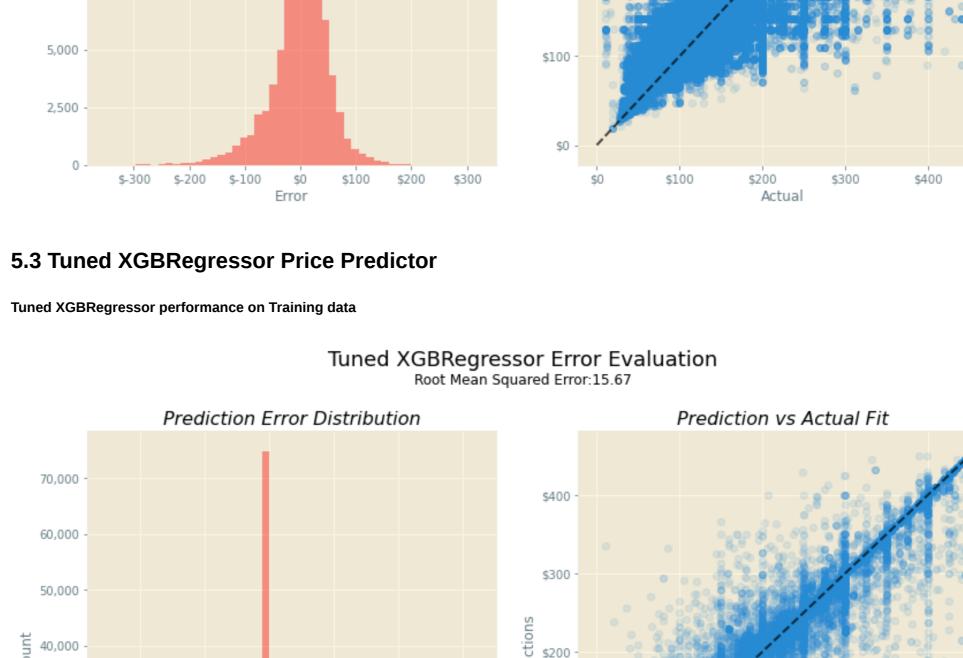
IN PROGRESS Click here to toggle on/off the raw code. 1. Introduction Airbnb is easily one of the most successful start-ups in recent memory. Airbnb's business model has been a disruptive force in the short-term rental market and was valued at \$31 billion prior to the Corona Virus epidemic (<u>source</u>%20%2D%20Airbnb%20Inc's%20new,the%20matter%20said%20on%20Tuesday.&text=Airbnb%20was%20valued%20at%20%2431,recent%20private% However, like many other businesses, COVID-19 had dire impacts on the company's business model. Due to the massive sweeps of lockdowns all across the world, the industry titan's valuation plummeted to nearly half of what it once was (source). In this analysis, I will be exploring the Airbnb rental market in San Francisco from 7/1/2018 through 6/7/2020 to develop a better understanding about the impact COVID-19 is having specifically on SF market and what it means for the company moving forward in 2020. 1.1 Objectives • Conduct an analysis of the impact COVID-19 has had on Airbnb's daily gross rental income in the San Francisco area since the lockdown began on 3/19/2020. Forecast 2020 Q2, Q3, and Q4 losses if lockdown conditions remain the same for the rest of the year. • Create a per night rental price predictor for Airbnb properties in San Francisco. Extract out the most pertinent features that predict prices best. Create a guide for first-time hosts that illustrates features and amenities that can be added to Airbnb rental to increase per night rental price. 1.2 Impact Understanding the impact COVID-19 has had on daily gross rental income will allow us to forecast losses for the rest of the year. Creating a worst-case scenario of earnings for the rest of 2020 will allow company leadership to manage expectations for shareholders and table/prioritize certain projects over others for the foreseeable future. Developing an accurate price predictor model will provide Airbnb hosts a more accurate representation of how much their property is worth on the rental platform. Doing so will increase the likelihood tourists and long-term renters will want to rent out their property. Additionally, providing an accurate guide of features and amenities for first-time hosts to consider including with their rental will not only allow for more successful first-time experiences, but also increase the number of regular Airbnb hosts throughout the year. 2. Table of Contents 1. Introduction 1.1 Objectives 1.2 Impact 2. Table of Contents 3. Data Overview 3.1 About the Data 3.2 Preview Data 3.3 Data Overview - Metrics 4. Impact of COVID 19 on Gross Rental Revenue 4.1 Methodology 4.2 Results 4.3 Decision Recommendations 5. Airbnb Rental Prices in San Francisco 5.1 Methodology 5.2 Base Model Decision Tree Model 5.3 Tuned XGBRegressor Model 5.4 Model Discussion and New Host Recommendations 6. Conclusion 6.1 Takeaways 6.2 Future Research 3. Data Overview 3.1 About the Data Airbnb Listings data was acquired from Inside Airbnb (no affiliation with Airbnb). Inside Airbnb is an independent, non-commercial set of tools and data that allows you to explore how Airbnb is really being used in cities around the world. By analyzing publicly available information about a city's Airbnb's listings, Inside Airbnb provides filters and key metrics so you can see how Airbnb is being used to compete with the residential housing market. 3.2 Preview Data Cleaned Data Shape: (126258, 48) accommodates amenities availability_30 availability_365 availability_60 availability_90 bathrooms bed_type bedrooms beds calculat TV, Cable TV 1.0 Real Bed 1.0 2.0 ,Internet,Wifi,Kitchen, Pets liv... TV,Internet,Wifi,Kitchen, Free 30 1.0 1.0 4.0 Real Bed 5 TV,Internet,Wifi,Kitchen,Heating, 1.5 Real Bed 2.0 2.0 Family/kid ... TV, Cable TV 20 2.0 3.0 1.0 Real Bed ,Internet,Wifi,Kitchen, Free par... 3 TV, Cable TV ,Internet,Wifi, Paid 131 19 1.0 Real Bed 1.0 2.0 3.3 Data Overview - Metrics **Data Description** Out[17]: 25% **50**% mean max 1.000000 2.000000e+00 2.000000e+00 4.000000e+00 1.600000e+01 accommodates 126258.0 3.217230e+00 1.947620e+00 availability_30 126258.0 9.534984e+00 1.049743e+01 0.000000 0.000000e+00 5.000000e+00 1.700000e+01 3.000000e+01 **availability_365** 126258.0 1.430715e+02 1.260210e+02 0.000000 2.500000e+01 1.070000e+02 2.630000e+02 3.650000e+02 **availability_60** 126258.0 2.290179e+01 2.104954e+01 0.000000 1.000000e+00 1.800000e+01 4.100000e+01 6.000000e+01 **availability_90** 126258.0 3.827428e+01 3.168983e+01 0.000000 5.000000e+00 3.500000e+01 6.600000e+01 9.000000e+01 0.000000 1.000000e+00 1.000000e+00 1.500000e+00 1.400000e+01 1.341998e+00 8.540870e-01 bathrooms 126258.0 **bedrooms** 126258.0 1.354290e+00 9.212903e-01 0.000000 1.000000e+00 1.000000e+00 2.000000e+00 1.400000e+01 1.768339e+00 1.206100e+00 0.000000 1.000000e+00 1.000000e+00 2.000000e+00 1.600000e+01 **calculated_host_listings_count** 126258.0 1.164335e+01 3.751799e+01 1.000000 1.000000e+00 2.000000e+00 4.000000e+00 3.010000e+02 calculated_host_listings_count_private_rooms 105994.0 2.726702e+00 7.322385e+00 0.000000 0.000000e+00 0.000000e+00 2.000000e+00 8.700000e+01 calculated_host_listings_count_shared_rooms 105994.0 4.762628e-01 2.893476e+00 0.000000 0.000000e+00 0.000000e+00 0.000000e+00 3.600000e+01 0.000000 3.500000e+01 7.500000e+01 1.200000e+02 1.350000e+03 cleaning_fee 126258.0 8.963388e+01 7.931858e+01 **extra_people** 126258.0 1.860098e+01 3.113538e+01 0.000000 0.000000e+00 5.000000e+00 2.500000e+01 3.000000e+02 guests_included 126258.0 1.797233e+00 1.407876e+00 1.000000 1.000000e+00 1.000000e+00 2.000000e+00 1.600000e+01 host_has_profile_pic 126258.0 9.942895e-01 7.535227e-02 0.000000 1.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00 28.000000 3.677879e+06 1.628587e+07 6.464682e+07 3.452837e+08 host_id 126258.0 4.915301e+07 6.890265e+07 1.000000e+00 1.000000e+00 1.000000e+00 host_identity_verified 126258.0 5.709816e-01 4.949379e-01 0.000000 0.000000e+00 host_is_superhost 126258.0 5.115636e-01 4.998682e-01 0.000000 0.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00 host_response_rate 126258.0 8.534871e+01 3.282895e+01 0.000000 9.700000e+01 1.000000e+02 1.000000e+02 1.000000e+02 958.000000 6.495015e+06 1.709399e+07 1.140872e+07 1.771798e+07 2.584326e+07 4.348882e+07 instant_bookable 126258.0 3.895436e-01 4.876487e-01 0.000000 0.000000e+00 0.000000e+00 1.000000e+00 1.000000e+00 is_location_exact 126258.0 8.863676e-01 3.173655e-01 0.000000 1.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00 37.704170 3.774793e+01 3.776351e+01 3.778060e+01 3.781031e+01 latitude 126258.0 3.776334e+01 2.287561e-02 -122.513065 -1.224454e+02 -1.224277e+02 -1.224120e+02 -1.223670e+02 -1.224327e+02 2.740284e-02 1.000000 2.800000e+01 9.000000e+01 1.125000e+03 2.147484e+09 maximum_nights 126258.0 1.747421e+04 6.043664e+06 1.072605e+01 1.590519e+01 1.000000 2.000000e+00 3.000000e+00 3.000000e+01 9.990000e+02 700.000000 2.500000e+03 3.500000e+03 4.999000e+03 2.500000e+04 monthly_price 19067.0 4.177245e+03 2.740558e+03 6.153323e+01 8.585735e+01 0.000000 6.00000e+00 2.600000e+01 8.400000e+01 8.210000e+02 price 126258.0 2.130139e+02 3.284934e+02 10.000000 9.900000e+01 1.490000e+02 2.390000e+02 1.000000e+04 0.000000e+00 0.000000e+00 1.000000e+00 require_guest_phone_verification 126258.0 5.468168e-02 2.273588e-01 0.000000 0.00000e+00 require_guest_profile_picture 126258.0 4.589016e-02 2.092477e-01 0.000000 0.000000e+00 0.000000e+00 0.000000e+00 1.000000e+00 9.941865e-01 7.602461e-02 0.000000 1.000000e+00 1.000000e+00 1.000000e+00 1.000000e+00 20.000000 9.500000e+01 9.800000e+01 1.000000e+02 1.000000e+02 **review_scores_rating** 126258.0 9.570775e+01 6.869996e+00 2.048854e+00 2.088294e+00 0.000000 4.200000e-01 1.290000e+00 3.080000e+00 4.107000e+01 0.000000 0.000000e+00 1.500000e+02 5.000000e+02 5.000000e+03 **security_deposit** 126258.0 3.419327e+02 6.260590e+02 1.297547e+03 9.815052e+02 250.000000 7.500000e+02 1.000000e+03 1.500000e+03 9.999000e+03 **zipcode** 126258.0 9.411478e+04 1.315900e+01 94013.000000 9.410900e+04 9.411400e+04 9.412100e+04 9.496500e+04 4. Impact of COVID 19 on Gross Rental Revenue 4.1 Methodology Due to the relatively new impact of COVID-19 shutdown, there is not enough data to accurately model an accurate time series forecast. To circumvent this, I calculated the average daily percentage loss based on the data we presently have concerning the lockdown (March 19/2020 – 6/6/2020) and then extrapolated out that loss against our previous 2019 data for Q2, Q3, and Q4 of 2020. It is safe to assume that 2020 would have resulted in similar(if not slightly greater) revenue than 2019, so this extrapolation will serve as a worst-case scenario. 4.2 Results **Losses Due to Initial Shutdown** The loss due to COVID-19 shutdown for the data we have concerning 2020 is below, data ranges from 3/19/2020 - 6/7/2020 Out[18]: Daily Gross Renatal Income Comparing 2020 Shutdown to 2019 Statistics 2020 Gross -- 2019 Gross \$80,000 \$60,000 Net Daily Income 2020 during Shutdown(3/19/2020 - 6/7/2020) Total Loss: \$-2426995.0 Percent Loss:-88.21% 2020 Income 2020 Losses Forecasting losses through Q2, Q3, and Q4 in 2020. Out[19]: Daily Rental Income 2019 Q2 Vs 2020 Q2 Projections — Q2 2019 Gross — Q2 2020 Projection \$100,000 \$60,000 Total Projected Loss for Q2: \$-2,692,339.16 \$-40,000 \$-60,000 \$-80,000 2020 Q2 Projected Income 2020 Q2 Projected Loss \$-100,000 Daily Rental Income 2019 Q3 Vs 2020 Q3 Projections \$120,000 — Q3 2019 Gross — Q3 2020 Projection \$40,000 \$20,000 Total Projected Loss for Q3: \$-3,402,457.29 \$-60,000 \$-80,000 2020 Q3 Projected Income \$-100,000 2020 Q3 Projected Loss Daily Rental Income 2019 Q4 Vs 2020 Q4 Projections — Q4 2019 Gross - Q4 2020 Projection \$500,000 \$400,000 \$300,000 \$200,000 \$100,000 Total Projected Loss for Q4: \$-4,097,722.34 \$-100,000 \$-200,000 \$-300,000 \$-400,000 2020 Q4 Projected Income 2020 Q4 Projected Loss \$-500,000 **4.3 Decision Recommendations** San Francisco Daily Gross Income Analysis Results: Q2 Projected Loss: \$2,692,339.16 Q3 Projected Loss: \$3,402,457.29 Q4 Projected Loss: \$4,097,722.34 Projected Net 2020 Losses: \$10,192,518.79 Looking at the results, things do not look great should the lockdown remain in place as currently constructed for the rest of 2020 in San Francisco. Due to concerns from COVID 19, it would be a wise decision to only offer listings where the full-property is rented to a single-person/family. No single rooms should be permitted. Airbnb does already require cleaning to be done after each rental is completed (which is included in the price), so consumers will feel more at ease should we make this mandate. Somethings to consider for Airbnb as well is moving into the long-term rental market. The most profitable/popular listings on the platform belong to long-term rental companies who provide fully furnish dwellings in San Francisco. A potential strategy to ensure secured revenue would be to work with local apartments in SF and use our platform/advertising as a means to attempt to book empty dwellings. Airbnb thus could then capture either a commission or a percentage of the monthly rent in long-term deals. **5. Airbnb Rental Prices in San Francisco 5.1 Methodology** After aggregating and cleaning Airbnb listings data in San Francisco, created a preprocessing pipeline that scaled numeric and encoded categorical variables. Created a base model decision tree as initial rental price predictor to serve as a baseline for testing out ensemble and boosting methods. After testing a few more complex models, settled on optimizing an XGBRegressor model using GridSearchCV and additionally extracted the most important features to the model's performance to serve as a reference to new Airbnb hosts. **5.2 Base Model Decision Tree** Out[20]: Base Tree Regressor Model Error Evaluation Root Mean Squared Error: 48.86 Prediction Error Distribution Prediction vs Actual Fit 17,500 \$400 15,000 12,500 \$300 ₩ 10,000 7,500 5,000

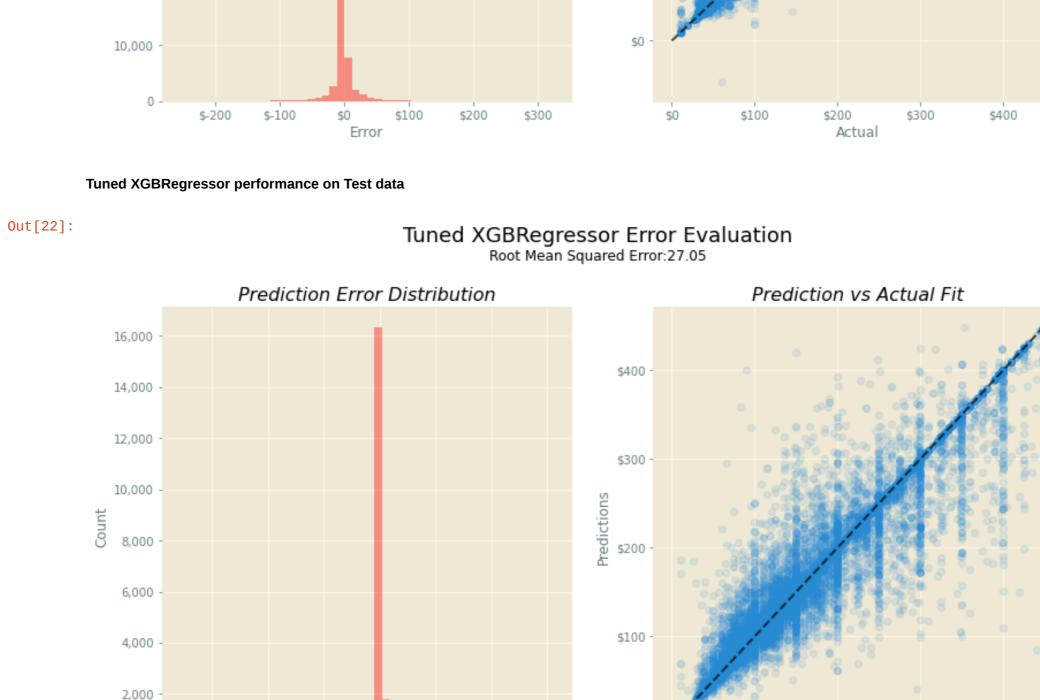


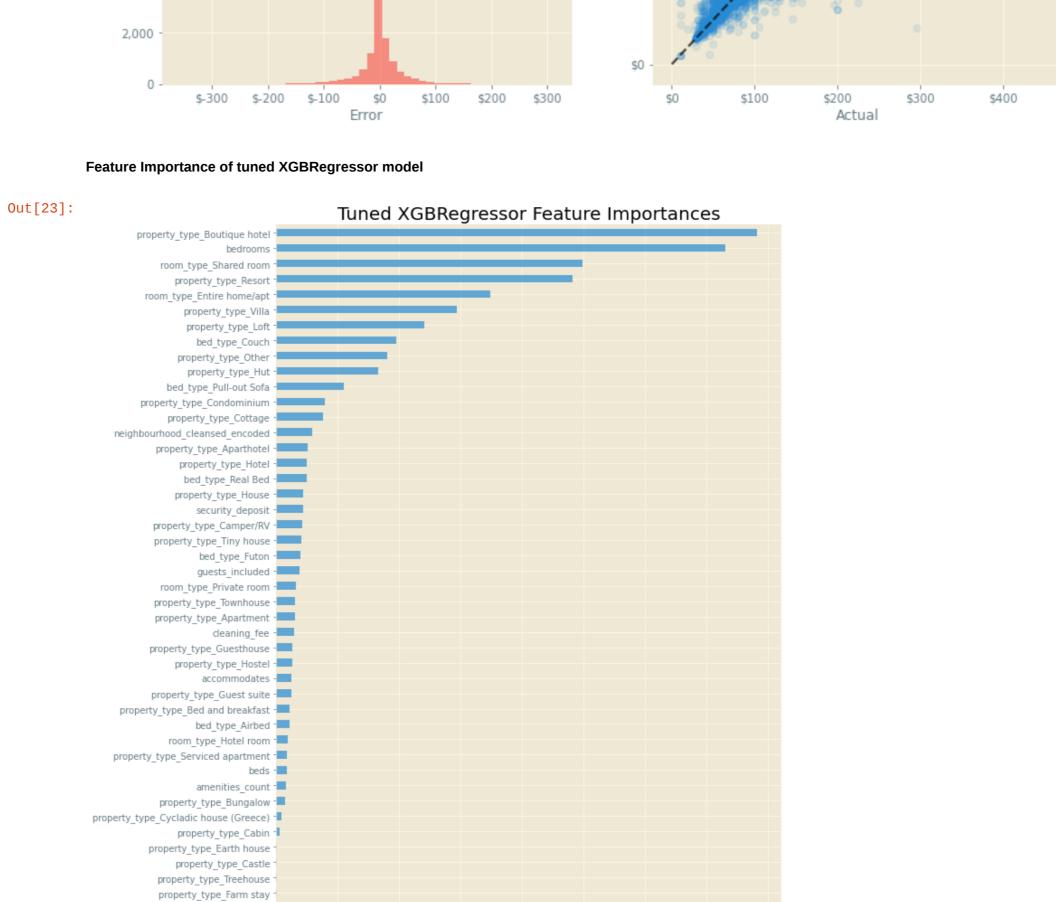
\$200

Out[21]:

30,000

20,000





0.00 0.02 0.04 0.06 0.08 0.10 0.12 0.14 0.16 5.4 Model Discussion and New Host Recommendations **Model Discussion**

property_type_In-law = property_type_Bus property_type_Dome house -

high or too low for what is provided.

Strengths

property type Boat

given property (+/- \$27.00) Weaknesses • Model does not do well when prices are closer to the minimum value (\$10). This could be due to users pricing their property too low (users are at liberty to price their property as they see fit) • Model under-estimates at increments of \$50. This may be due to user behavior (they're more likely to price at a whole number).

• Overall, model performs quite nicely. There are some listings where it over/under predicts but this may be a consequence of the user setting a price too

• Because the model can accommodate almost any feature one can think of when listing a property for rent, it gives a very accurate range of prices for any

Feature Importance

- Big Picture Aplication of model Due to COVID-19, it may be time to consider setting ranges on what prices can be set based on the model given the input about the property from the Airbnb host. Though users have a lot of flexibility to set prices as they git, this ensures each rental is more optimal/likely to be rented by a consumer on the site. **Property Recommendations for New Airbnb Hosts**
- Consider offering the entire property or an entire section of the property such as an floor or side of the property.
- More bedrooms and more privacy for renters was important in how much renters were willing to pay in order to rent the property. Providing a bed for your visitor is not a hyper-critical feature of the listing. Couches, pull-out beds and futons do not have a detrimental effect on the value of the listing

Some critical components of a successful Airbnb listing are the amenities offered that can make a listing really feel comfortable for a visitor. Though we do not have specifics, more amenities provided does appear to correlate positively to a higher rental price. 6. Conclusion

• Increase the number of amenities your property provides.

6.1 Takeaways

to grow, it may be worth considering entering the long-term rental market to serve as a middle-man between empty rental properties and those looking for long-term housing. The top hosts who generate the most revenue in SF are in fact long-term rental companies that provide furnished rental properties to users in SF. This may be a market to consider entering. Airbnb presently allows users to price their property as they see fit. With the impact of the Corona Virus, to guarantee fair pricing and increase the likelihood of

COVID-19 has had an immense impact on Airbnb's business model. It has exposed a weakness that we can grow and learn from. As the company continues

listings being rented, it may be worth considering using the updated model that calculates per night prices given features of the hosts' property and allow the host to select the price within a specified range. This way, prices are fair for potential travelers and hosts would have to provide additionally amenities to their property in order to charge higher prices.

6.2 Future Research In the future, the most important feature to know from each listing would be the square footage of the property being offered (be it a single room or an entire house). I intend on conducting a user reviews analysis to isolate the most important words and phrases used that indicated positive/negative experiences to provide new users with new information about the do's and don'ts of good hosting.