

DataFest 2017

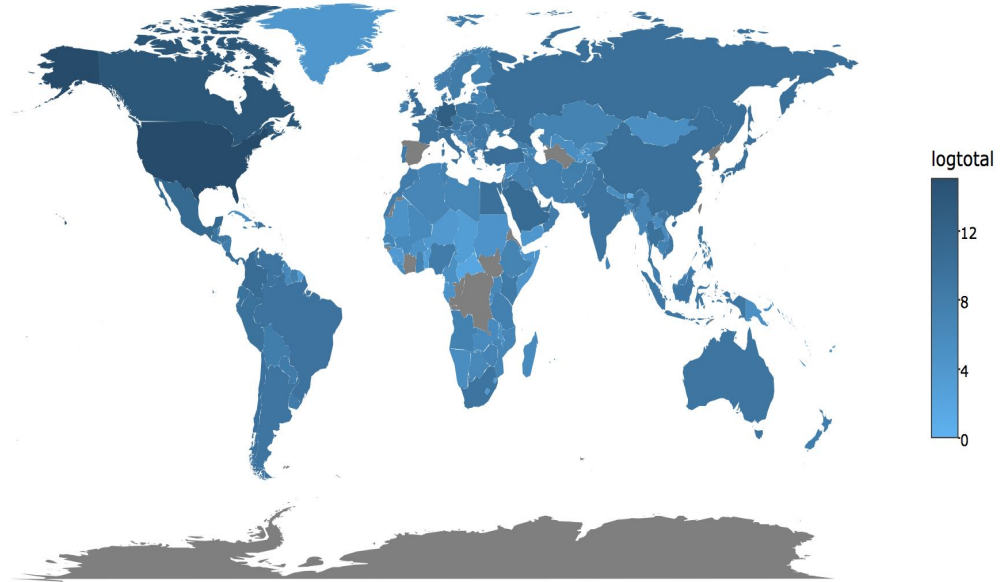
Smith College

Team: **Normally Distributed**

April 2nd, 2017

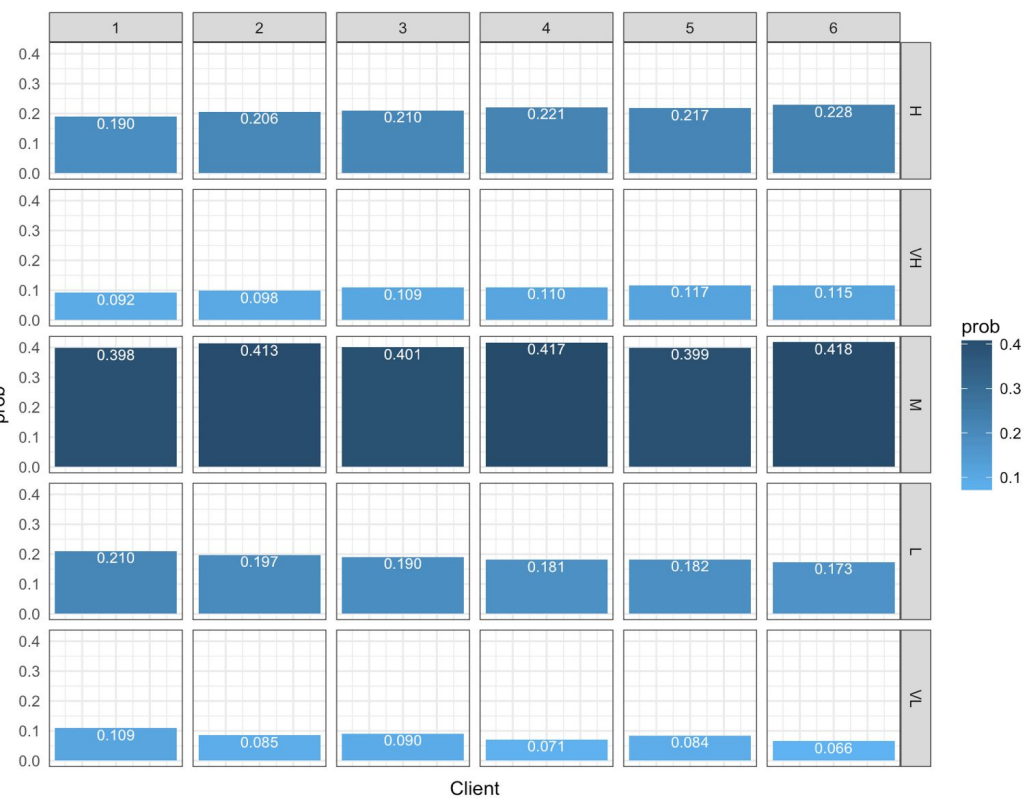


Customer Distribution of Expedia



91.19% of clicks do not convert to hotel purchases

85% of site visitors click on only one hotel and leave the site

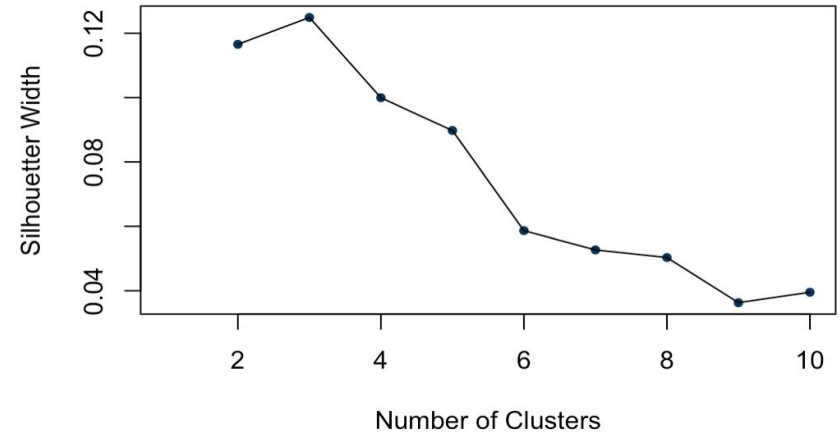


6 types of Customers' choice about different price range of hotel

Most popular months for hotel booking: **June - August**

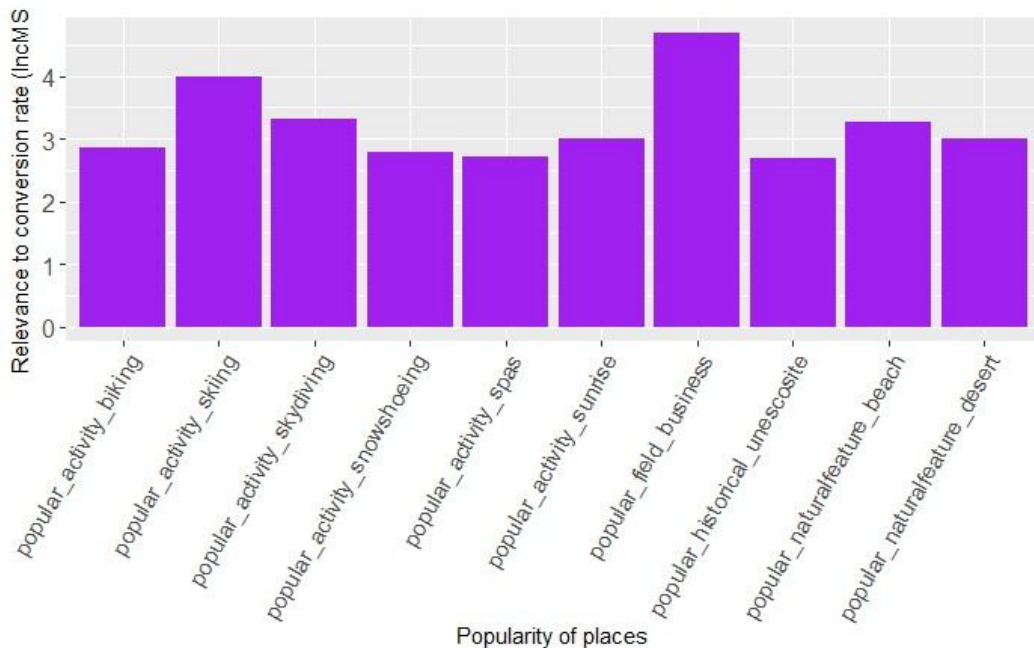
Most popular time for hotel booking: **10am-1pm**

Cluster Analysis: three groups of customers with different preferences



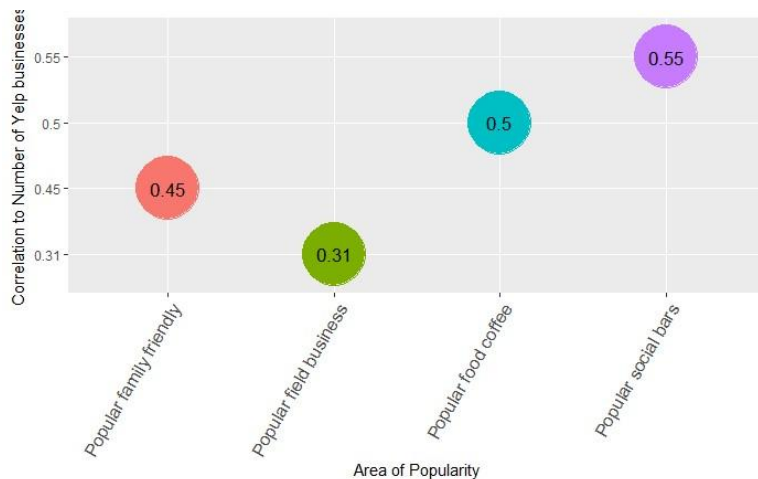
What drives popularity and purchase rate of a city?

Random forest model to predict purchase rate of a destination based on its popularity score

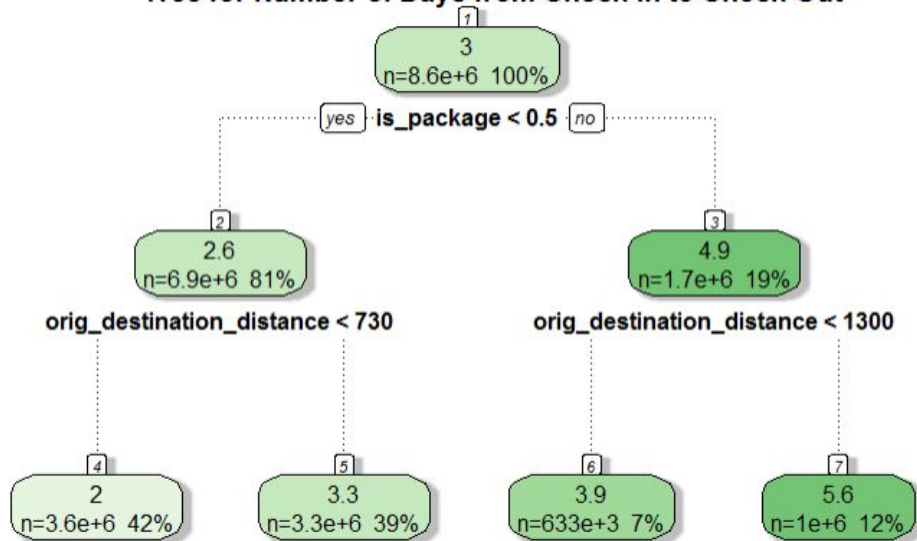


Root MSE: 0.018. Data range from 0-0.25

Correlation between the number of Yelp businesses and popularity score

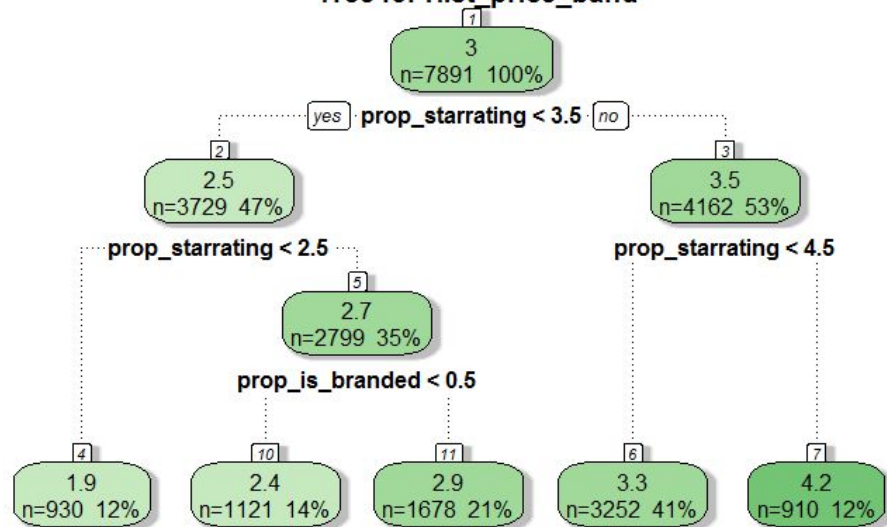


Tree for Number of Days from Check-In to Check-Out



KEY: Very Low=1, Low=2, Medium=3, High=4, Very High=5

Tree for Hist_price_band



KEY: Very Low=1, Low=2, Medium=3, High=4, Very High=5