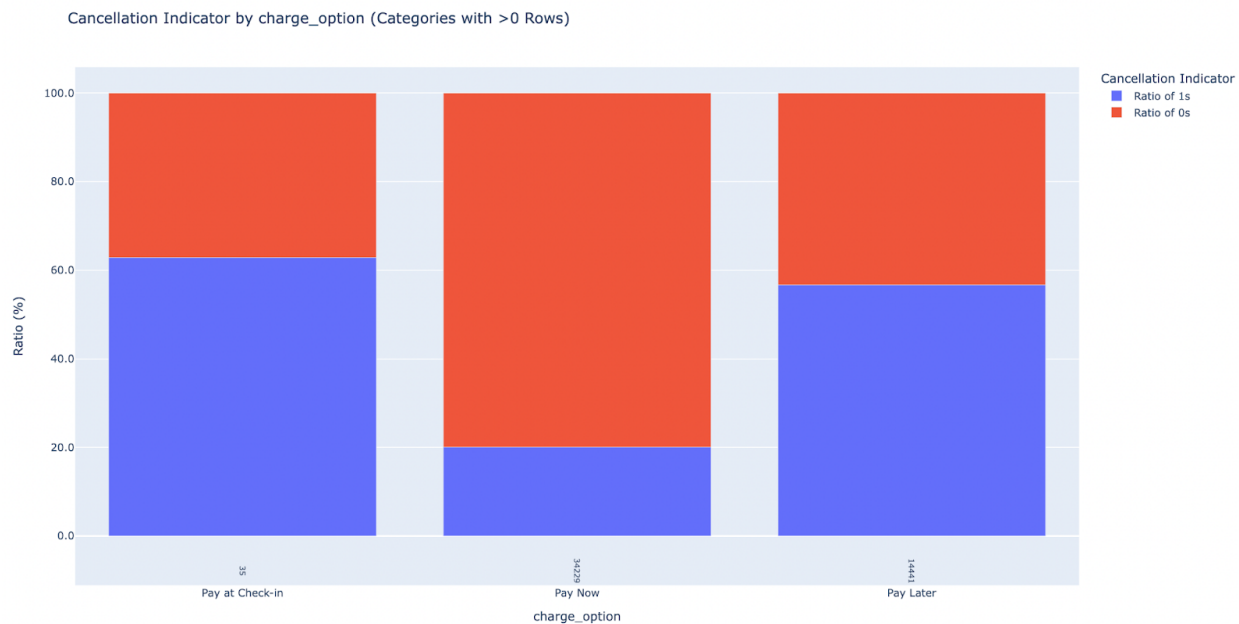
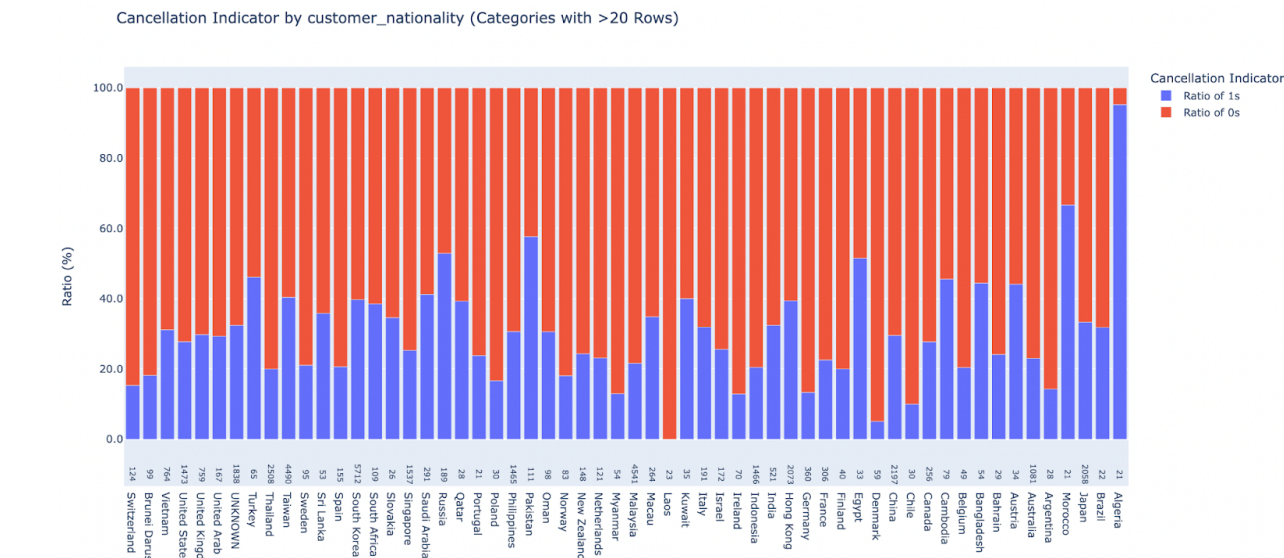


the most relevant feature for cancellation prediction we found were:  
charge option , customer nationality, Cancellation policy code

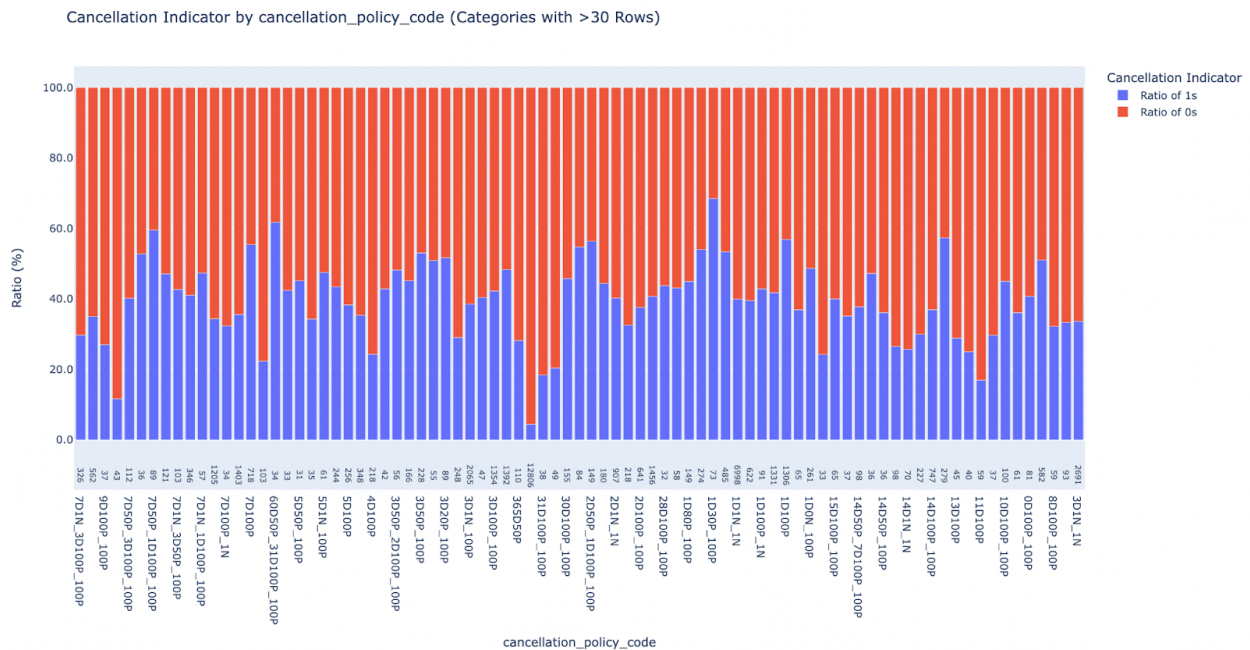
The charge option was one of the first features that stood out to us as a relevant feature.  
We found that customers who pay upfront tend to cancel less often than other customers as we can see in the following graph



The costumers nationality shows us that there are regions that are more prone to cancellations such as Algeria as opposed to switzerland that is less likely



the Cancellation policy code feature had a huge impact on the customer. Since if there is a very harsh policy it is much less likely to cancel as can be seen with the 31D100P which has a very large amount of people that canceled due to the very harsh restrictions on the costumer as opposed to the more lax policies that don't scare the customers away



We can also see by SelectKBest with chi2/f\_classif that the following are have great impact:

#### SelectKBest with score\_func=chi2

1st Feature: original\_selling\_amount, Score: 432626.012551094411

2nd Feature: booking\_to\_checkin, Score: 251393.4513224351

3rd Feature: charge\_option\_Pay Later, Score: 4475.098761031088

#### SelectKBest with score\_func=f\_classif

1st Feature: charge\_option\_Pay Now, Score: 7355.364402004186

2nd Feature: charge\_option\_Pay Later, Score: 7316.508024585366

3rd Feature: cancellation\_policy\_code\_365D100P\_100P, Score: 6477.45661343796

4th Feature: booking\_to\_checkin, Score: 3984.843761507652

Which also compatible with the result of Lasso with lambda from 0.1 to 1 with jumps of 0.1,

Gave us the following result for all the lambdas, Top contributing features:

booking\_to\_checkin

original\_selling\_amount

hotel\_star\_rating