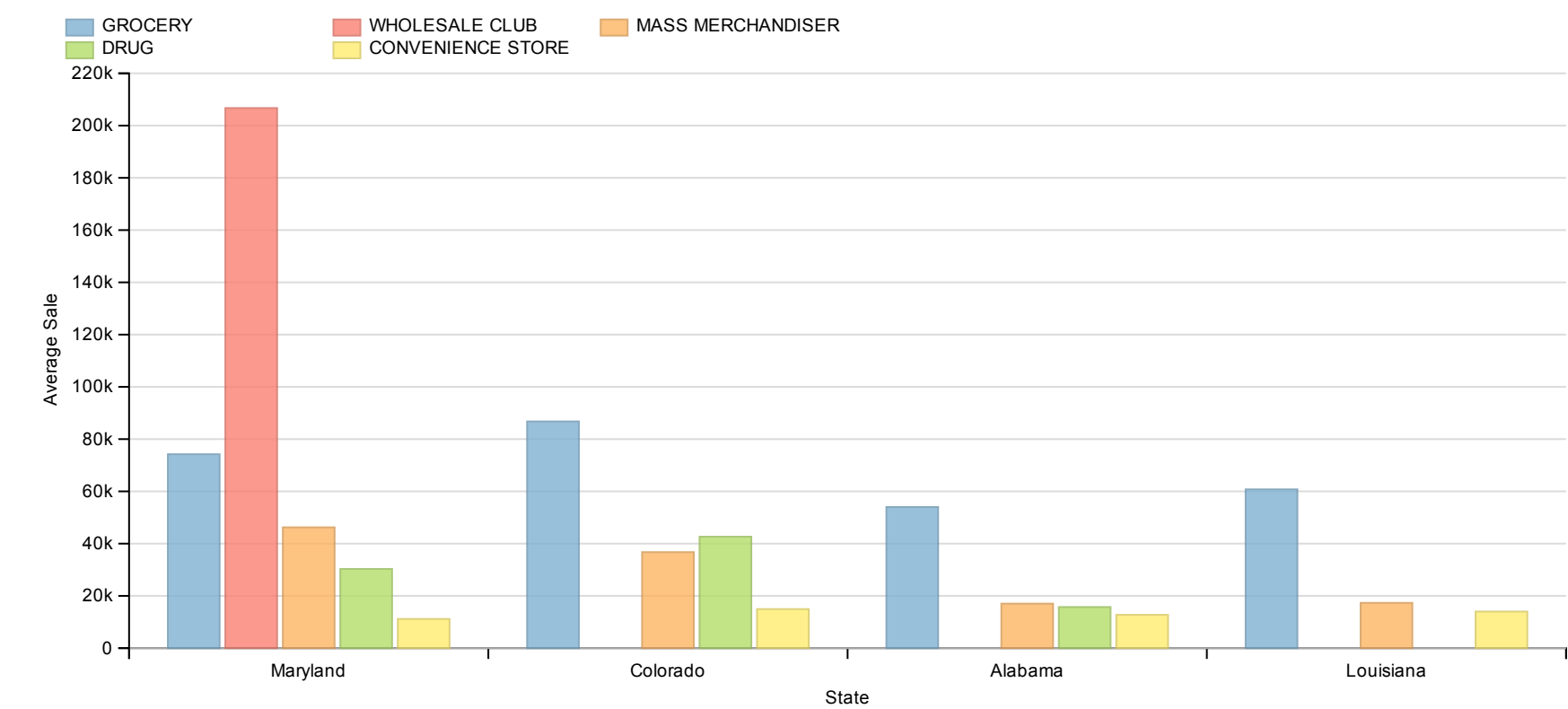


Football Games and Candy Sales in Four States

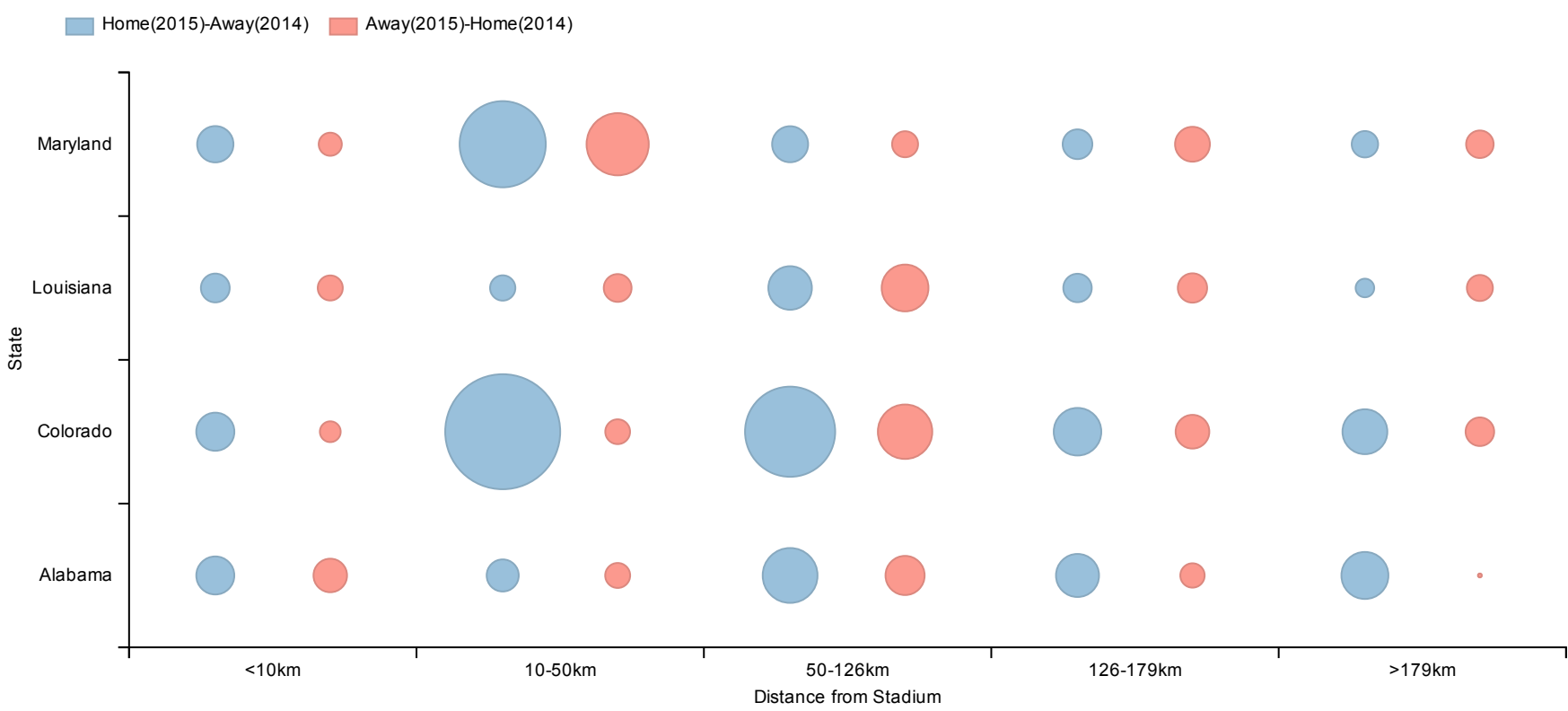
Data -- This data has been provided by the Hershet Chocolate company. The data contains the sales of candy stores during 15 consecutive Saturdays (with football games), their distance from the stadium where the football took place, the sales difference between 2014 and the same date in 2015, as well as the avergae of weekly sales in both years. The data also indicates whether there was an away game or home game in either 2014 or 2015 during that speciic week.

The goal of this visualization is to discover this dataset to see how different market channels and states affect the sales. Also, we want to see if there is any difference between those weeks with home game (2015) and away game (2014) versus the opposite.



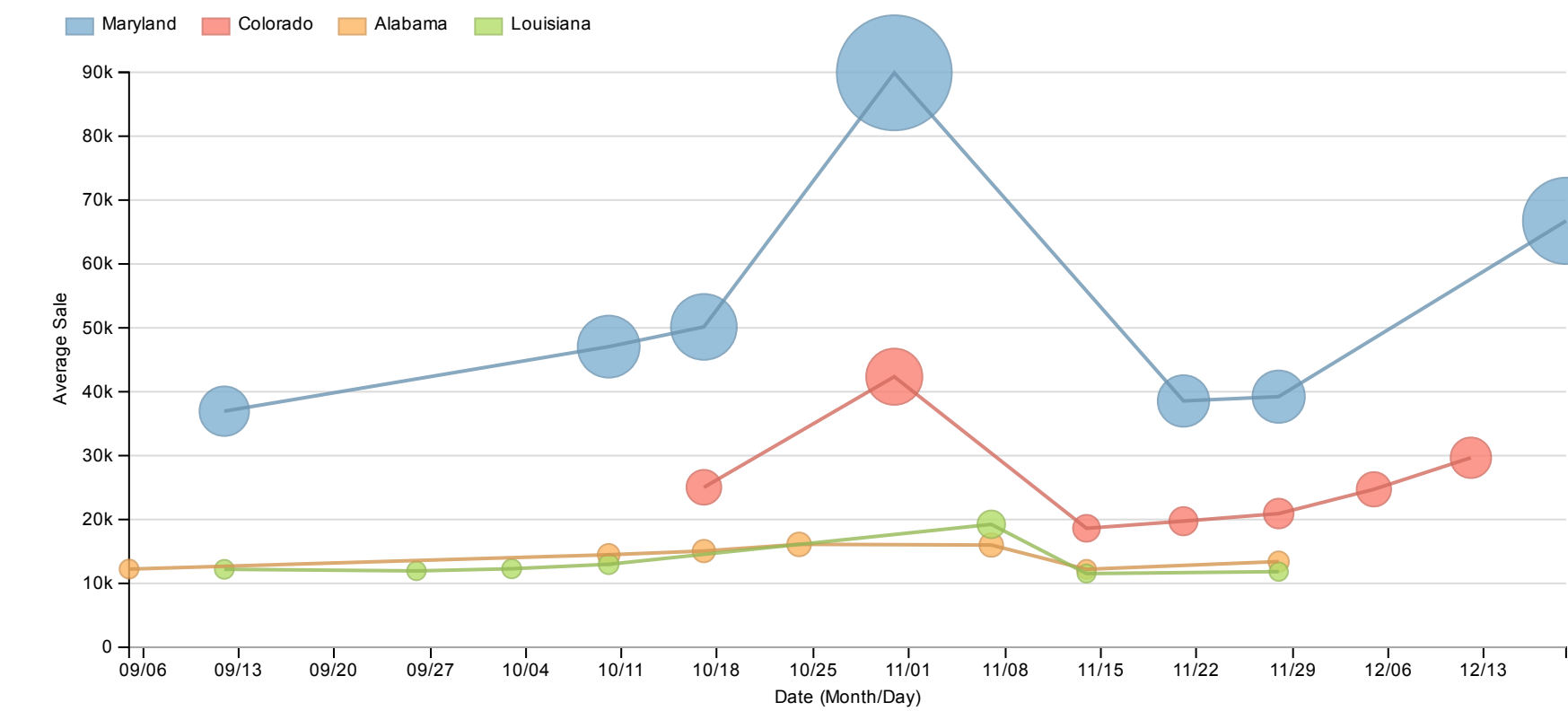
The bar plot above is intended to show to what extent different states and channels are different in terms of average sale. For this case, bar plot is one of the best options, as we only have two categorical variables (i.e. state and channel) and one numerical (average sale). One categorical variable can be presented in one axis and the other can be presented by coloring different levels of the variable. As we can see here, we have managed to show that:

- 1- The whole sale club sells the highest on average and it can only be found in Maryland.
- 2- The average sale in Maryland and Colorado are higher than th eother two
- 3- Grocery stores are the most successful channels except for Maryland's case



Visualizing three variables in the most simple way is a bit more challenging. In the bisualization above, we have the games variable, the state as well as the distance level from stadium. Dimple js. has the bubblechart option which enables us to visualize this type of data in the most simple way. In this visualization we can see that:

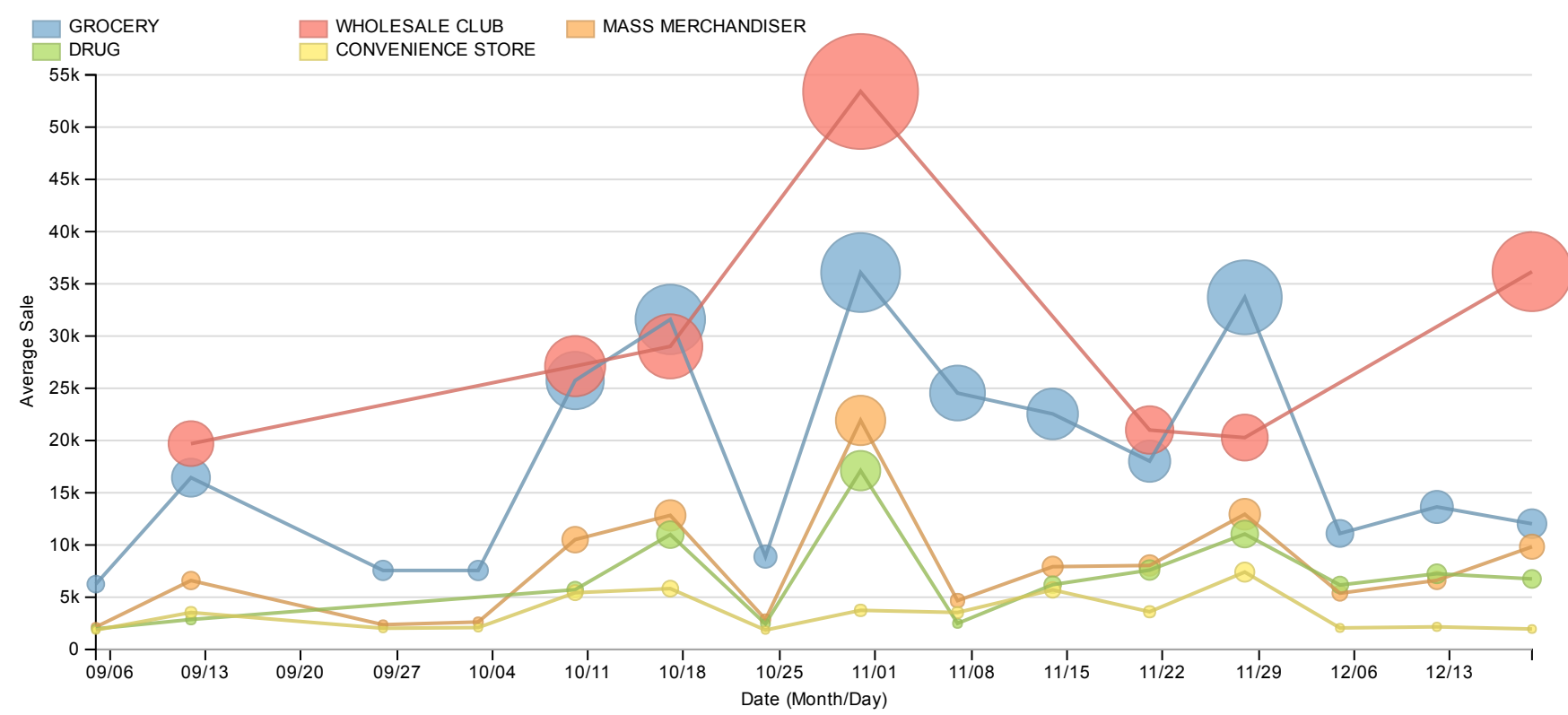
- 1- stores within 10 km of the stadium do not show that much f a difference in sales in neither of states. Although we can see that Maryland and Colorado show a slightly higher sale within 10km of the stadium when there was a home game in that stadium.
- 2- For both colorado and Maryland the difference is more evident in those stores that are within 10-50 kms from the stadium. Louisiana barely shows any difference during the away and home games. In case of alabama, distance is that influential and all the stores do constantly better during the home games.



We also wanted to investigate the trend of sales in the four states in this 15 weeks. Line chart is a good option when it comes to time series. Our surveys indicated that lines alone while indicative of overall trend, cannot fully show the extent of difference between break points. To address this, I added a bubble chart to accentuate the average sale in each data for each state. the chart above successfully shows the following:

- 1- The hierarchy of sales for different states is more evident here.
- 2- We can see that the week of October's end there is a jump in average sale (probably because of Hollween)

3- The jumps for Maryland and Colorado are more noticable compared the other two states.



It would be interesting to also check out the performance of different channels. Since the variables are rather the same, we use the same type of visualization. This chart shows the following:

- 1- Generally we can see how sales' trends for different channels compare to one another.
- 2- Convenience store are not as successful during the Holloween week compared to other channels.
- 3- Unlike other channels, Wholesale clubs are not as successful during the last week of November. Although, recall that our datapoints were less for Wholesale clubs as they can only be found in Maryland.