Airbnbs in Dublin:

Analysing Customer Patterns and Sentiment

Colin Yee, Rohan Mistry, Rahul Joshi, Alex Ngo

Tools Used

TEAM COLLAB

- Deepnote
- GitHub
- Google Workspace

DATA ANALYTICS

- Alteryx Designer
- Python (Pandas, Matplotlib, Seaborn, NumPy, Scikit-Learn)

Deepnote

*More details on our data cleaning & analysis can be found on Devpost.

Primary Goals

Understand

Break down search data so hosts can **understand** customer search patterns and preferences

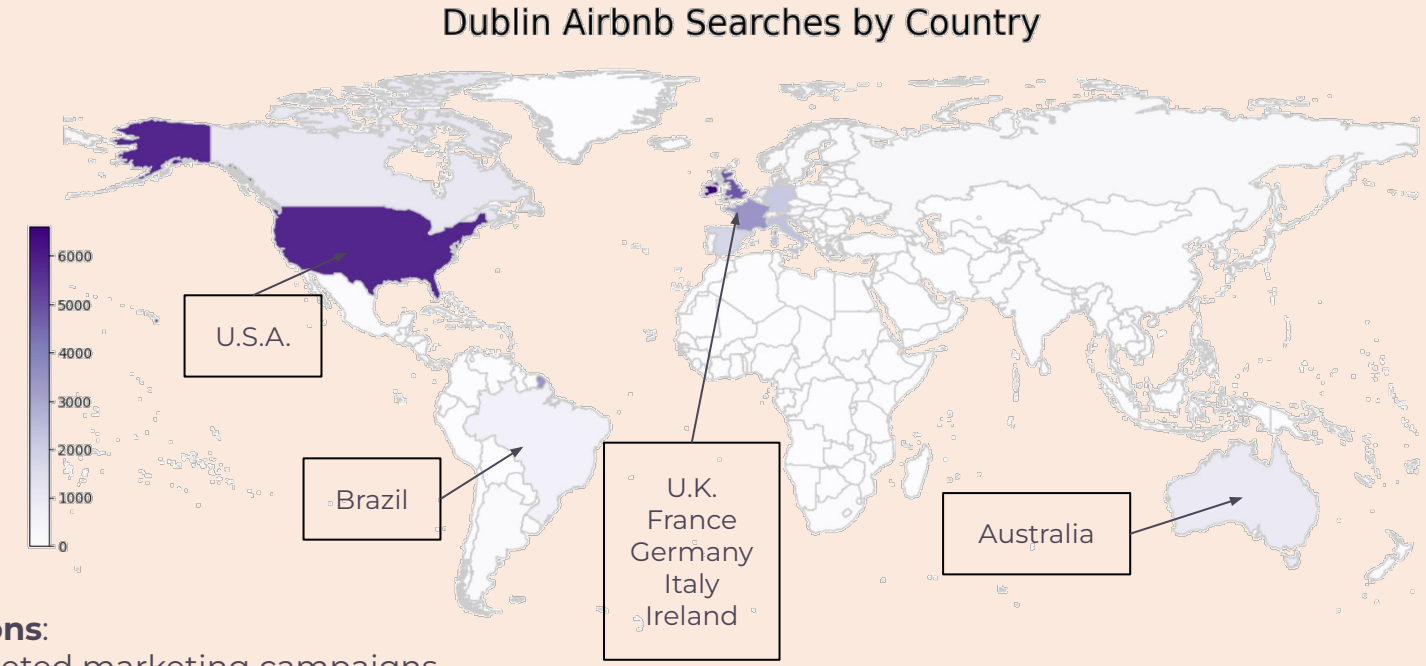
Suggest

Suggest solutions to hosts to make an optimal booking experience for everyone (including guests :D)

Outline of Analysis

1. Neighborhood and Room Type
2. Pricing
3. Guest Information
4. Timing of Check-In

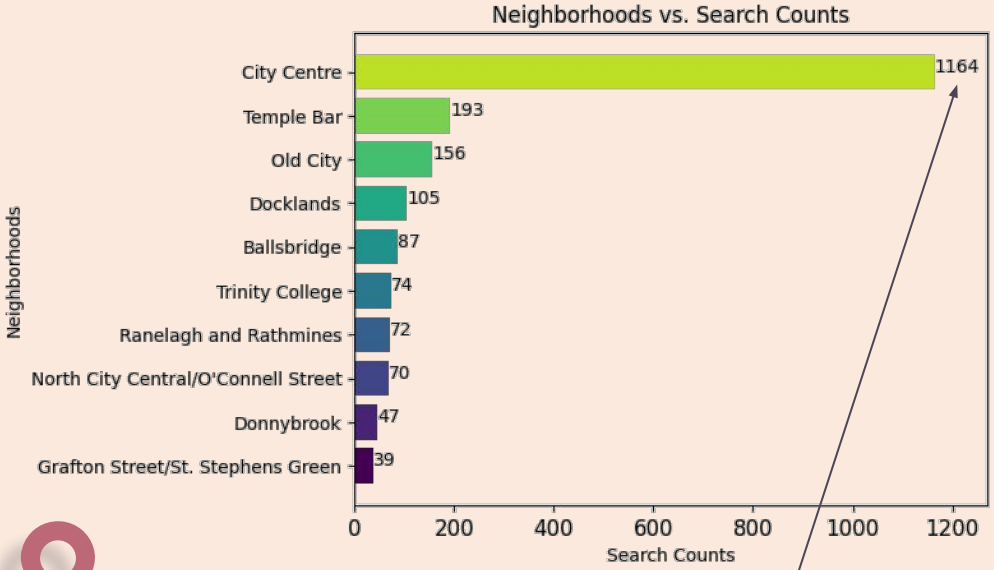
Worldwide Searches



Suggestions:

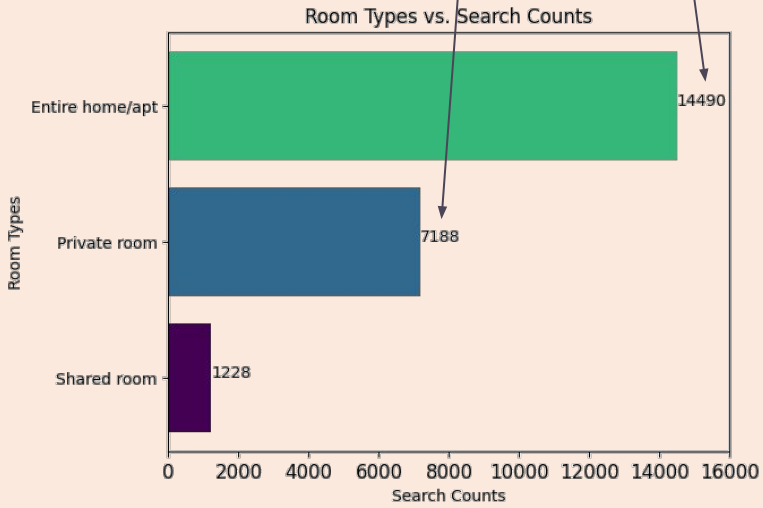
- Targeted marketing campaigns
- Language/cultural-specific accommodations
- Partnerships with travel agencies, airlines, tour operators, etc. for local promotions

Geographic Attention



City Centre: Main attraction in Dublin. Lots of things to do

Popular Types: Entire Home/Apt, Private Room



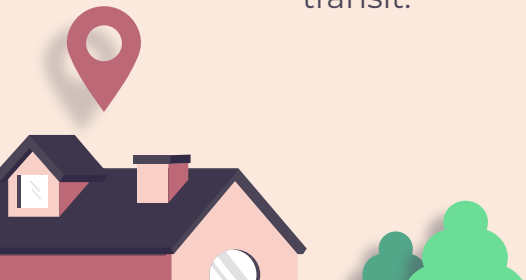
Suggestion: Convert shared rooms into more 'private' rooms if struggling to receive inquiries.

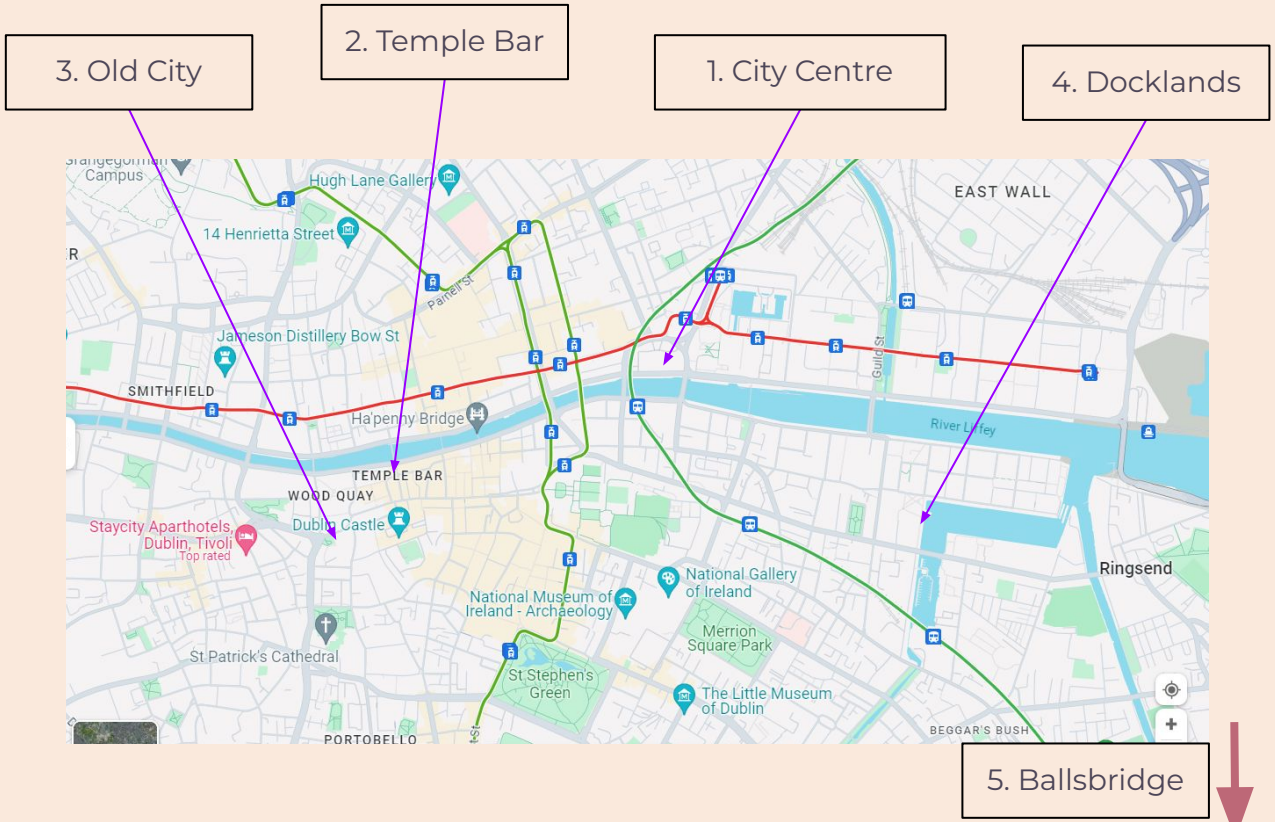
Top 5 Neighborhoods

Neighborhood demands follow proximity towards City Centre.

Suggestion:

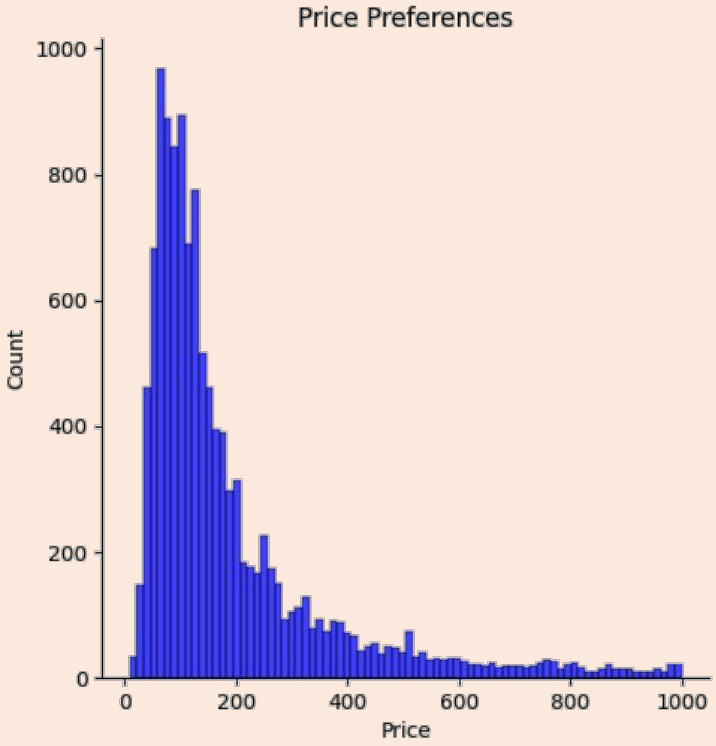
- Advertise easy/quick transportation to City Centre
  - Might sway those who aren't familiar with Ireland's transit.

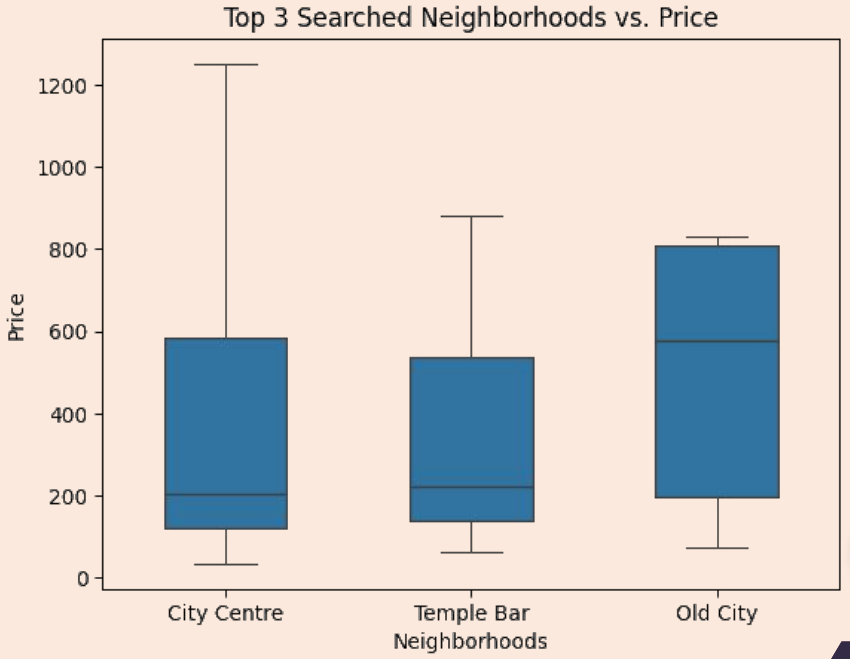




Requested Price Distribution

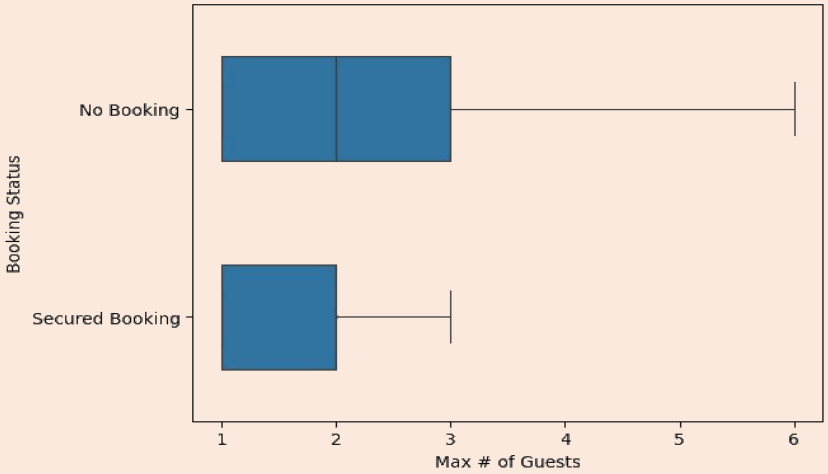
Suggestion: Have lower prices for places further away from City Centre

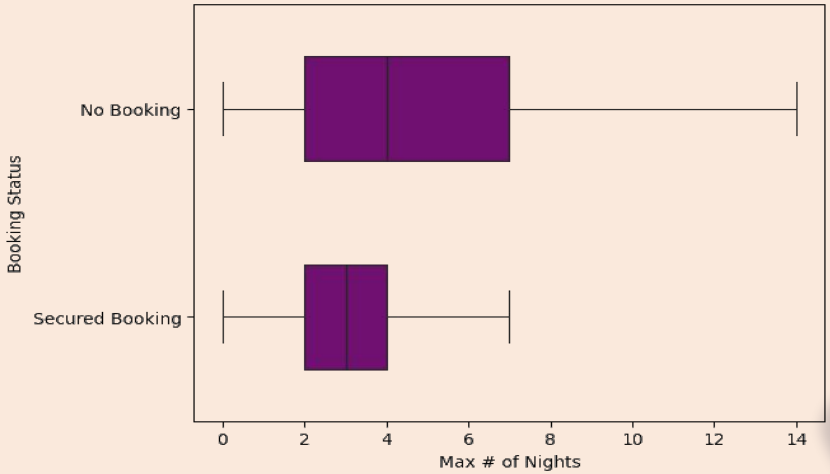




Bookings Based on Guests & Nights

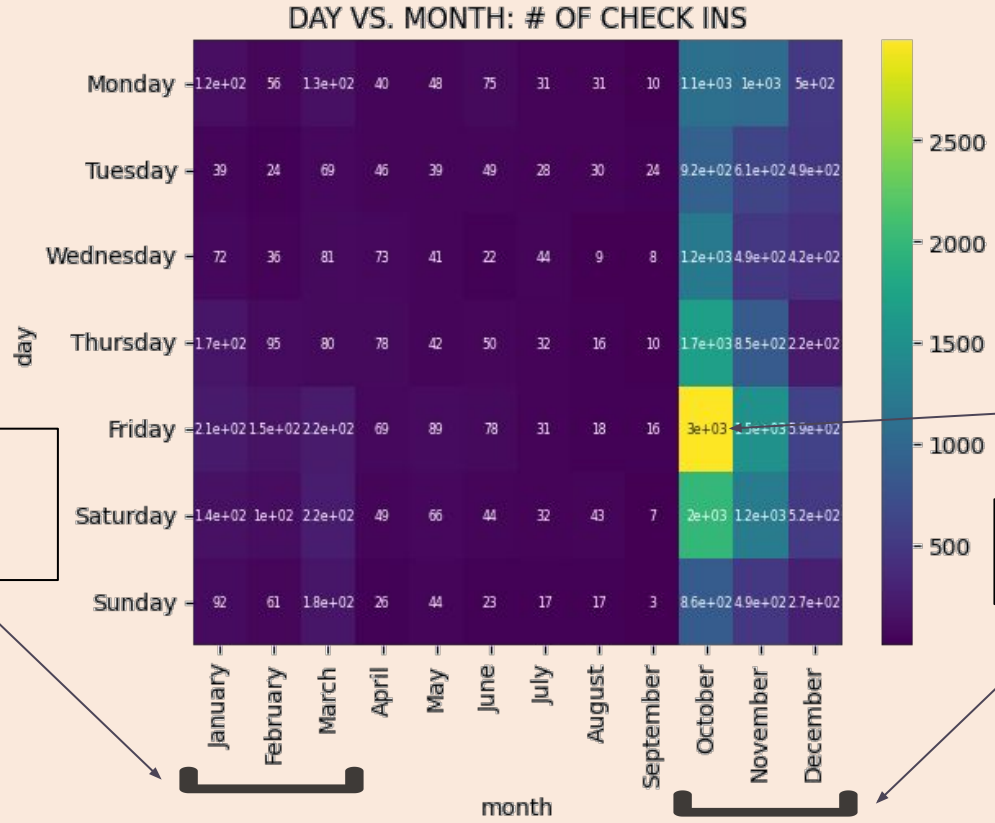
Most successful bookings are from inquiries of at most 2 guests and at most 4 nights.
Inquiries with larger parties and # of nights seem to be at a disadvantage.
Suggestion: Implement a ‘deposit’ incentive for larger inquiries (both sides could benefit)





Timing of Check-Ins

Suggestion: Lower prices during off peak months to help gain business



Winter/Spring Months: Weekends only

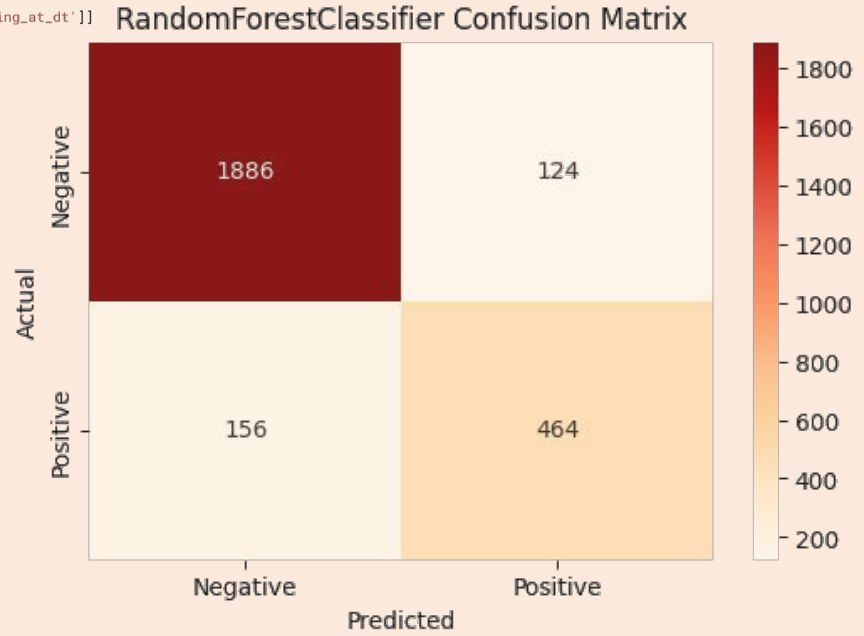
October 31st: Halloween originated in Ireland

Fall/Winter Months: Festivals

89.35% accuracy

11. RandomForestClassifier Confusion Matrix

```
search_input = searches_contacts[['n_nights', 'n_guests_max', 'filter_price_max', 'n_messages', 'ts_booking_at_dt']]

for i in range(len(search_input)):
    if pd.isna(search_input.loc[i, 'ts_booking_at_dt']):
        search_input.loc[i, 'ts_booking_at_dt'] = 0
    else:
        search_input.loc[i, 'ts_booking_at_dt'] = 1

search_input['ts_booking_at_dt'] = search_input['ts_booking_at_dt'].astype(int)
search_input = search_input.dropna()
search_input['filter_price_max'] = search_input['filter_price_max'].astype(int)
inputs = search_input[['n_nights', 'n_guests_max', 'filter_price_max', 'n_messages']]
outputs = search_input['ts_booking_at_dt']

X_train, X_test, y_train, y_test = train_test_split(inputs, outputs, test_size=0.2, random_state=42)
rf = RandomForestClassifier(random_state=42)
rf.fit(X_train, y_train)
y_pred = rf.predict(X_test)

cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(6, 4))
sns.heatmap(cm, annot=True, cmap="OrRd", fmt='g',
            xticklabels=['Negative', 'Positive'],
            yticklabels=['Negative', 'Positive'])
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('RandomForestClassifier Confusion Matrix')
plt.show()

accuracy = accuracy_score(y_test, y_pred)
print("\nAccuracy:", accuracy)

mse = mean_squared_error(y_test, y_pred)
print("Mean Squared Error:", mse)
```



THANK YOU!

Overview

OUR GUESTS

- **18.6k** potential guests from Oct. 1st, 2014 to Oct. 14th, 2014
- **Top 5** Countries are USA (1), Ireland (2), Great Britain (3), France (4), and Germany (5)

OUR SEARCHES

- **Entire Home/Apt** and **Private Rooms** are very popular
- Most guests search for AirBnbs that are available **within the next 2 months**
- Average guest size is **2**. Median # of nights is **3 days**
- **City Centre** is the most requested among searches with specified neighborhoods
- Among searches with specified prices, the median is **$200**