Avocado Rise

Due to the rise of vegetarianism, I was wondering if the price of the avocado, depending on the varieties increased over time and if it has an impact on the total volume sold as well as the volume sold for three different PLU codes:PLU4046:

- Hass smallPLU4225
- Hass largePLU4770
- Hass extra large

I downloaded the data from the years 2015 to 2021 on the website click here: HassAvocadoBoard (primary data source)

I then compiled them in one single sheet and cleaned the data, formatted the cells, and cleaned the region data as we had duplicates like “BaltimoreWashington” and “Baltimore/Washington” so I put all that in one single style. I then uploaded the dataset in my cluster on Databricks and started the analysis

QUESTIONS TO ANSWER

1. Does the price have an influence on the volume?

2. Which areas consumed avocados the most from 2015 to 2021?

3. The plu4770 being the biggest Hass variety, is it also the one consumed the most?

4. Is the trend in the most consuming city a mirror image of the general trend in the U.S?

HYPOTHESIS

1. Yes, the price has an influence on the volume and vice versa

2. Mostly the highly developped areas because of the vegetarian and vegan trend

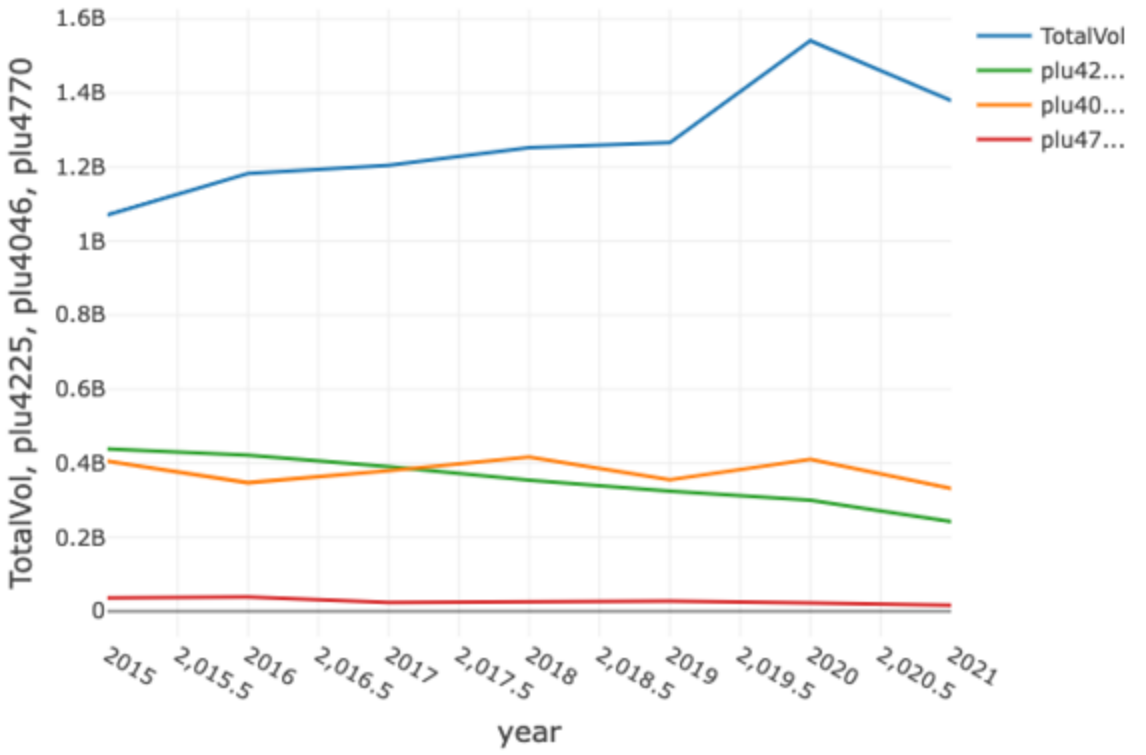
3. The consumer always going for the biggest, the cleanest, the most beautiful product, is more likely to choose the biggest avocado variety

4. Yes

DATA ANALYSIS

I decided to go for the metrics that interested me the most: the date, average price, the total volume, and the volume per PLU codes .I also noticed a problem with the “region” field: there are not only cities but also general areas like TotalUS, West, California, SouthCentral, Northeast, Southeast, GreatLakes, Midsouth, Plains so they had to be removed (or added) depending on what we wanted to know.First, let's see the general volume sold as well as the sales for each variety throughout the years.

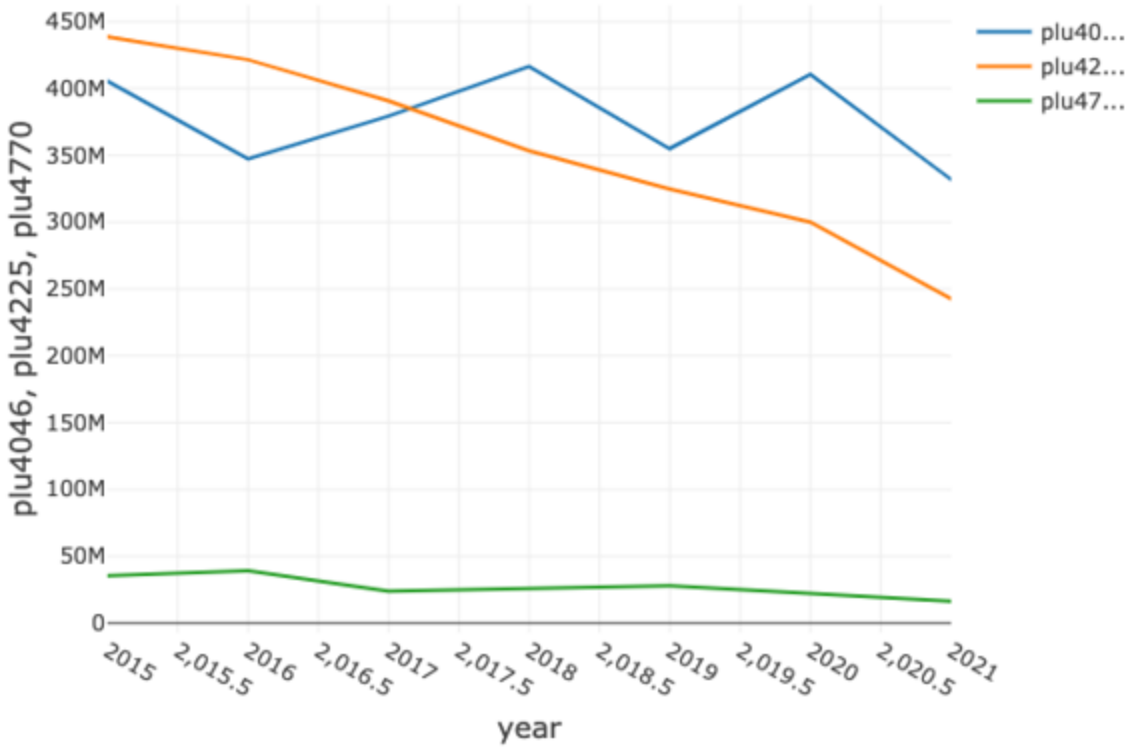
Total sales per year



It seems that the PLU4225(Hass big size) and PLU4046 (Hass small size) are chasing each other until 2018 when plu 4046 (the small variety of Hass avocado) became the most sold.Fun fact: it takes 70 liters to grow ONE avocado so on average. In 2020 for example, 1.54 billion avocados were sold, so 107;8 billion liters of water were needed to

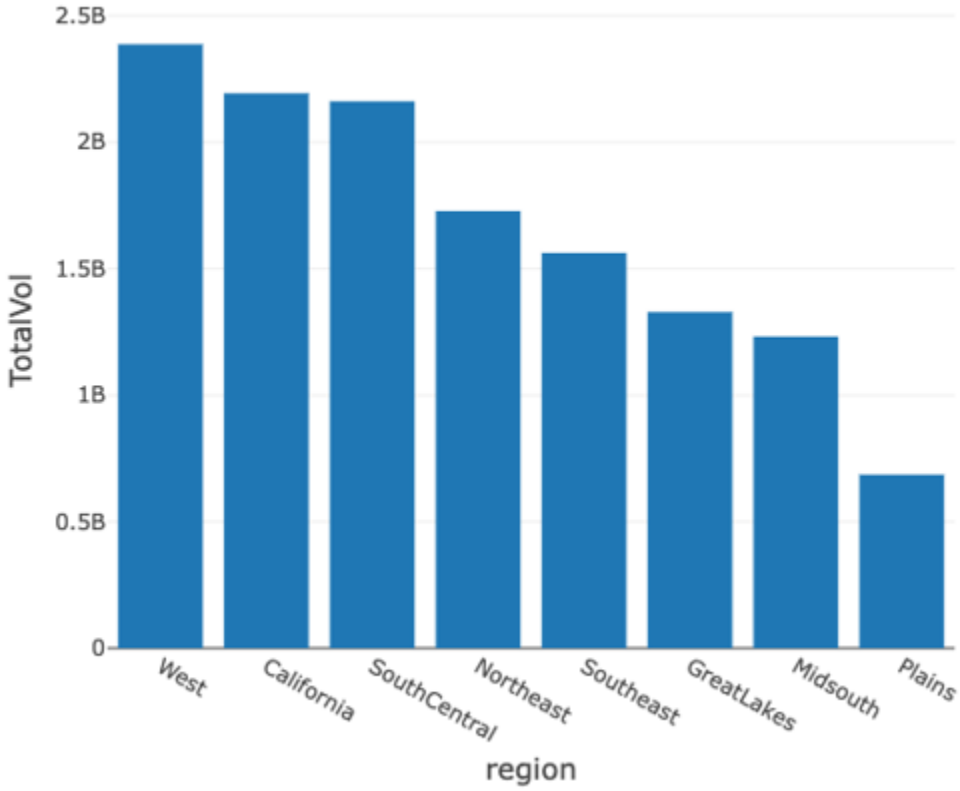
produce those.Source click hereLet’s have a proper look at each variety of avocados . It’s clear that the plu4046 is ahead of the game now.

Variety sales per year

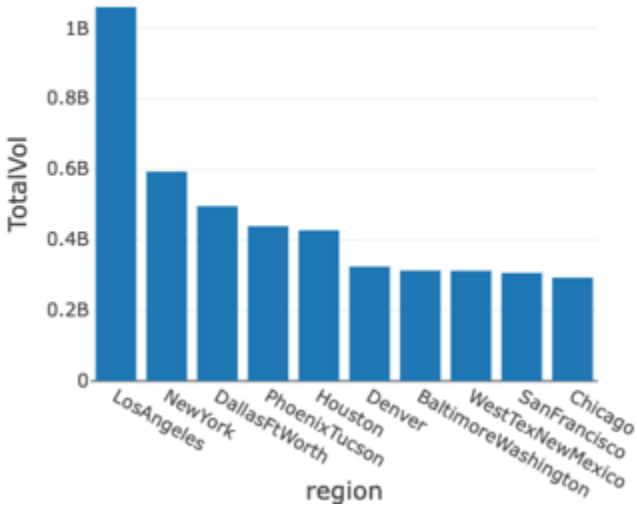


Now let's see and compare the areas, as I'd mentioned, in our dataset, we've got the data per city, as well as per region so we'll compare all that.

Total volume per region

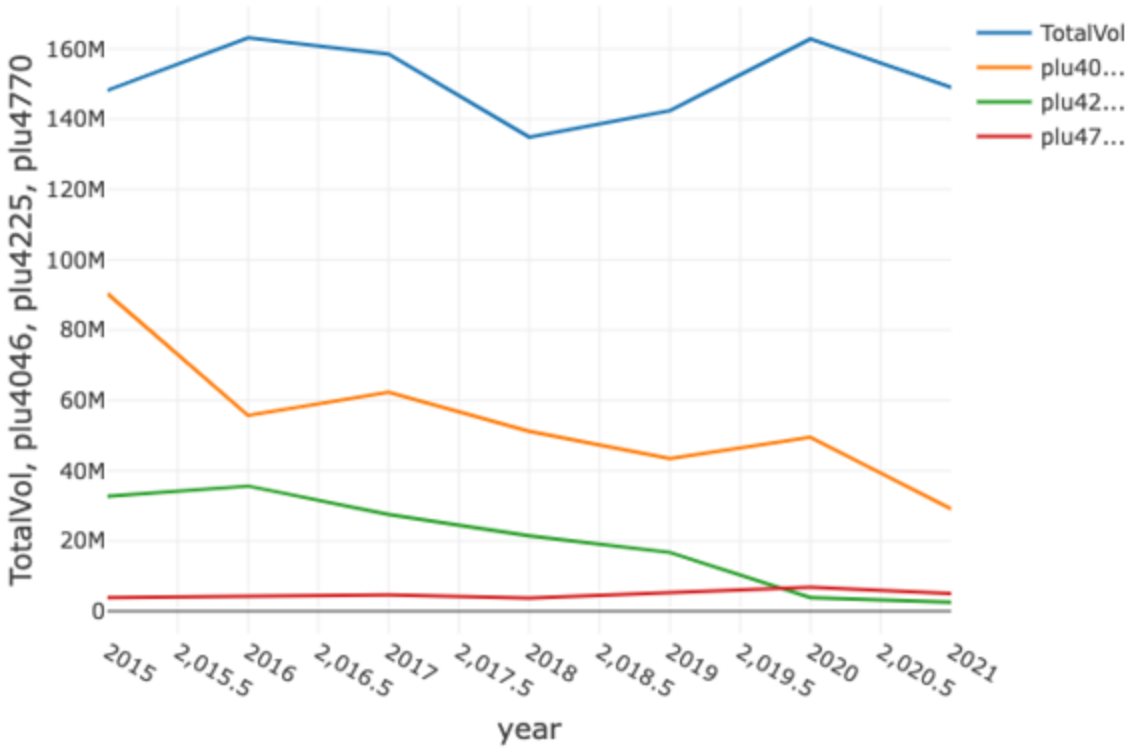


Total volume per city



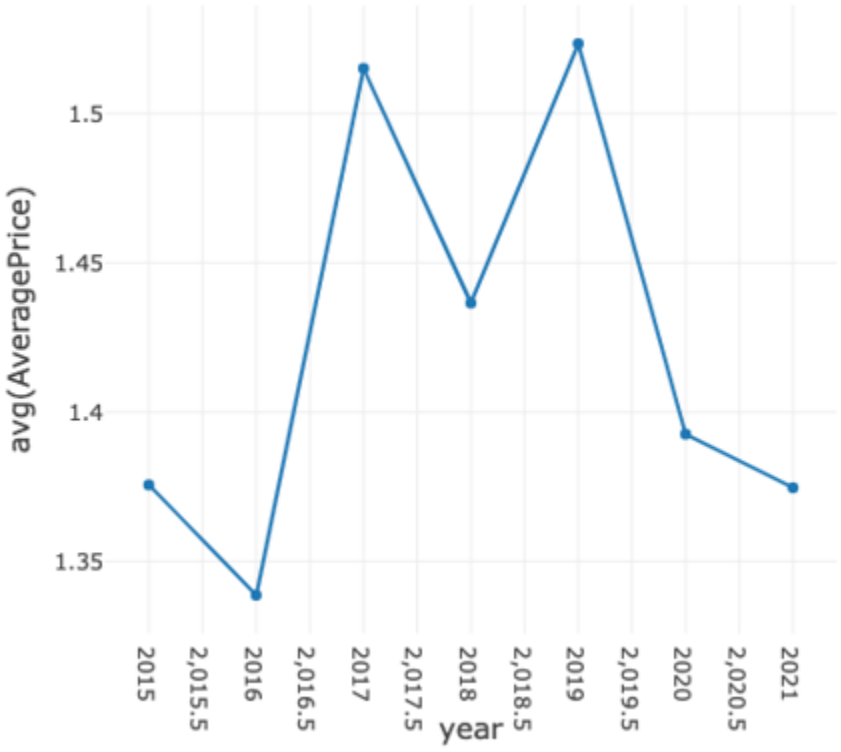
I also want to see if the sales from the biggest avocado consuming city: Los Angeles, is a mirror image of the general sales of the avocados in the U.S. When comparing the graphs below with the first one shown previously., we see it isn't the case

Units sold in Los Angeles

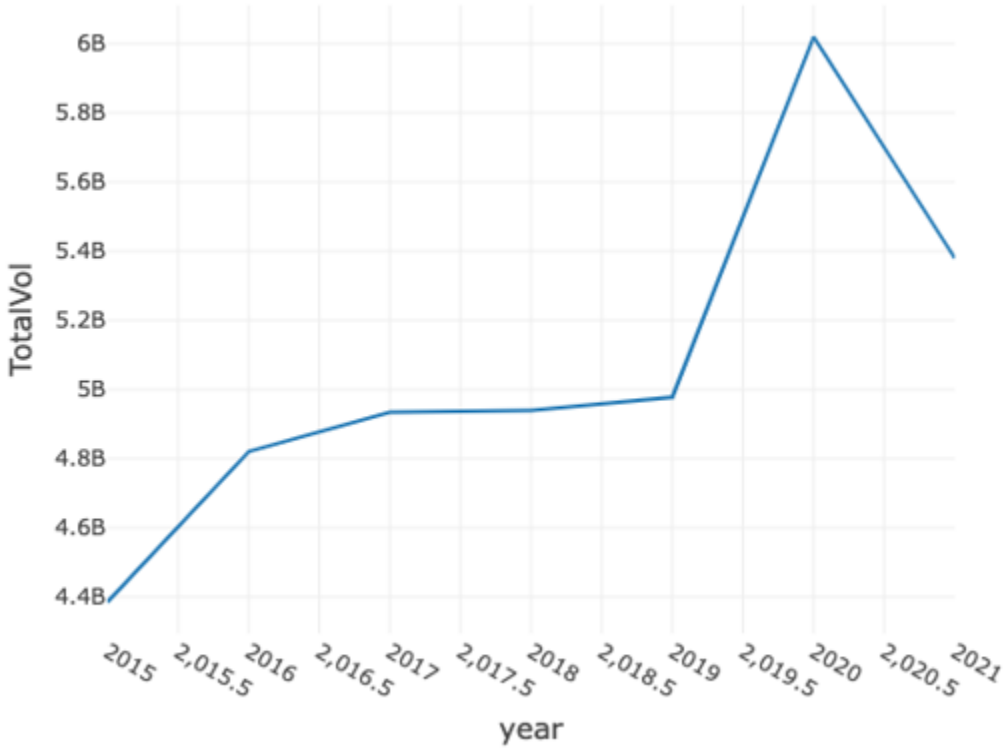


Lastly, I want to check if the price follows the sales trend.

Average price evolution



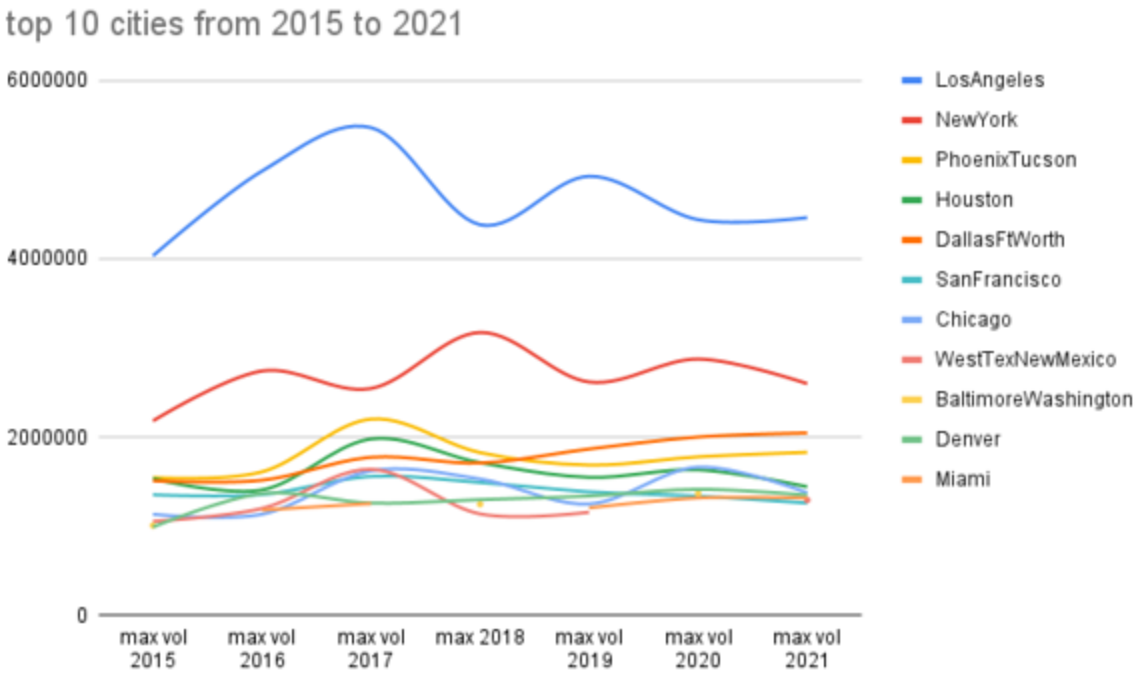
Unit sales evolution



What is the most amazing is from 2019 to 2020. the average price went from $1.5 to $1.3 but the sales went from 5 billions to 6 billions. From 2020 to 2021, the price continued to drop so one would think the consumption would continue to rise but instead it dropped. So we can conclude the price alone is not sufficient enough to predict the sales, the quality of the product, the trends, the taste of people, in 2019 the COVID appeared so that surely played an effect on the sales of avocados.

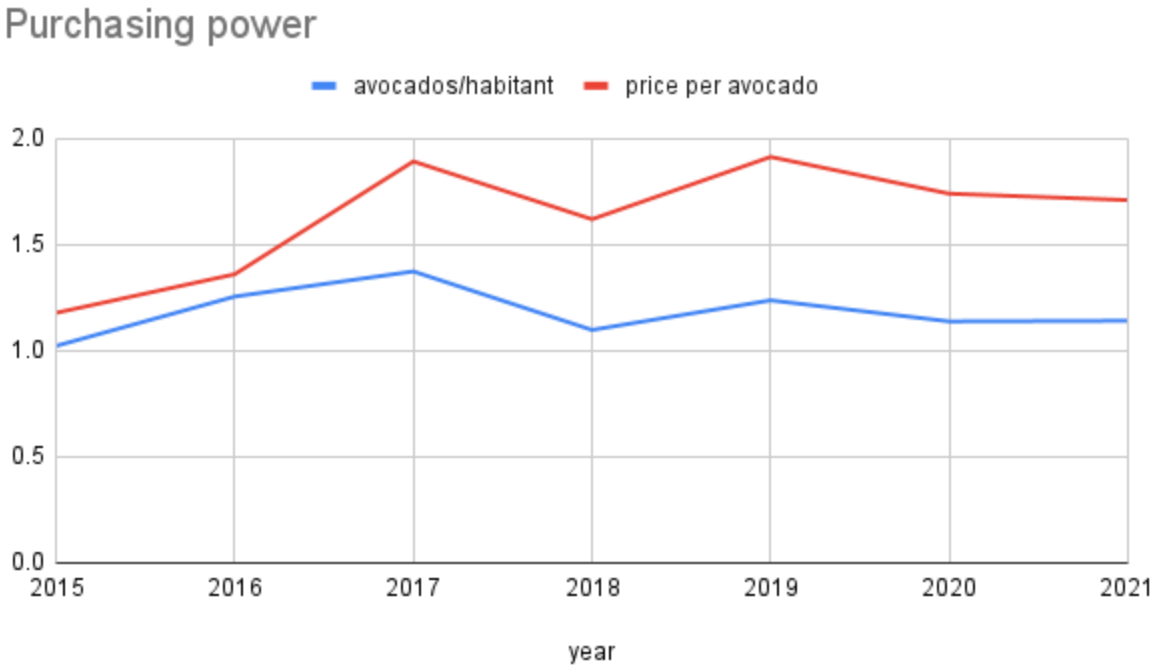
DEEPER ANALYSIS

I checked the top 10 cities in terms of volume from 2015 to 2021 and came up with the following graph on which we can clearly see that Los Angeles and New York are the two majors contributors to the sales of avocados in the U.S.



Finally, I wondered if the purchasing power of the inhabitants changed as the volume sold grew over the years. So I found the population in Los Angeles from 2015 to 2021 on census.gov (primary source) and made some calculations in order to figure out the number of avocados bought on average per habitant, and the number of avocados that can be bought

year	avg(AveragePrice)	avg(TotalVolume)	habitants	avocados/habitant	price per avocado
2015	1.1525	4031949.04	3943215	1.022502968	1.17843467
2016	1.08375	4982700.11	3969262	1.255321546	1.360454725
2017	1.377358491	5470227.08	3982002	1.373737904	1.892129565
2018	1.475334091	4381075.68	3990456	1.097888482	1.619752306
2019	1.546266184	4924667.45	3979576	1.237485463	1.913481925
2020	1.528983699	4437369.67	3898747	1.138152763	1.740217021
2021	1.49832849	4461087.4	3909274	1.141154956	1.709824982


Purchasing power
avocados/habitant
price per avocado
2.0
1.5
1.0
0.5
0.0
2015
2016
2017
2018
2019
2020
2021
year