PREDICTING HOUSE PRICE USING MACHINE LEARNING

## TEAM MEMBER 810021106038:KARTHIK.R

PHASE-1 Document Submission

Project : House Price Prediction



# Data Source :

## ~A good data source for house price prediction using a machine learning should be accurate,complete,covering the geographic area of intrests,accessible.

**Dataset Link: (**[**https://www.kaggle.com/datasets/vedavyasv/usa-housing**](https://www.kaggle.com/datasets/vedavyasv/usa-housing)**)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Avg. Area House Age** | **Avg. Area Number**  **of Rooms** | **Avg. Area Number of**  **Bedrooms** | **Avg. Area Income** | **Area Population** | **Price** | **Address** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 86754.2 | 6.60444 | 6.252455 | 4.02 | 43017.44 | 1662495 | 91863 Curtis Point  New Richard, AK 99996- 7554 |
| 74399.84 | 6.382453 | 7.252665 | 6.36 | 41084.66 | 1417820 | 03819 Lee Junction Suite 046  Mooneyborough, WA 19656 |
| 49408.2 | 5.82592 | 5.831739 | 3.32 | 26881.13 | 549976.1 | 7796 Joseph Burg  Danielsside, RI 70370 |
| 62279.79 | 6.10845 | 6.306116 | 4.12 | 35263.65 | 1054771 | 173 Mendoza Land  West John, NJ 57690 |
| 73078.2 | 5.923906 | 6.445457 | 3.32 | 54915.96 | 1415648 | 3288 Lee Pass  South Julieton, PW 02759-  4964 |
| 72942.71 | 4.786222 | 7.319886 | 6.41 | 24377.91 | 948788.3 | 86908 Marshall Port Suite  252  Scotttown, NM 69143 |
| 63819.62 | 5.949839 | 8.022469 | 4.09 | 27825.57 | 1159597 | 951 Bryant Dale Suite 735  Lake Jacqueline, NH 95266 |
| 73265.45 | 8.314762 | 7.425597 | 3.24 | 21030.97 | 1547133 | 00301 Bradshaw Avenue  Suite 833  Suzannefurt, IN 00654-8754 |
| 68488.13 | 6.116112 | 7.182527 | 5.08 | 18267.95 | 1186689 | 84473 Ochoa Pines Apt. 808  Jamesport, KY 70207-2955 |
| 55193.86 | 7.186121 | 5.096917 | 4.01 | 32537.82 | 772112 | 13706 Morgan Turnpike  Suite 378  Hohaven, NC 42699 |
| 77434.69 | 6.309271 | 5.219754 | 3.06 | 36252.34 | 1172730 | 4872 Delgado Ramp  North James, IA 36544 |
| 66158.88 | 4.476429 | 6.911743 | 2.28 | 37098.74 | 1111085 | 73202 Christopher Tunnel  New Roberttown, MN  08925-1536 |
| 60502.91 | 7.533381 | 5.731824 | 4.23 | 33579.63 | 1022781 | 8781 Olivia Port Apt. 225  Thompsonside, VT 22884 |
| 60910.89 | 5.635467 | 7.325974 | 4.2 | 43347.8 | 1274475 | PSC 9354, Box 0703  APO AA 58696-1278 |
| 73931.98 | 6.394108 | 4.58084 | 4.32 | 36543.07 | 1213531 | 271 Johnson Hills Apt. 001  Mcculloughfurt, FM 70847 |
| 59539.95 | 6.01859 | 7.007676 | 5.43 | 58600.83 | 1411730 | 122 Russo Neck  South Kevin, KY 20712-5282 |
| 56547.51 | 5.435415 | 6.51515 | 3.13 | 37585.27 | 858685.6 | 85225 Christopher Inlet Apt.  627  West Calebberg, KS 76460 |
| 65950.35 | 5.476513 | 6.717844 | 3.28 | 40110.85 | 1200962 | 7125 Mullins Cliff  Maryborough, WY 66971 |
| 74533.16 | 6.679353 | 5.919231 | 4 | 49481.57 | 1520234 | 905 Lane Pines Suite 348  Brownborough, DE 57196-  3319 |
| 66422.92 | 7.122072 | 7.078584 | 6.36 | 31019.32 | 1360908 | 61090 Griffith Ridges  Craigstad, AK 28807-8602 |
| 74334.49 | 5.419013 | 6.261535 | 4.12 | 41640.44 | 1360921 | 270 Jennifer Loop Suite 343  South Jesse, ND 60570-1483 |
| 63538.35 | 4.764499 | 7.168662 | 3.27 | 43282.18 | 1146532 | 10973 Clark Trafficway  Meyersbury, LA 63237 |
| 83953.1 | 7.385135 | 6.898847 | 3.08 | 37283.36 | 1789099 | 91733 Baker Orchard  South Kyle, KS 37301 |
| 57279.06 | 5.118109 | 6.56152 | 3.17 | 44023.79 | 852099.5 | 024 William Course  East Charlotteview, DE  69073-3104 |
| 56553.55 | 5.691129 | 7.021252 | 3.15 | 29682.41 | 746096.7 | PSC 4455, Box 3268  APO AP 45454 |
| 75795.58 | 5.78678 | 7.327325 | 6.25 | 33197.78 | 1534480 | 562 Brown Junction Suite  282  Christopherborough, KS  20719 |
| 70848.79 | 5.282325 | 6.766445 | 3.27 | 40458.74 | 1215609 | USCGC Thompson  FPO AA 13237-3887 |

# Problem Definition:

~The goal is to create a machine learning model that accurately predicts house prices based on various features. This model should assist potential homebuyers and sellers in making informed decisions about property transactions.

# Design Thinking:

## ~By clearly defining the problem and considering the needs of various stakeholders, you set the foundation for developing a robust and effective machine learning solution for house price prediction.

1. EMPATHIZE:

* Understand the user's needs. In this case, it's potential homebuyers or sellers who want accurate price predictions.
* Consider factors like location, size, amenities, and market trends.

1. DEFINE:
   * Clearly define the problem. For example, "How might we provide accurate and user-friendly house price predictions?"
   * List the key features users would want to know when predicting house prices.
2. IDEATE:
   * Brainstorm potential features that could contribute to accurate predictions (e.g., location, number of bedrooms, square footage, recent sales data).
   * Think about how to present this information in an understandable and user-friendly way.
3. PROTOTYPE:
   * Develop a simple machine learning model using a dataset of house prices and relevant features.
   * Create a basic interface to input features (e.g., a form for users to input details about a house)
4. TEST:
   * Gather feedback on the prototype. What do users like? What's confusing?
   * Test the accuracy of predictions against real market data.
5. ITERATE:
   * Based on feedback, refine the model and the user interface.
   * Consider adding more features or tweaking existing ones to improve accuracy.
6. IMPLEMENT:
   * Develop the final product incorporating all the feedback and improvements.
   * Ensure it's user-friendly and accessible.
7. LAUNCH:
   * Release the tool to users, perhaps as a web application.
   * Monitor user interactions and gather additional feedback for future updates.
8. EVALUATE:
   * Continuously assess the model's accuracy as it interacts with real-world data.
   * Collect user reviews and make necessary adjustments.
9. EVOLVE:

* Stay updated on real estate trends and adjust the model as needed.
* Consider expanding the tool to include additional features or cater to different markets.

