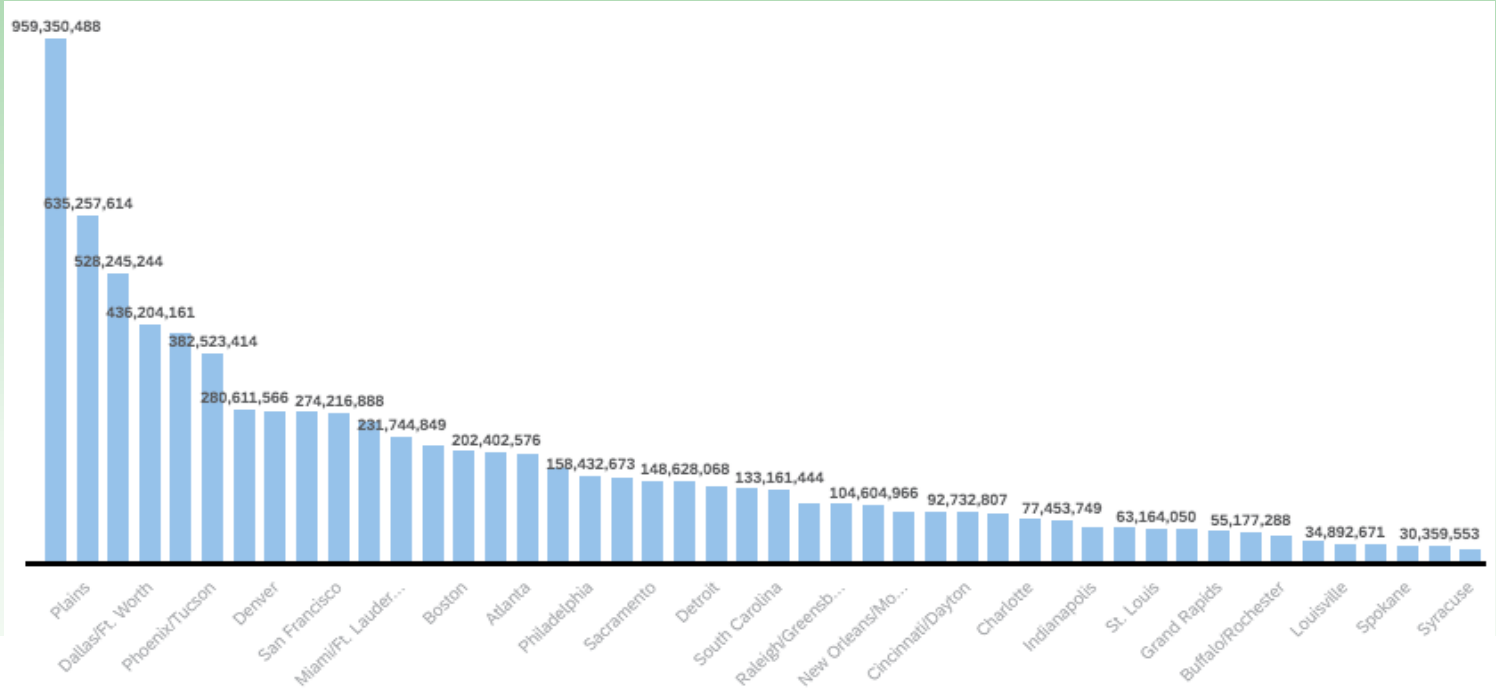


# Data Visualization

## Time Series

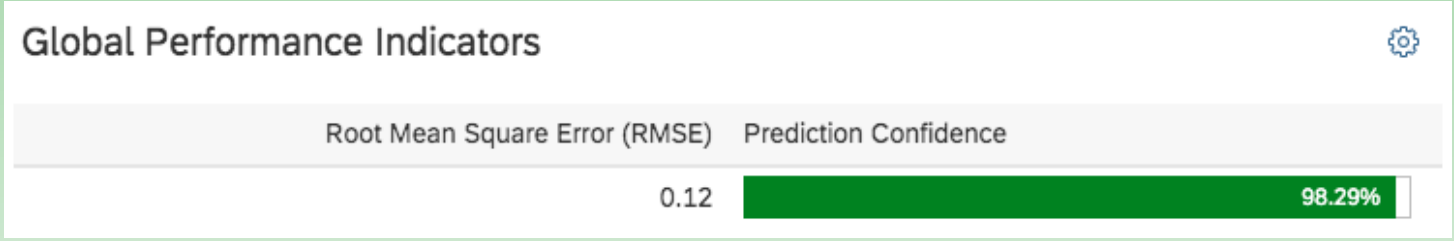
# Data Visualization

## Total Volume Purchased By Geography



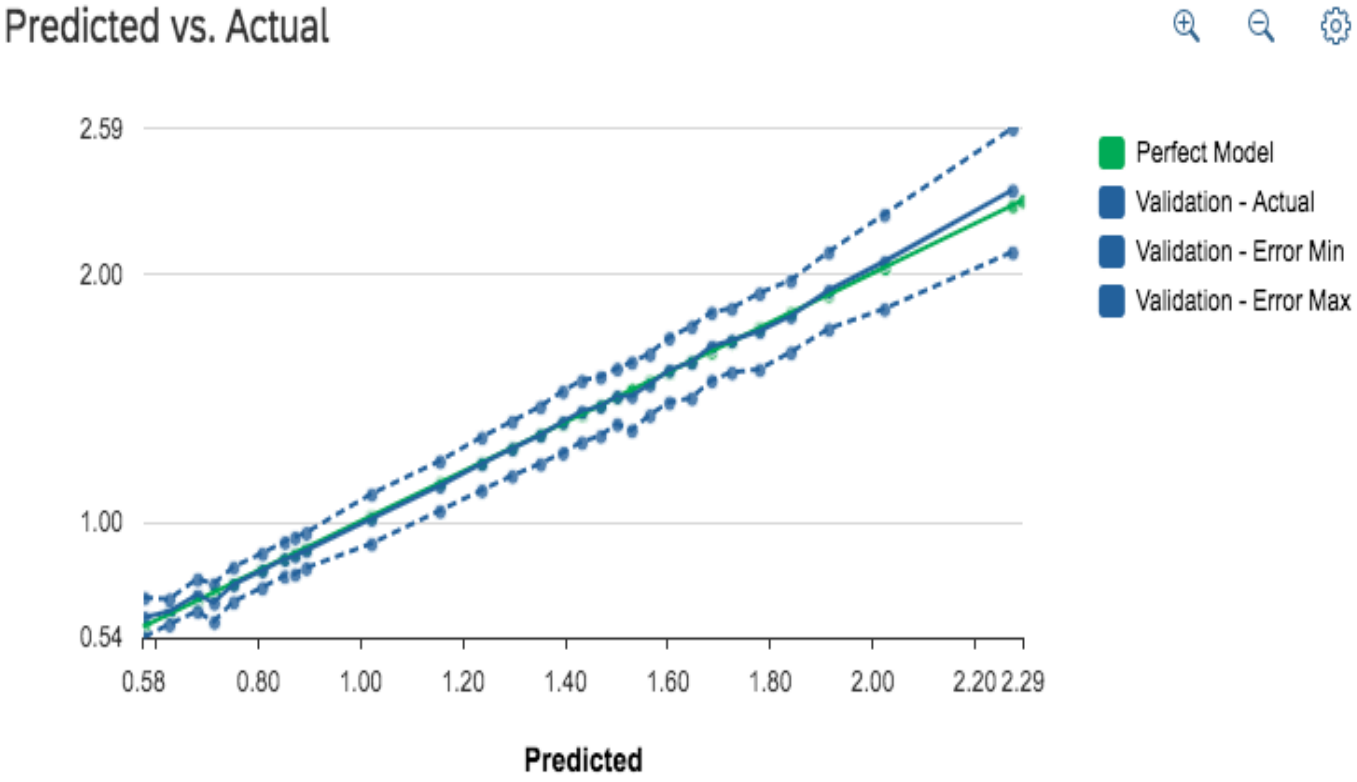
# Data Visualization

## Predictive Price Model – Regression Analysis




- 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



# Key Findings

- The price of organic avocadoes is on average 35-40% higher than conventional avocadoes.
- The sales volume of conventional avocadoes 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
- Climate, health-conscious actions, and geographical location play 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/timmate/avocado-prices-2020

# THANK YOU