

## Avocado Project

### Data Wranglers

Jovany G. ,Rashmi C. , Jorge M.

We were able to get a data set for avocados that had information for the years 2015 – 2020.

When we first came across the data set, we thought there was not enough data to play around with, were we wrong about that. We ended up dropping some incomplete data and decided to focus on the complete data instead.

When I first started to play with the data, I started to split it by year, size, and price to practice using Python. After I split the data, it made it easier to make visualizations. I was able to see the price spread between years, the sales by year and month, and even what regions and cities sold the most. Before the start of the project, I did not know there was a harvesting shortage with avocados which caused the prices to go up. Even with the higher prices it was concluded that the sales were not impacted all that much and people still went out to buy avocados, the highest price that I noticed was almost two dollars, now that is impressive.

Did Small avocados sell more than Large avocados based of Regions?

I was able to determine that there were more sales of small avocados than large avocados by a mean difference of 639,417. Even with the harvesting shortage, 2017 did not fall behind and the sales kept rising to 2019.