Write a summary of the exploratory data analysis above. What numerical or categorical features were in the data? Was there any pattern suggested of a relationship between state and ticket price? What did this lead us to decide regarding which features to use in subsequent modeling? What aspects of the data (e.g. relationships between features) should you remain wary of when you come to perform feature selection for modeling? Two key points that must be addressed are the choice of target feature for your modelling and how, if at all, you're going to handle the states labels in the data.

This EDA started with us having 2 data frames. One data frame containing our cleaned resort data and the other containing data state data. The state data was important for getting an idea of how the resorts related to their state in terms of density of resorts and population coverage. No states seemed to hold a particular advantage to our target feature of ticket pricing, though there were a handful of outliers within statistical categories. So far it appears that state data may only play a role in determining price when comparing to in state or nearby state competition. Some potential standout features appear to be related to elevation, snow making, runs, and run length when comparing for ticket price on the weekends. Elevation characteristics are highly correlating features that we’ll need to be wary of when looking for correlations among the broader data set. We should also be wary of things such as the numbers of chairs and runs. More runs may indicate the need for more chairs but that when isolated do not seem to necessarily correlate with price.