The dataset provided is a selection of 2200 bookings that came for a single voyage. There were additional bookings made that comprise a holdout dataset that we will run through your answer to #1.

The dataset provide is a small selection of the available data fields we have at our disposal at Holland America. As such we don’t expect a perfect model, but instead wish to see how you would approach this problem, the creativity you have in order to solve an open problem, and your ability to construct R models and code.

We will run your R code against our holdout dataset upon submission.

Instructions:

- Use R to answer the 2 questions below.

- Include all code you use to answer the questions and provide general comments through your code; including any code you use for data discovery.

- Active bookings can be identified by BKNG\_STATUS = 'B', cancelled bookings can be identified by BKNG\_STATUS = 'C'

Data:

1) Create a model that predicts whether a person will cancel a booking

2) What kind of people book suites (meta in ('S','D')) (Active bookings only)?