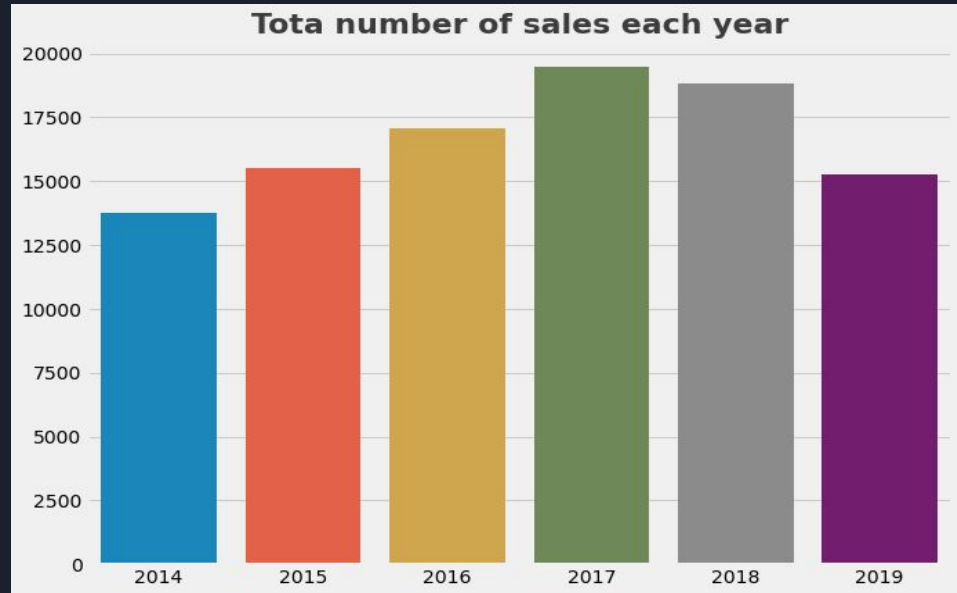# Real Advisor France

Contents:
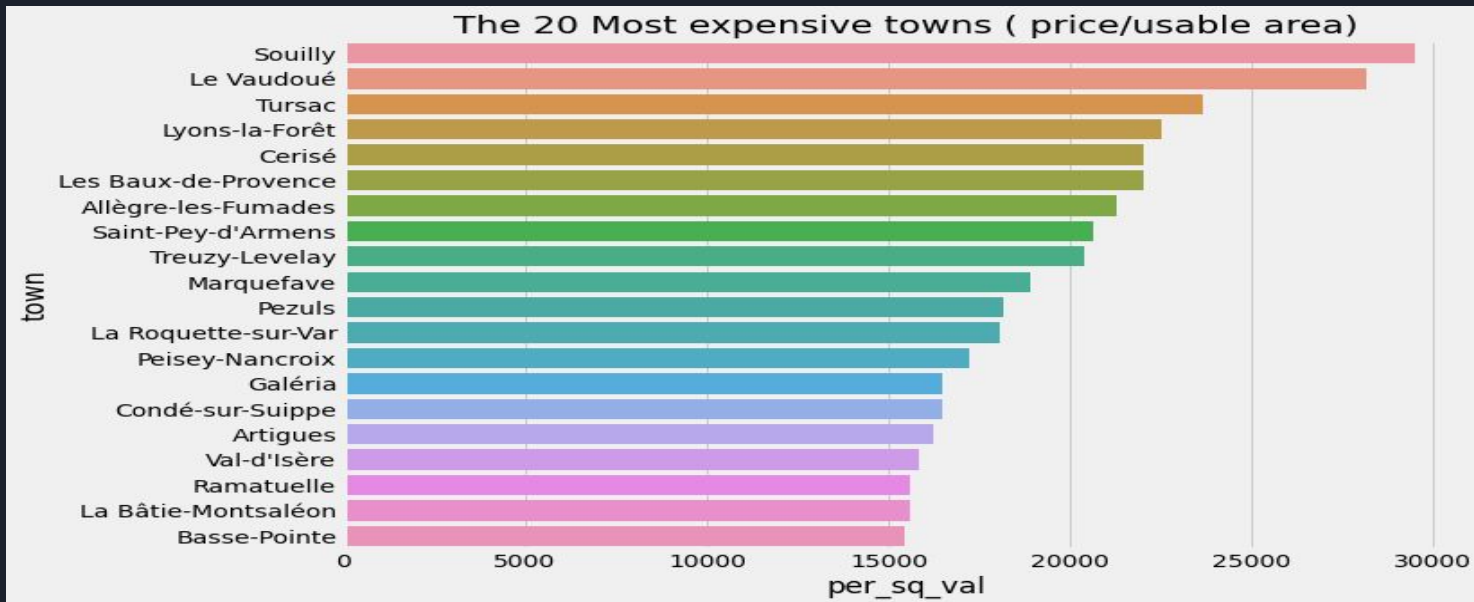- EDA
- ML Model
- ML Explainability

# Total Sales each year

- 2017 was the busiest year in the French real estate market.



**Tota number of sales each year**
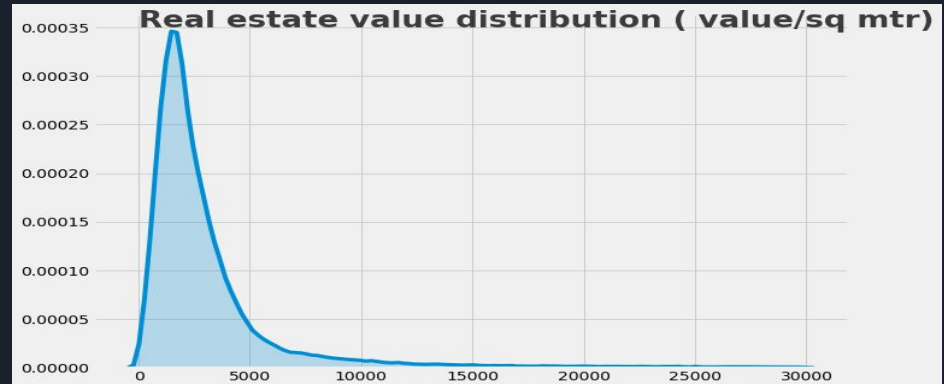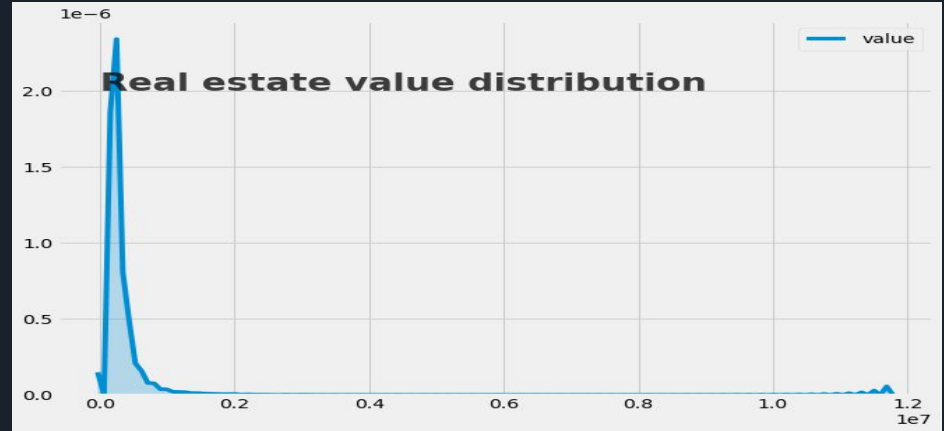
# Most expensive towns



The 20 Most expensive towns ( price/usable area)

# Most popular region ( postcode)



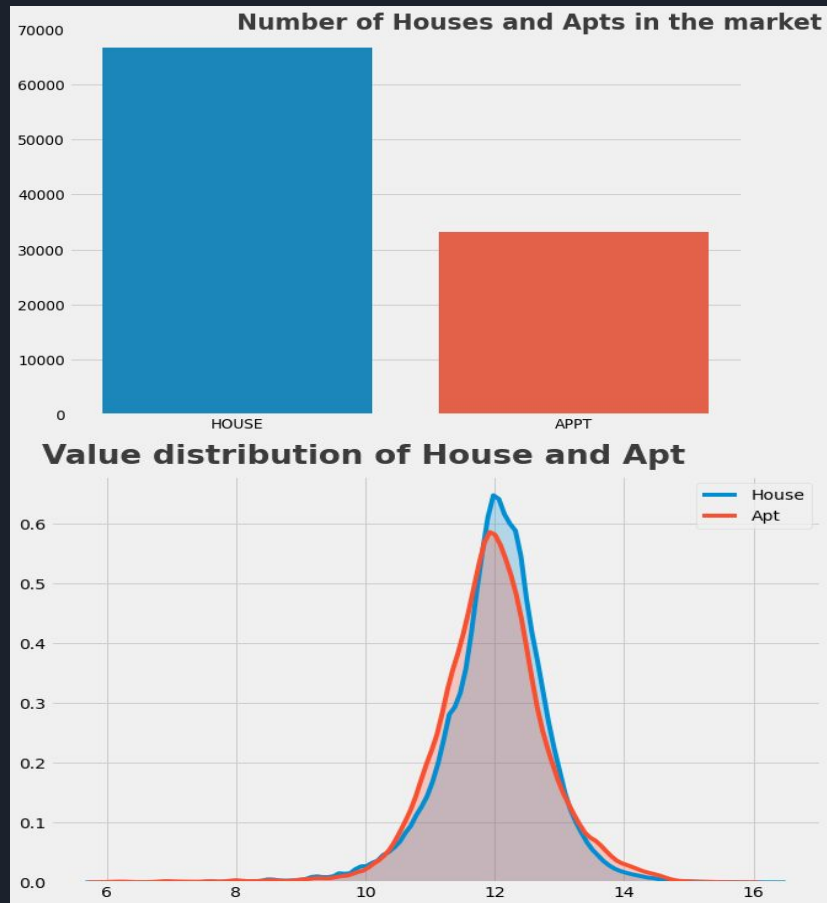Treemap of House Property counts across different postcodes

# Property value distribution

The plot shows that most of the real estate properties sold in France are towards the affordable budget price range.
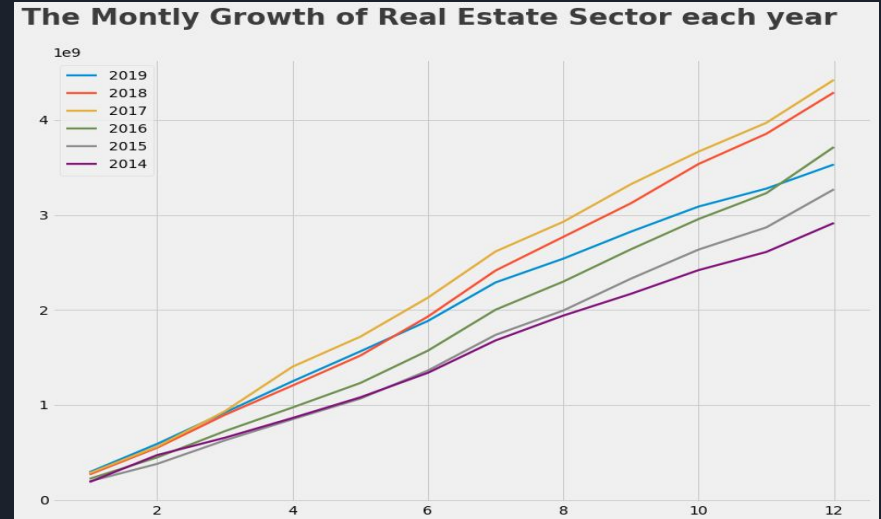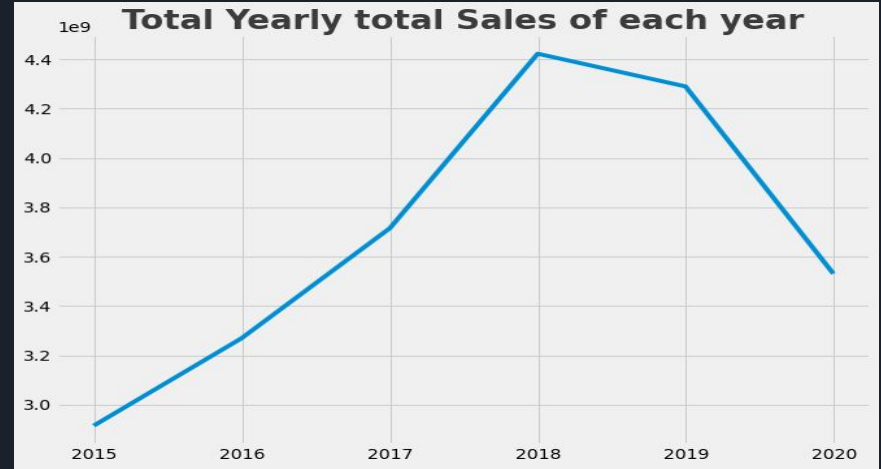
# Houses vs Apartments

- There are more houses in the market that apartments.
- The price distribution is almost the same.

# Yearly Sales

- We can observe that the real estate market in France grew at a steep rate until the end of 2017.
- From 2018 the total sales in the market have fallen.



Total Yearly total Sales of each year



The Montly Growth of Real Estate Sector each year

# Total monthly sales of each year

- In all years from 2014, it is observed that the month of july records a good peak in total monthly property sales .
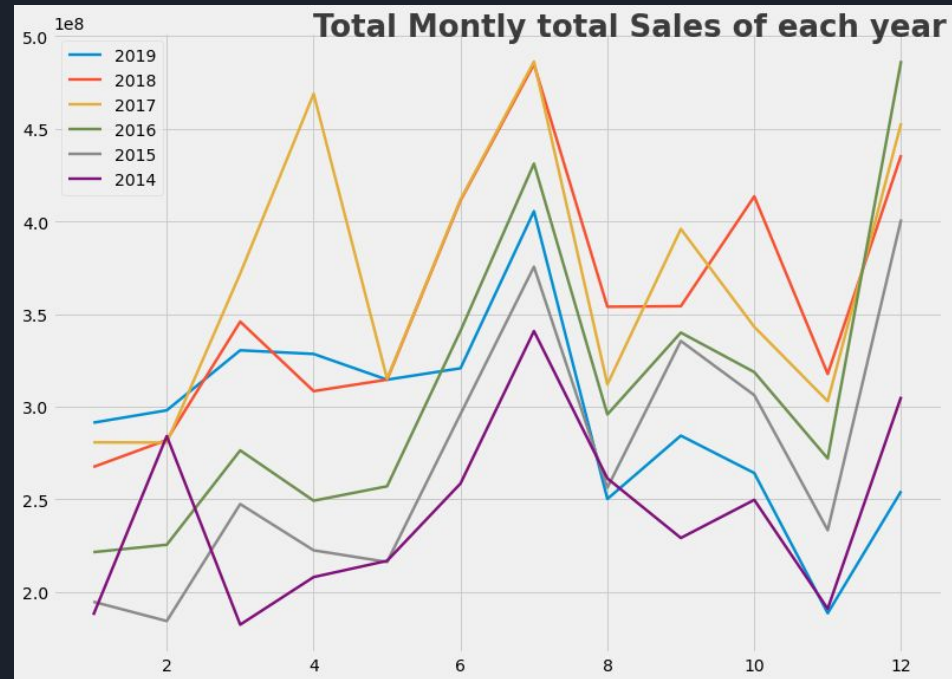
# Yearly sales by property type

- The total sales value of houses in France has been greater than apartments all over the years.
- This can be a strong indicator that people choose house more than buying an apartment.



The yearly Growth of Sector by property type



Total Monthly Sales by property type over the time period
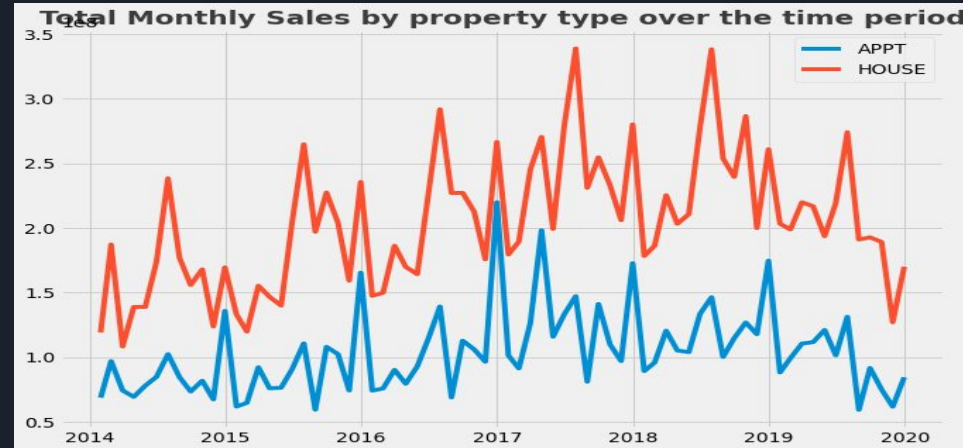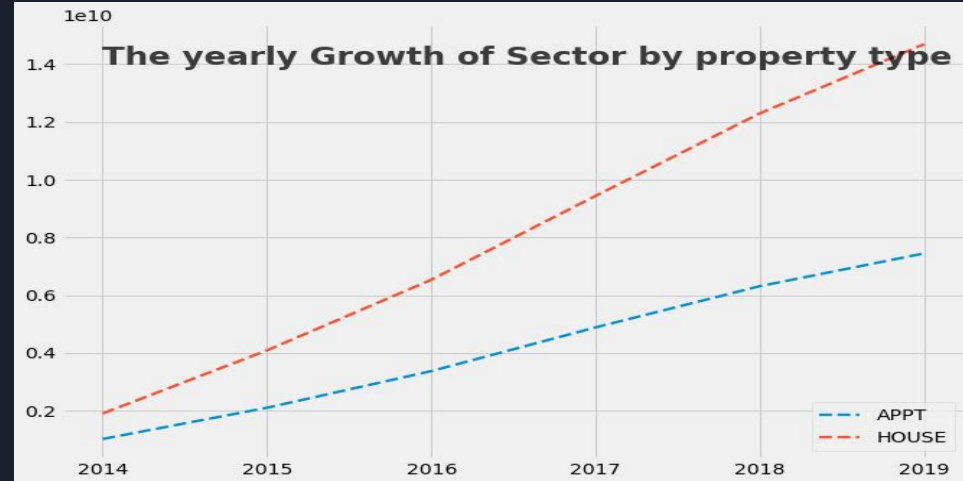
# Reason?

- The mean price per surface area of apartment is much higher than house.
- This can be the reason why people choose house over apartments leading to better house sales.



The Mean price/area of property type over years
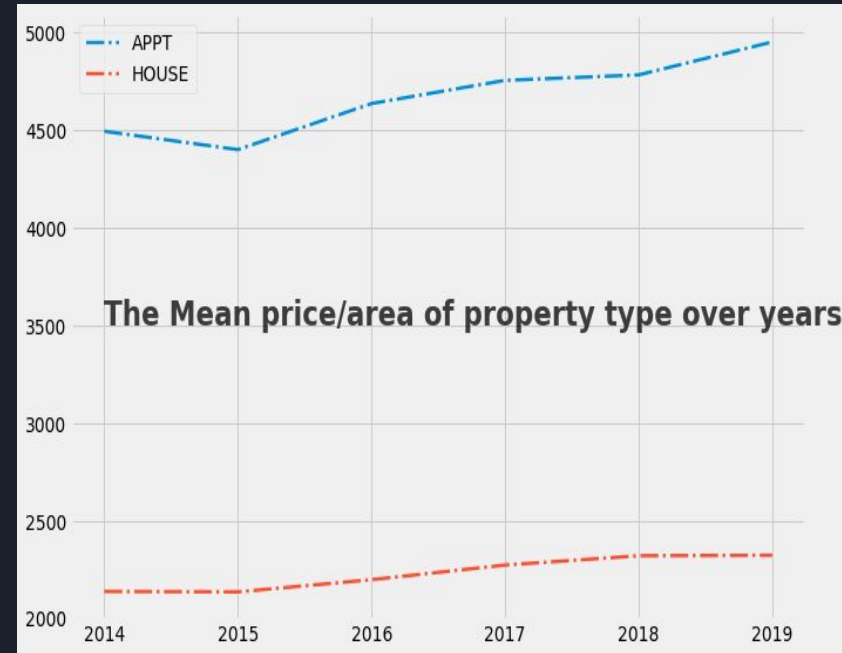
# Market segments

- Categorized into affordable, middle class and luxury.
- Identified using grouping by postcode and binning to corresponding category using property value.
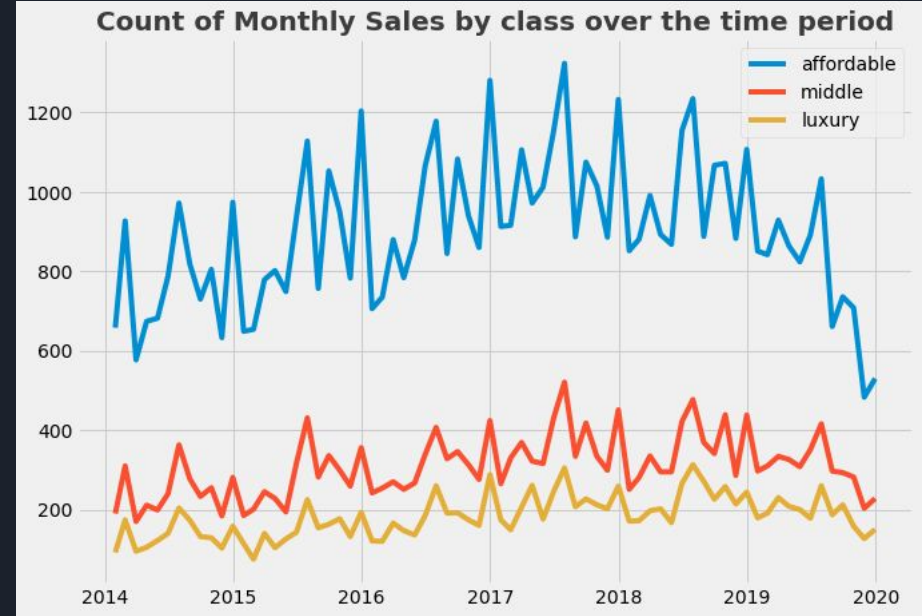
# Geographical analysis
Property by value per usable area over France

- The coastal areas seems to have high property value than other.

# Key insights

- 2017 was the best year for real estate market in France.
- Most of the properties has a value less than 3000 per/sq meter.
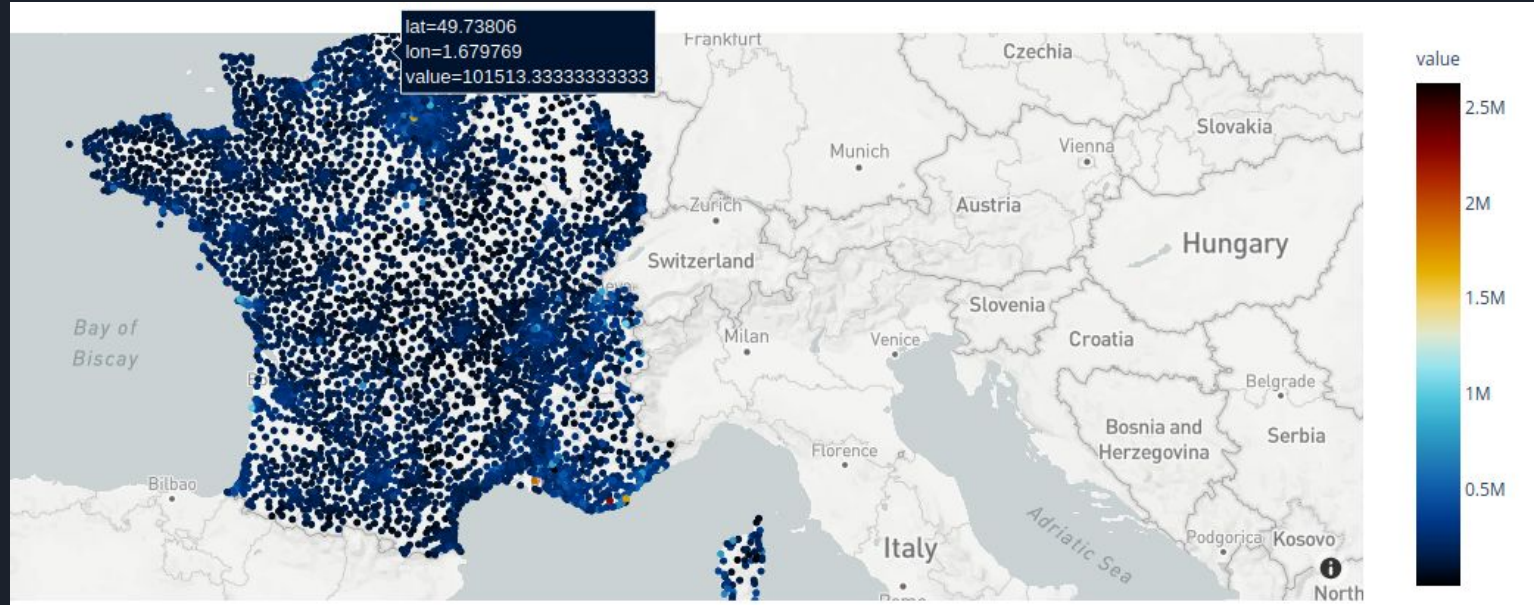- There are more houses for sale than apartments.
- The total sales has declined from the year 2017.
- The month of july is the hottest month for real estate sales in France.
- The price/usable area of apartments has always been much higher than the houses.
- The different market segment grows differently.
- Coastal areas have more property value compared to others.

# ML Pipeline

Data preparation pipeline

- Label encode
- Lag features
- Datetime features
- Count features for high cardinality features

Train - Validation Split

- Consider Out of time sample for validation
- Train sample - 2014-01-01 to 2018-12-31
- Validation sample - 2019-01-01 to 2019-12-31
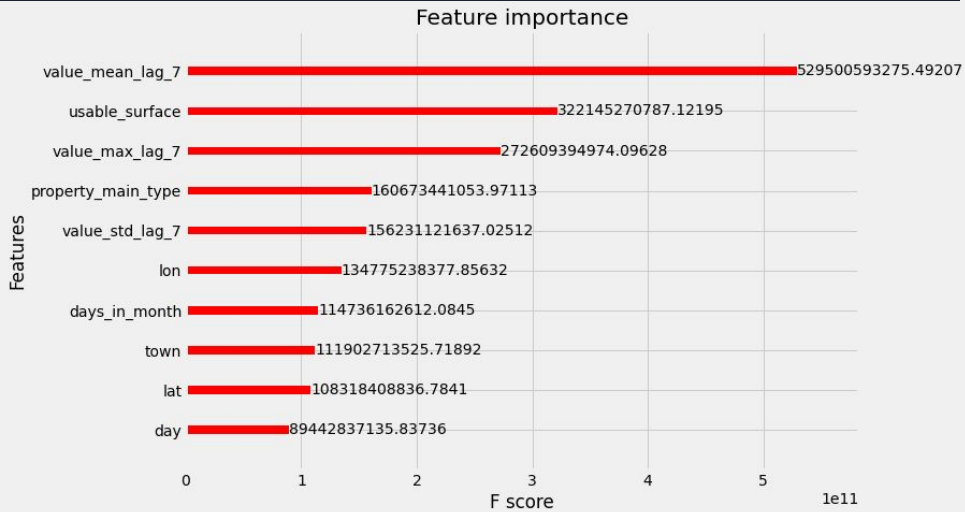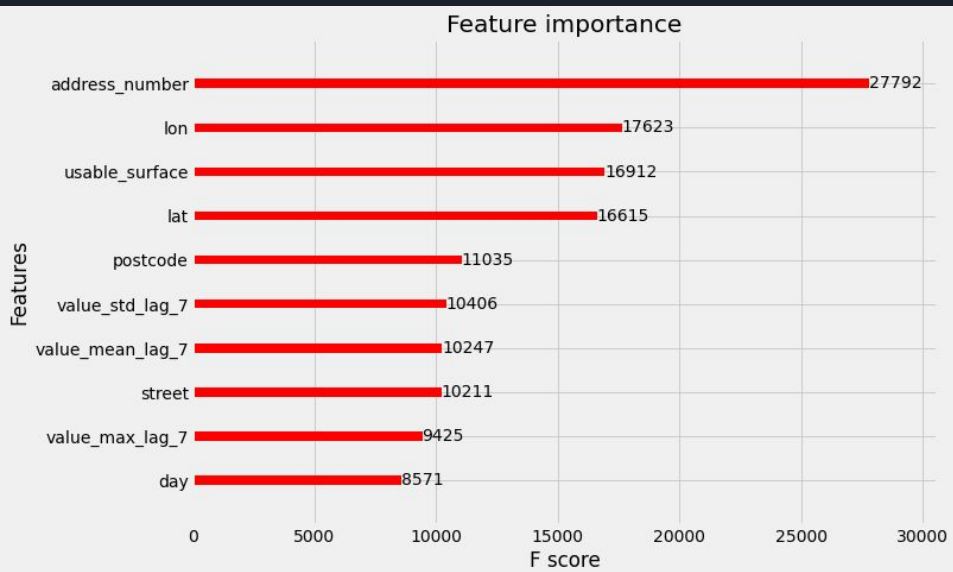
XGBoost Model

- XGBoost Regressor
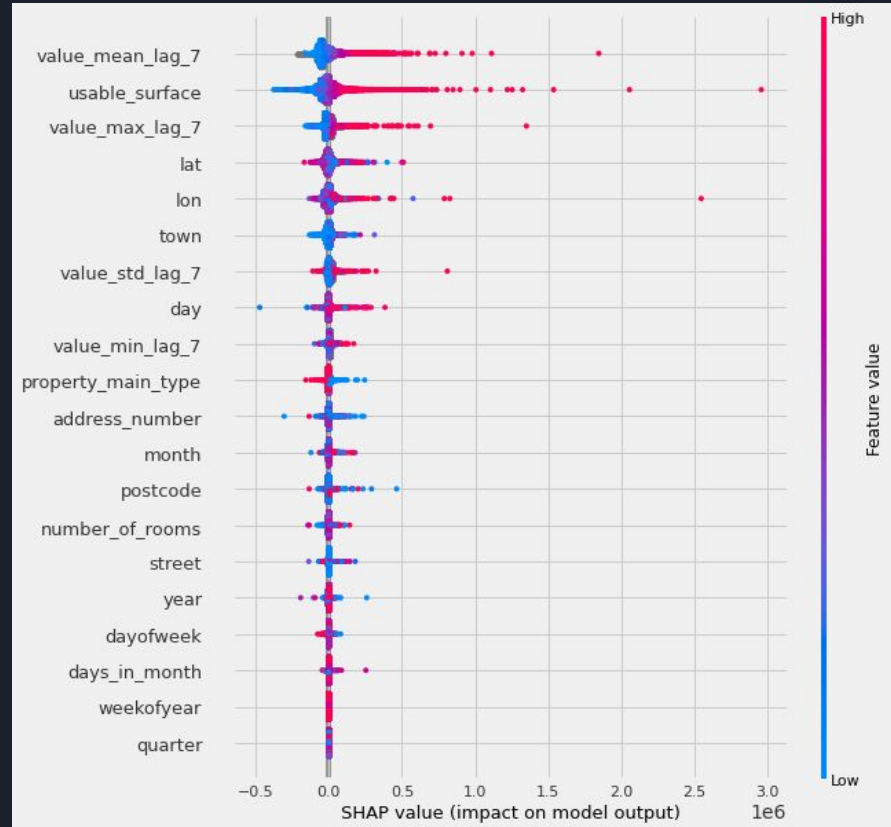- Evaluation metric - Root mean squared error

# ML Explainability

## Feature Importances

- Weight importance type
- Gain importance type



Feature importance

| Features | F score |
|----------|---------|
| address_number | 27792 |
| lon | 17623 |
| usable_surface | 16912 |
| lat | 16615 |
| postcode | 11035 |
| value_std_lag_7 | 10406 |
| value_mean_lag_7 | 10247 |
| street | 10211 |
| value_max_lag_7 | 9425 |
| day | 8571 |



Feature importance

| Features | F score |
|----------|---------|
| value_mean_lag_7 | 529500593275.49207 |
| usable_surface | 322145270787.12195 |
| value_max_lag_7 | 272609394974.09628 |
| property_main_type | 160673441053.97113 |
| value_std_lag_7 | 156231121637.02512 |
| lon | 134775238377.85632 |
| days_in_month | 114736162612.0845 |
| town | 111902713525.71892 |
| lat | 108318408836.7841 |
| day | 89442837135.83736 |

# SHAP Values

- The plot uses SHAP values to show the distribution of the impacts each feature has on the model output.
- For example, increase in value of usable area increases the property value.

Thank you.