

Targeting Expedia's Niche Market in Travel Packages via Analytical and Predictive Modeling

Ji Won Chung, Zainab Rizvi, Subashini Sridhar, Ji Young Yun

Smith College

Data

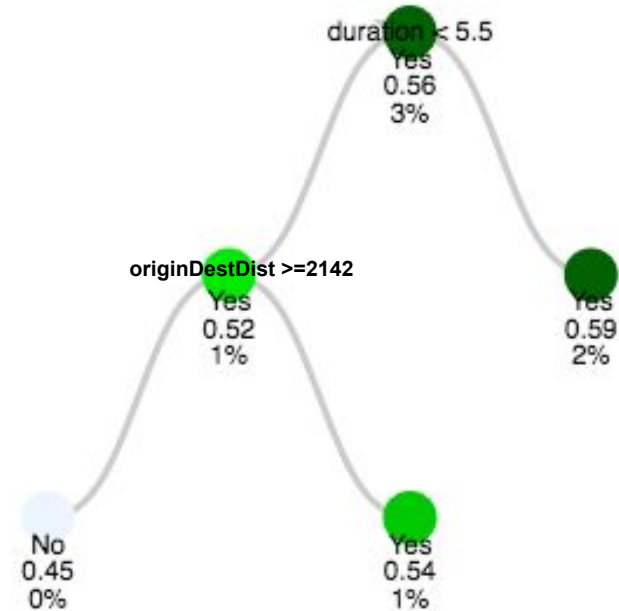
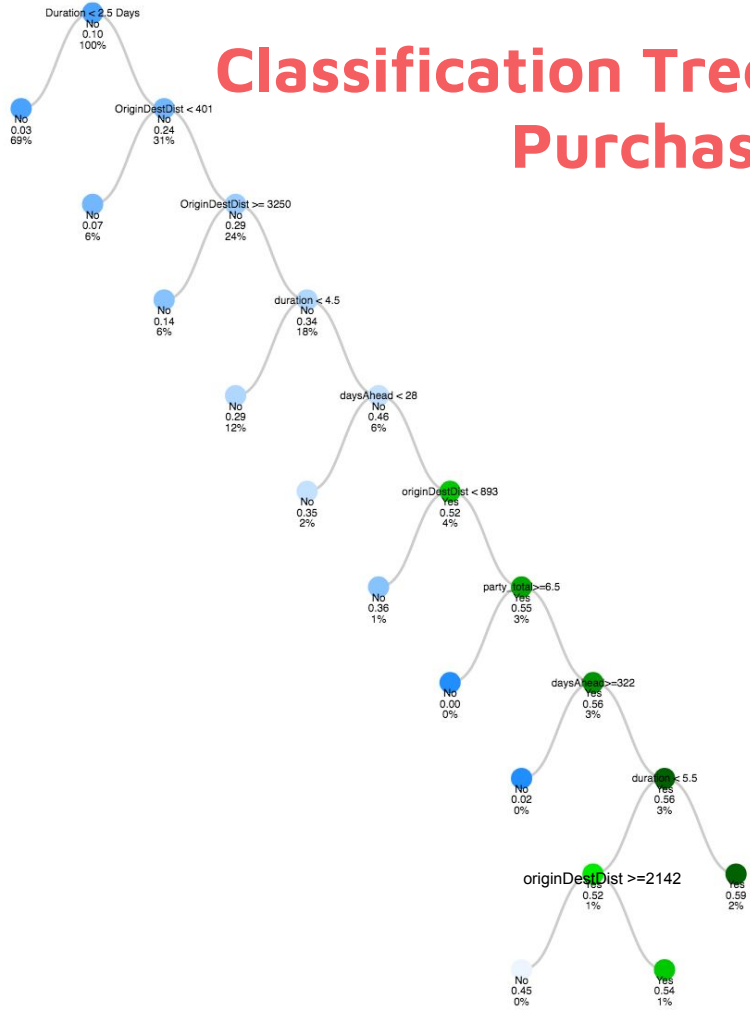
- The dataset was given as part of the Five College ASA 2017 Datafest competition and was provided by Expedia Inc.
- From a 10 million records, we filtered to analyze customers who booked through Expedia (isBooking == true)

Dimensions: 680,615 customer records

Response variable: is_package (binary)

Explanatory variables: duration, daysAhead, distance from origin, party total

Classification Tree to Model Expedia Customers Purchasing Travel Packages



Results: Confusion Matrix to Assess Accuracy

pred_tree	Predicted: 0 (package not bought)	Predicted: 1 (package bought)
Actual: 0 (package not bought)	614674	57256
Actual: 1 (package bought)	7643	10402

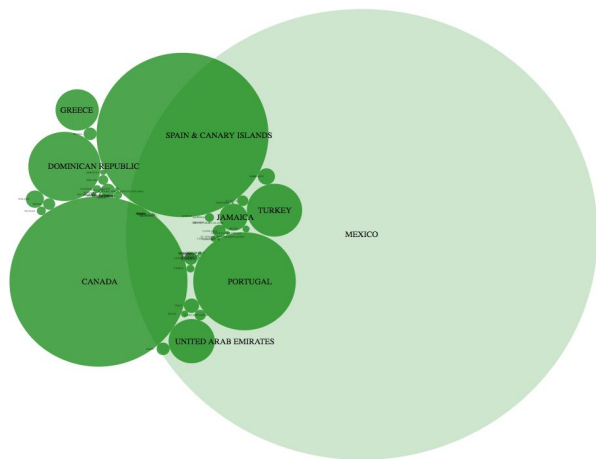
- Classification Tree has 91.8% accuracy
- 8.2% Training Error

Visualizing Data

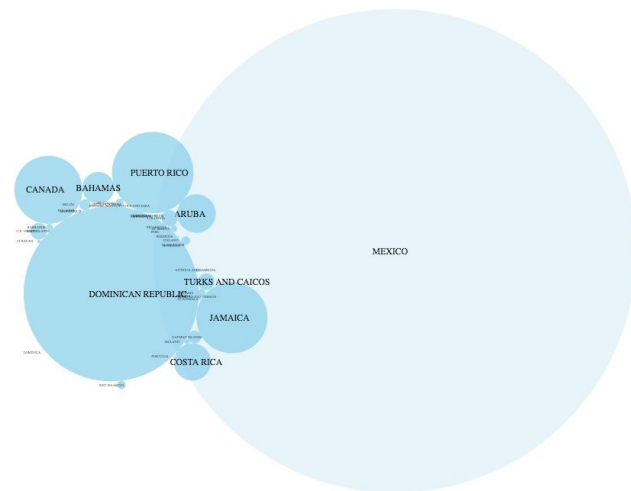
- D3.js for Bubble Chart Visualization
 - Allows for easy stylization
 - Repeat visualizations
 - Compatible with large data sets
 - Clean, simple visualizations

US Origin

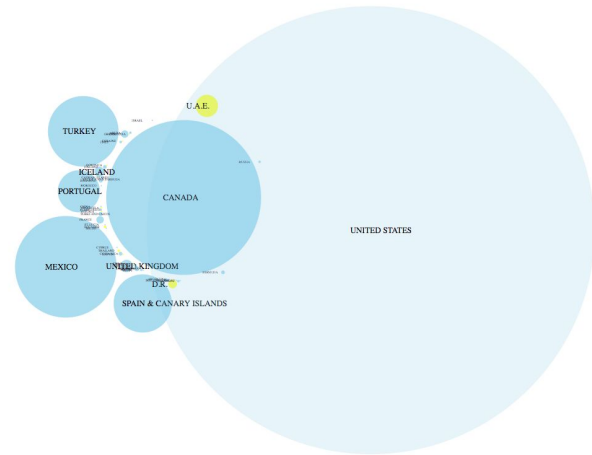
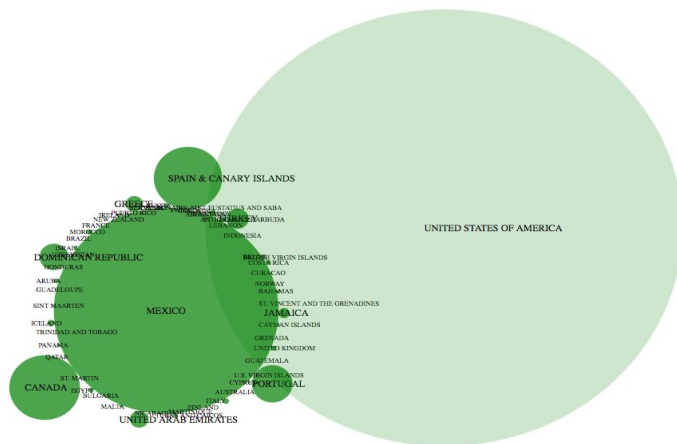
Buy the Package



Don't Buy the Package



Non-US Origin

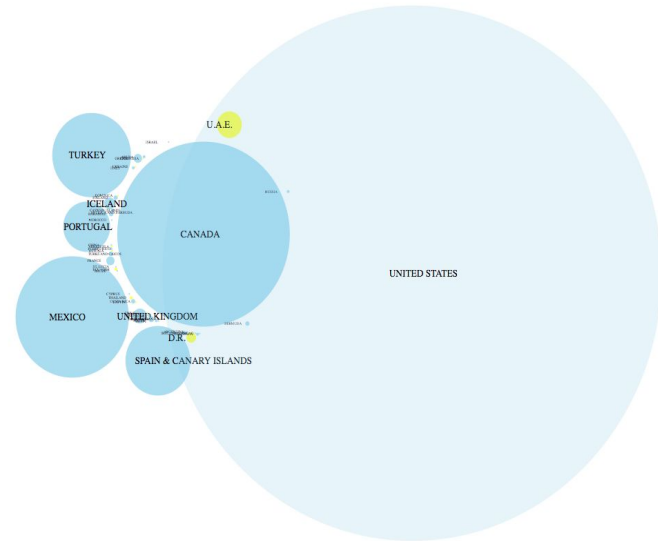
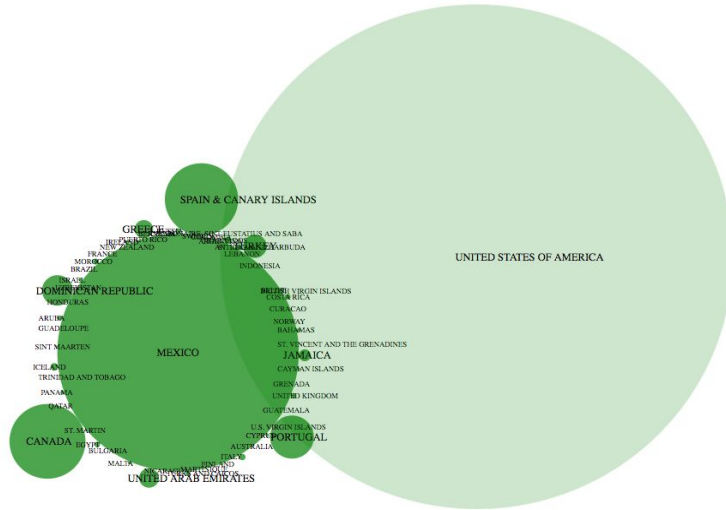


Understanding the Bubble Chart

- Filtered dataset by travel destination
 - Record the countries people traveled to and from the most
- Size represents the number of travellers
 - Big circles are the countries we don't want to focus
 - Big circles = established market
- United States (less opacity)
 - most number of travellers
- Other countries (regular opacity)
- Yellow - target countries = niche market to target

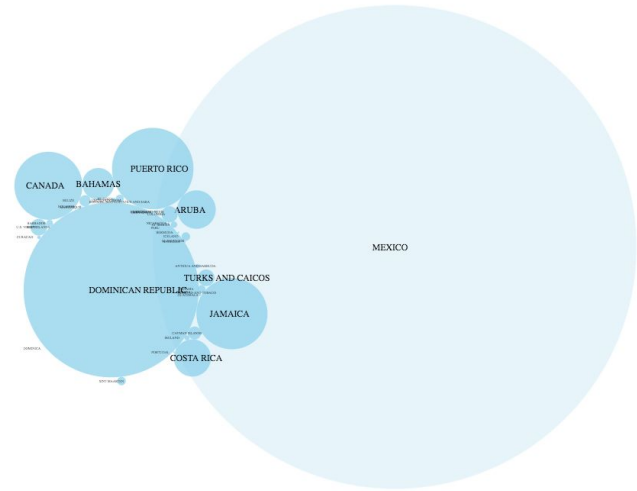
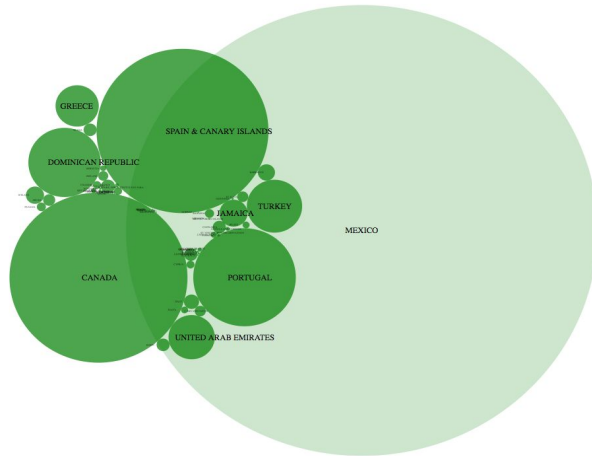
Bubble Chart Analysis

- Top countries with the most number of people traveling to the U.S.
 - Canada, Germany, Mexico



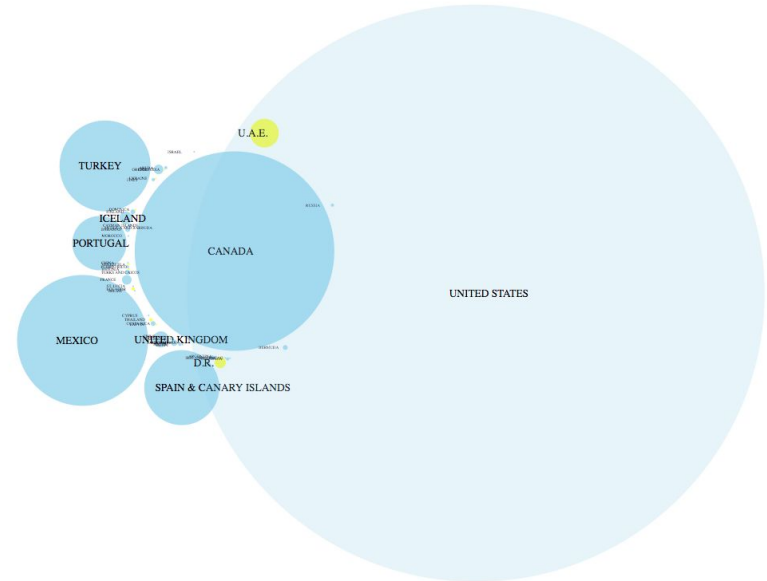
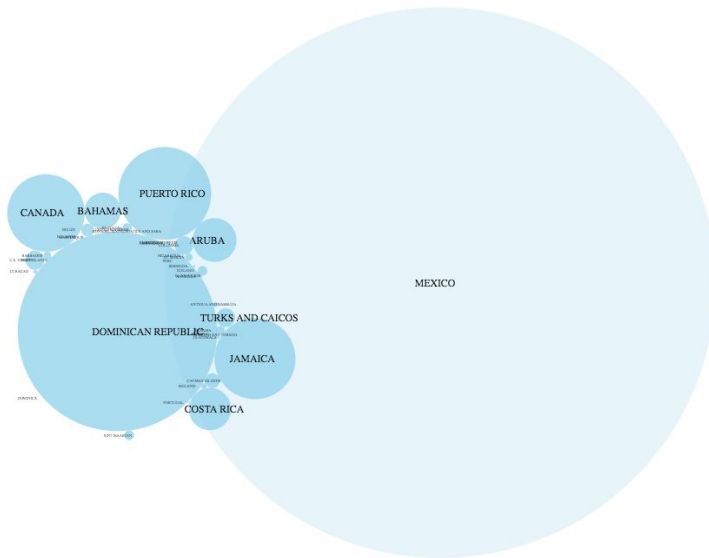
Bubble Chart Analysis

- Top countries with the most number of people traveling from the U.S.
 - Mexico, Dominican Republic, Puerto Rico



Locating Our Target

- Countries with small market = small circles
- No market exists , but almost bought packages = potential customers

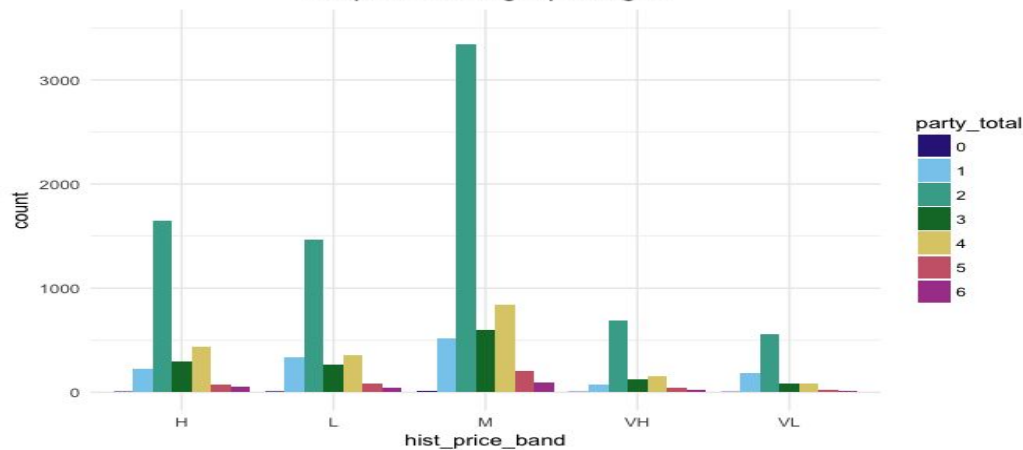


Flaws with Visualization

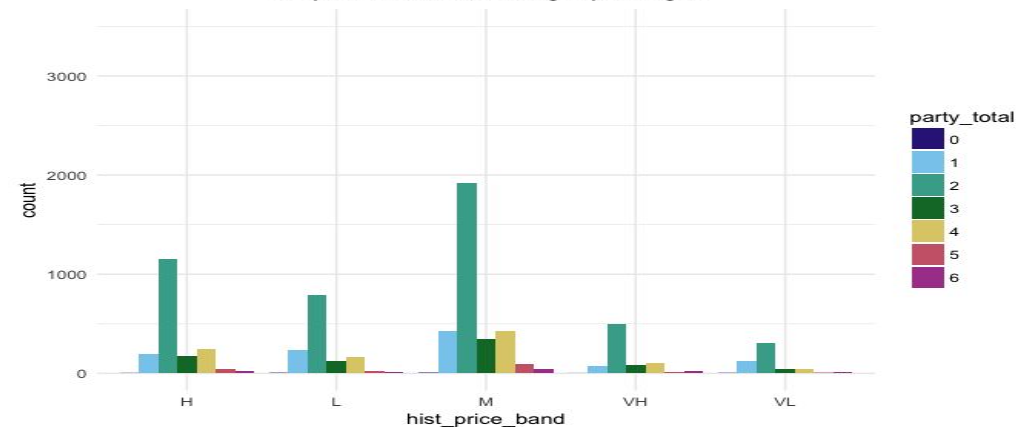
- Radius represents the population - to enhance the differences
- Area could be a better representation

Hotel Data Analysis

People who bought packages



People who almost bought packages




Insights


- + Groups are likely to choose hotels with medium price
- + Solo travellers are least likely to choose VH bands
- + As group size increases beyond 3, they prefer M and H bands and shy away from VL
- + Groups under 3 prefer M and H range
- + There are many more groups of 2 in the graph for people who bought packages than those who almost bought packages


Recommendations

Unreal Deal


Get 1 night free by booking together.





 Limited Time: **Free self parking**

GR Solaris Cancun & Spa - All Inclusive ★★★★★ 

Hotel Zone
1-800-222-0892 • Expedia Rate

 **Free Cancellation within 24 hours!**


 +  Economy/Coach flights and Economy car included in price

Booked in the last 20 hours

Very good! 4.2/5
(6,468 reviews)

~~\$1,242~~ \$1,111

per person includes flight + hotel + car, taxes and fees

 Earn 13,333 Expedia+ points

Destination: Cancun (3,412 mi)
Party size: 5

Duration: 6 days

Booking ahead: 29 days

Rating: 4 star

20,785 people * \$1,111 = \$23 million in revenue

Recommendations:

- Hotel packages should be tiered according to party sizes
- Focus on different regions for US and non-US travellers

Thank You