To Buy or Not to Buy?

An analysis of crime's impact on rent prices in the San Francisco Bay Area

By: Meng-Kang Kao, Brian Tung, Nic Brathwaite, and Kevin Cahillane

Good evening everyone. Thank you for coming to our presentation. We are Scott, Brian, Nic and Kevin and this is To Buy or Not to Buy, an analysis of crime's impact on rent prices in the San Francisco Bay Area.

Research Question

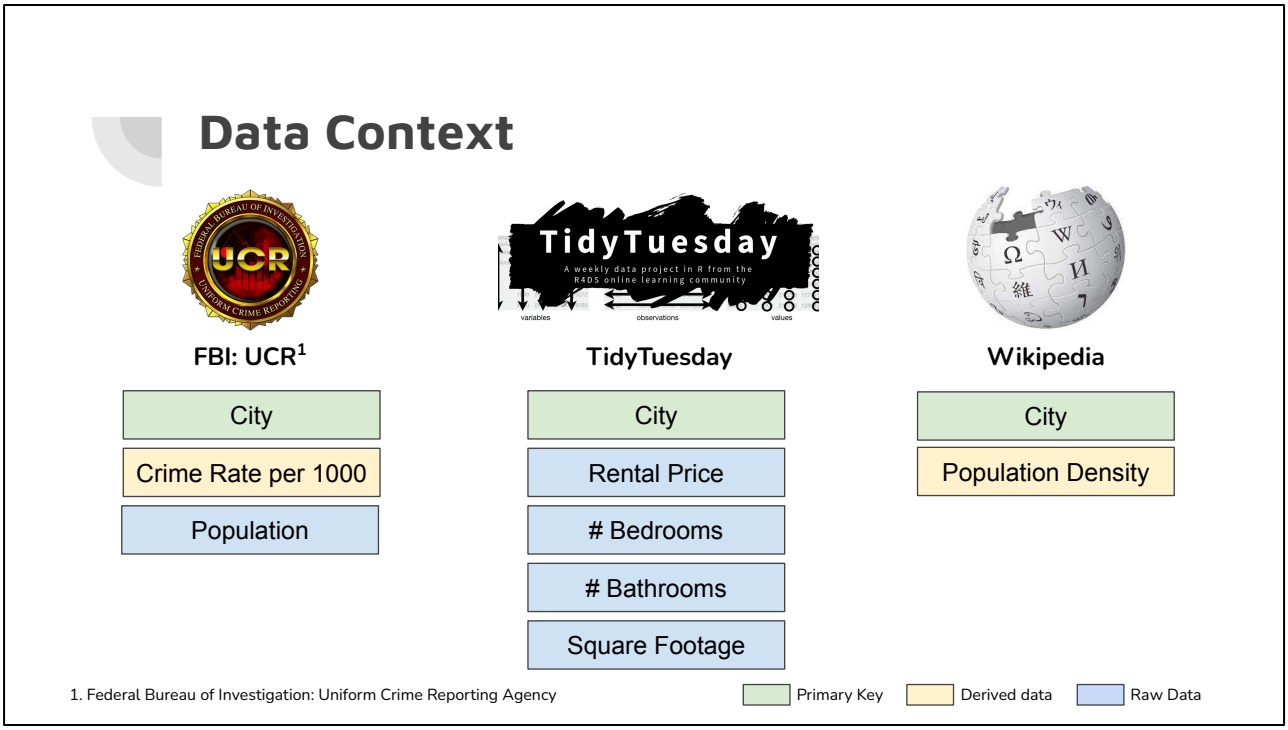
Research Question: Does the **city-level crime** rate have an impact on **home rental prices** in the Bay Area?

2 minutes max

Brian: Home ownership has always been a core component in the American Dream and a milestone of prosperity. In fact, many Americans prioritize it over other life goals such as career, family, and college. Further a common way for Americans to build their wealth through home ownership is to build out a portfolio of rental properties as a landlord. This is most evident in a wealth-concentrated location like the Bay Area where second-home sales jumped 54% in 2021 from pre-pandemic levels per Mercury News.

However, today's competitive housing market and astronomical prices force buyers to make compromises in purchasing decisions, such as neighborhood safety. Experts have found that properties in high crime areas may be subject to frequent property damage, low-quality renters, and lower property values. For home buyers with a tight budget, uncertainty in navigating real estate opportunities may likely lead to a poor investment decision.

The core objective of our study is to leverage rental listing data in the Bay Area along with city safety information in hopes to provide guidance for home buyers to make the best purchasing decision and maximize ROI. To provide better guidance for potential home buyers and property investors, we aim to utilize regression models to answer the question: Does the **city-level crime** rate have an impact on **home rental prices** in the Bay Area?



2-3 minutes

Kevin: The data used in our study consists of datasets from three separate sources.

The first dataset is part of "San Francisco Rentals" dataset on TidyTuesday, which was web-scraped from Bay Area Craigslist rental listings using the internet archive repository, Wayback Machine. This dataset contains detailed property information on city, county, rental price, number of bedrooms, number of bathrooms, and square footage. To avoid any time series complications, we chose to limit our model just 2012 which was the year with the most listings in our rentals dataset.

Our second dataset is a violent crimes report from the FBI's Uniform Crime Reporting agency, which includes categorical crime rates and population data by city. The violent crimes dataset was collected via the Hierarchy Rule, which only counts the most serious offense in a multiple-offense criminal incident.

The third dataset is the city area in square miles from Wikipedia. We then derived two more fields: population density which is population divided by the city square mileage and crime per 1000 residents which is the total crime count divided by the population multiplied by 1000. We refer to the latter as the city crime rate. All of the datasets used in this study are observational.

Regression Results

Model Progression

1. Initial LM omitting population density
2. Adding population density
3. Omitting square feet

Table 1: Estimated Regressions

	Output Variable: Rental Cost		
	(1)	(2)	(3)
Number of bedrooms	−89.67	36.40	245.37***
	(71.73)	(67.05)	(26.97)
Number of bathrooms	511.53***	486.17***	689.61***
	(98.04)	(89.51)	(57.22)
Square feet	0.70**	0.61**	
	(0.23)	(0.21)	
Population density(K)		130.22***	137.28***
		(4.70)	(4.65)
Crime per 1000 population	88.70***	−125.23***	−128.98***
	(11.22)	(8.39)	(9.08)
Constant	417.99***	34.81	−81.39
	(84.21)	(83.16)	(94.08)
Observations	3,374	3,374	3,374
Adjusted R^2	0.36	0.50	0.45
Residual Std. Error	979.43 (df = 3369)	866.84 (df = 3368)	912.68 (df = 3369)
Note:			*p<0.05; **p<0.01; ***p<0.001

2 minutes minimum

Brian: Over the course of running linear regressions on our data, we created 3 distinct models that substantially vary in coefficient and p-value for independent variables. We began with a linear model that includes basic rental listing information, such as number of bedrooms, number of bathrooms, square footage, and our key variable of interest for this study, crime rate. For the second model, we added in population density as a proxy for the urbanness of a city. Lastly, in model 3, we removed square footage and just regressed against the number of beds, baths, population density and crime rate. Next, we'll review our reasoning behind this model progression as well as interpretation of the results from each model.

Regression Results

Model	Observation
1	Crime rate is statistically significant with a positive coefficient
2	Pop. density has a positive impact and crime rate has a negative impact; both are statistically significant
3	Pop. density has a positive impact and crime rate has a negative impact; We believe this model is the best representation of the real-world.

Table 1: Estimated Regressions

	Output Variable: Rental Cost		
	(1)	(2)	(3)
Number of bedrooms	−89.67	36.40	245.37***
	(71.73)	(67.05)	(26.97)
Number of bathrooms	511.53***	486.17***	689.61***
	(98.04)	(89.51)	(57.22)
Square feet	0.70**	0.61**	
	(0.23)	(0.21)	
Population density(K)		130.22***	137.28***
		(4.70)	(4.65)
Crime per 1000 population	88.70***	−125.23***	−128.98***
	(11.22)	(8.39)	(9.08)
Constant	417.99***	34.81	−81.39
	(84.21)	(83.16)	(94.08)
Observations	3,374	3,374	3,374
Adjusted R^2	0.36	0.50	0.45
Residual Std. Error	979.43 (df = 3369)	866.84 (df = 3368)	912.68 (df = 3369)
Note:			*p<0.05; **p<0.01; ***p<0.001

2 minutes minimum

Scott's write-up: Kevin talking

For the first model where we had omitted a measure of urbanness we saw that the crime rate seemed to have a positive, statistically significant affect on rental prices increasing rent $88 for every additional crime per 1000 residents. This went against both our intuition and prior academic literature on this topic and suggested an omitted variable bias which prompted us to run a second regression including population density.

In model 2 we see that population density has an even more positive and significant affect on price and even more importantly crimes coefficient has flipped to now have a significant negative impact on price. 1 incremental crime per 1000 residents now reduces rent by $125 which is certainly not negligible from a landlord's standpoint. Additionally, we noticed that the number of bedrooms were not statistically significant in either of the models and suspected this might be because it was likely strongly correlated with the square footage. We figured that from a renter's perspective, the number of bedrooms is more important than the total square footage so we decided to develop a third model where we removed square footage.

Scott Talking

We believe Model 3 is the best representation of how crime affects the price of rent. In the absence of square footage, we see that the number of bedrooms becomes statistically significant in estimating the price of rent. For potential renters, the number of rooms often serves as the most significant constraint they consider when searching for homes. This model

clearly reflects that showing that an incremental bedroom will raise rent by $245 per month and an incremental bathroom will raise rent by nearly $700. , crime still decreases the overall price of rent in confirmation with our initial belief.

Additionally, we see that the removal of the square footage variable has no impact on the direction or significance of the other variables. Population density is still statistically significant increasing by rent $137 for an additional 1000 people per square mile. Even more importantly, our interpretation of the impact of crime is not weakened as now an incremental crime per 1000 people will reduce rent by $129 per month. In total, we believe that this third model offers the most practical significance in describing how the variables of interest affect rent pricing.

Conclusion

Target Audience: Landowners and prospect property buyers

Limitations: Omitted Variables, available data on houses, craigslist

Key Takeaways: What should land owners and property owners consider most in buying homes? How impactful is crime on the price of rent?

Future Research

1-2 minutes

Nic: This study would be useful to landowners and prospective property buyers. There are some potential limitations to the interpretation and application of our findings. To start, we do not believe that our dataset is IID, since Craigslist data is not able to capture the entire rental market. As highlighted by the author of the dataset, there is reason to expect that the dataset fails to capture the highest end of the market, where housing transactions are primarily managed by real estate agencies.

We also anticipate our model could be subject to omitted variable bias by not including city-level income data due to a lack of reliable data sources. We expect income to have a negative effect on crime rate meaning the omitted variable bias is pushing the negative coefficient of crime towards zero. Therefore, the inclusion of income in our model would not reduce the significance of our crime rate coefficient.

In future research within this field, we would advise researchers to consider including average income if possible. Additionally, we would like to see supplemental analysis done regarding how crime types affect the cost of rent. A potential subquestion could be: does violent crime affect rent prices more than property crime?

Overall, however, we believe this study provides useful information for prospective property buyers regarding how local crime negatively affects their potential return via rent.