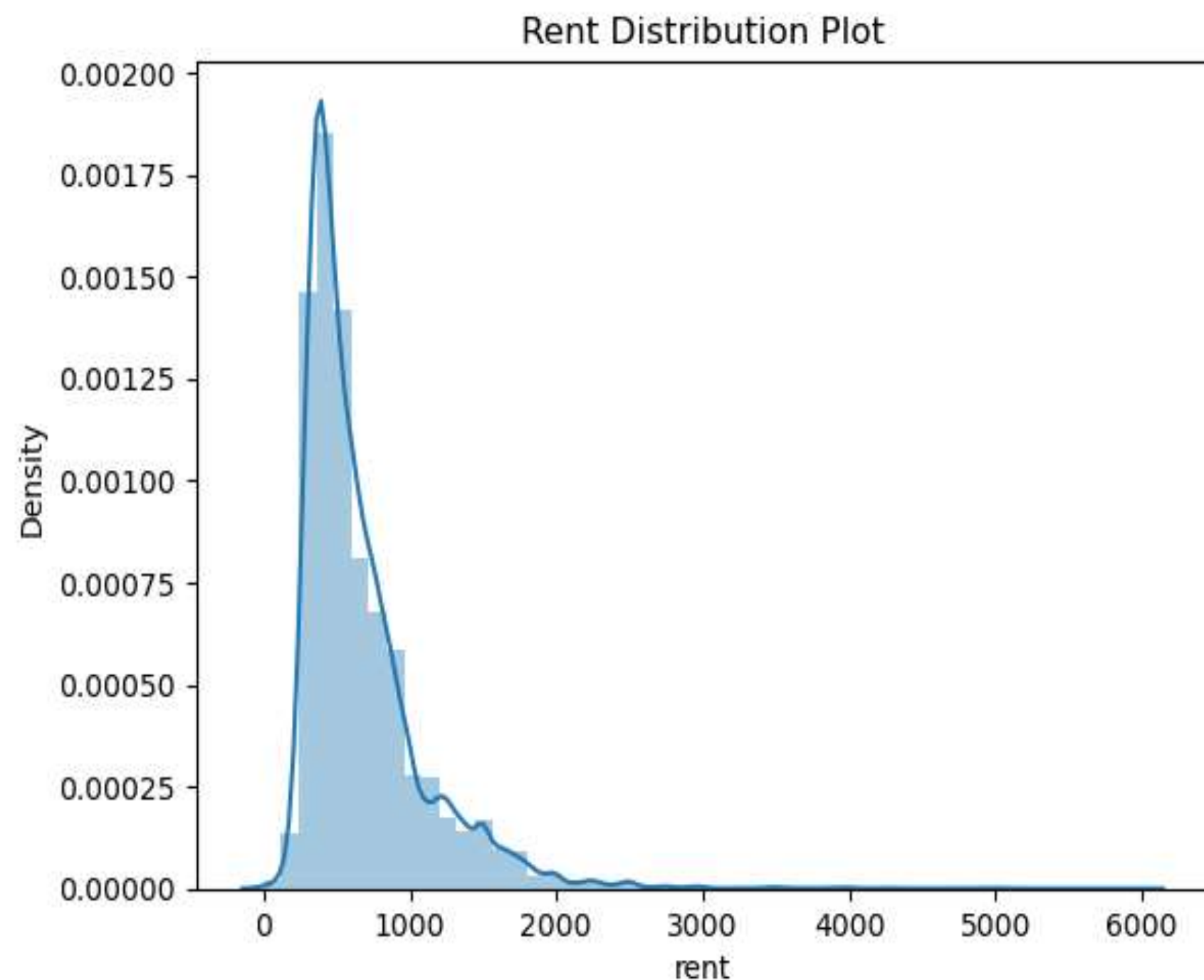


Real Estate AI-ML Case Study

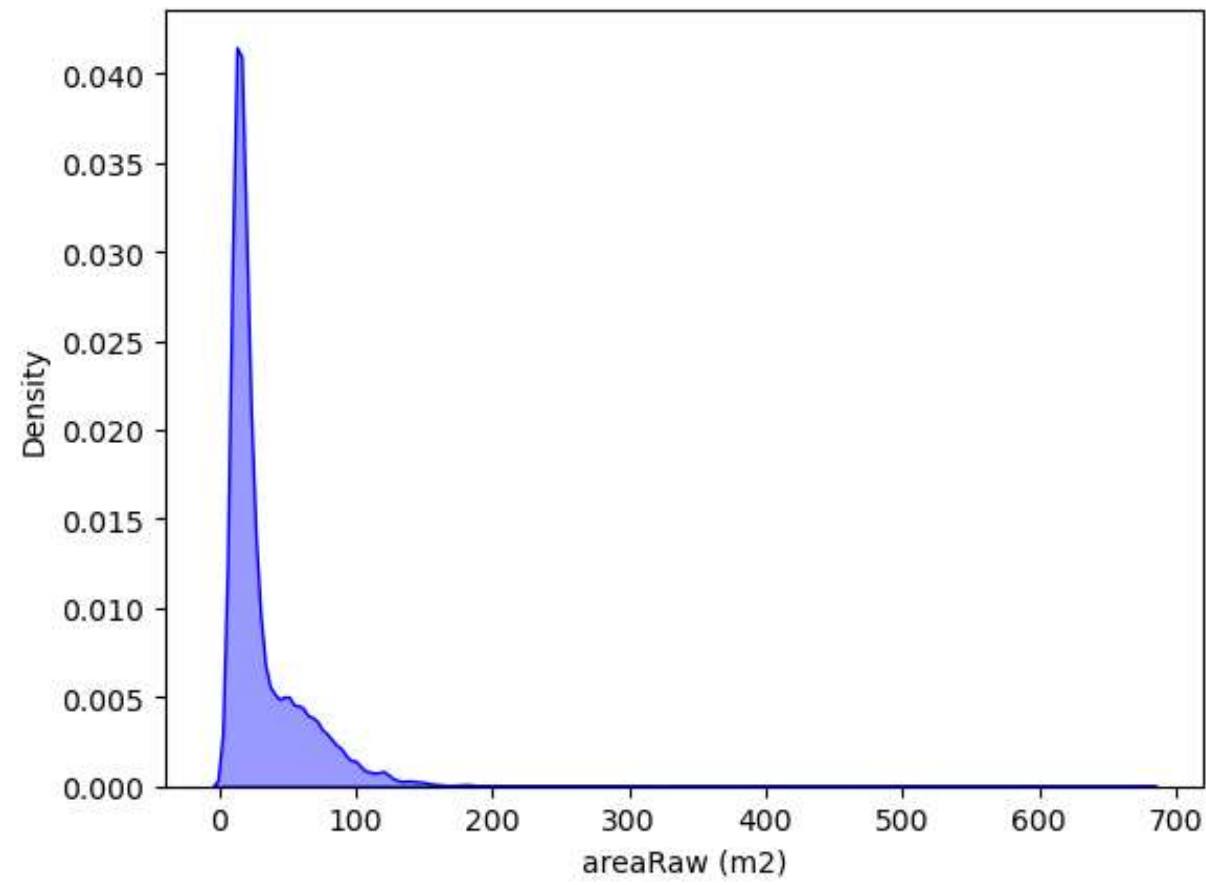
The purpose of this data is to be able to investigate trends and patterns in the real-estate rental market in Netherlands. A real estate company in Netherlands wants to improve their property valuation process and increase the accuracy of property prices. They currently rely on manual property valuation methods, which can be time-consuming and prone to errors. The company wants to leverage Auto-ML to automate the process and improve accuracy.

The aim of this study is to predict the rent of a property and find the factors influencing the rent using Auto-ML.

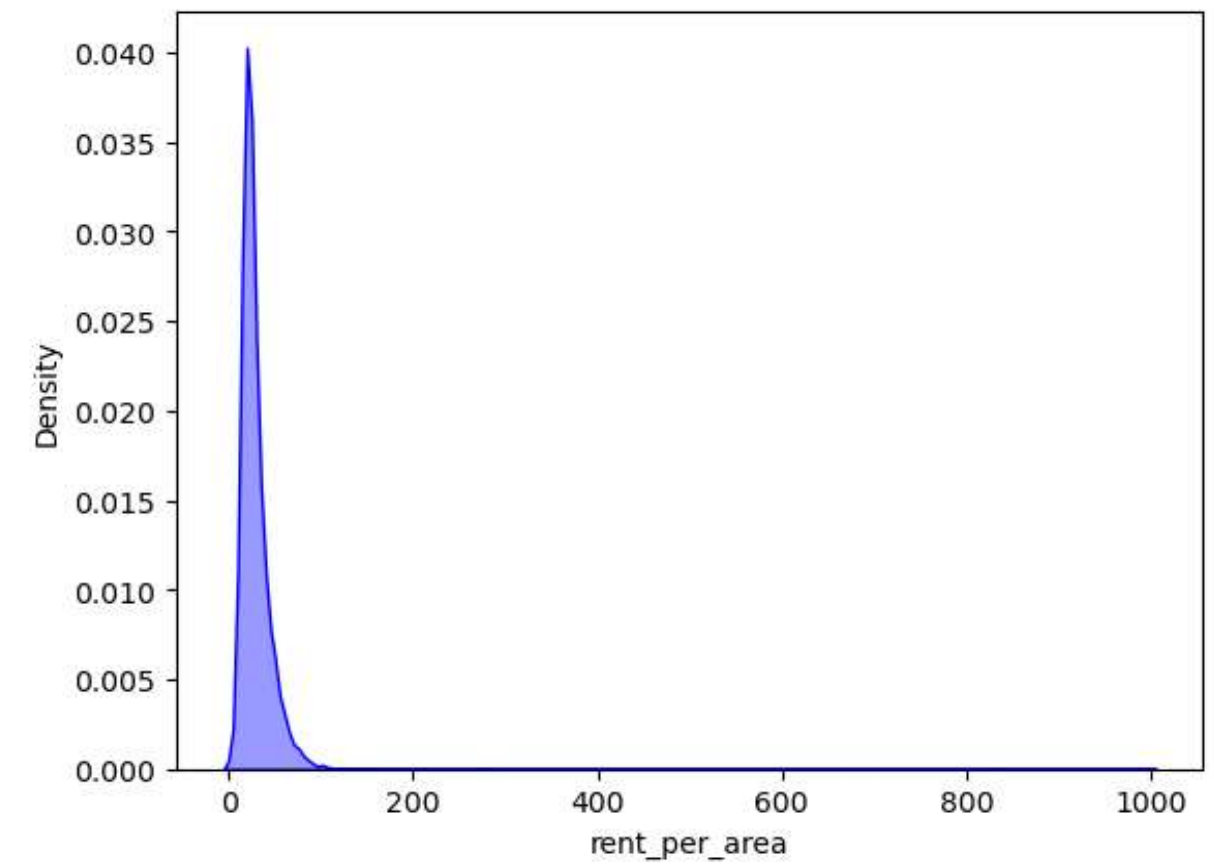


- The business feature used here is Rent.
- Rent feature refers to the rent of the property.
- The distribution plot shows that most of the rent is below 2000 euro.

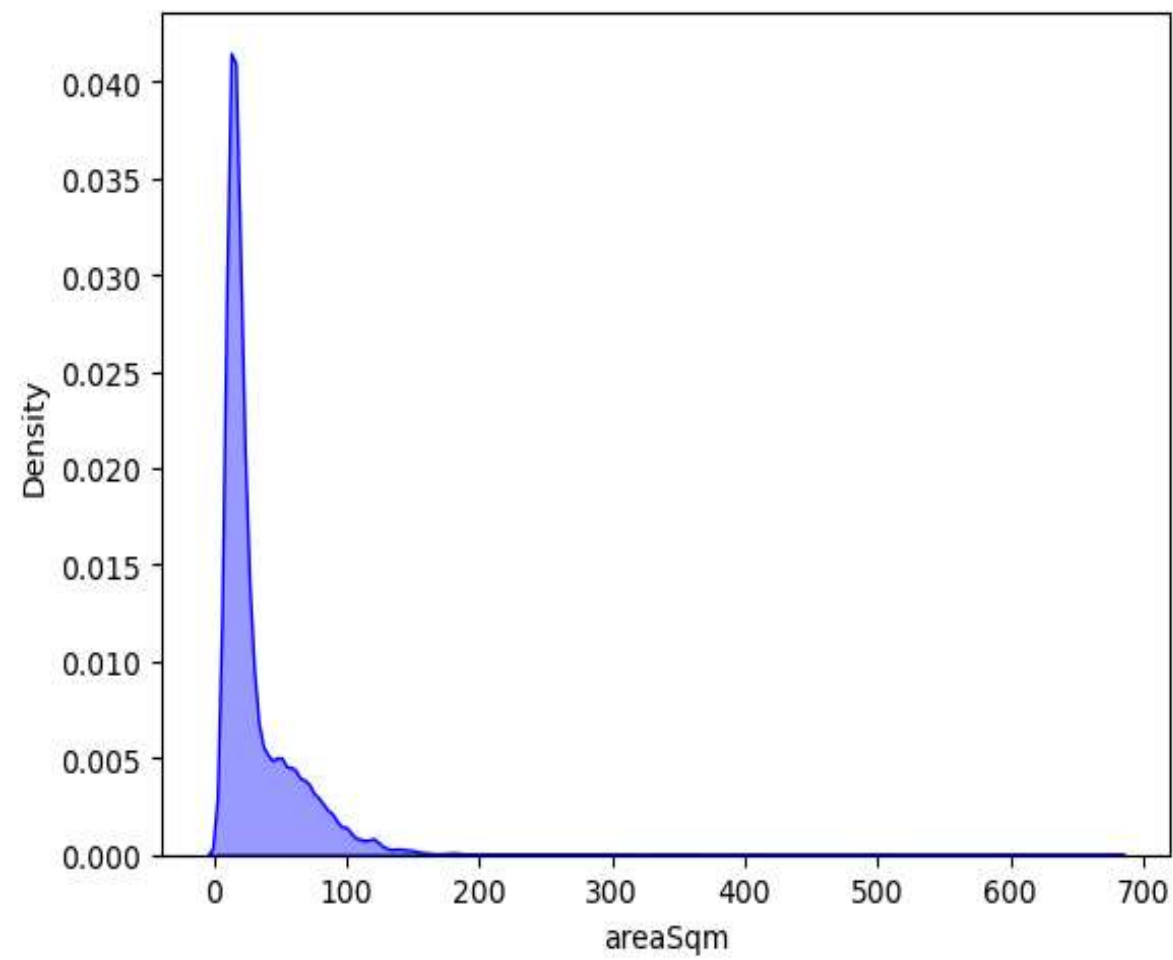
Features Responsible



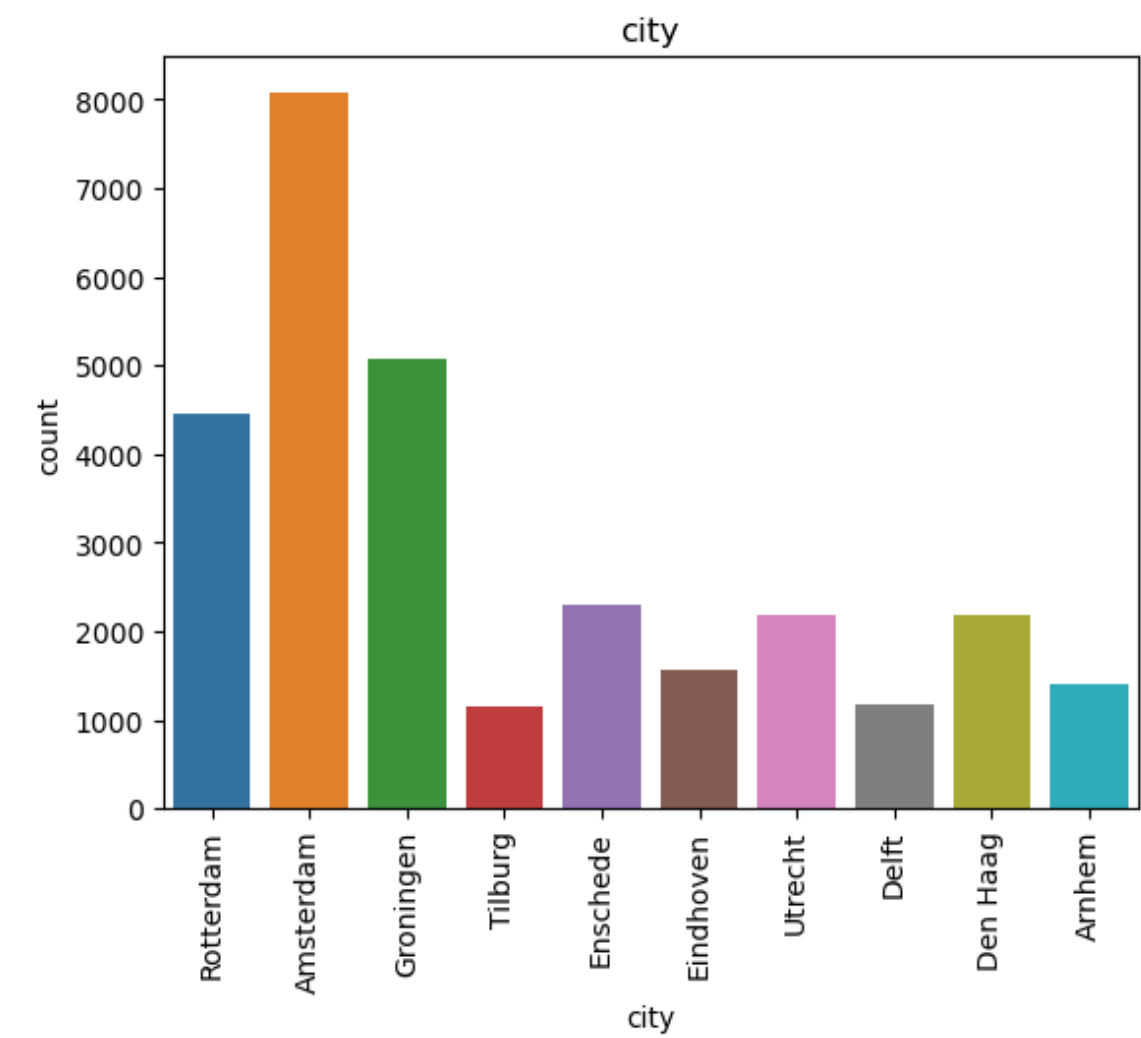
- **areaRaw (m2)** : The meter square of raw land



- **rent_per_area** : The rent per area square meter (rent/areaSqm)



- **areaSqm** : The square meter of the rental property



- **city** : name of the city where the rental property is in
- top 10 cities where people prefer to rent properties in

Auto-ML Methodology Results

Algorithms	Test Accuracy (25 percentile)	Test Accuracy (50 percentile)	Test Accuracy (75 percentile)	Test Accuracy (90 percentile)
Lasso	66.3	67.6	69.4	69.7
Random Forest	99.6	99.5	99.5	99.5
XGBoost	99.7	99.7	99.7	99.7
MLP	97.2	98.8	98.6	99.3
RNN	54	55.8	53.1	55.8
Total Features	5	10	15	18
Avg. Accuracy	83.36	84.28	84.06	84.8

- Based on our observation from the standard ML algorithms, 90th percentile has the best average accuracy.
- XGBoost was the best performing algorithm with 99.7% accuracy in all percentile.

Conclusion

Data science and Auto-ML are solving the Real Estate industry in several ways. Auto-ML can help solve several problems in the real estate industry related to optimizing operations, improving customer experience, and increasing revenue. Auto-ML algorithms can analyze historical data and trends to predict property valuations. By predicting property valuations, real estate companies can optimize their pricing strategies, ensure fair pricing for buyers and sellers, and increase revenue. The dataset has 46,622 records with 1 Categorical Features and 19 Numerical Features.

For regression, models were created with algorithms using Auto-ML techniques like Lasso, Random forest, XGBoost, Multilayer Perceptron and Recurrent Neural Network. With these models, performance measurement values were obtained for feature sets of 5, 10, 15 and 18. The Auto-ML algorithms were able to predict the rent of the property with an average accuracy between 83% – 85% and helped to identify factors that determine the rent. The major factors include areaRaw (m2), rent_per_area, areaSqm and city. The Random forest with 99.6 % accuracy in 25th percentile where tree showed a threshold of rent per area \leq €38.21 units and areaRaw (m2) \geq 105.5 m2 units which leads to lowest rent.