

Model: RandomForestRegressor (scikit-learn)
Split: 80/10/10 (chronologically - train/val/test, where test is very latest data)
Data: Base HDB data (no location features)

Hyperparameters:

```
{'bootstrap': True,  
 'ccp_alpha': 0.0,  
 'criterion': 'squared_error',  
 'max_depth': 22,  
 'max_features': 18,  
 'max_leaf_nodes': None,  
 'max_samples': None,  
 'min_impurity_decrease': 0.0,  
 'min_samples_leaf': 2,  
 'min_samples_split': 2,  
 'min_weight_fraction_leaf': 0.0,  
 'n_estimators': 200,  
 'n_jobs': -1,  
 'oob_score': False,  
 'random_state': 42,  
 'verbose': 0,  
 'warm_start': False}
```

Results:

--- Test Set ---

Mean Absolute Error: ... 7.357105190379599
Mean Squared Error:..... 104.88
RMSE: 10.241221739791756
Coeff of det (R^2):..... 0.705

--- Val Set ---

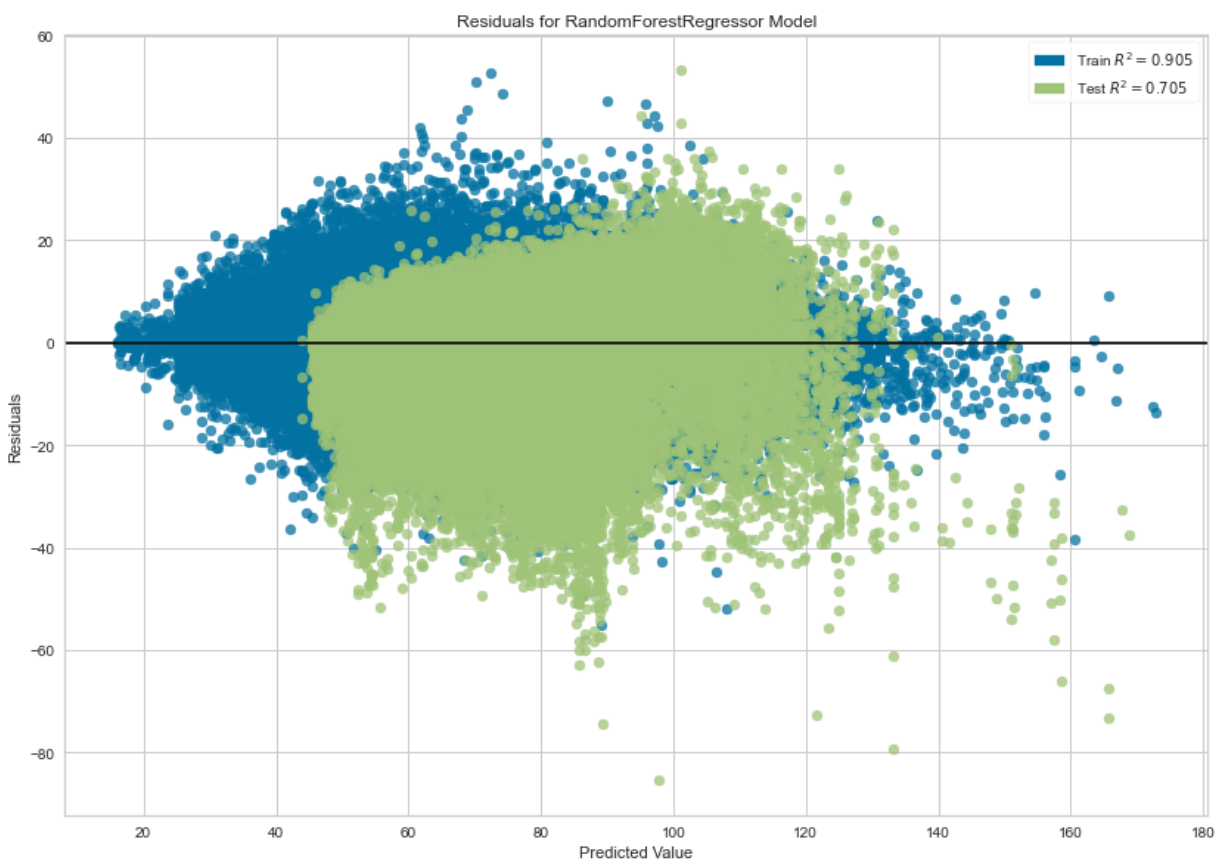
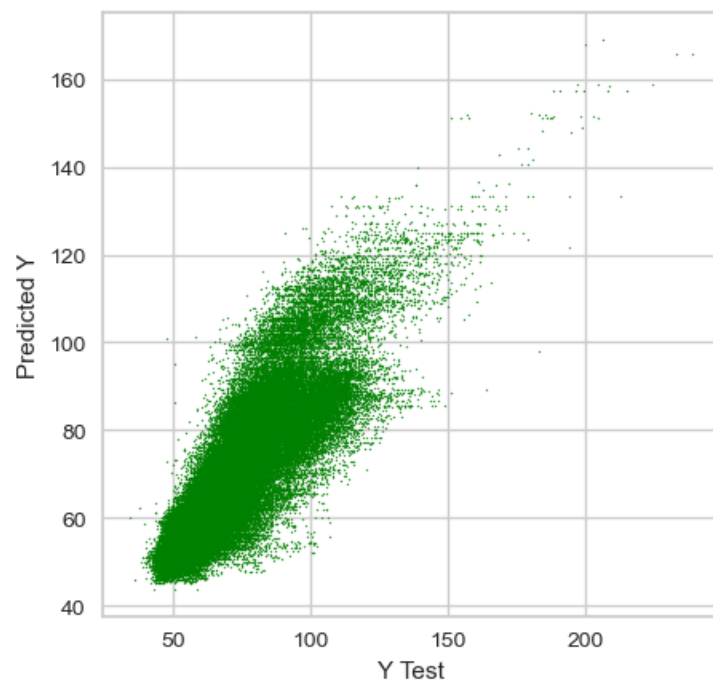
Mean Absolute Error: ... 5.641605598502287
Mean Squared Error:..... 61.92
RMSE: 7.869156277441714
Coeff of det (R^2):..... 0.848

--- Train Set ---

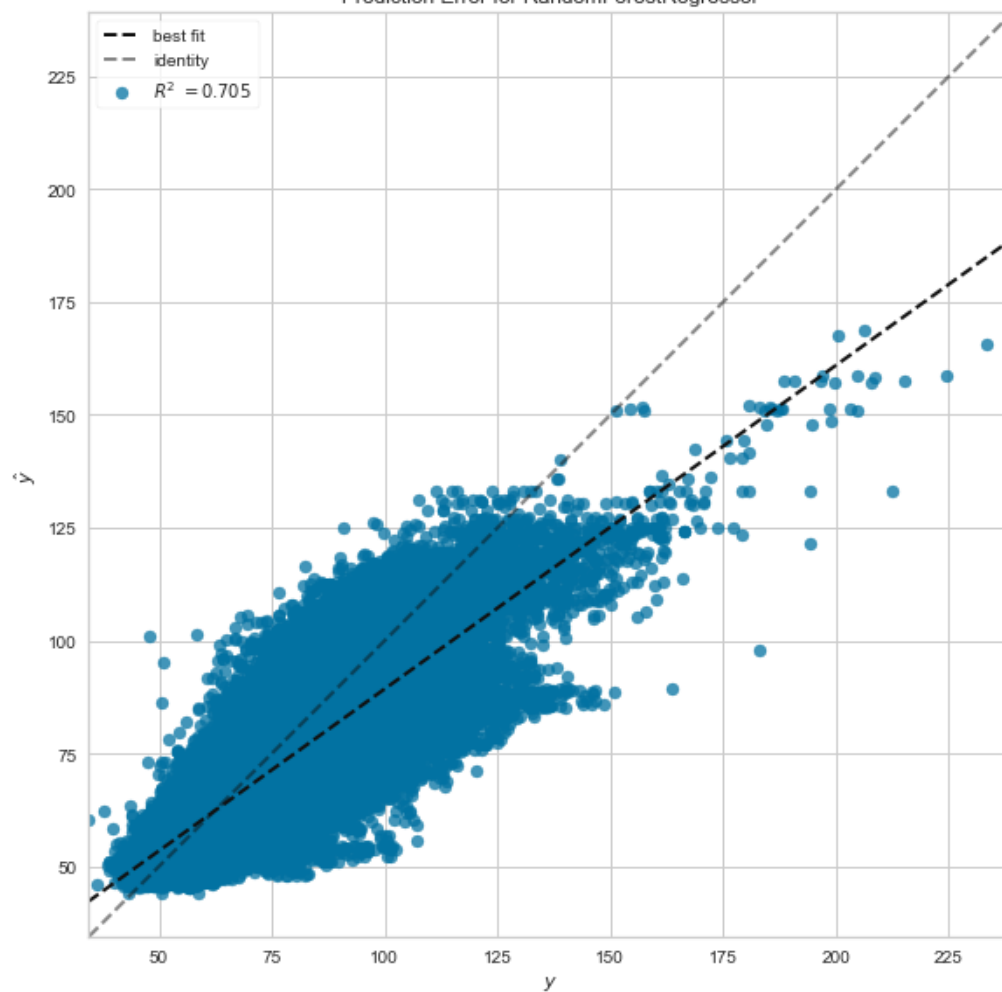
Mean Absolute Error: ... 3.1527835875700263
Mean Squared Error:..... 19.86
RMSE: 4.456422803668449
Coeff of det (R^2):..... 0.905

Feature ranking:

1. remaining_lease_years (0.367498)
2. region_Central (0.122463)
3. lease_commence_date (0.100029)
4. floor_area_sqm (0.078654)
5. n_rooms (0.042052)
6. region_North (0.029872)
7. region_West (0.021646)
8. storey_range_max (0.015075)
9. flat_type_3 Room (0.013889)
10. storey_range_min (0.013531)
11. storey_range_avg (0.013468)
12. town_Clementi (0.013035)
13. town_Woodlands (0.012928)
14. town_Jurong West (0.011618)
15. town_Marine Parade (0.011298)
16. flat_type_5 Room (0.009114)
17. flat_model_Standard (0.008182)
18. region_East (0.007804)
19. flat_model_Improved (0.007487)
20. flat_type_4 Room (0.007349)
21. town_Bukit Merah (0.006243)
22. town_Central Area (0.005663)
23. town_Ang Mo Kio (0.005363)
24. town_Geylang (0.005334)
25. flat_type_Executive (0.004637)
26. flat_model_New Generation (0.004471)
27. town_Bishan (0.004466)
28. flat_model_Terrace (0.004381)
29. region_North-East (0.004297)
30. town_Queenstown (0.004269)
31. flat_model_Model A (0.003833)
32. town_Toa Payoh (0.003448)
33. town_Tampines (0.002844)
34. town_Choia Chu Kang (0.002713)
35. town_Bukit Batok (0.002687)
36. town_Jurong East (0.002667)
37. town_Bedok (0.002540)
38. town_Yishun (0.002536)
39. town_Kallang/Whampoa (0.002473)
40. flat_type_2 Room (0.002389)
41. flat_model_Maisonette (0.002050)
42. town_Serangoon (0.001890)
43. flat_model_Simplified (0.001736)
44. town_Bukit Panjang (0.001708)
45. town_Hougang (0.001678)
46. town_Sengkang (0.001292)
47. town_Pasir Ris (0.001096)
48. flat_type_1 Room (0.000852)
49. flat_model_Apartment (0.000839)
50. flat_model_Premium Apartment (0.000698)
51. town_Bukit Timah (0.000548)
52. flat_model_Adjoined Flat (0.000445)
53. town_Sembawang (0.000424)
54. town_Punggol (0.000157)
55. flat_model_Model A2 (0.000111)
56. flat_model_Model A-Maisonette (0.000105)
57. town_Lim Chu Kang (0.000051)
58. flat_type_Multi Generation (0.000024)
59. flat_model_Multi Generation (0.000022)
60. flat_model_Improved-Maisonette (0.000018)
61. flat_model_Premium Maisonette (0.000008)
62. flat_model_2-Room (0.000005)
63. flat_model_Dbss (0.000000)
64. flat_model_Type S1 (0.000000)
65. flat_model_Type S2 (0.000000)
66. flat_model_Premium Apartment Loft (0.000000)



Prediction Error for RandomForestRegressor



Model: RandomForestRegressor (scikit-learn)
Split: 80/20 (chronologically - train/test)
Data: Base HDB data (no location features)

Hyperparameters:

Optimal Hyper-parameters:

```
{'bootstrap': True,  
 'ccp_alpha': 0.0,  
 'criterion': 'squared_error',  
 'max_depth': 22,  
 'max_features': 18,  
 'max_leaf_nodes': None,  
 'max_samples': None,  
 'min_impurity_decrease': 0.0,  
 'min_samples_leaf': 2,  
 'min_samples_split': 3,  
 'min_weight_fraction_leaf': 0.0,  
 'n_estimators': 500,  
 'n_jobs': -1,  
 'oob_score': False,  
 'random_state': 42,  
 'verbose': 0,  
 'warm_start': False}
```

Results:

--- Test Set ---

Mean Absolute Error: ... 6.499360358414414

Mean Squared Error:..... 83.40

RMSE: 9.132537759432319

Coeff of det (R^2):..... 0.773 (have been able to get as high as 0.815 with variable tweaking)

--- Train Set ---

Mean Absolute Error: ... 3.152783587570027

Mean Squared Error:..... 19.86

RMSE: 4.45642280366845

Coeff of det (R^2):..... 0.905

Feature ranking:

1. remaining_lease_years (0.367498)
2. region_Central (0.122463)
3. lease_commence_date (0.100029)
4. floor_area_sqm (0.078654)
5. n_rooms (0.042052)
6. region_North (0.029872)
7. region_West (0.021646)
8. storey_range_max (0.015075)
9. flat_type_3 Room (0.013889)
10. storey_range_min (0.013531)
11. storey_range_avg (0.013468)
12. town_Clementi (0.013035)
13. town_Woodlands (0.012928)
14. town_Jurong West (0.011618)
15. town_Marine Parade (0.011298)
16. flat_type_5 Room (0.009114)
17. flat_model_Standard (0.008182)
18. region_East (0.007804)
19. flat_model_Improved (0.007487)
20. flat_type_4 Room (0.007349)
21. town_Bukit Merah (0.006243)
22. town_Central Area (0.005663)
23. town_Ang Mo Kio (0.005363)
24. town_Geylang (0.005334)
25. flat_type_Executive (0.004637)
26. flat_model_New Generation (0.004471)
27. town_Bishan (0.004466)
28. flat_model_Terrace (0.004381)
29. region_North-East (0.004297)
30. town_Queenstown (0.004269)
31. flat_model_Model A (0.003833)
32. town_Toa Payoh (0.003448)
33. town_Tampines (0.002844)
34. town_Chua Chu Kang (0.002713)
35. town_Bukit Batok (0.002687)
36. town_Jurong East (0.002667)
37. town_Bedok (0.002540)
38. town_Yishun (0.002536)
39. town_Kallang/Whampoa (0.002473)
40. flat_type_2 Room (0.002389)
41. flat_model_Maisonette (0.002050)
42. town_Serangoon (0.001890)
43. flat_model_Simplified (0.001736)
44. town_Bukit Panjang (0.001708)
45. town_Hougang (0.001678)
46. town_Sengkang (0.001292)
47. town_Pasir Ris (0.001096)
48. flat_type_1 Room (0.000852)
49. flat_model_Apartment (0.000839)
50. flat_model_Premium Apartment (0.000698)
51. town_Bukit Timah (0.000548)
52. flat_model_Adjoined Flat (0.000445)
53. town_Sembawang (0.000424)
54. town_Punggol (0.000157)
55. flat_model_Model A2 (0.000111)
56. flat_model_Model A-Maisonette (0.000105)
57. town_Lim Chu Kang (0.000051)
58. flat_type_Multi Generation (0.000024)
59. flat_model_Multi Generation (0.000022)
60. flat_model_Improved-Maisonette (0.000018)
61. flat_model_Premium Maisonette (0.000008)
62. flat_model_2-Room (0.000005)
63. flat_model_Dbss (0.000000)
64. flat_model_Type S1 (0.000000)
65. flat_model_Type S2 (0.000000)
66. flat_model_Premium Apartment Loft (0.000000)

