

Lab 1

Carmen Canedo

6 October 2020

Exercise 1

Read in `avocadoes.csv` and save into data frame.

Exercise 2

Change column names to small, large, and extra large.

Exercise 3

The data is not tidy, I will need to use `pivot_longer()` to tidy it. The size variable is split into small, large, and extra large.

Exercise 4

Create a function to read in the avocado data, renames columns, writes the transformed, and returns transformed data

Exercise 5

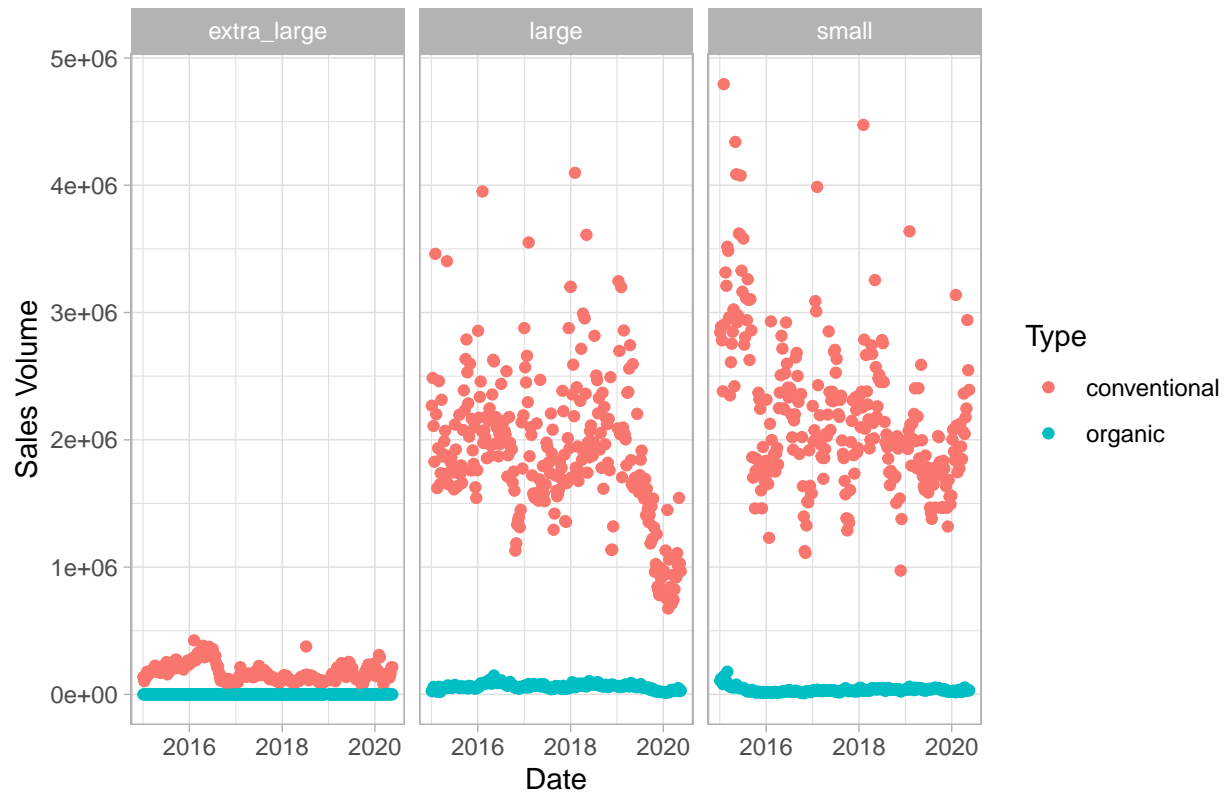
Create table with top 5 geographies with highest total yearly sales of large avocados in 2019

Region	Conventional	Organic	Overall
Northeast	177207881	2012480.9	179220362
California	89348979	2745852.6	92094831
West	72804707	2538921.5	75343628
New York	59386225	963466.0	60349691
South Central	52490049	180508.9	52670558

Exercise 6

Create a plot showing daily sales volume in California for different avocado types and sizes.

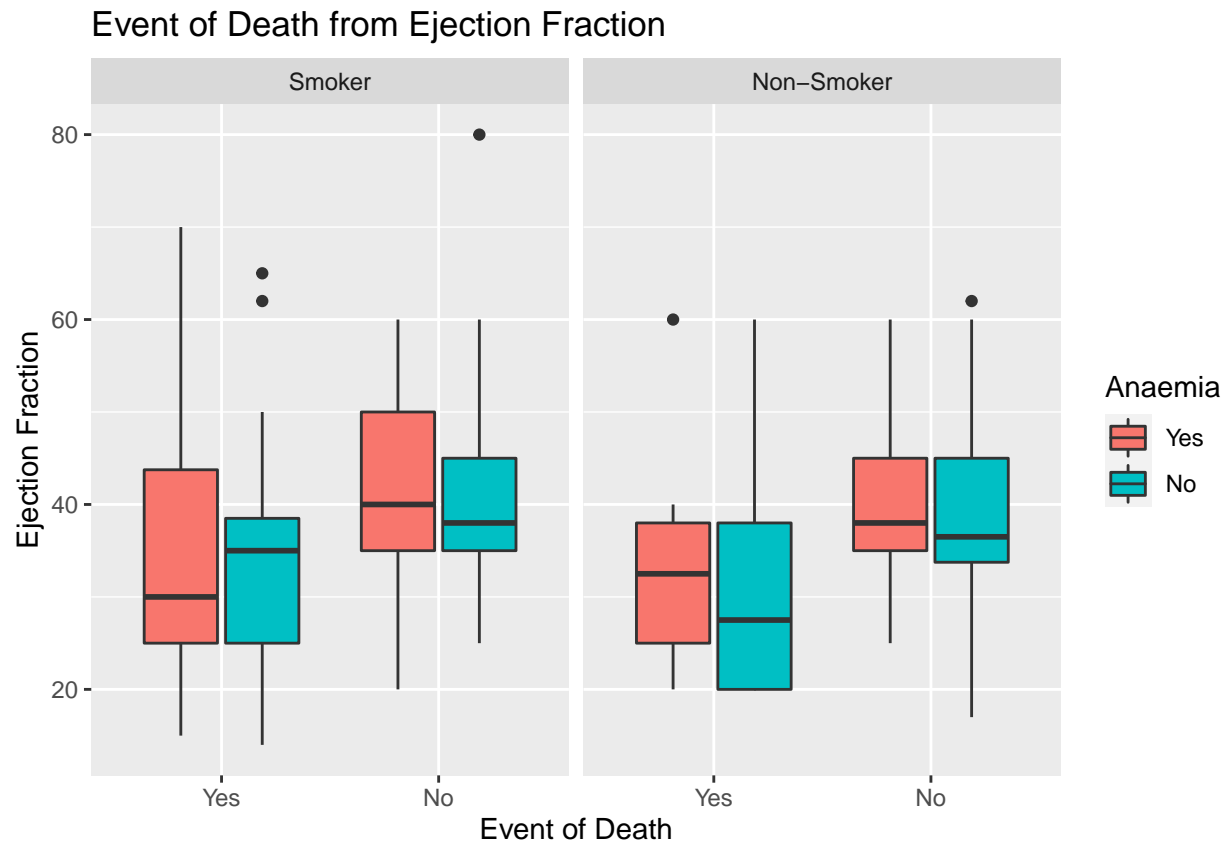
Daily Sales Volume in California by Size and Type



Exercise 7

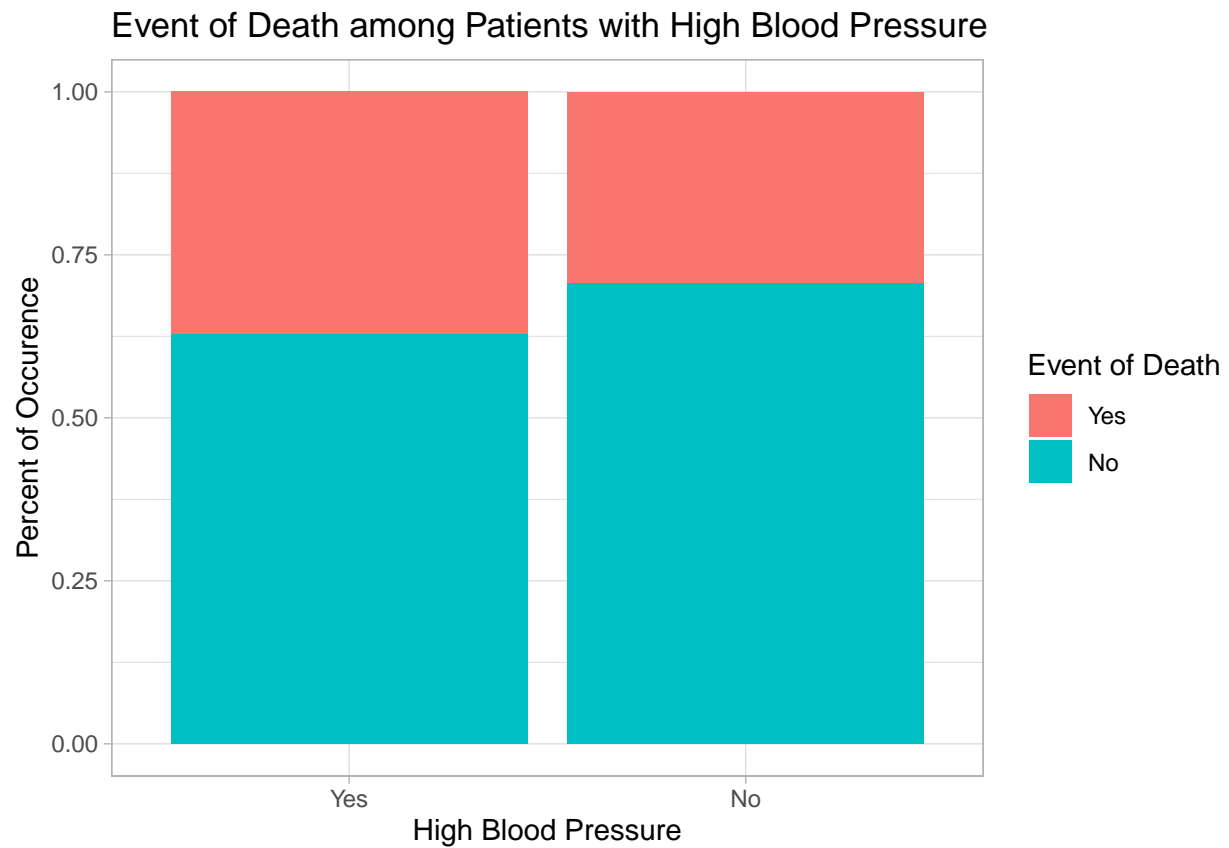
Load `heart.txt`

Exercise 8



Exercise 9

Create a segmented bar chart where the bars represent one of the factor variables and `DEATH_EVENT` is the fill.



Takeaway: From this plot we can see that patients with high blood pressure are more likely to die than patients who did not have high blood pressure.