Analytics Case study

Over the last five years, we have witnessed an intriguing trend that suggests a correlation between the number of property images associated with a listing and the number of bookings it attracts. We have also noticed an overwhelming number of listings being redundant, due to lack of any associated images.

1. **You need to help the management decide upon a minimum number of images to be made mandatory for a listing that would ensure bookings.**
2. **Also, come up with an optimal number of images that we can suggest the host to post along with a listing that would attract the most bookings and ensure success.**

For this exercise you have to use the three datasets provided *(see the downloadable links and the description below)* to develop your argument.

Write your solutions in the adjacent text box and support your arguments in 4-5 powerpoint or keynote slides which you have to upload as an attachment.

Do include a description of the methodology and the software packages employed to arrive at your results, as well as any assumptions you made. Your submission will be evaluated on your dexterity in data manipulation, modeling, visualization, logical assumptions and communication.

About the data

(This is not real data and is only provided to you for the purpose of this case. Download the .csv files by clicking on the dataset links below)

Dataset 1 **Listings.csv**

provides a random data sample of 500 listings posted by various hosts (including Superhosts) in last 5 years from various locations, along with their associated number of property images and the number of bookings they attracted.

Dataset 2 **Open_Listings.csv**

provides data for over a year, that shows number of open listings for each date. Where open listings mean the property listings that were available but did not attract any booking by end of the day. The listings have been classified according to the number of associated images.

Dataset 3 **Redundant_Listings.csv**

provides data as on August 31, 2019 for the Total Listings and the Redundant Listings in each category. Redundant listings here means the listings that have not attracted even a single booking in last 1 year. The categories here are classified according to the associated number of property images.

A glossary of the variables in each dataset is as follows:

Dataset 1: Listings.csv

Variable: Description

- Listing_Id: Id of the property listing
- Posting_Date: Listing posted on a random date in last 5 years
- Posting_Time: UTC Time when the listing as posted
- Location: Location of the property
- Images: Number of property images associated with the listing
- Bookings: Number of booking the listing has attracted until Aug 31, 2019, since it was first posted.
- Host_Type: Posted by a regular or a Superhost (host status as on Aug 31, 2019)

Dataset 2: Open_Listings.csv

Variable: Description

- Date: Each date between Aug 1, 2018 to Aug 31, 2019
- Open_Listings_0_2: Number of listings which were available for the mentioned date but did not attract a booking by end of the day even on that specific date and have 0 to 2 associated property images.
- Open_Listings_3_5: Number of listings which were available...date and have 3 to 5 associated property images.
- Open_Listings_6_10: Number of listings which were available...date and have 6 to 10 associated property images.
- Open_Listings_11_15: Number of listings which were available...date and have 11 to 15 associated property images.
- Open_Listings_16: Number of listings which were available...date and have more than 16 associated property images.

Dataset 3: Redundant_Listings.csv

Variable: Description

- Property_Images: Range for the number of associated images for the property, posted along with the listing.
- Total_Listings: Total number of listings, with the associated number of property images in the specified range, that are active.
- Redundant_Listings: Number of listings that are active, having the associated number of property images, but did not attract even a single booking in last 1 year.

(Note: You must treat all information in this case analysis as strictly confidential.)

Best Wishes!

Team Airseva