Group DataBest: Ji Won Chung, Ji Young Yun, Subashini Sridhar, Van Nguyen, Zainab Rizvi

Slide 1: Binary classification tree
Shows the binary classification tree for all bookers. The response variable is if someone buys a package and the explanatory variables are duration of the trip, distance from origin, size of travelling party, how many days ahead the trip was booked. These variables were selected after doing a logistic regression for the same response.

We will focus our analysis in this presentation on the last leaf node on the left. These are the people who almost booked a package and are thus, prime for conversion. We will now look at other factors that distinguish these travellers from those that did buy packages. Tamra is our persona who represents these travellers. We will now consider the countries that Tamra is most likely to travel to as well as the type of hotels (price range, star rating, brand) she prefers. This insight will allow Expedia to customize packages for Tamra and capture this critical market segment.

Slide 2: Cluster of countries
This slide shows the cluster of countries that people who buy packages are likely to travel to as well as countries that Tamra is most likely to travel to. We have divided these clusters in US to non-US travel and non-US to everywhere else travellers. We have highlighted countries that are specific to the group represented by Tamra.

Top countries for people traveling to the US is Canada, Germany, Mexico. Top countries people choose to travel from the US is Mexico, Dominican Republic and Puerto Rico.

Slide 3: Hotel analysis (brand, star rating, distance band) for people
The bar charts differentiate between the two groups again. Groups with smaller number of people tend to choose hotels with lower range prices. We also found that there are many more groups of 2 in the graph for people who bought packages than those who almost bought packages. Analysis we conducted (not shown on the slides) shows that people tend to choose branded hotels with star ratings of 4 or 5.

Slide 4: Summary and recommendations
Our two main recommendations are:
1. Hotel packages should be tiered according to party sizes
2. Focus on different regions for US and non-US travellers

We did an Expedia search today with the following criteria to follow Tamra:
New York => Cancun, duration: 6 days, booking: 29 days in advance, dates: 05/02/17 - 05/08/17, rating: 4 star hotel

The discounted rental car + hotel + flights package was $1,111. The last node that we showed in slide 1 was 20,785 people who are prime for conversion. Targeting this market segment would approx add another $23 million in revenue for Expedia.