

The background of the slide is a light blue-grey color. In the center, there is a large, semi-transparent globe. The globe is composed of many small, light-grey cubes that are arranged to form its shape. Some cubes are slightly offset from the main surface, giving it a 3D, pixelated appearance. The text is centered over this globe.

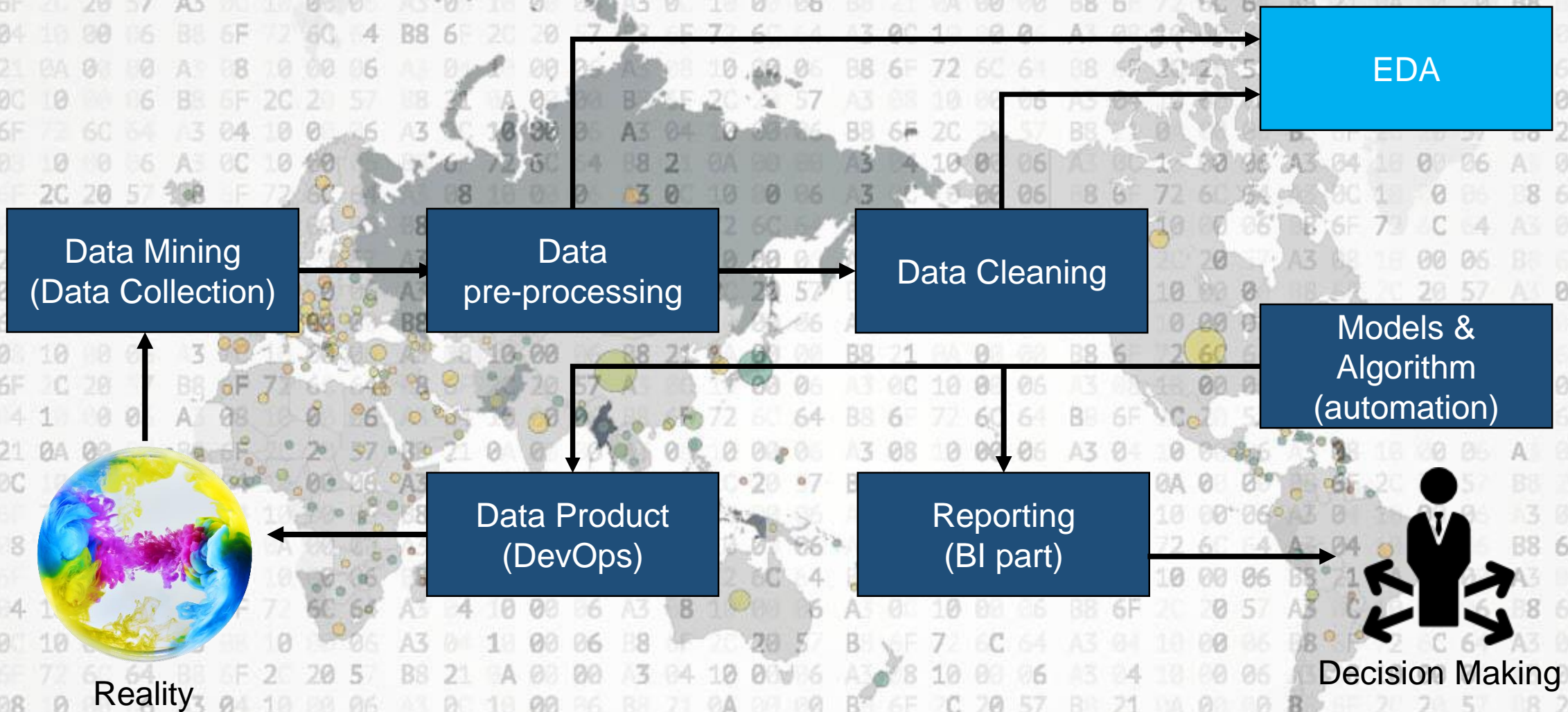
Artificial Intelligence and Data Science

for researches of REAL ESTATE

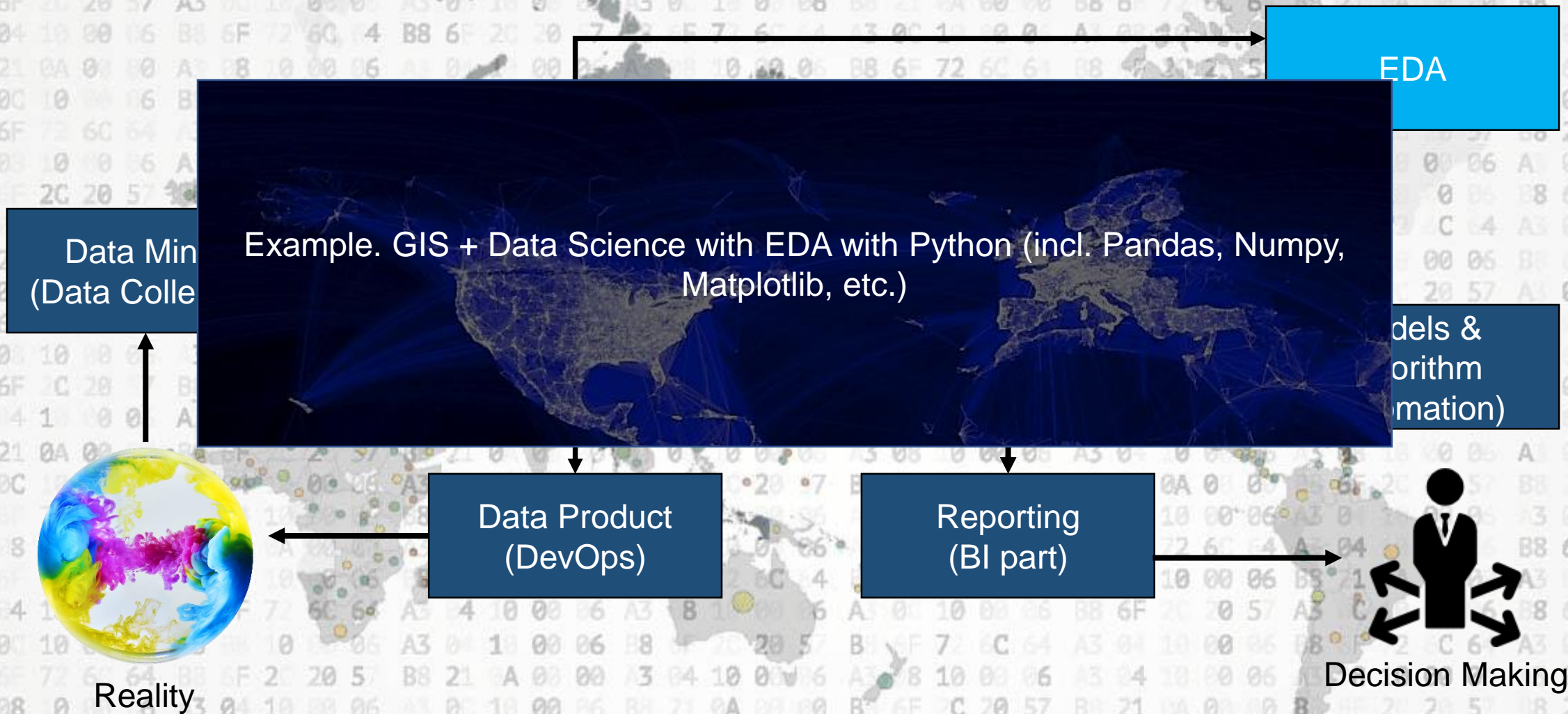
EDA vs. Machine Learning



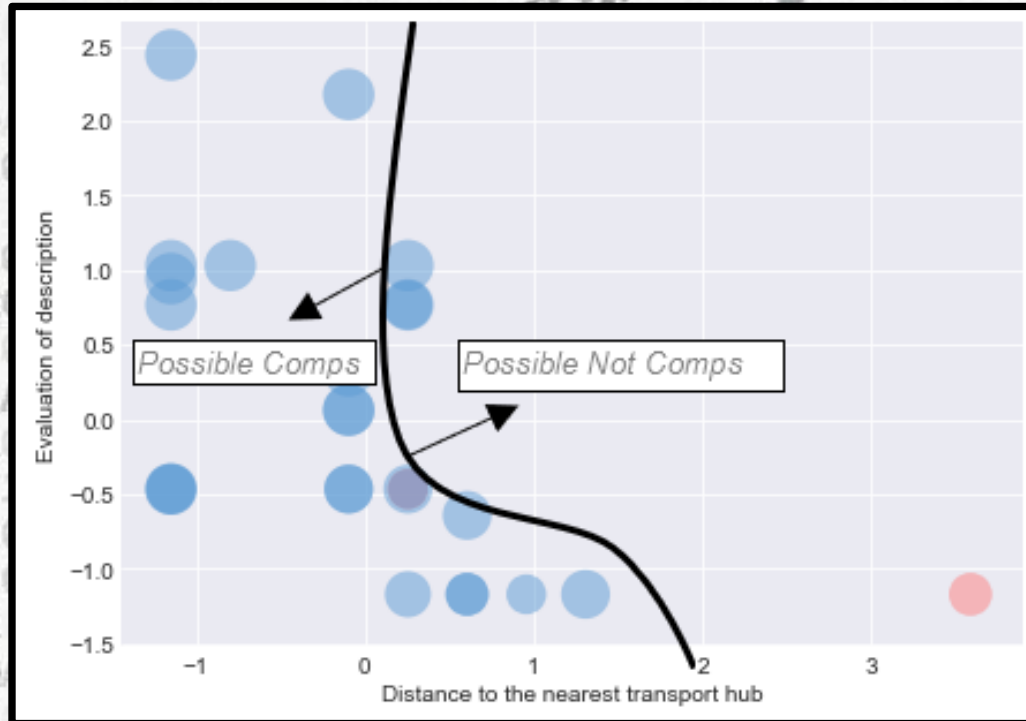
EDA vs. Machine Learning



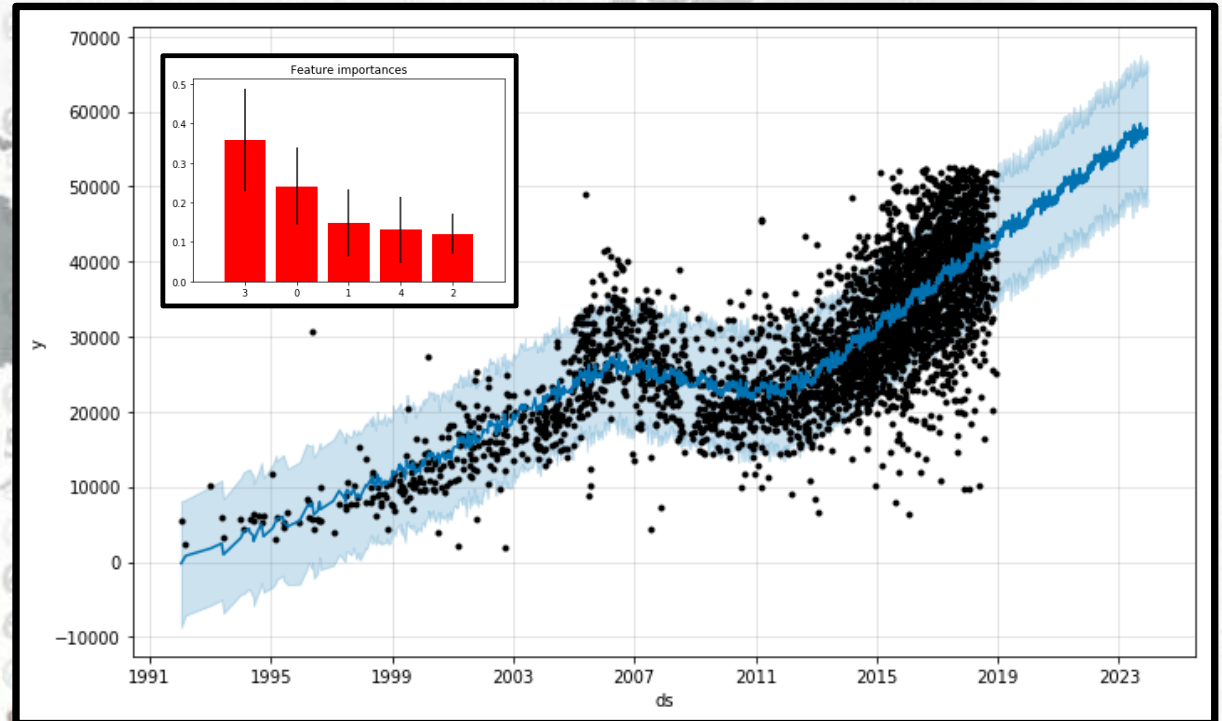
EDA vs. Machine Learning



EDA vs. Machine Learning

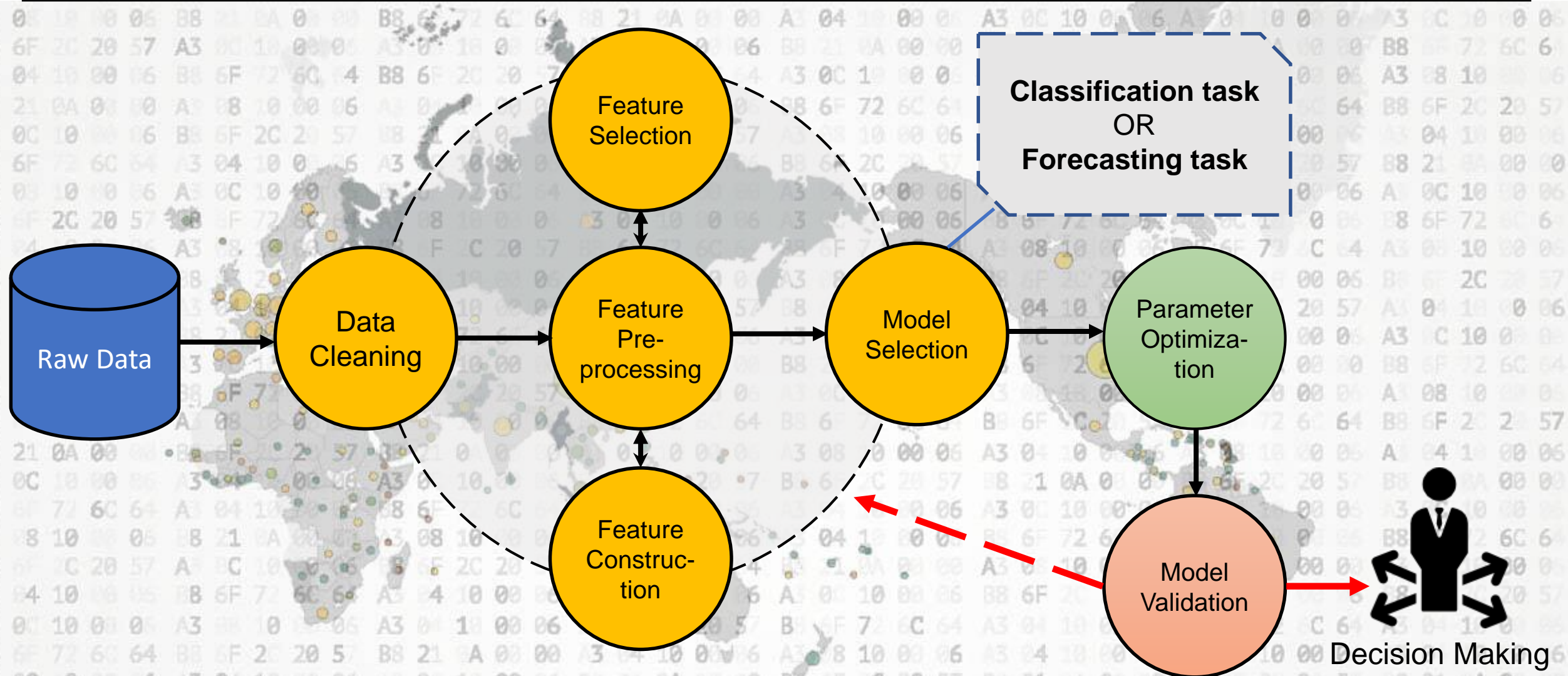


Classification Task



Regression (Forecasting) Task

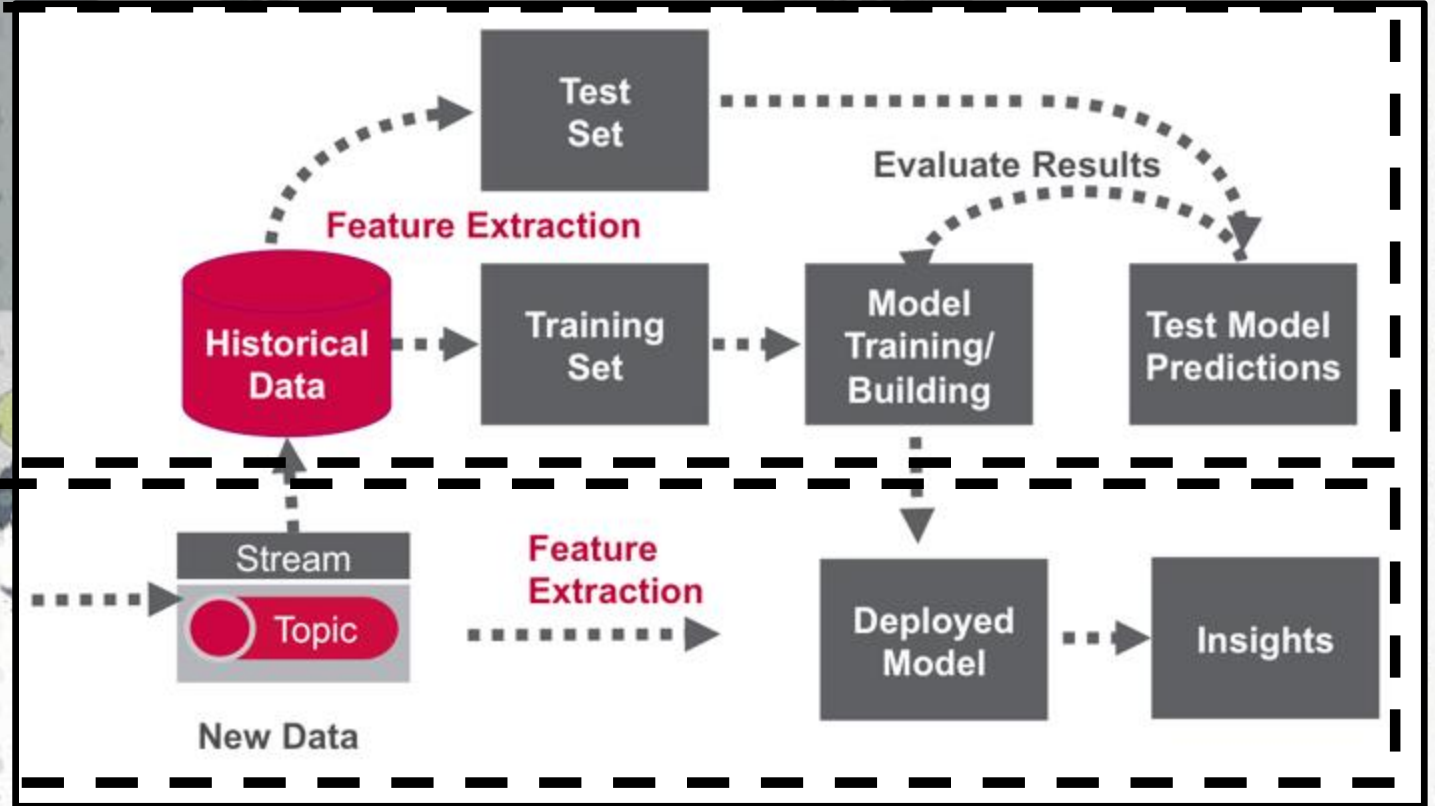
EDA vs. Machine Learning



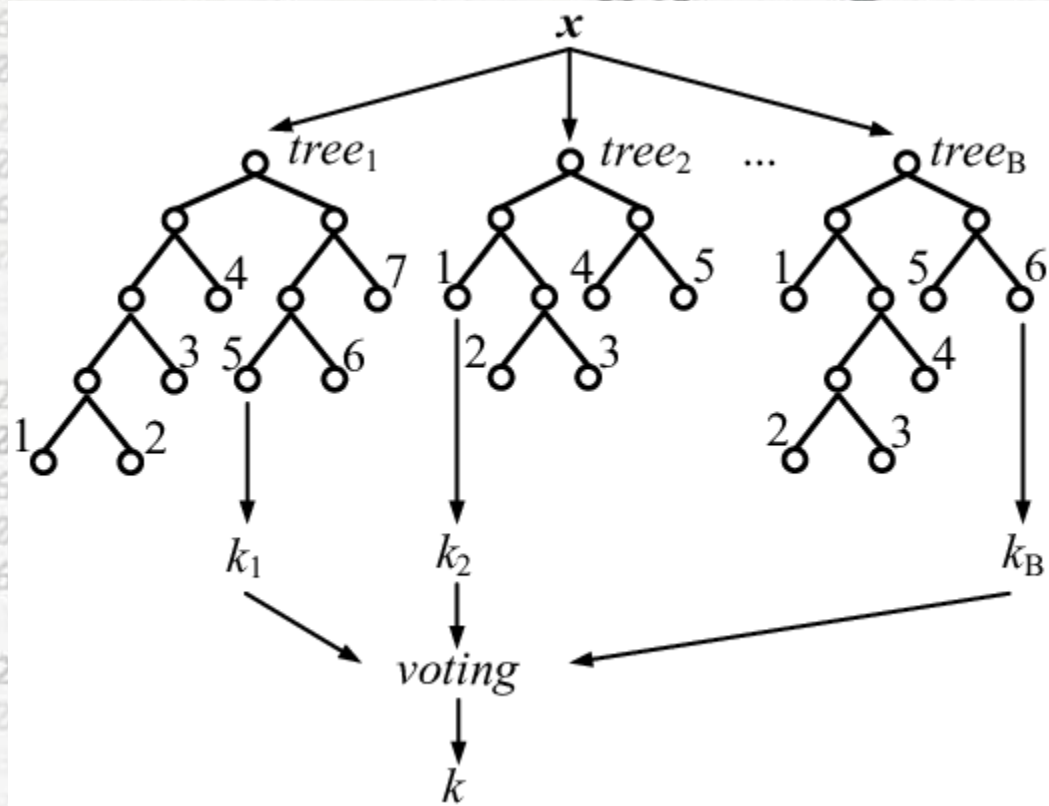
EDA vs. Machine Learning

MODEL BUILDING

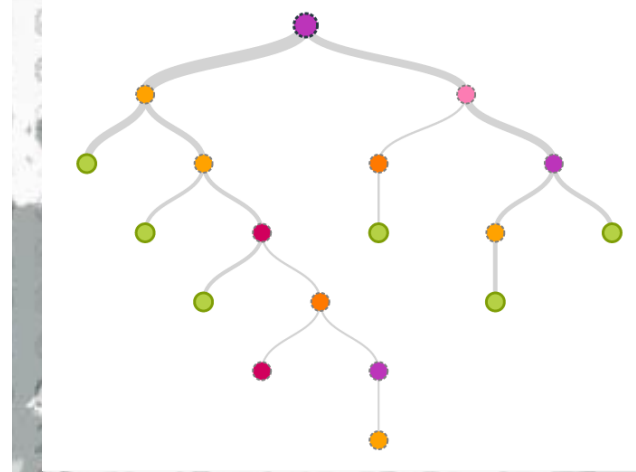
**DATA MINING +
WAREHOUSING**



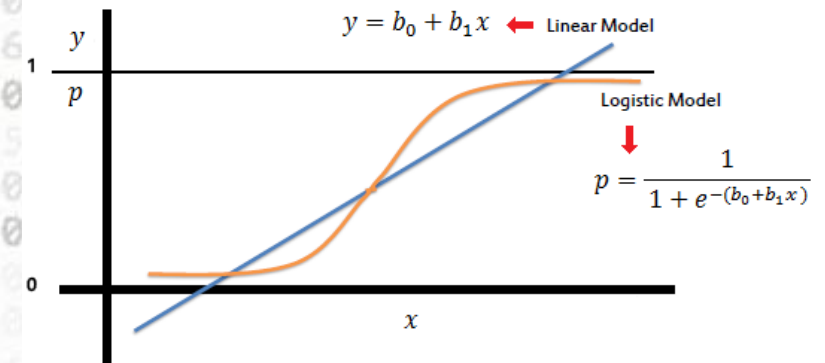
MODELS for EDA vs. Machine Learning



Random Forest



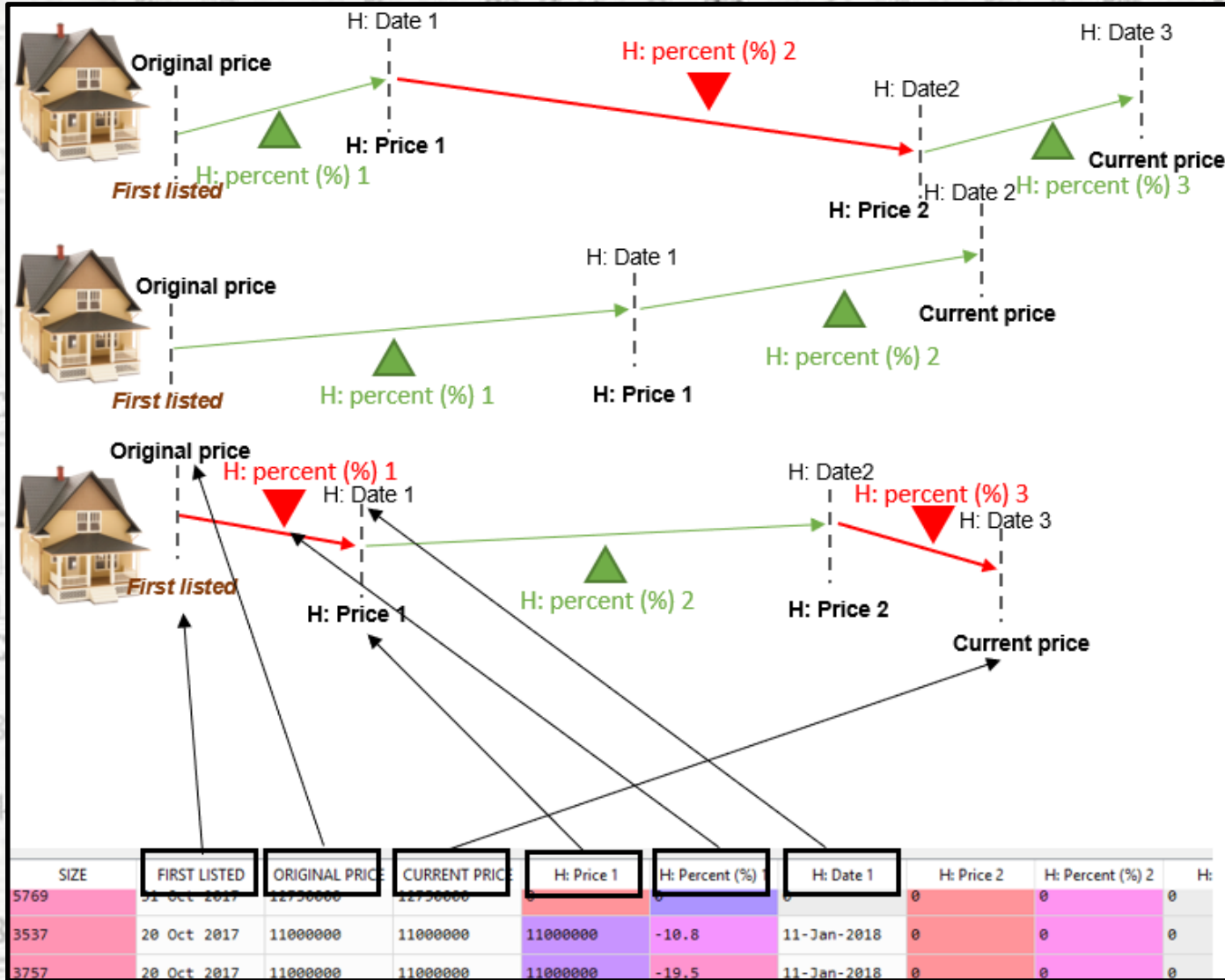
Decision Tree



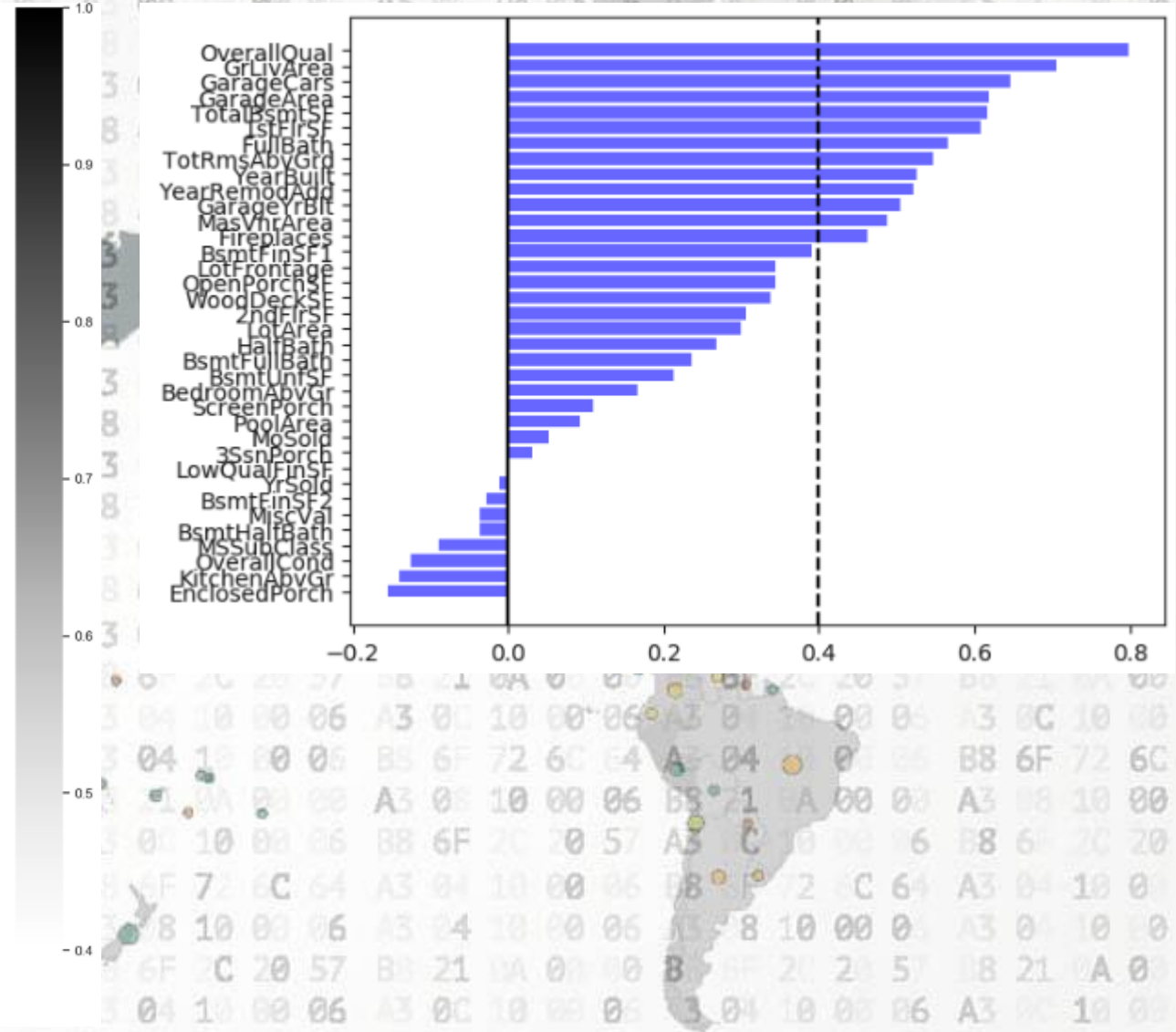
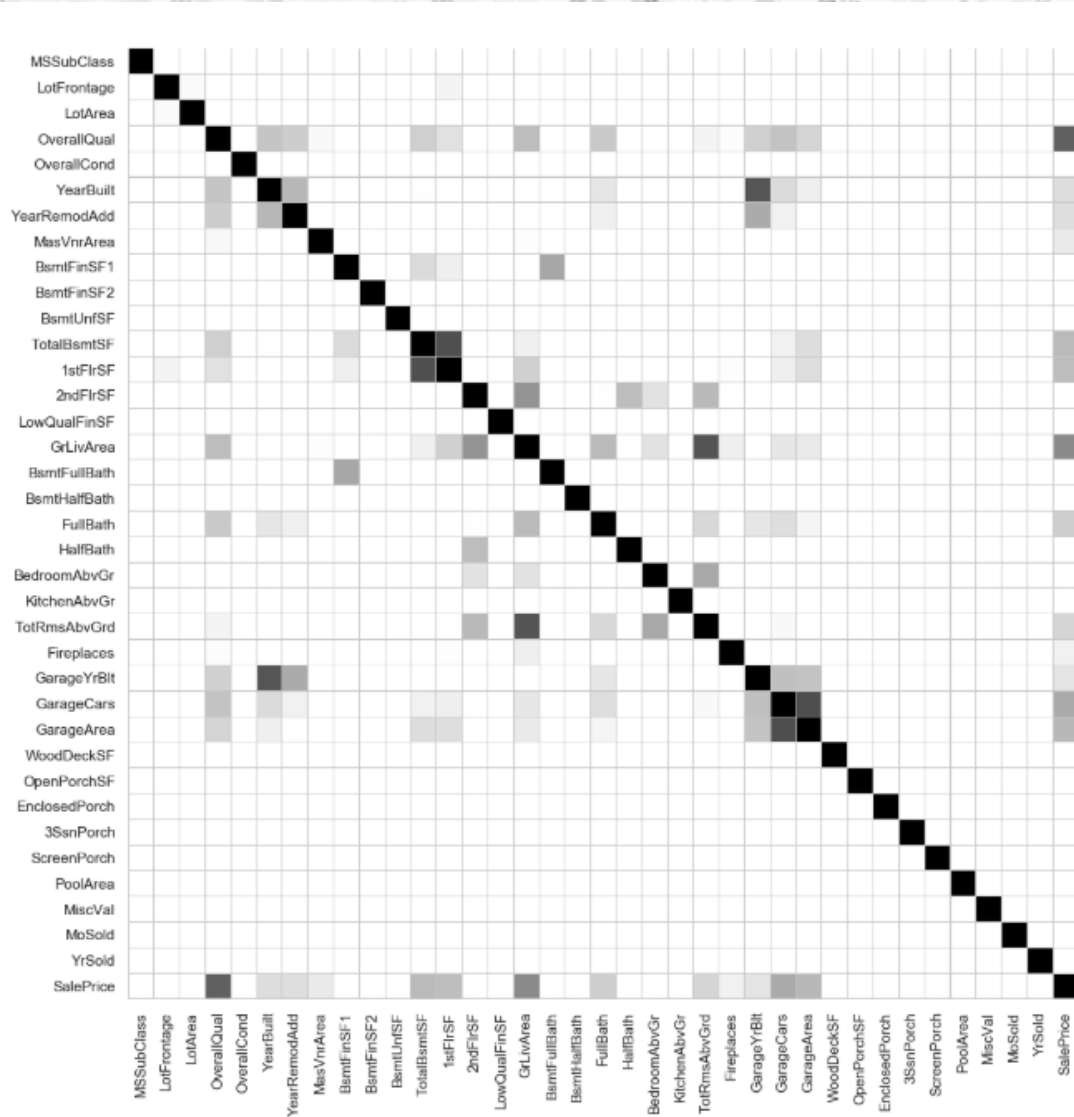
Logistic Regression

Or many more!

DATA MINING for EDA vs. Machine Learning

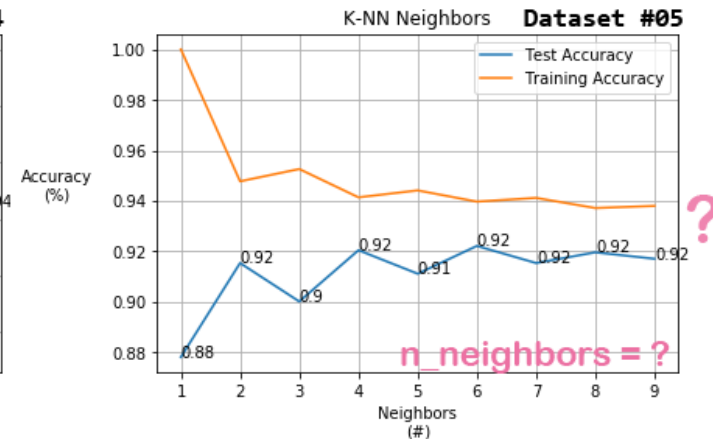
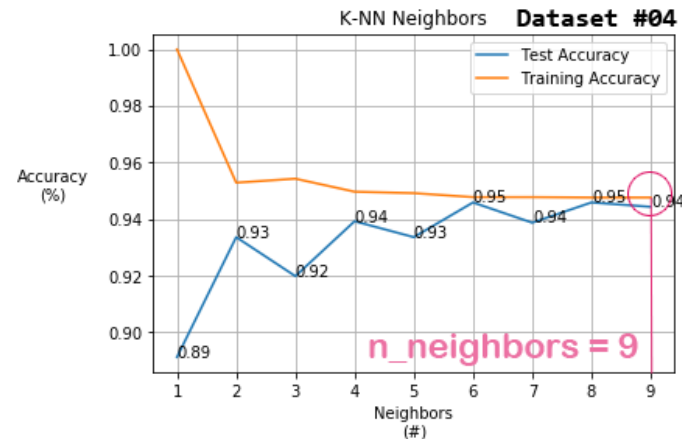
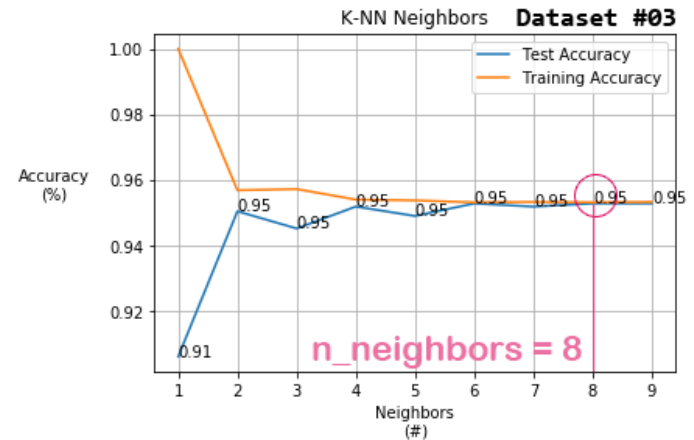
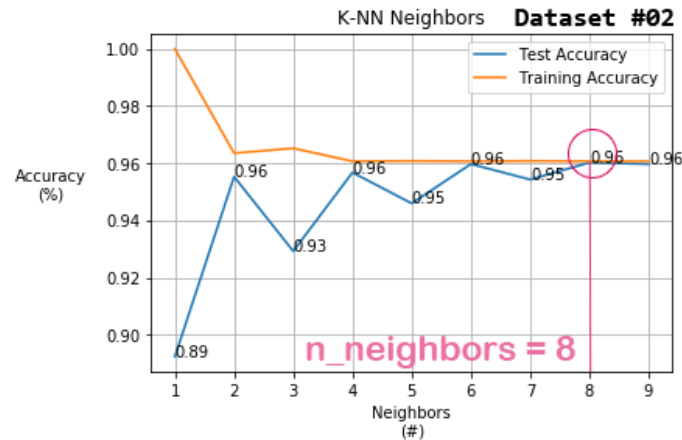
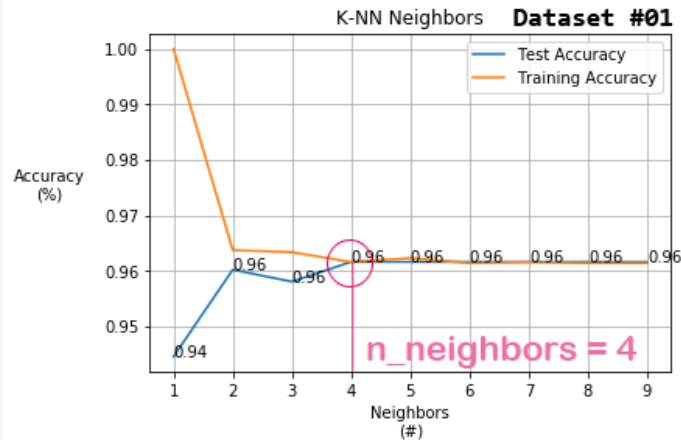


FEATURE ENGINEERING for EDA vs. Machine Learning



EDA vs. Machine Learning

!!! IMPORTANT !!!



Prepared by Vytautas Bielinskas

Hyperparameter Tuning is one of Solutions.

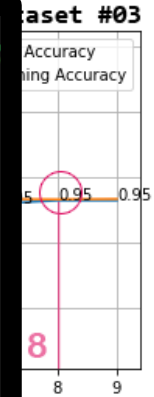
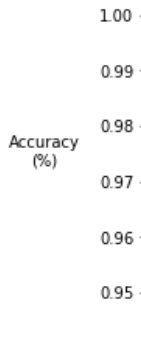
EDA vs. Machine Learning

!!! IMPORTANT !!!

```

29 # ..... HYPERPARAMETER TUNING .....
30 # Setup the parameters and distributions to sample from: param_dist
31 param_dist = {"max_depth": [3, None],
32               "min_samples_leaf": randint(1, 9),
33               "criterion": ["gini", "entropy"]}
34
35 # Instantiate a Decision Tree classifier: tree
36 tree = DecisionTreeClassifier()
37
38 # Instantiate the RandomizedSearchCV object: tree_cv
39 tree_cv = RandomizedSearchCV(tree, param_dist, cv=5)
40
41 # Fit it to the data
42 tree_cv.fit(X_train, y_train)
43
44 # Print the tuned parameters and score
45 print("Tuned Decision Tree Parameters: {}".format(tree_cv.best_params_))
46 print("Best score is {}".format(tree_cv.best_score_))
47 # .....

```



Hyperparameter Tuning is one of Solutions.

EDA vs. Machine Learning

| OLD DATA | | | | | | | | | | | | | | | | | | |
|-----------|----------|-------|-------------------------------------------------------|---|---|---|-----|--------------|---------|---------|----|-----------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|-----|-----|-----|-----|-----|
| 30-Sep-17 | 40330393 | Iford | Gloicester Road, London E12 | 4 | 2 | 1 | | 11-May-16 | 435,000 | 450,000 | | | https://www.zoopla.co.uk/new-homes/details/40330393 | 517 | 4 | 0 | 0 | 0 |
| 30-Sep-17 | 40340530 | Iford | Grantham Road, London E12 | 4 | | | | 6-May-16 | 415,000 | | | | https://www.zoopla.co.uk/new-homes/details/40340530 | 522 | 4 | 0 | 0 | 0 |
| 30-Sep-17 | 38823361 | Iford | Hickling Road, Iford, Essex IG1 | 3 | 2 | 2 | | 2-Dec-15 | 254,000 | | | | https://www.zoopla.co.uk/new-homes/details/38823361 | 678 | 3 | 0 | 0 | |
| 30-Sep-17 | 38818553 | Iford | Hickling Road, Iford, Essex IG1 | 3 | 2 | 2 | 519 | 5-Dec-15 | 380,000 | | | | https://www.zoopla.co.uk/new-homes/details/38818553 | 679 | 3 | 519 | 519 | |
| 30-Sep-17 | 31459444 | Iford | Mortlake Road, Iford IG1 | 2 | 1 | 1 | | 10-Jul-15 | 235,000 | 245,000 | | | https://www.zoopla.co.uk/new-homes/details/31459444 | 823 | 2 | 0 | 0 | 0 |
| 27-Jan-18 | 48021459 | Iford | Iford Hill, Iford IG1 | 2 | 2 | 1 | 0 | 6th Jan 2018 | 370000 | 370000 | 0% | | https://www.zoopla.co.uk/new-homes/details/48021459 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 47944631 | Iford | Raphel House, 250 High Road, Iford IG1 | 2 | 2 | 1 | 0 | 6th Apr 2018 | 375000 | 375000 | 0% | https://www.zoopla.co.uk/new-homes/details/47944631 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 47772244 | Iford | Horizon, Iford IG1 | 2 | 2 | 1 | 0 | 5th May 2018 | 445000 | 445000 | 0% | https://www.zoopla.co.uk/new-homes/details/47772244 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 47847736 | Iford | Prior Road, Iford IG1 | 2 | 2 | 1 | 0 | 2nd Jan 2018 | 339000 | 339000 | 0% | https://www.zoopla.co.uk/new-homes/details/47847736 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 47833114 | Iford | Axon Place, Iford, Essex IG1 | 2 | 2 | 1 | 0 | 6th Jan 2018 | 280000 | 280000 | 0% | https://www.zoopla.co.uk/new-homes/details/47833114 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 41240339 | Iford | Cranbrook Road, Iford IG1 | 2 | 2 | 1 | 872 | 3rd Aug 2016 | 444000 | 444000 | 0% | 509.17 | https://www.zoopla.co.uk/new-homes/details/41240339 | 2 | 872 | 1 | 0 | |
| 27-Jan-18 | 47115173 | Iford | Valentine House, Iford IG1 | 2 | 2 | 1 | 0 | 6th Apr 2018 | 375000 | 375000 | 0% | https://www.zoopla.co.uk/new-homes/details/47115173 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 44814802 | Iford | Centrecrofts Axon Place, Iford IG1 | 2 | 2 | 1 | | 4th Aug 2017 | 355000 | 355000 | 0% | https://www.zoopla.co.uk/new-homes/details/44814802 | 2 | 0 | 0 | 1 | 0 | |
| 27-Jan-18 | 48033390 | Iford | Horizon, Iford IG1 | 2 | 2 | 1 | 0 | 8th Jan 2018 | 385000 | 385000 | 0% | https://www.zoopla.co.uk/new-homes/details/48033390 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 47542012 | Iford | Hainsalt Street, Iford IG1 | 2 | 2 | 1 | 0 | 8th May 2018 | 325000 | 325000 | 0% | https://www.zoopla.co.uk/new-homes/details/47542012 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 47841535 | Iford | Horizon, Iford, Essex IG1 | 2 | 2 | 1 | 0 | 8th May 2018 | 400000 | 400000 | 0% | https://www.zoopla.co.uk/new-homes/details/47841535 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 46083331 | Iford | Hainsalt Street, Iford IG1 | 2 | 2 | 1 | 0 | 8th Dec 2017 | 310000 | 310000 | 0% | https://www.zoopla.co.uk/new-homes/details/46083331 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 47455637 | Iford | Horizon Apartments, 51-63 Iford Hill, Iford IG1 | 2 | 2 | 1 | 0 | 0th Oct 2017 | 410000 | 410000 | 0% | https://www.zoopla.co.uk/new-homes/details/47455637 | 2 | 0 | new | 1 | 0 | |
| 27-Jan-18 | 46718863 | Iford | Carlton House, High Street, Iford IG1 | 2 | 2 | 1 | 0 | 2nd Mar 2018 | 300000 | 300000 | 0% | https://www.zoopla.co.uk/new-homes/details/46718863 | 2 | 0 | new | 0.9 | 0.1 | |
| 27-Jan-18 | 47970350 | Iford | Hainsalt Bridge Parade, Hainsalt Street | 2 | 2 | 1 | 0 | 2th Jan 2018 | 335000 | 335000 | 0% | https://www.zoopla.co.uk/new-homes/details/47970350 | 2 | 0 | new | 0.9 | 0.1 | |
| 27-Jan-18 | 46409151 | Iford | 2 Bedroom 2 Bathroom, City View Apartments, Iford IG1 | 2 | 2 | 1 | 0 | 5th Mar 2018 | 320000 | 320000 | 0% | https://www.zoopla.co.uk/new-homes/details/46409151 | 2 | 0 | new | 0.8 | 0.2 | |
| 27-Jan-18 | 47815672 | Iford | Buckingham Road, Iford, Essex IG1 | 2 | 2 | 1 | 0 | 1st May 2018 | 275000 | 275000 | 0% | https://www.zoopla.co.uk/new-homes/details/47815672 | 2 | 0 | new | 0.8 | 0.2 | |
| 27-Jan-18 | 47735325 | Iford | Iford Hill, Iford IG1 | 2 | 2 | 0 | 0 | 5th May 2018 | 134750 | 134750 | 0% | https://www.zoopla.co.uk/new-homes/details/47735325 | 2 | 0 | new | 0.8 | 0.2 | |
| 27-Jan-18 | 45705588 | Iford | Francis Avenue, Iford IG1 | 2 | 2 | 2 | 0 | 3th Nov 2017 | 330000 | 330000 | 0% | https://www.zoopla.co.uk/new-homes/details/45705588 | 2 | 0 | new | 0.7 | 0.3 | |
| 27-Jan-18 | 46474561 | Iford | Carlton House, 501 High Road, Iford, Essex IG1 | 2 | 2 | 1 | 646 | 4th Feb 2018 | 239000 | 239000 | 0% | 462.85 | https://www.zoopla.co.uk/new-homes/details/46474561 | 2 | 646 | new | 0.7 | 0.3 |
| 27-Jan-18 | 47370839 | Iford | Hainsalt Bridge Parade, Hainsalt Street | 2 | 2 | 1 | 0 | 3rd Apr 2018 | 239000 | 239000 | 0% | https://www.zoopla.co.uk/new-homes/details/47370839 | 2 | 0 | new | 0.6 | 0.4 | |
| 27-Jan-18 | 47871038 | Iford | Centrecrofts, City View, Axon Place, Iford IG1 | 2 | 2 | 1 | 0 | 1st Jan 2018 | 239000 | 239000 | 0% | https://www.zoopla.co.uk/new-homes/details/47871038 | 2 | 0 | new | 0.6 | 0.4 | |
| 27-Jan-18 | 45215623 | Iford | Iford, Essex IG1 | 2 | 2 | 2 | | 9th Oct 2017 | 370000 | 370000 | 0% | https://www.zoopla.co.uk/new-homes/details/45215623 | 2 | 0 | 0 | 0.6 | 0.4 | |
| 27-Jan-18 | 47891395 | Iford | Chamberlain Close, Iford IG1 | 2 | 2 | 1 | 0 | 2nd Mar 2018 | 314395 | 314395 | 0% | https://www.zoopla.co.uk/new-homes/details/47891395 | 2 | 0 | new | 0.5 | 0.5 | |
| 27-Jan-18 | 47233173 | Iford | Carlton House, High Street, Iford IG1 | 2 | 2 | 1 | 0 | 6th Apr 2018 | 315000 | 315000 | 0% | https://www.zoopla.co.uk/new-homes/details/47233173 | 2 | 0 | new | 0.5 | 0.5 | |
| 27-Jan-18 | 46431156 | Iford | Ley Street, Iford IG1 | 2 | 2 | 1 | 0 | 6th Feb 2018 | 280000 | 280000 | 0% | https://www.zoopla.co.uk/new-homes/details/46431156 | 2 | 0 | new | 0.4 | 0.6 | |
| 27-Jan-18 | 48079502 | Iford | City View, Axon Place, Iford, Essex IG1 | 2 | 2 | 1 | 0 | 1st Jan 2018 | 299395 | 299395 | 0% | https://www.zoopla.co.uk/new-homes/details/48079502 | 2 | 0 | new | 0.3 | 0.7 | |

If there is green colored cell, that means that this property is checked before and comps value is taking from history records

New properties are sorted by it's probability to being a comps

EDA vs. Machine Learning

MC

Machine Learning for classify new properties from listings for new investments
based on HISTORICAL BUSINESS DECISIONS

(Decision Tree / Random Forest / XgBoost)

Test Model
Predictions

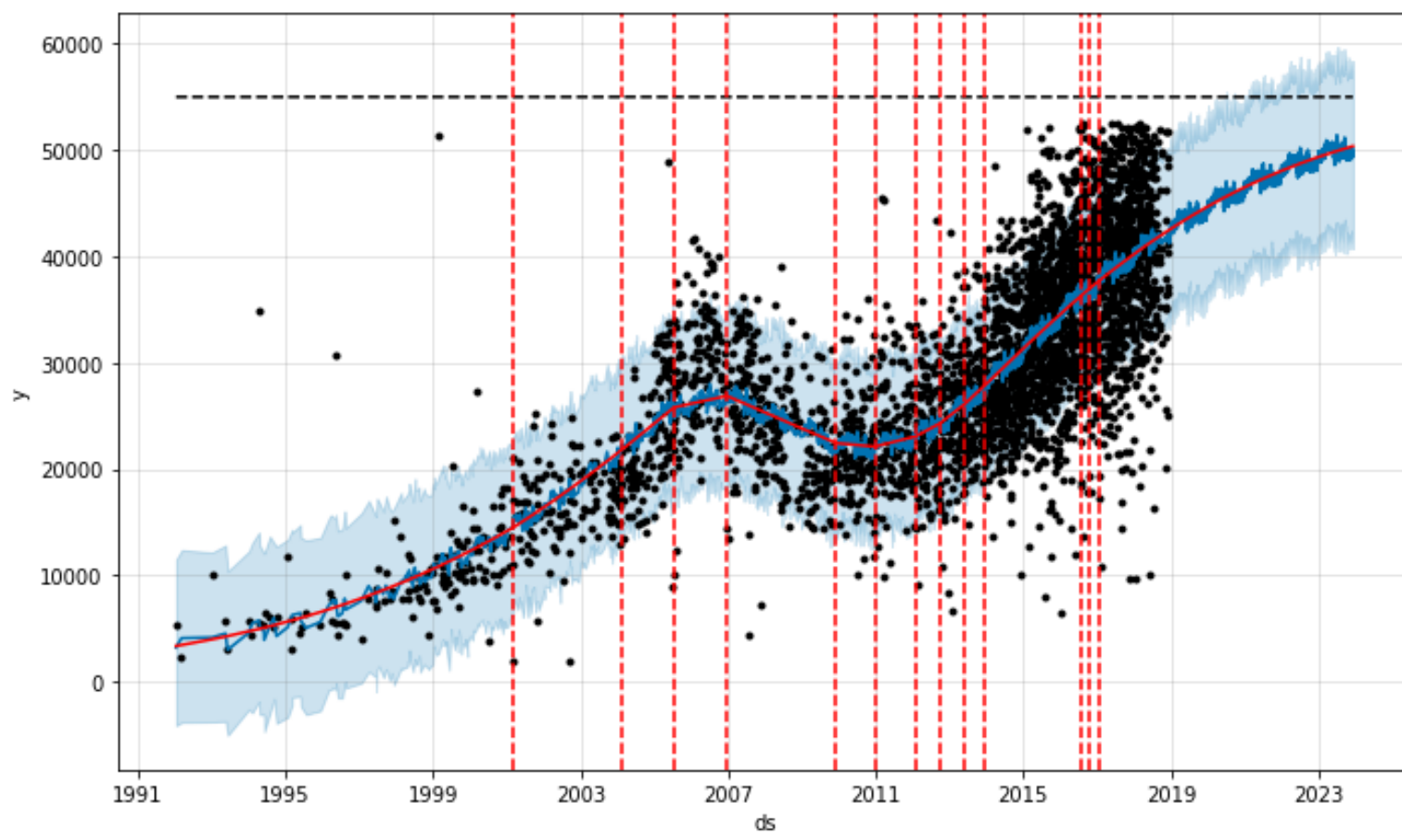
**DATA MINING +
WAREHUOSING**



Statistics + Machine Learning in Data Science

facebook

provides useful tools for perform Real Estate price level predictions.



Statistics + Machine Learning in Data Science

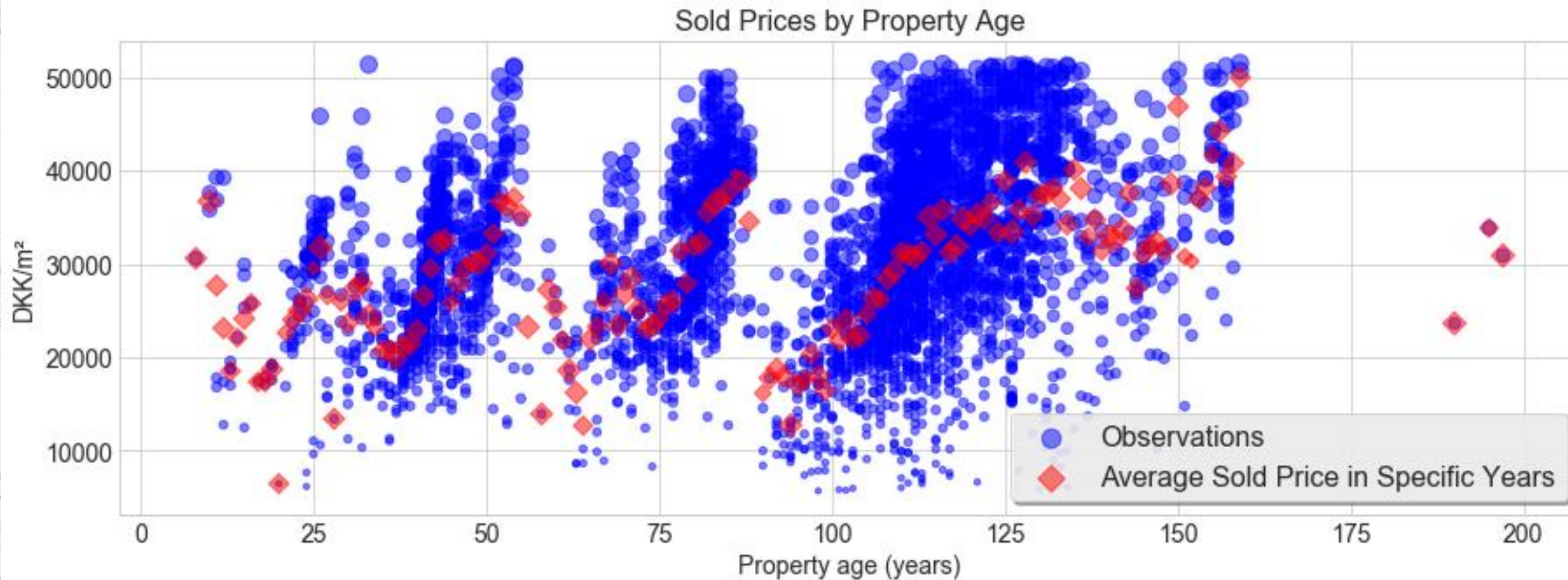
facebook

provides useful tools for perform Real Estate price level predictions.

Prophet Forecast Implementation (Predicted vs. Real Values)
Zone: Copenhagen N

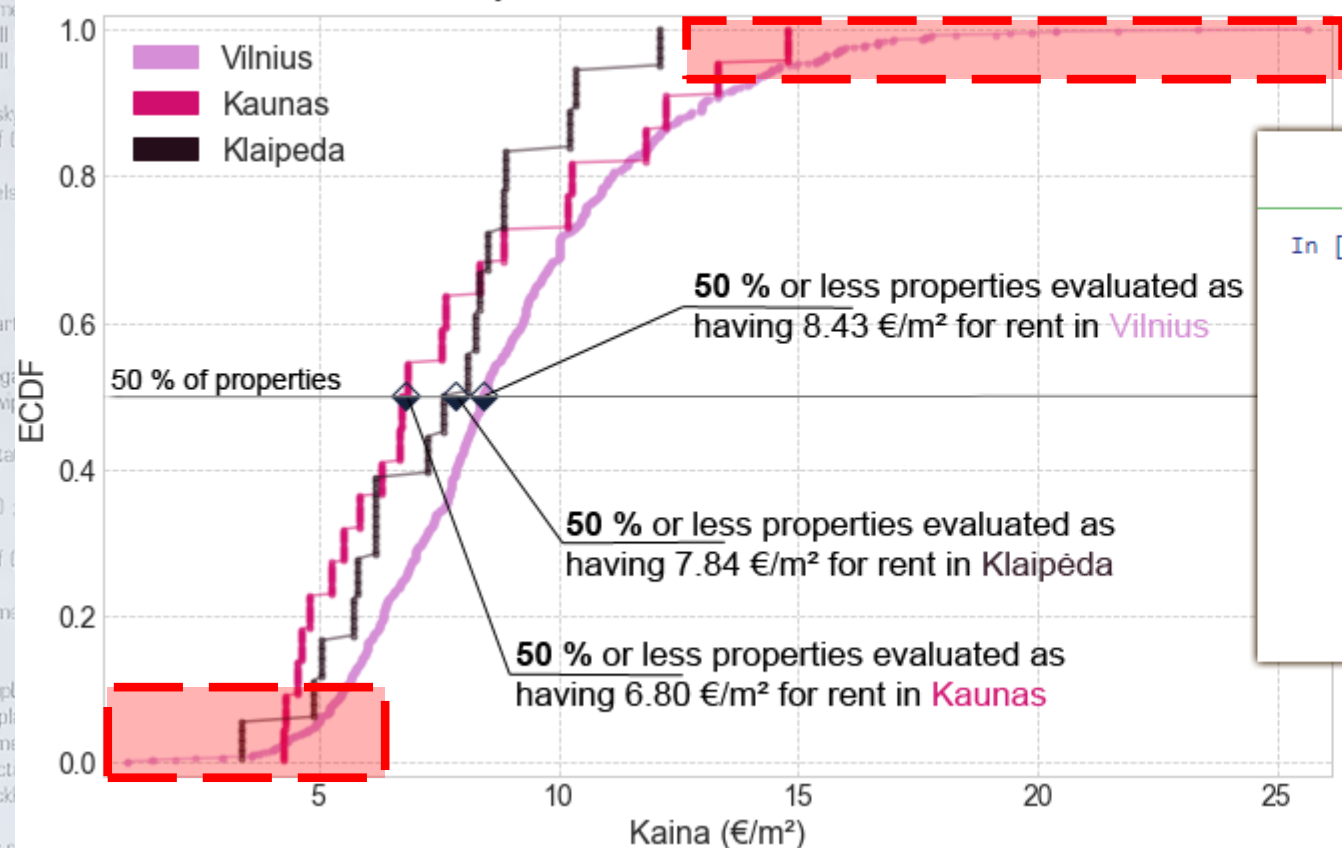


Statistics + Machine Learning in Data Science



Statistics in Data Science

Empirical cumulative distribution function for Statistical models in Real Estate



ECDF function

```
In [1]: # 01 Calculate ECDF for Vilnius, Kaunas and Klaipėda distribution model
def ecdf(data):
    """Compute ECDF for a one-dimensional array of measurements."""

    # Number of data points: n
    n = len(data)

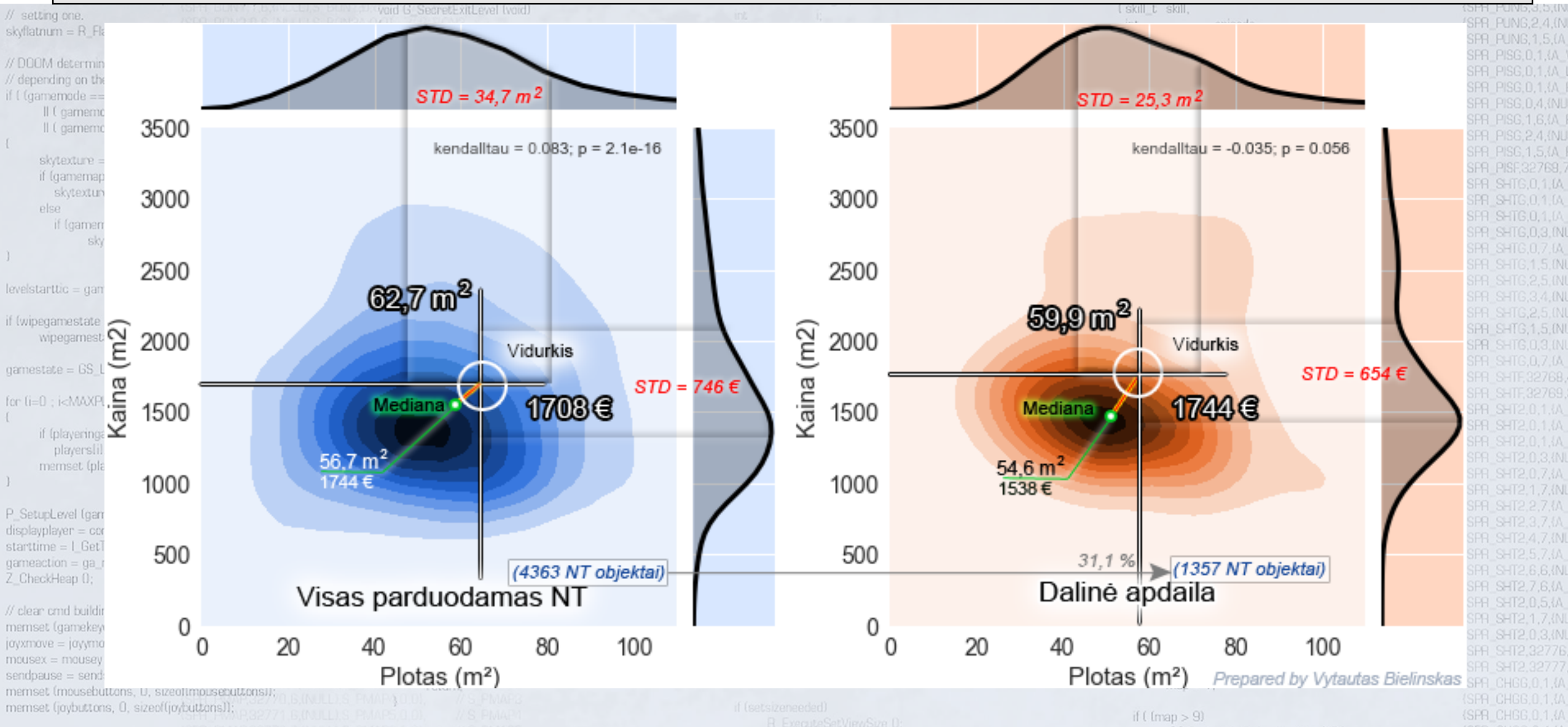
    # x-data for the ECDF: x
    x = np.sort(data)

    # y-data for the ECDF: y
    y = np.arange(1, len(x)+1) / n

    return x, y
```

Prepared by Vytautas Bielinškas

Statistics in Data Science

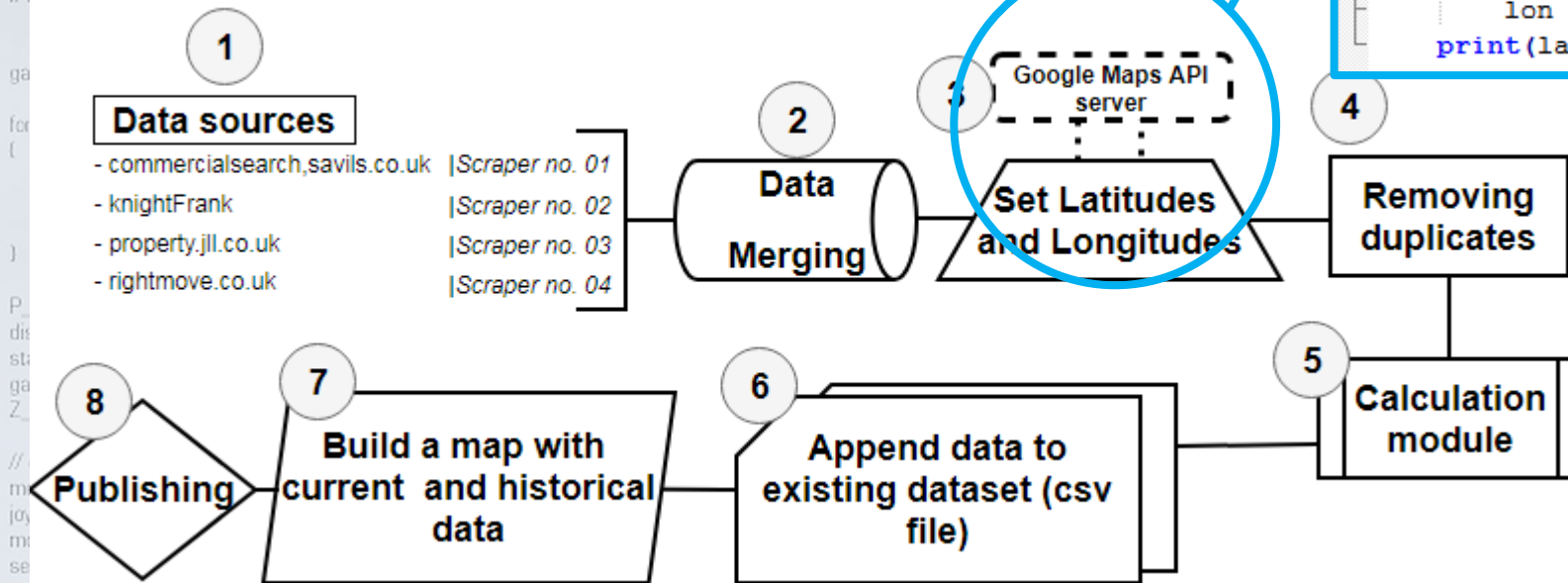


Real Project Processing Pipeline

Google Maps APIs

```
for i in range(0, len(df), 1):
    address = df.iat[i, 1]
    print(i, address)
    geocode_result = gmaps_01.geocode(address)

    try:
        print(i, "Trying API")
        lat = geocode_result[0]["geometry"]["location"]["lat"]
        lon = geocode_result[0]["geometry"]["location"]["lng"]
        df.iat[i, df.columns.get_loc("LAT")] = lat
        df.iat[i, df.columns.get_loc("LON")] = lon
    except:
        lat = 0
        lon = 0
    print(lat, lon)
```

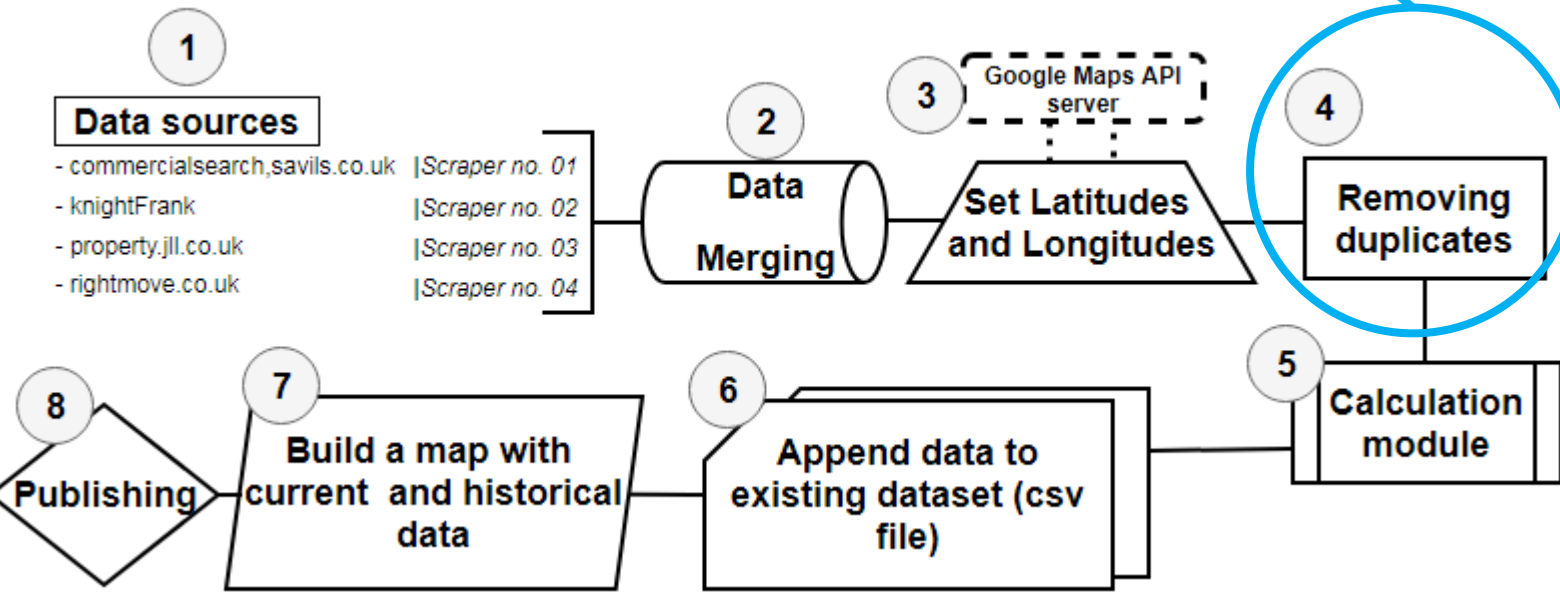


Real Project Processing Pipeline

```

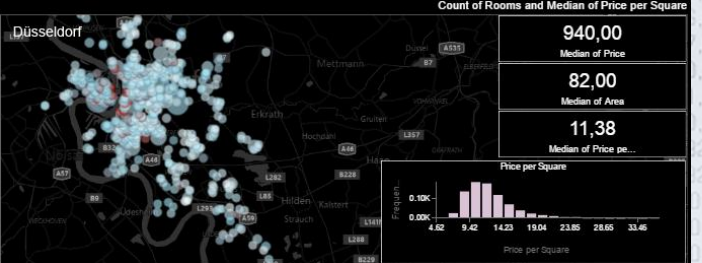
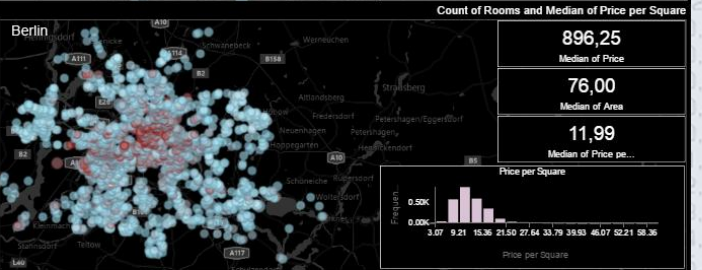
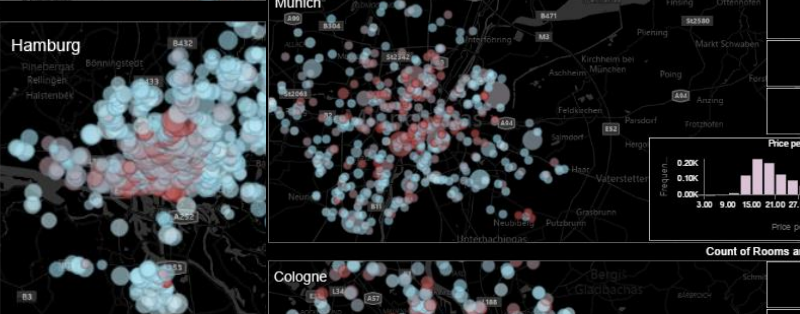
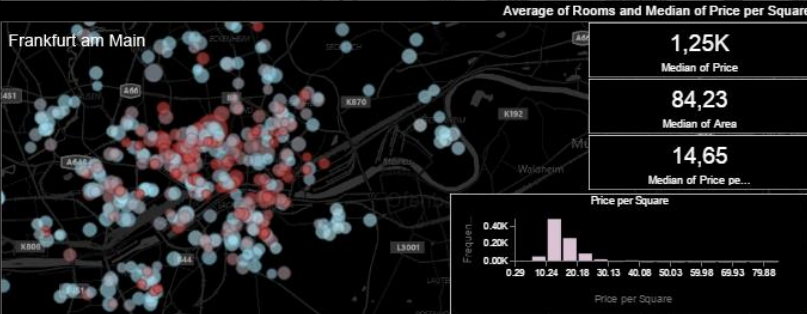
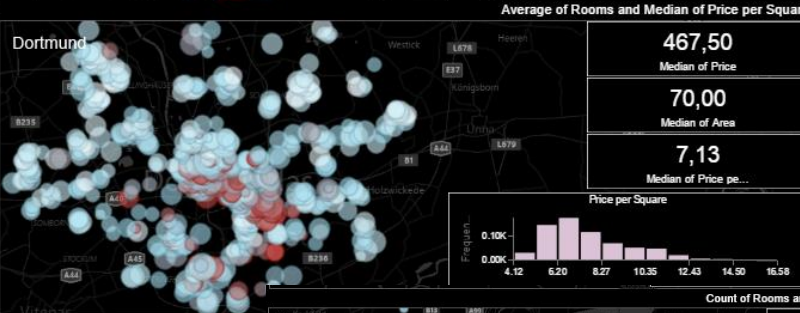
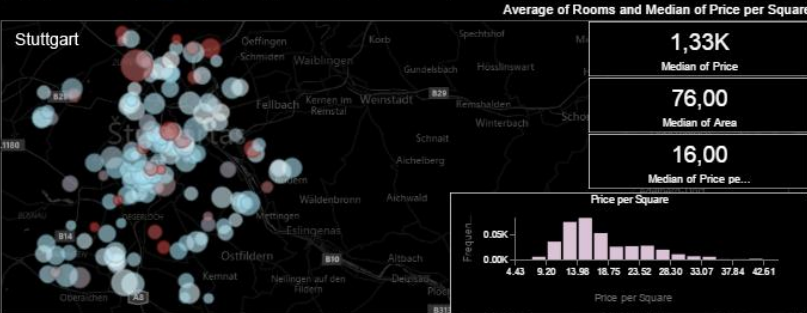
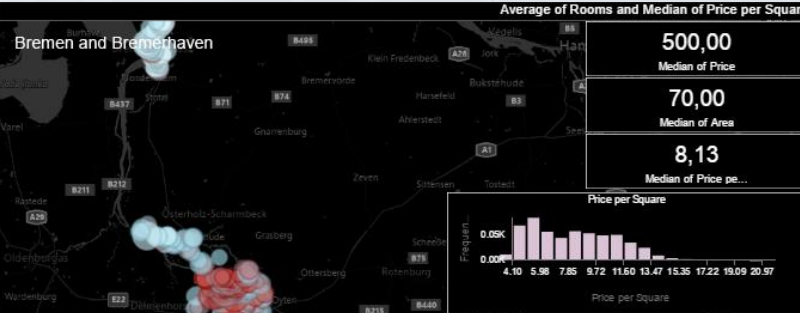
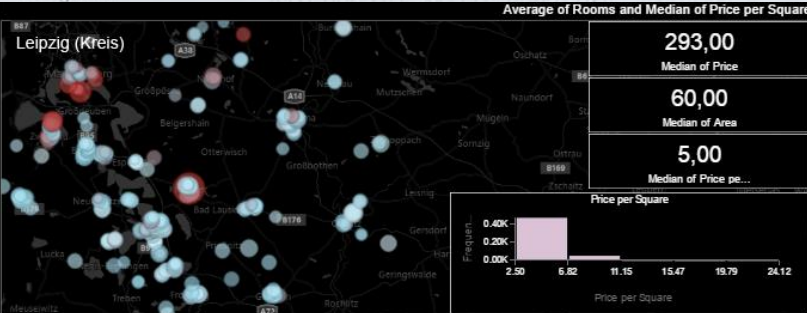
if item < len(df)-1:
    for next_items in range(item+1, len(df), 1):
        if min_size == df.iat[next_items, 5]:
            if max_size == df.iat[next_items, 6]:
                if price == df.iat[next_items, 3]:
                    if differ(lat, df.iat[next_items, 8]) < 0.025:
                        if differ(lon, df.iat[next_items, 9]) < 0.025 :
                            df.iat[next_items, df.columns.get_loc("Duplicate")] = 1

```



```
// This was quite messy with SPECIAL and commented parts.
// Supposedly hacks to make the latest edition work.
// It might not work properly.
if (episode < 1)
    episode = 1;
```

Reporting by frameworks of Business Intelligent



```

// clear cmd building stuff
memset(gamekeydown, 0, sizeof(gamekeydown));
joyxmove = joyymove = 0;
mousex = mousey = 0;
sendpause = sendsave = paused = false;
memset(mousebuttons, 0, sizeof(mousebuttons));
memset(joybuttons, 0, sizeof(joybuttons));

```

Reporting by frameworks of Business Intelligent

```
// setting one.  
skyflatnum = R_FlatNumForName (SKYFLATN/ME);  
// DOOM determines the sky texture to be  
// depending on the current episode, and t  
if (gamemode == commercial)  
    if (gamemode == pack_tnt)  
    if (gamemode == pack_plut)
```

```
skytexture = R_TextureNumForName  
if (gamemap < 12)  
    skytexture = R_TextureNumForName  
else  
    if (gamemap < 21)  
        skytexture = R_TextureNum
```

```
levelstarttic = gametic; // for time ca  
if (wipegamestate == GS_LEVEL)  
    wipegamestate = -1; // for earst  
gamestate = GS_LEVEL;  
for (i=0 ; i<MAXPLAYERS ; i++)  
{  
    if (playeringame[i] && players[i].play  
        players[i].playerstate = PST_REE  
        memset (players[i].frags,0,sizeof(pl
```

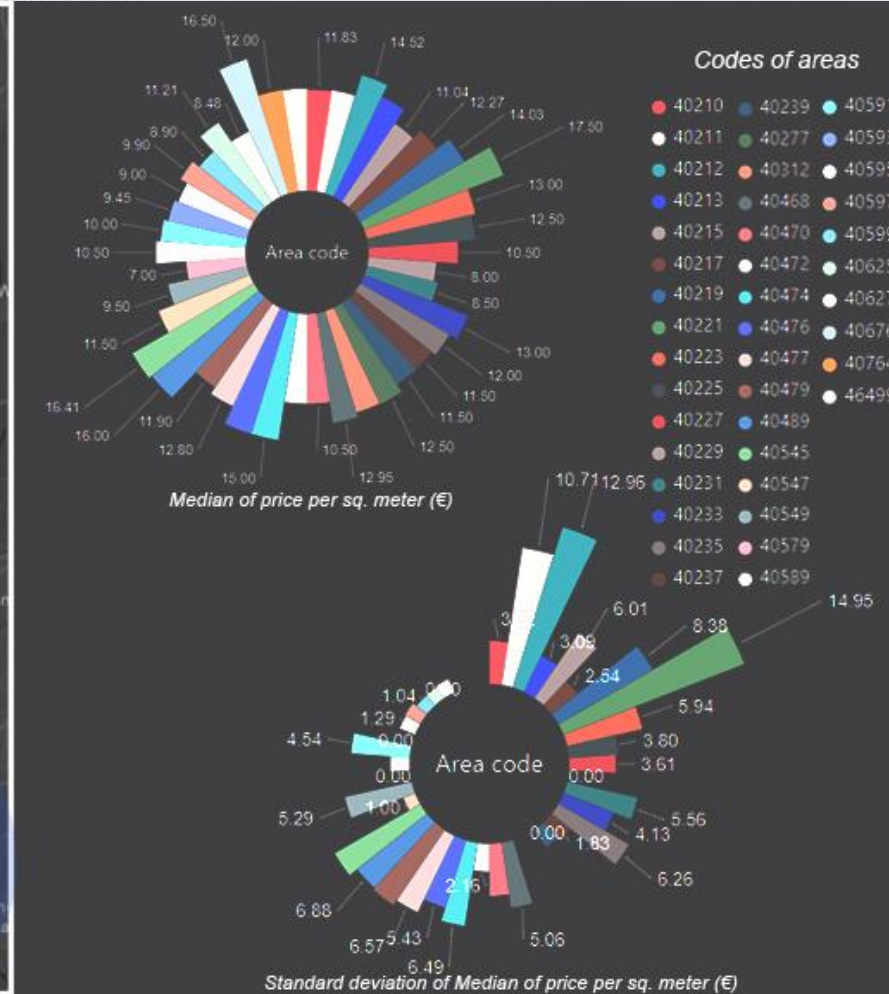
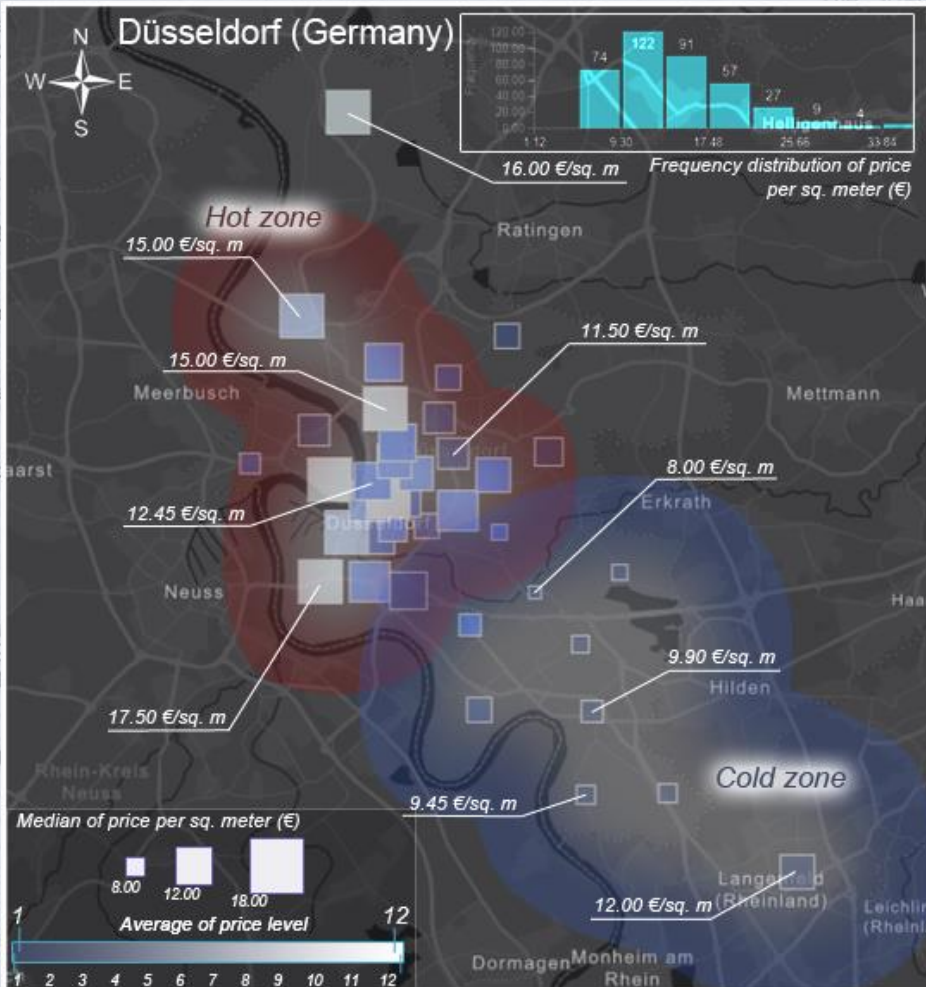
```
P_SetupLevel (gameepisode, gamemap, 0  
displayplayer = consoleplayer;  
starttime = I_GetTime (0);  
gameaction = ga_nothing;  
Z_CheckHeap (0);  
// clear cmd building stuff  
memset (gamekeydown, 0, sizeof(gamekey  
joymove = joymove = 0;  
mousex = mousey = 0;  
sendpause = sendsave = paused = false;  
memset (mousebuttons, 0, sizeof(mousebut  
memset (joybuttons, 0, sizeof(joybutton
```

```
gameaction = ga_victory;  
return;  
if (setsizeneeded)  
    R_ExecuteSetViewSize (0);  
if ( (map > 9)
```

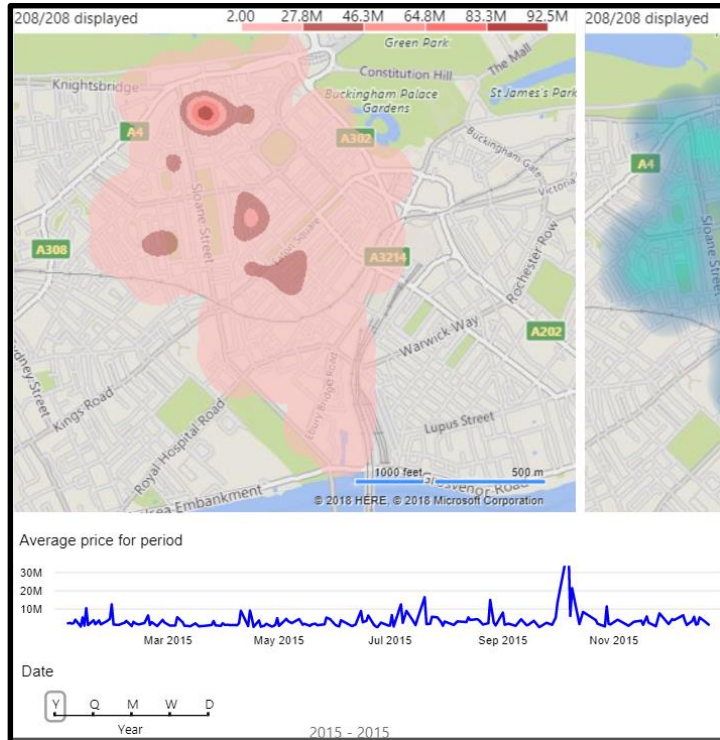
```
if (map < 1)  
    map = 1;  
if (map > 9)
```

```
if (map < 1)  
    map = 1;  
if (map > 9)
```

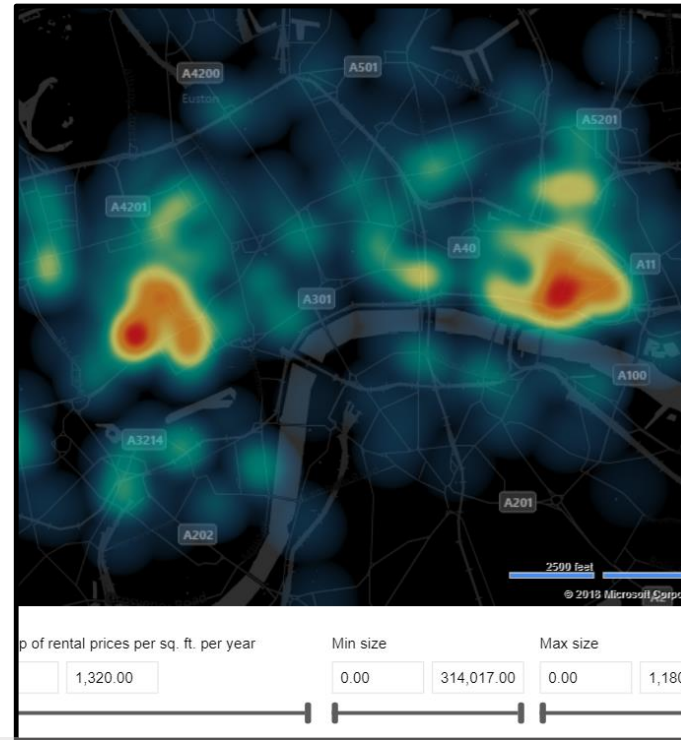
```
if (map < 1)  
    map = 1;  
if (map > 9)
```



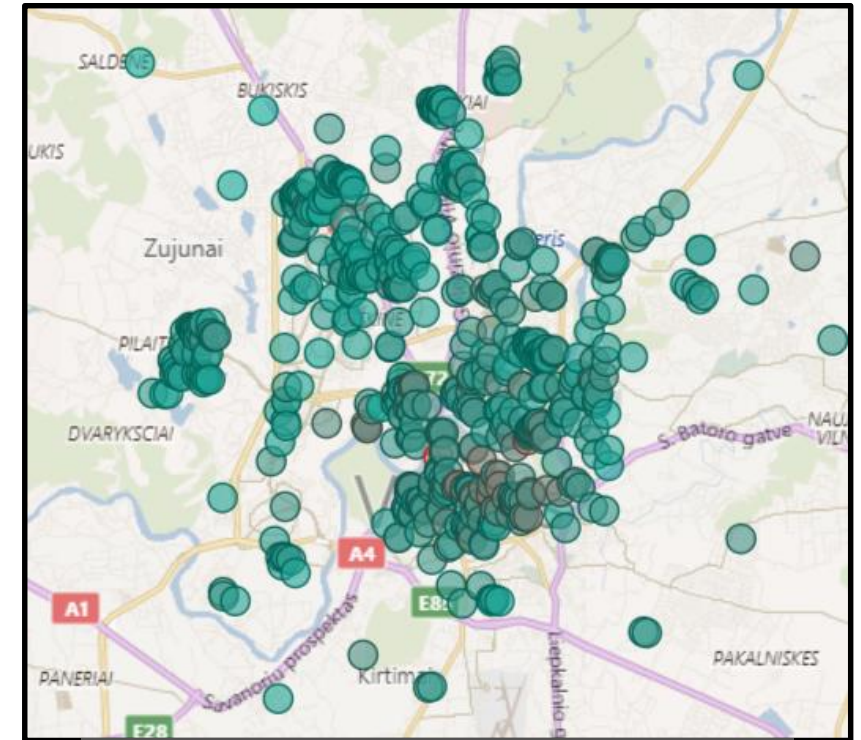
Examples



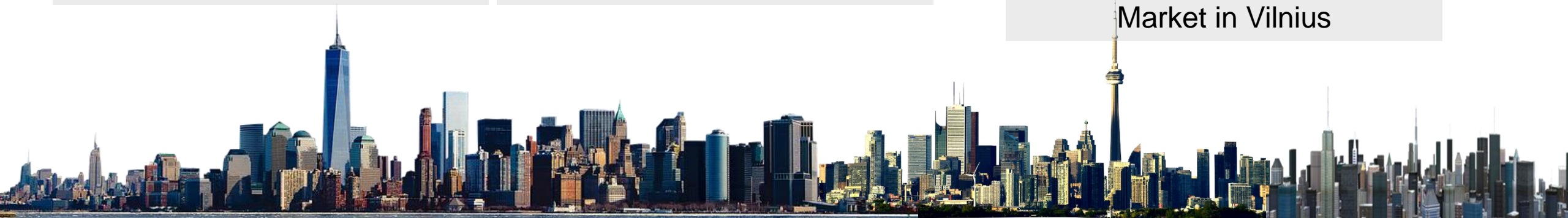
UK Land Registry Data report



London Offices Rental Market



Apartments Rental Flats
Market in Vilnius

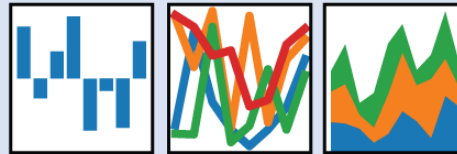


Tools and Libraries



matplotlib

pandas
 $y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$



Seaborn





***Vytautas Bielinskas,
Ekistics Property Advisors LLP***

Thank you for the attention!

