Data Challenge: Market Analysis in Dublin

Challenge Description

Congratulations! Your team has been hired as a new city Airbnb city manager for Dublin, Ireland! Your job requires you to better understand:

- What guests are searching for in Dublin
- Which inquiries hosts tend to accept

Based on the findings, you will try to boost the number and quality of hosts in Dublin to fit the demands from guests. The goal of this challenge is to analyze, understand, visualize, and communicate the demand / supply in the market. For example you may want to look at the breakdown of start date day of the week, or number of nights, or room type that is searched for, and how many hosts accepted the reservation. In particular, we are interested in:

- What the gaps are between guest demand and host supply that the new city manager could plug to increase the number of bookings in Dublin
- What other data would be useful to have to deepen the analysis and understanding

Analyze the provided data to the best of your abilities. Include any relevant tables/graphs/visualization to explain what you have learnt about the market.

Data Description

The data is given through 2 .tsv files (tab-separated values), their data descriptions below:

1. **searches.tsv : Contains a row for each set of searches that a user does for Dublin.**
    a. **ds :** Date of the search
    b. **id_user :** Alphanumeric user_id
    c. **ds_checkin :** Date stamp of the check-in date of the search
    d. **ds_checkout :** Date stamp of the check-out date of the search
    e. **n_searches :** Number of searches in the search set
    f. **n_nights :** The number of nights the search was for
    g. **n_guests_min :** The minimum number of guests selected in a search set
    h. **n_guests_max :** The maximum number of guests selected in a search set
    i. **origin_country :** The country the search was from
    j. **filter_price_min :** The value of the lower bound of the price filter, if the user used it
    k. **filter_price_max :** The value of the upper bound of the price filter, if the user used it
    l. **filter_room_types :** The room types that the user filtered by, if the user used the room_types filter
    m. **filter_neighborhoods :** The neighborhoods types that the user filtered by, if the user used the neighborhoods filter
2. **contacts.tsv : Contains a row for every time that an assigned visitor makes an inquiry for a stay in a listing in Dublin.**
    a. **id_guest :** Alphanumeric user_id of the guest making the inquiry
    b. **id_host :** Alphanumeric user_id of the host of the listing to which the inquiry is made
    c. **id_listing :** Alphanumeric identifier for the listing to which the inquiry is made
    d. **ts_contact_at :** UTC timestamp of the moment the inquiry is made.
    e. **ts_reply_at :** UTC timestamp of the moment the host replies to the inquiry, if so

f. **ts_accepted_at :** UTC timestamp of the moment the host accepts the inquiry, if so
g. **ts_booking_at :** UTC timestamp of the moment the booking is made, if so
h. **ds_checkin :** Date stamp of the check-in date of the inquiry
i. **ds_checkout :** Date stamp of the check-out date of the inquiry
j. **n_guests :** The number of guests the inquiry is for
k. **n_messages :** The total number of messages that were sent around this inquiry

EDIT: Strongly Recommended Prompts

If you need a place to start, it may be good to ask yourself a few questions and evaluate which variables are most relevant to your research questions and goals. For example, here are a couple of questions you can start with:

1. When are the most common times of the year that visitors and guests try to book an Airbnb in Dublin? Likewise, what about when bookings are searched? How long is the average stay? Which country are most visitors booking from?
2. What patterns emerge from inquiries that lead to bookings? What patterns emerge from those that didn't lead to bookings?
3. What could be the external factors that influence any patterns that emerge from your analysis? Any National Holidays, etc..?
4. How can your analysis help Airbnb hosts in Dublin fulfill guest demands?
5. To take it further, how could this analysis be applied to a different host city?

Want to see examples? Check out our Datathon @ UCI Resource Sheet for project examples and more!

If you want to take this data challenge as an opportunity to learn about machine learning, go ahead! Check out our Datathon @ UCI Resource Sheet for our compilation of different ML resources.