```
In [1]: import pandas as pd
         from pandasai import SmartDataframe
         from pandasai.llm import OpenAI
 In [2]: import numpy as np
         import matplotlib.pyplot as plt
         from sklearn.tree import DecisionTreeRegressor, export text
         from sklearn.model_selection import train_test_split
         from sklearn.metrics import mean_squared_error
In [3]: from sklearn.metrics import r2_score
         from sklearn.model_selection import GridSearchCV
In [4]: house_price = pd.read_csv('Real_estate _1.csv')
         house price.head()
Out[4]:
                                                                                          γ
                                      X3 distance
                         X1
                                 X2
                                                  X4 number of
                                                                                      house
                                           to the
                                                                      X5
                                                                                X6
            No
                 transaction
                              house
                                                   convenience
                                                                                     price of
                                         nearest
                                                                 latitude longitude
                        date
                                                         stores
                                age
                                                                                        unit
                                     MRT station
                                                                                       area
                                        84.87882
                                                            10 24.98298
         0
              1
                    2012.917
                                32.0
                                                                         121.54024
                                                                                        37.9
         1
              2
                    2012.917
                                19.5
                                       306.59470
                                                             9 24.98034
                                                                          121.53951
                                                                                        42.2
         2
              3
                    2013.583
                                13.3
                                       561.98450
                                                             5 24.98746
                                                                         121.54391
                                                                                        47.3
                    2013.500
                                13.3
                                       561.98450
                                                             5 24.98746
                                                                         121.54391
                                                                                        54.8
              5
                    2012.833
                                                                                        43.1
         4
                                 5.0
                                       390.56840
                                                                24.97937
                                                                          121.54245
In [8]:
         house_price.columns = house_price.columns.str.strip().str.lower().str.replace(" ",
In [9]: house_price.rename(columns={'no': 'transaction_id'}, inplace=True)
In [10]: def cap_outliers(series, upper_percentile=0.97):
             upper_bound = series.quantile(upper_percentile)
             return series.clip(upper=upper bound)
In [11]: house_price['x3_distance_to_the_nearest_mrt_station'] = cap_outliers(house_price['x
         print(f"{'x3_distance_to_the_nearest_mrt_station'}: capped at 97th percentile = {ho
        x3_distance_to_the_nearest_mrt_station: capped at 97th percentile = 4435.03305000000
In [13]: outlier_values = [78.3, 117.5, 78.0]
         house_price = house_price[~house_price['y_house_price_of_unit_area'].isin(outlier_v
In [14]: house_price = house_price.drop(columns=['transaction_id', 'x1_transaction_date'])
         house price.head()
```

 Out[14]:
 x2_house_age
 x3_distance_to_the_nearest_mrt_station
 x4_number_of_convenience_stores

 0
 32.0
 84.87882
 10

 1
 19.5
 306.59470
 9

 2
 13.3
 561.98450
 5

561.98450

390.56840

5

5

In [16]: llm = OpenAI(api_token="sk-proj-KtQvt_jEawaTmmBFpqxRE5fXxd7dTJfPGgsmwGYJpNW19XAv3cH
house_price = SmartDataframe(house_price, config ={"llm": llm})

In [17]: house_price.chat('use wisestep to find the best predictors for y_house_price_of_uni

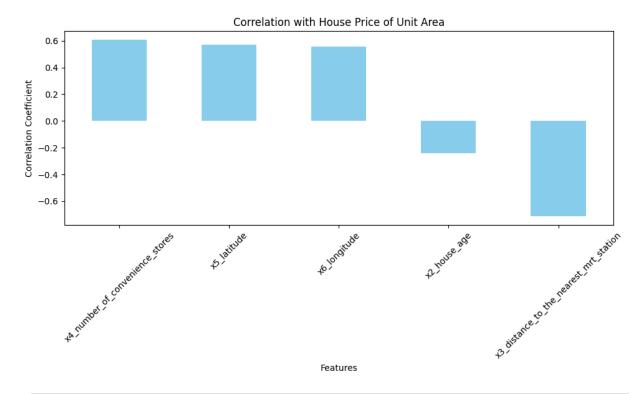
Out[17]:y_house_price_of_unit_areax4_number_of_convenience_stores0.605853x5_latitude0.571849

3

4

13.3

5.0



In [18]: house_price.chat('define x2_house_age, x3_distance_to_the_nearest_mrt_station, x4_n

	x2_house_age	ge x3_distance_to_the_nearest_mrt_station \			
0	32.0	84.87882			
1	19.5		30	6.59470	
2	13.3		56	1.98450	
3	13.3		56	1.98450	
4	5.0		39	0.56840	
	• • •			• • •	
409	13.7		408	2.01500	
410	5.6		9	0.45606	
411	18.8		39	0.96960	
412	8.1		10	4.81010	
413	6.5		9	0.45606	
	v4 numbon of	convenience stance	vE latituda	vC longitudo	
0	x4_number_of_	convenience_stores		x6_longitude	
0	x4_number_of_	10	24.98298	121.54024	
1	x4_number_of_	10 9	24.98298 24.98034	121.54024 121.53951	
1 2	x4_number_of_	10 9 5	24.98298 24.98034 24.98746	121.54024 121.53951 121.54391	
1	x4_number_of_	10 9	24.98298 24.98034	121.54024 121.53951	
1 2	x4_number_of_	10 9 5	24.98298 24.98034 24.98746	121.54024 121.53951 121.54391	
1 2 3	x4_number_of_	10 9 5 5	24.98298 24.98034 24.98746 24.98746	121.54024 121.53951 121.54391 121.54391	
1 2 3 4	x4_number_of_	10 9 5 5 5	24.98298 24.98034 24.98746 24.98746 24.97937	121.54024 121.53951 121.54391 121.54391 121.54245	
1 2 3 4	x4_number_of_	10 9 5 5 5	24.98298 24.98034 24.98746 24.98746 24.97937	121.54024 121.53951 121.54391 121.54391 121.54245	
1 2 3 4 409	x4_number_of_	10 9 5 5 0	24.98298 24.98034 24.98746 24.98746 24.97937 24.94155	121.54024 121.53951 121.54391 121.54391 121.54245 121.50381	
1 2 3 4 409 410	x4_number_of_	10 9 5 5 5 0 9	24.98298 24.98034 24.98746 24.98746 24.97937 24.94155 24.97433	121.54024 121.53951 121.54391 121.54391 121.54245 121.50381 121.54310 121.53986	

[411 rows x 5 columns]

Out[18]: x2_house_age x3_distance_to_the_nearest_mrt_station x4_number_of_convenience_stores

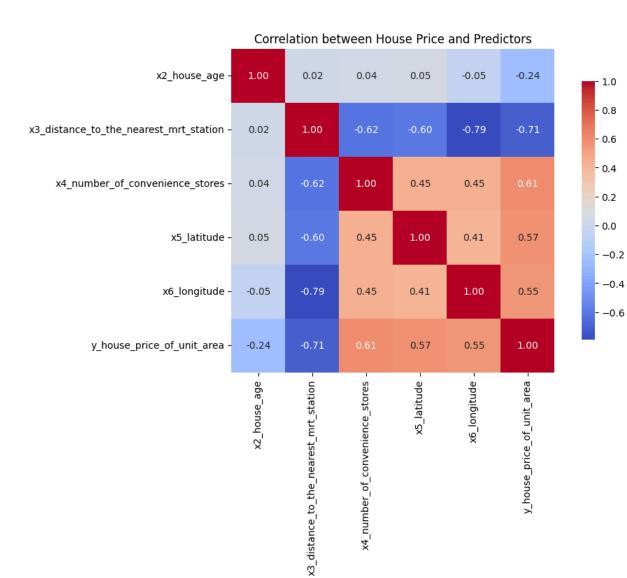
X	70_ui5taiice_to_tiie_iicui c5t_iiii t_5tatioii	x=ouse_uge	
1(84.87882	32.0	0
Č	306.59470	19.5	1
Ē	561.98450	13.3	2
į	561.98450	13.3	3
Ĩ	390.56840	5.0	4
			•••
(4082.01500	13.7	409
ç	90.45606	5.6	410
7	390.96960	18.8	411
Ę	104.81010	8.1	412
ć	90.45606	6.5	413

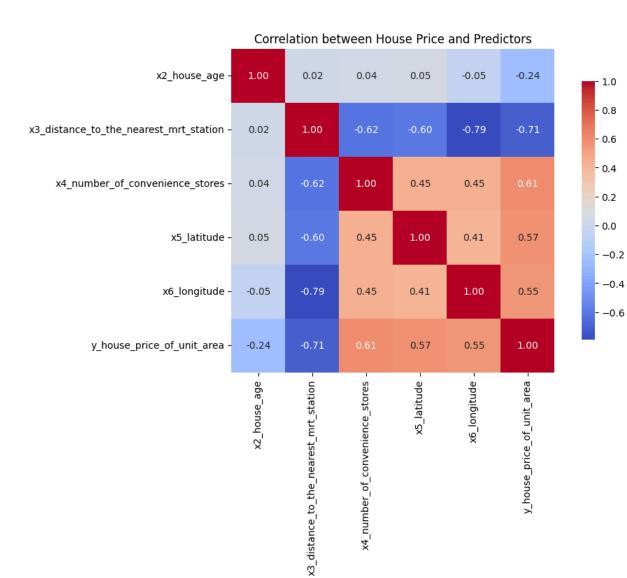
411 rows × 5 columns

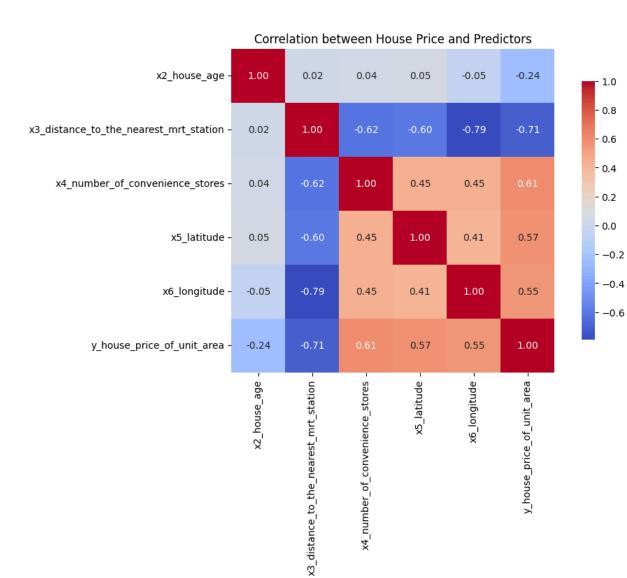
Out[20]:		x2_house_age	x3_distance_to_the_nearest_mrt_station	x4_number_of_convenience_stores
	0	32.0	84.87882	1(
	1	19.5	306.59470	ć
	2	13.3	561.98450	Ē
	3	13.3	561.98450	Ē
	4	5.0	390.56840	Ē
	•••			
	409	13.7	4082.01500	(
	410	5.6	90.45606	ć
	411	18.8	390.96960	7
	412	8.1	104.81010	Ē
	413	6.5	90.45606	Č

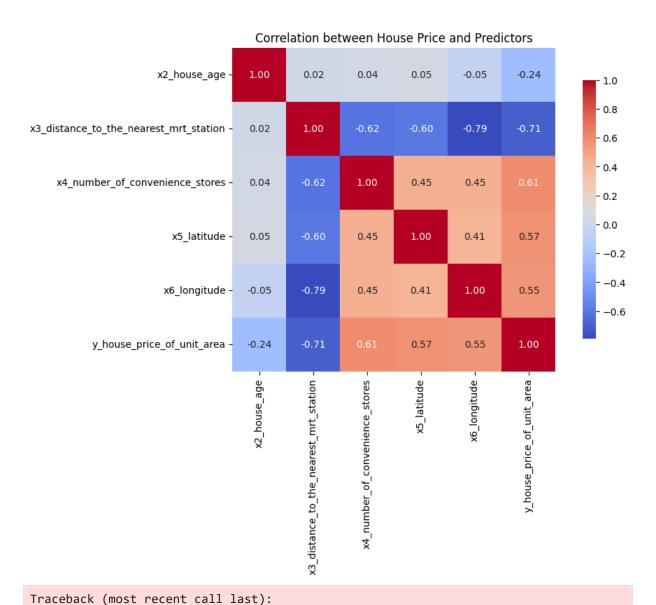
411 rows × 6 columns

In [21]: house_price.chat('visualize correlation between outcome and best predictors')









File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa ges\pandasai\pipelines\chat\generate_chat_pipeline.py", line 335, in run
).run(input)

^^^^^^^

File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa ges\pandasai\pipelines\pipeline.py", line 137, in run raise e

File "C:\Users\ $\ddot{O}RNEK$ A $\ddot{I}LES\dot{I}\AppData\Local\Programs\Python\Python312\Lib\site-packa ges\pandasai\pipelines\pipeline.py", line 101, in run$

step_output = logic.execute(

File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa ges\pandasai\pipelines\chat\code_execution.py", line 113, in execute raise e

File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa ges\pandasai\pipelines\chat\code_execution.py", line 97, in execute raise InvalidOutputValueMismatch(

pandasai.exceptions.InvalidOutputValueMismatch: Value type <class 'str'> must match
with type plot

Out[21]: "Unfortunately, I was not able to answer your question, because of the following e rror:\n\nValue type <class 'str'> must match with type plot\n"

Splittig data

In [32]: house_price.chat('divide the original data frame house_price into two data frames w

```
Independent Variables (X):
     x2_house_age x3_distance_to_the_nearest_mrt_station \
0
             32.0
                                                  84.87882
             19.5
1
                                                 306.59470
2
             13.3
                                                 561.98450
3
             13.3
                                                 561.98450
4
              5.0
                                                 390.56840
              . . .
                                                        . . .
. .
             13.7
                                                4082.01500
409
             5.6
                                                  90.45606
410
411
             18.8
                                                 390.96960
              8.1
412
                                                 104.81010
413
              6.5
                                                  90.45606
     x4_number_of_convenience_stores x5_latitude x6_longitude
                                          24.98298
                                                       121.54024
0
                                   10
1
                                    9
                                          24.98034
                                                       121.53951
2
                                    5
                                          24.98746
                                                       121.54391
3
                                    5
                                          24.98746
                                                      121.54391
4
                                    5
                                          24.97937
                                                      121.54245
                                  . . .
                                               . . .
                                                             . . .
                                                   121.50381
121.54310
409
                                    0
                                          24.94155
410
                                    9
                                          24.97433
411
                                    7
                                          24.97923
                                                       121.53986
412
                                    5
                                          24.96674
                                                       121.54067
413
                                    9
                                          24.97433
                                                      121.54310
[411 rows x 5 columns]
Dependent Variable (y):
       37.9
1
       42.2
2
       47.3
3
       54.8
4
       43.1
       . . .
409
       15.4
410
       50.0
       40.6
411
412
       52.5
413
       63.9
Name: y_house_price_of_unit_area, Length: 411, dtype: float64
Independent Variables (X):
     x2_house_age x3_distance_to_the_nearest_mrt_station \
0
             32.0
                                                  84.87882
             19.5
                                                 306.59470
1
2
             13.3
                                                 561.98450
3
             13.3
                                                 561.98450
4
              5.0
                                                 390.56840
              . . .
                                                4082.01500
             13.7
409
410
             5.6
                                                  90.45606
             18.8
411
                                                 390.96960
412
              8.1
                                                 104.81010
413
              6.5
                                                  90.45606
```

```
x4_number_of_convenience_stores x5_latitude x6_longitude
0
                                           24.98298
                                                         121.54024
1
                                     9
                                           24.98034
                                                         121.53951
2
                                     5
                                           24.98746
                                                         121.54391
3
                                     5
                                           24.98746
                                                         121.54391
4
                                     5
                                           24.97937
                                                         121.54245
                                                . . .
409
                                     0
                                           24.94155
                                                         121.50381
                                     9
                                           24.97433
410
                                                         121.54310
                                     7
                                           24.97923
                                                         121.53986
411
412
                                     5
                                           24.96674
                                                         121.54067
                                     9
                                           24.97433
413
                                                         121.54310
[411 rows x 5 columns]
Dependent Variable (y):
       37.9
       42.2
1
2
       47.3
3
       54.8
4
       43.1
       . . .
409
       15.4
       50.0
410
411
       40.6
412
       52.5
413
       63.9
Name: y_house_price_of_unit_area, Length: 411, dtype: float64
Independent Variables (X):
     x2_house_age x3_distance_to_the_nearest_mrt_station \
0
             32.0
                                                    84.87882
1
             19.5
                                                   306.59470
2
             13.3
                                                   561.98450
3
             13.3
                                                   561.98450
4
              5.0
                                                   390.56840
              . . .
409
             13.7
                                                 4082.01500
410
              5.6
                                                   90.45606
411
             18.8
                                                   390.96960
412
              8.1
                                                  104.81010
413
              6.5
                                                   90.45606
     x4_number_of_convenience_stores x5_latitude x6_longitude
0
                                    10
                                           24.98298
                                                         121.54024
1
                                     9
                                           24.98034
                                                         121.53951
2
                                     5
                                           24.98746
                                                         121.54391
3
                                     5
                                           24.98746
                                                         121.54391
4
                                     5
                                           24.97937
                                                         121.54245
. .
                                   . . .
                                                . . .
                                                               . . .
                                           24.94155
                                                         121.50381
409
                                     0
                                     9
410
                                           24.97433
                                                         121.54310
                                     7
                                           24.97923
411
                                                         121.53986
412
                                     5
                                           24.96674
                                                         121.54067
413
                                     9
                                           24.97433
                                                         121.54310
```

[411 rows x 5 columns]

```
Dependent Variable (y):
              37.9
       0
       1
              42.2
       2
              47.3
       3
              54.8
              43.1
               . . .
       409
              15.4
              50.0
       410
              40.6
       411
       412
              52.5
       413
              63.9
       Name: y_house_price_of_unit_area, Length: 411, dtype: float64
       Traceback (most recent call last):
          File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\generate_chat_pipeline.py", line 335, in run
            ).run(input)
              ^^^^^
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\pipeline.py", line 137, in run
          File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\pipeline.py", line 101, in run
            step_output = logic.execute(
                         ^^^^^^
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\code_execution.py", line 113, in execute
            raise e
          File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\code_execution.py", line 85, in execute
            result = self.execute code(code to run, code context)
                    ^^^^^
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\code_execution.py", line 175, in execute_code
            raise NoResultFoundError("No result returned")
       pandasai.exceptions.NoResultFoundError: No result returned
Out[32]: 'Unfortunately, I was not able to answer your question, because of the following e
         rror:\n\nNo result returned\n'
In [29]: house_price.chat('split y_house_price_of_unit_area and best predictors into training
       Training data shape: (246, 6)
       Validation data shape: (165, 6)
       Training data shape: (246, 6)
```

Validation data shape: (165, 6)

```
Traceback (most recent call last):
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\generate_chat_pipeline.py", line 335, in run
    ).run(input)
      ^^^^^
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\pipeline.py", line 137, in run
    raise e
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\pipeline.py", line 101, in run
    step_output = logic.execute(
 File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code_execution.py", line 113, in execute
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code_execution.py", line 85, in execute
    result = self.execute_code(code_to_run, code_context)
            ^^^^^^
 File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code_execution.py", line 175, in execute_code
    raise NoResultFoundError("No result returned")
pandasai.exceptions.NoResultFoundError: No result returned
```

Out[29]: 'Unfortunately, I was not able to answer your question, because of the following e rror:\n\nNo result returned\n'

Best Hyperparameters

In []: # We will use Hyperparameters parameters that can be fine-tuned to improve the accu
In [30]: # Checking for Best Hyperparameters
house_price.chat('look at the best hyperparameter combination of max_depth, min_sam

```
arn\model_selection\_validation.py:528: FitFailedWarning:
720 fits failed out of a total of 2160.
The score on these train-test partitions for these parameters will be set to nan.
If these failures are not expected, you can try to debug them by setting error_score
='raise'.
Below are more details about the failures:
720 fits failed with the following error:
Traceback (most recent call last):
 File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\sklearn\model_selection\_validation.py", line 866, in _fit_and_score
   estimator.fit(X_train, y_train, **fit_params)
 File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\sklearn\base.py", line 1382, in wrapper
   estimator. validate params()
 File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\sklearn\base.py", line 436, in _validate_params
   validate_parameter_constraints(
 File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\sklearn\utils\_param_validation.py", line 98, in validate_parameter_constraints
   raise InvalidParameterError(
sklearn.utils._param_validation.InvalidParameterError: The 'max_features' parameter
of RandomForestRegressor must be an int in the range [1, inf), a float in the range
(0.0, 1.0], a str among {'sqrt', 'log2'} or None. Got 'auto' instead.
 warnings.warn(some_fits_failed_message, FitFailedWarning)
C:\Users\\ORNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packages\skle
arn\model_selection\_search.py:1108: UserWarning: One or more of the test scores are
non-finite: [
                 nan
                            nan
                                                           nan
                                                                     nan
                nan
                          nan
                                     nan
       nan
                                                nan
                                                          nan
                         nan
       nan
                nan
                                     nan
                                                nan
                                                          nan
       nan
                nan
                          nan
                                     nan
                                              nan
                                                          nan
                 nan
                           nan
                                     nan
                                               nan
                                                          nan
       nan
       nan
                 nan
                           nan
                                     nan
                                                nan
                                                          nan
 0.83008139 0.82618807 0.82302974 0.81892816 0.8184234 0.81322589
 0.79577817 0.78991245 0.79276693 0.79111162 0.79102203 0.78713341
 0.82150463 0.82481318 0.81463409 0.81237812 0.81396504 0.80897802
 0.80176877 0.80537533 0.8046106 0.82453139 0.82651971 0.8202324
 0.81947375 0.81601267 0.81334362 0.80820602 0.80471164 0.80455327
 0.82982119 0.82651411 0.8239287 0.82050459 0.82074811 0.81644803
 0.79570255 0.79289969 0.79412681 0.79128027 0.78806808 0.78716633
 0.82201881 0.8248137 0.81818862 0.8167128 0.81774719 0.81073448
 0.80798828 0.80498673 0.80639542 0.82875709 0.82631475 0.82051849
 0.81582827 0.81708482 0.8119455 0.80609625 0.80286359 0.8042052
                         nan
                nan
                                              nan
       nan
                                     nan
                                                          nan
                nan
                          nan
                                     nan
                                               nan
                                                          nan
       nan
                nan
                                     nan
                                                          nan
       nan
                          nan
                                               nan
       nan
                 nan
                           nan
                                      nan
                                                nan
                                                          nan
                 nan
                           nan
                                      nan
                                                nan
                                                          nan
                 nan
       nan
                           nan
                                     nan
                                                nan
                                                          nan
 0.82633769 0.82777903 0.8218108 0.81947208 0.81891324 0.81617359
 0.80675022 0.80869544 0.80356731 0.79573689 0.79935594 0.80137387
```

C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packages\skle

```
0.79214023 0.79448161 0.79366645 0.7871792 0.79016348 0.78497814
 0.82143914 0.8219534 0.82153031 0.81055681 0.81590478 0.81144615
0.80391774 0.80846275 0.80142565 0.82351078 0.82581967 0.82105303
 0.81969248 0.81415624 0.81236743 0.80496293 0.80524377 0.80570014
 0.82646231 0.8285453 0.81992706 0.81892201 0.81965132 0.81536428
 0.80651154 0.8017928 0.8030058 0.80057649 0.80230443 0.8016478
 0.79745392 0.79667195 0.79258557 0.78850377 0.78726229 0.79379069
0.82376414 0.82017963 0.81939714 0.81281844 0.81747576 0.81053313
0.80481227 0.80107984 0.80479062 0.82299226 0.82549796 0.82360246
 0.81996241 0.81431264 0.81101939 0.80759652 0.80961219 0.80064549
       nan
                  nan
                             nan
                                        nan
                                                   nan
                                                             nan
       nan
                  nan
                             nan
                                        nan
                                                  nan
                                                             nan
       nan
                  nan
                             nan
                                        nan
                                                  nan
                                                             nan
       nan
                  nan
                             nan
                                        nan
                                                  nan
                                                             nan
       nan
                  nan
                             nan
                                                             nan
                                        nan
                                                   nan
       nan
                  nan
                             nan
                                        nan
                                                   nan
                                                             nan
0.83058054 0.82605932 0.81951695 0.81878006 0.81839811 0.81238774
 0.80556028 0.80590953 0.80340711 0.79596973 0.80089614 0.79792412
 0.78989939 0.79564613 0.7939421 0.78833417 0.79062119 0.79321095
 0.82075308 0.82190132 0.82043086 0.81558915 0.8119523 0.81262339
 0.80393138 0.8065429 0.8038433 0.82879852 0.82465638 0.82411909
 0.81914943 0.81726023 0.81126224 0.80507838 0.80333522 0.80019723
 0.82664167 0.82465983 0.8240187 0.82208521 0.8239796 0.80889864
 0.79925811 0.80660806 0.80258977 0.80278912 0.80511943 0.79889164
 0.79428141 0.7915417 0.79355413 0.7859805 0.79131408 0.78552655
 0.82479356 0.82500239 0.82001576 0.81217775 0.81775976 0.81452941
0.80701288 0.80152315 0.80225209 0.82560588 0.82482494 0.8198719
 0.82248329 0.81859698 0.81319551 0.8047644 0.80525052 0.80281958
       nan
                  nan
                             nan
                                        nan
                                                  nan
                                                             nan
                  nan
                             nan
                                        nan
                                                             nan
       nan
                                                   nan
       nan
                  nan
                             nan
                                        nan
                                                  nan
                                                             nan
       nan
                  nan
                             nan
                                        nan
                                                  nan
                                                             nan
       nan
                  nan
                             nan
                                        nan
                                                   nan
                                                             nan
       nan
                  nan
                             nan
                                        nan
                                                   nan
                                                             nan
0.82939413 0.82622125 0.82436517 0.8235258 0.81710395 0.81272576
 0.80294061 0.80284522 0.80156009 0.80009413 0.8004597 0.80381489
 0.79420082 0.79549828 0.79677372 0.78574933 0.78897111 0.79472319
0.82239243 0.81607222 0.82234748 0.8134314 0.81355304 0.81446441
 0.82074294 0.82015714 0.81351266 0.80510856 0.80331144 0.8019615
 0.83222408 0.82627213 0.81950084 0.81778363 0.82471391 0.81035461
0.80046761 0.80384716 0.80535885 0.79532565 0.7996502 0.80082225
 0.79545341 0.79245568 0.79450247 0.7897818 0.79296477 0.78906178
 0.82071833 0.82306083 0.82120571 0.81856395 0.8130703 0.80778973
 0.80396573 0.80582828 0.80244896 0.8260285 0.82653997 0.82361196
0.81764595 0.81769215 0.81228967 0.80539801 0.80603591 0.79980132]
 warnings.warn(
{'type': 'string', 'value': "The best hyperparameter combination is: {'max_depth': 3
0, 'max_features': 'log2', 'max_leaf_nodes': None, 'min_samples_leaf': 1, 'min_sampl
es split': 2}."}
```

In []: # max_depth: It denotes the tree's maximum depth. It supports any int value or "Non
min_samples_split: It refers to the minimum number of samples needed to split an

```
# min_samples_leaf: It refers to the minimum no. of samples required at the leaf no
# max_features: It indicates the number of features to be considered in order to fi
```

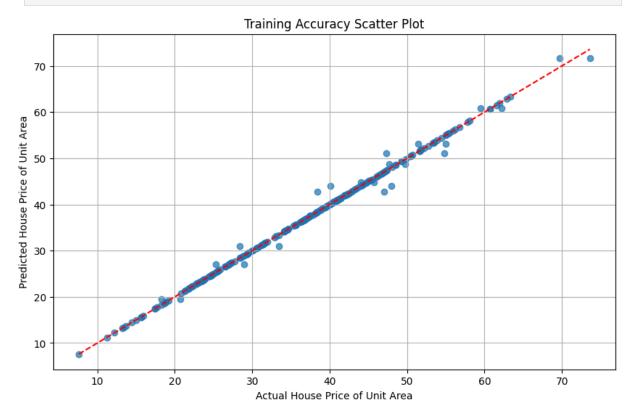
In [40]: house_price.chat('create an object of DecisionTreeRegressor with max_depth =30, min

Out[40]: 'DecisionTreeRegressor object created successfully.'

Training accuracy

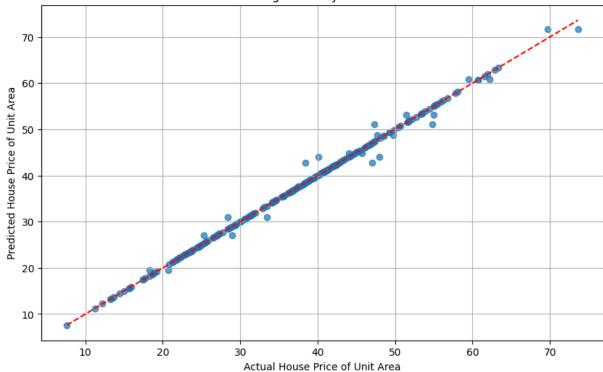
```
In [41]: house_price.chat('alculate the training accuracy using the R2 score')
Out[41]: 0.9962342622663247
In [ ]: # training accuracy is very hight = 99,62%
         # Both spreads are almost completely overlapping one another, indicating that train
```

house_price.chat('use a scatter plot to see the training accuracy') In [42]:



Training Accuracy (R2 Score): 0.9962342622663247





```
Traceback (most recent call last):
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\generate_chat_pipeline.py", line 335, in run
    ).run(input)
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\pipeline.py", line 137, in run
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\pipeline.py", line 101, in run
    step output = logic.execute(
                 ^^^^^^
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code_execution.py", line 113, in execute
    raise e
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code execution.py", line 85, in execute
    result = self.execute_code(code_to_run, code_context)
            ^^^^^^
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code_execution.py", line 175, in execute_code
    raise NoResultFoundError("No result returned")
pandasai.exceptions.NoResultFoundError: No result returned
```

Out[42]: 'Unfortunately, I was not able to answer your question, because of the following e rror:\n\nNo result returned\n'

```
In [43]: house_price.chat('calculate regression statistic')
# this time pandasAI calculated testing accuracy againg and gave us different resul
```

Out[43]: 'R2 Score: 0.987795704926137, RMSE: 1.4089038781531908'

Testing accuracy

```
In [47]: house_price.chat('show the testing accuracy')

Out[47]: 'The testing accuracy (R2 score) is 0.6771, MAE: 5.6497, RMSE: 7.5989.'

In []: # Testing accuracy is # There is overfitting in the model. Our training accuracy is between 99-98% while

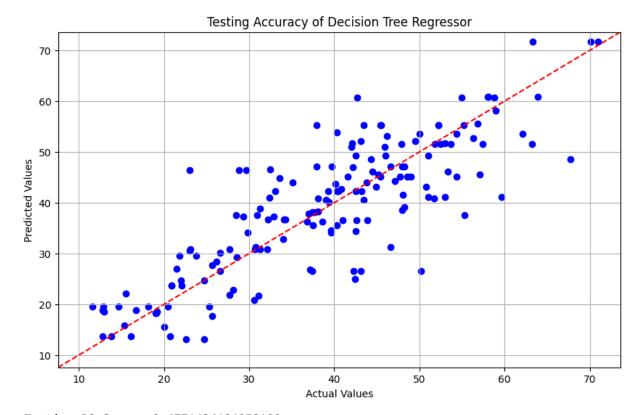
In [48]: house_price.chat('calculate ME, MPE, MAPE for testing dataset ')

Out[48]: ME MPE MAPE

O -3.759124 -10.0 10.0

In [50]: house_price.chat('visualize the testing accuracy of decision tree regressor')
```

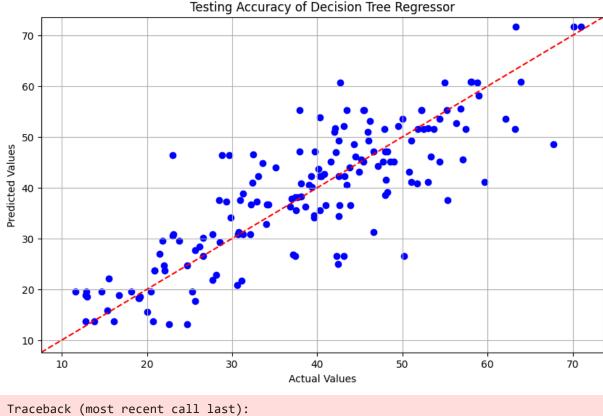
Testing R2 Score: 0.6771436184958102



Testing R2 Score: 0.6771436184958102



Testing R2 Score: 0.6771436184958102



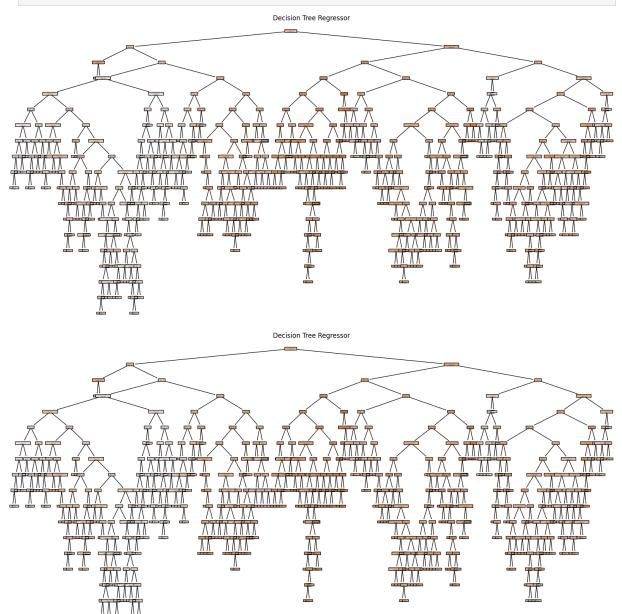
```
File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\generate_chat_pipeline.py", line 335, in run
    ).run(input)
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\pipeline.py", line 137, in run
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\pipeline.py", line 101, in run
    step output = logic.execute(
                 ^^^^^^
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code_execution.py", line 113, in execute
    raise e
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code execution.py", line 85, in execute
    result = self.execute_code(code_to_run, code_context)
            ^^^^^^
  File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
ges\pandasai\pipelines\chat\code_execution.py", line 175, in execute_code
    raise NoResultFoundError("No result returned")
pandasai.exceptions.NoResultFoundError: No result returned
```

Out[50]: 'Unfortunately, I was not able to answer your question, because of the following e rror:\n\nNo result returned\n'

Visualizing Regression Decision Tree with Graphviz

In [52]: # downloading Grahpviz package
 from sklearn import tree
 import graphviz

In [51]: house_price.chat('visualize the decision tree itself by using the tree module of sk



```
Traceback (most recent call last):
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\generate_chat_pipeline.py", line 335, in run
            ).run(input)
             ^^^^^
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\pipeline.py", line 137, in run
            raise e
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\pipeline.py", line 101, in run
            step_output = logic.execute(
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\code_execution.py", line 113, in execute
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\code_execution.py", line 85, in execute
            result = self.execute_code(code_to_run, code_context)
                    ^^^^^^
         File "C:\Users\ÖRNEK AİLESİ\AppData\Local\Programs\Python\Python312\Lib\site-packa
       ges\pandasai\pipelines\chat\code_execution.py", line 175, in execute_code
            raise NoResultFoundError("No result returned")
       pandasai.exceptions.NoResultFoundError: No result returned
Out[51]: 'Unfortunately, I was not able to answer your question, because of the following e
         rror:\n\nNo result returned\n'
```

In []: