Ask

The questions asked in this analysis are:
- What is the distribution of Average Daily Rates for non-canceled rooms?
- Do rates have anything to do with the reason for canceling?
- How have prices changed for each reserved room over the years?

Prepare

This dataset is found on Kaggle. The dataset details 119390 observations for a city and a resort hotel between July 1st, 2015, and August 31st, 2017. The Data is from Hotel Booking Demand Datasets, written by Nuno Antonio, Ana Almeida, and Luis Nunes for Data in Brief, Volume 22, February 2019

Process

After cleaning the data in Google sheets, I imported the data to Google Collaboratory and checked for nulls. In this case, we had a lot of nulls and in the country section, I just dropped the null values because it had the least amount of nulls. The 'Agent' and 'Company' had a lot of null values so I just replaced the null values with 'No Agent' and 'No Company' respectively. '

Analyze

While researching the first question, I found that there is nothing special with the Average Daily Rates. They have the greatest density in the 70 - 120 dollar mark. I have lived in a Big City for my whole life and these prices are not that different from the prices we offer.
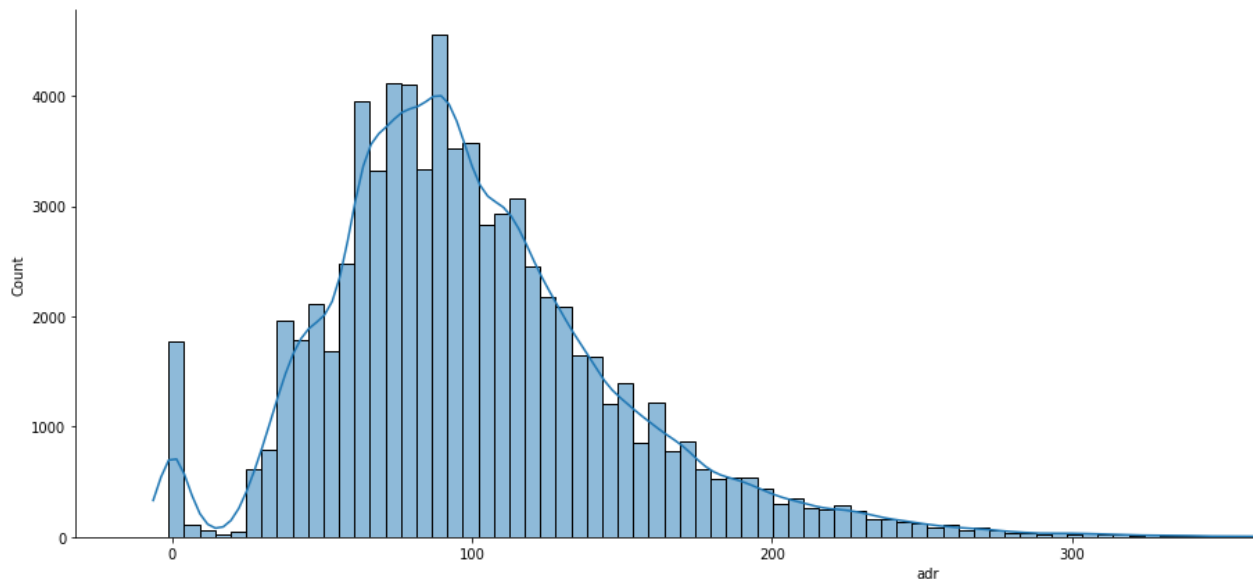
For the second question, we see that as the years go on, there is no difference in the distribution of price between canceled and not canceled rooms. In 2015, there is a lot more people who arrived in their approximately 50-70 dollar room than there are who canceled. The following years they seem to be getting more and more equal.
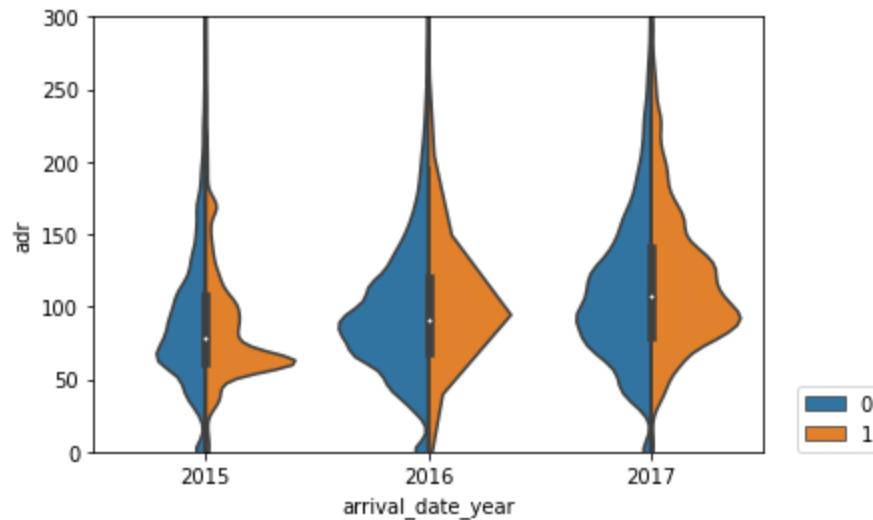
This makes sense because generally there is no difference in rates for canceled and not canceled reservations. As shown in this KDE plot
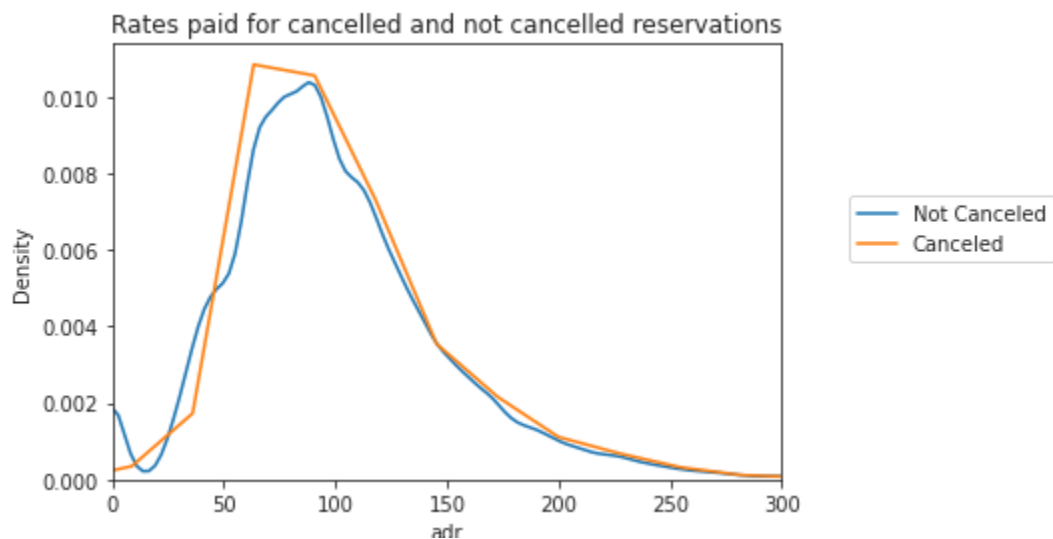
Over the years we've seen an increase in pricing in all the reserved room letters. This is no surprise however we see in most rooms in 2017 have a drastic increase in price when compared to their growth in 2016 from 2015 year. I hypothesize that it has to do with the inflation rate, in 2016 the inflation rate was 1.26%, while in 2017 the inflation rate was 2.13%.