Predicting Rio Negro real estate tax

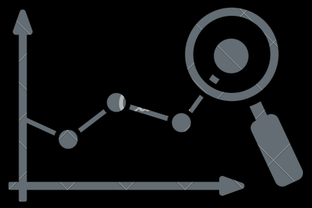
Sara Gonzalez, David Arbelaez, Andres Caballero y Oscar Rodriguez

Highlights

- Determine the next income from the real estate taxes.
- Construct a trustworthy database for the forecasting of the variables
- Build an app to show the results of the predictions of the model

Background

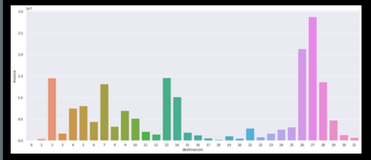
One of the incomes of a country or a city comes from taxes. Rio Negro, Colombia's municipality, every year needs to know how much money will have the next years to organize their expenses for those years. For that reason, the aim of this project is to predict the next incomes coming from the real estate taxes.



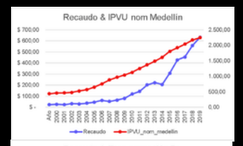
Business Impact

Predicting the income from taxes, it is believed that efficient tax administration can highly improve the quality of life of the citizens and the infrastructure of the city. Also, helps the administration to make a better administration of their expenses.

Exploratory Data Analysis



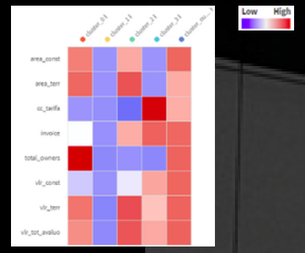
On average the destination 27, properties dedicated to build educational centers, pays the highest taxes.



The IPVU nom Medellin has a major correlation with the dependent variable (Recaudo). The occupancy rate presents a big correlation with the Y and 3 other variables. And the last variable selected for the model is the TRM. Also, the behaviour of he real estate tax and the IPVU seems like an exponential graph.

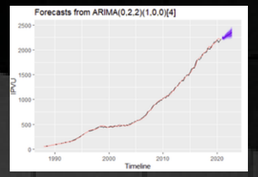
Model

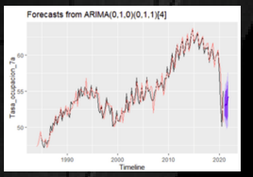
First, it was made a segmentation model (Kmean, k =4) having in count the characteristic of the properties, to have a better prediction. But this one can be used in the future when the Id of the property is related to their bill.



Model

The final model was constructed with Mulltiple Linear Regression and for the prediction of the next year, it was made different models for each feature.





Tasa de Cambio Nominal		
Analistas Locales		
	Alianza Valores	3.750
	ANIF	n.d.
	Banco de Bogotá	3.950
	Bancolombia	n.d.
	BBVA Colombia	3.430
	BTG Pactual	n.d.
	Corficolombiana	3.557
	Corredores Davivienda /2	n.d.
	Credicorp Capital /3	3.300
	Davivienda	n.d.
	Fedesarrollo	n.d.
	Itaú /1	3.950
	Ultraserfinco /4	n.d.
	Promedio	**3.658**
Analistas Externos		
	Citigroup	n.d.
	Deutsche Bank	n.d.
	Goldman Sachs	3.200
	JP Morgan	n.d.
	Promedio	**3.200**

For IPVU and occupancy rate, it used ARIMA models and for the TRM it used expert analysis. And those values went for the final model to determine the future incomes. Using Amazon EC2 and Plotly Dash was placed an app and in this, the user can suggest their own values. The app is in the next URL : http://18.218.156.175:8050/

Conclusions

- A specific prediction can be made using the segmentation model, however, it was not possible because there is not a relationship between the properties characteristics database and real estate tax.
- The prediction model was carried out through macroeconomic variables that helped to understand the historical economic situation of the country and describe the behavior of the real estate income.