Lab 2

In this lab, we'll work on reporting and interpreting some of the basic summary statistics, calculate and interpret z-scores, and apply the Empirical Rule when reasonable.

Getting the Data

In this lab, we'll be looking at the Avocados data set. Download the `avocados.csv` file from Canvas. Open the data set in jamovi, and we'll be ready to begin! Attached to this lab is some data documentation to explain what each column represents and how to read the measurement.

Tasks

Overall Price Summary

- Create a table of descriptive statistics for `AveragePrice`, reporting mean, median, standard deviation, inter-quartile range, and the 25th, 50th, and 75th percentiles.
- Create a histogram of the `AveragePrice`.

In Your Writeup

> Include the table. Interpret the mean, median, standard deviation, inter-quartile range, and the percentiles in the context of avocado prices.
>
> Are the mean and median different from each other? Why?
>
> Include your histogram for `AveragePrices`. Does it seem like the Empirical Rule will apply? What about the histogram leads you to believe that? What would the Empirical Rule say, specifically in this context?

Price by Type

We want to investigate the price of avocados based on their type: organic or not.

- Create a table of descriptive statistics for `AveragePrices` split by `type`, reporting the mean and standard deviation.
- Create histograms of `AveragePrice` split by `type`.
- Create boxplots of `AveragePrice` split by `type`.

In Your Writeup

> Include the table. Explain what the difference in mean and standard deviation mean between the two types of avocados.
>
> Include the histograms. Interpret those differences in mean and standard deviation in the context of your histograms.

> Include the boxplots. Make a similar comparison as above, noting the differences in center (median) and spread (IQR).

Price by Year

We want to investigate the price of avocados based on the year.

- Create a table of descriptive statistics for `AveragePrices` split by `year`, reporting the mean and standard deviation.
- Create histograms of `AveragePrice` split by `year.`
- Create boxplots of `AveragePrice` split by `year.`

In Your Writeup

> Include the table. Explain what the difference in mean and standard deviation mean between the different years. Are there any years where it seems like avocado prices acted differently? Which ones?
>
> Include the histograms. Interpret those differences in mean and standard deviation in the context of your histograms.
>
> Include the boxplots. Make a similar comparison as above, noting the differences in center (median) and spread (IQR).

Price by Year and Type

- Create a table of descriptives for `AveragePrices` split by both `year` and `type`. Again, report the mean and standard deviation.
- Create histograms and boxplots for `AveragePrice` split by `year` and `type`.

In Your Writeup

> Include both the table and the histograms or boxplots.
>
> Interpret what happens to the price of organic avocado prices from 2015-2018, using means and standard deviations.
>
> Interpret what happens to the price of conventional avocado prices from 2015-2018, using means and standard deviations.

Standardizing Prices (z-scores)

- In the Data pane, create a new variable for standardized prices.
  - Add a computed variable, named `AveragePrices_standardized`.
  - Use the formula `(AveragePrice - 1.41)/0.403`, where 1.41 is the overall mean of `AveragePrice` and 0.403 is the overall standard deviation of the `AveragePrice` (no matter what the region).
- Create a table of descriptives for `AveragePrices` and `AveragePrices_standardized` reporting the means and standard deviations, split by `type`.

In Your Writeup

> You do not need to include the whole table, but report the means of `AveragePrices_standardized` split by `type`. These are z-scores: interpret each. Why are these values not 0?
>
> Similarly, why are the standard deviations for each `type` not 1?
>
> Feel free to reference the overall mean and standard deviations that we used to calculate these z-scores, and reference the histograms of price by type (from earlier).

Writeup

Your writeup should include a short introduction into the overall tasks we'll be doing in this lab. You can organize your writeup with headings for each task. You can copy-paste tables, and save graphs as images by right-clicking on each and selecting "Export..." and saving as a .png file. These can be included into your writeup then.

Submit it as a physical (printed) copy in class with your name on it.

Column Names	Description
`Date`	The date of the observation
`AveragePrice`	The average price of a single avocado
`Total Volume`	Total number of avocados sold
`4046`	Total number of avocados sold with PLU 4046: non-organic small/medium Hass Avocados
`4225`	Total number of avocados sold with PLU 4225: non-organic largeHass Avocados
`4770`	Total number of avocados sold with PLU 4770: non-organic extra large Hass Avocados
`type`	Type of avocado: conventional or organic
`year`	The year of the observation
`region`	The city or region of the observation