Report on Capstone 2020

Chao Wang 19000151

6/4/2020
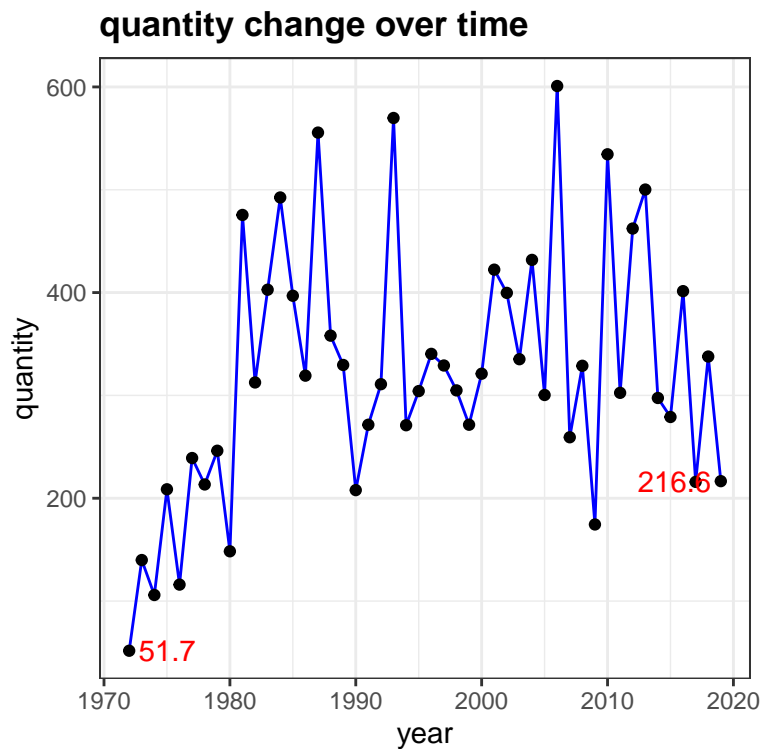
GBXQ3

In order to write a report on avocado data, I started an R project which is also stored in Github. 'https://github.com/Kennethws/capstone-stat-2020.git'

Question 1

First, I studied how the variables `quantity` and `price` vary over time, respectively.

On one hand, the following is a simple time series plot for `quantity` over time:

As shown in the plot above, the quantity increased from 51.7 million pounds in 1972 to 216.6 million pounds in 2019 with dramatic fluctuation. This implies the development of

productivity as well as the surging demand worldwide.

On the other hand, below is the time series plot for `price` over time:

price change over time
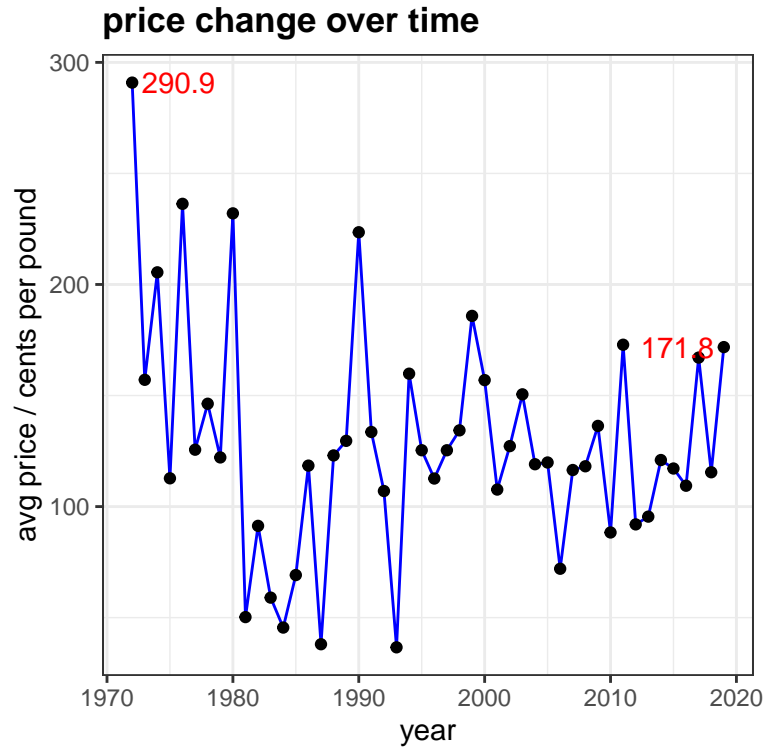
We can therefore conclude from the graph that the average price suffered a great loss from 290.9 to 171.8 cents per pound between 1970 and 2019, which was also accompanied by huge fluctuation.

Finally, I looked closer into the relationship between `quantity` and `price` by plotting a scatterplot supplemented by Pearson and Spearman correlation coefficients.

price vs. quantity

Pearson: −0.83
Spearman: −0.79

avg price / cents per pound

quantity / million pound
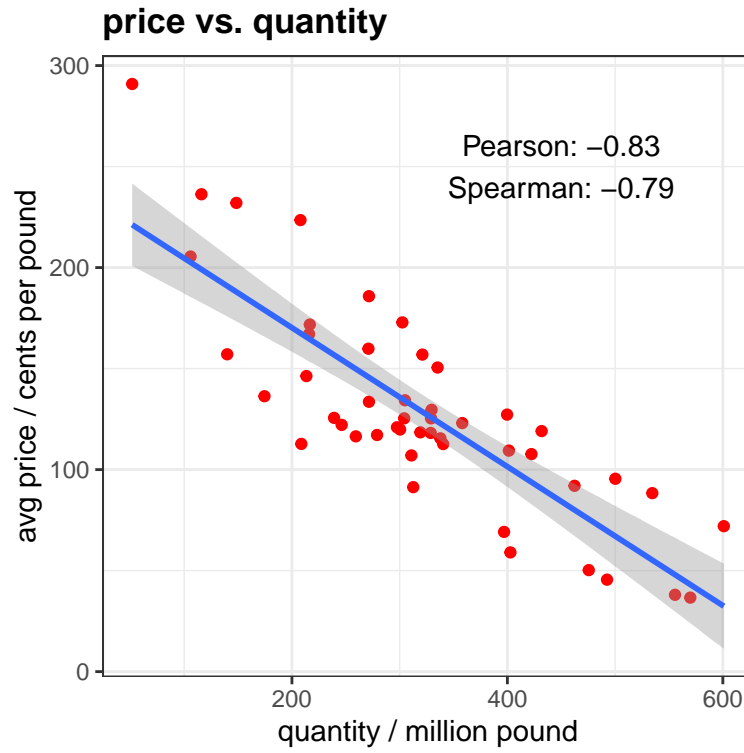
It seems safe to state that `quantity` and `price` are negatively correlated to a great extent.