



Project 3

— Zillow National Home Prices —



Project Scope

An analysis of pre-pandemic national housing price and days on market trends



Source: Zillow House Price Data from [Kaggle*](#)

1

Days on Market by State

- 1 CSV file, 53 records
- Collects how long listings remain on Zillow by state
- Gathers monthly from Jan 2010 - Feb 2020

2

Housing Price by Number of Bedrooms

- 3 CSV files: 1 bedroom, 14927 records; 2 bedroom 23396 records, 3 bedroom 25735 records
- Collects Zillow Home Value Index, which measures typical home value by city
- Gathers monthly from Jan 1996 - Mar 2020

*<https://www.kaggle.com/datasets/paultimothymooney/zillow-house-price-data>

Cleaning CSV Files

Sample code

- 1 Start with 4 CSV files
 - Days on market
 - 1, 2, and 3 bedroom housing price
- 2 Strip down and reformat columns to only gather data from the most recent 2 years
- 3 Add number of bedrooms column to each housing price dataframe
- 4 Append housing price datasets together
- 5 End with 2 CSV files
 - Days on market
 - Housing prices

```
# Strip down and reformat "Housing Prices - 1BD" df
housing_cols = ["RegionID", "RegionName", "State", "CountyName", "2018-03-31", "2018-04-30", "2018-05-31", "2018-06-30",
                "2018-07-31", "2018-08-31", "2018-09-30", "2018-10-31", "2018-11-30", "2018-12-31",
                "2019-01-31", "2019-02-28", "2019-03-31", "2019-04-30", "2019-05-31", "2019-06-30",
                "2019-07-31", "2019-08-31", "2019-09-30", "2019-10-31", "2019-11-30", "2019-12-31",
                "2020-01-31", "2020-02-29", "2020-03-31"]

bed_1_df_a = bed_1_df[housing_cols].copy()

# Rename columns to avoid confusion between Region IDs and Regions Names in datasets (RegionID and RegionName does not
bed_1_df_a.rename(columns = {"RegionID": "CityID", "RegionName": "CityName", "2018-03-31": "Mar2018", "2018-04-30": "Apr2018",
                             "2018-07-31": "Jul2018", "2018-08-31": "Aug2018", "2018-09-30": "Sep2018", "2018-10-31": "Oct2018", "2018-11-30": "Nov2018", "2018-12-31": "Dec2018",
                             "2019-01-31": "Jan2019", "2019-02-28": "Feb2019", "2019-03-31": "Mar2019", "2019-04-30": "Apr2019", "2019-05-31": "May2019", "2019-06-30": "Jun2019",
                             "2019-07-31": "Jul2019", "2019-08-31": "Aug2019", "2019-09-30": "Sep2019", "2019-10-31": "Oct2019", "2019-11-30": "Nov2019", "2019-12-31": "Dec2019",
                             "2020-01-31": "Jan2020", "2020-02-29": "Feb2020", "2020-03-31": "Mar2020"}, inplace = True)

# Restate number of days as integers
convert_dict = {"Mar2018": int, "Apr2018": int, "May2018": int, "Jun2018": int, "Jul2018": int, "Aug2018": int, "Sep2018": int, "Oct2018": int, "Nov2018": int, "Dec2018": int,
                "Jan2019": int, "Feb2019": int, "Mar2019": int, "Apr2019": int, "May2019": int, "Jun2019": int, "Jul2019": int, "Aug2019": int, "Sep2019": int, "Oct2019": int, "Nov2019": int, "Dec2019": int,
                "Jan2020": int, "Feb2020": int, "Mar2020": int}

bed_1_df_a = bed_1_df_a.astype(convert_dict)
# Add column for number of bedrooms
bed_1_df_a.insert(4, "NoOfBeds", 1)
# Preview dataframe
bed_1_df_a.head(2)
```

	CityID	CityName	State	CountyName	NoOfBeds	Mar2018	Apr2018	May2018	Jun2018	Jul2018	...	Jun2019	Jul2019	Aug2019	Sep2019
0	6181	New York	NY	Queens County	1	611603	613536	614251	614393	614446	...	606162	604457	603126	601366
1	12447	Los Angeles	CA	Los Angeles County	1	517180	522014	524605	526132	528795	...	531014	530916	530814	531586

Screenshot examples

C:\Users\School\Desktop> cd hskakernum > Project > NU-Project > Output > Housing Prices.csv

```

1 Index, City, State, County, State, County, HousPrice, Mar2018, Apr2018, May2018, Jun2018, Jul2018, Aug2018, Sep2018, Oct2018, Nov2018, Dec2018, Jan2019, Feb2019, Mar2019, Apr2019, May2019,
2 0.6181, New York, NY, Queens County, 1, 611663, 616336, 614251, 613403, 614446, 616565, 615024, 617478, 614794, 616408, 613446, 611683, 608969, 608813, 607523, 606162, 604947, 603312, 603615, 603129,
3 1.12447, Los Angeles, CA, Los Angeles County, 1, 151180, 152014, 152460, 152632, 152785, 153580, 153360, 153577, 153496, 153595, 153472, 1533170, 153408, 1533154, 153204, 153109, 153104, 1530916, 153086,
4 2.39051, Houston, TX, Harris County, 1, 199702, 100286, 100296, 100290, 101623, 101023, 101055, 102767, 103192, 103417, 104040, 104056, 104371, 104101, 104071, 104015, 103910, 103552, 103502, 103651,
5 1.17426, Chicago, IL, Cook County, 1, 240319, 239624, 239434, 239351, 239329, 238894, 239365, 239568, 240270, 240033, 239992, 238891, 238372, 238481, 237841, 238429, 239106, 239230, 239086, 2379,
6 0.6083, San Antonio, TX, Bexar County, 1, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549, 81549,
7 2.7217, Philadelphia, PA, Philadelphia County, 1, 282872, 284384, 284384, 284384, 284384, 285591, 286838, 287836, 289364, 289364, 291654, 291753, 291470, 290485, 290485, 290485, 288843, 287,
8 0.40326, Phoenix, AZ, Maricopa County, 1, 145970, 146376, 147964, 147044, 150348, 151453, 152673, 153562, 154543, 152991, 152570, 153676, 153748, 152855, 152872, 153569, 150168, 151651, 162170, 161270,
9 7.18959, Las Vegas, NV, Clark County, 1, 1144691, 116372, 117787, 119661, 121599, 123807, 125506, 126824, 127270, 128041, 128272, 129048, 129063, 129994, 130193, 130108, 129712, 129232, 128989,
10 8.54296, San Diego, CA, San Diego County, 1, 342602, 342910, 342596, 341310, 344117, 345716, 346535, 347500, 348138, 348455, 346691, 347549, 345212, 346796, 347474, 348972, 349970, 350783, 35,
11 9.38128, Dallas, TX, Dallas County, 1, 117255, 118321, 119434, 124586, 123116, 124510, 121983, 121634, 122679, 121825, 12322, 124713, 125439, 126166, 126448, 127201, 128248, 127201, 128248, 127201, 128248,
12 10.18319, Santa Cruz, CA, Santa Cruz County, 1, 186559, 189347, 603216, 613784, 626424, 608810, 626424, 61040, 62640, 62641, 61801, 61801, 608810, 608810, 608810, 608810, 608810, 608810, 608810, 608810,
13 12.25290, Jacksonville, FL, Duval County, 1, 94889, 97388, 99311, 101713, 101385, 102134, 103383, 104117, 105010, 105582, 105671, 105742, 105919, 106264, 106036, 106638, 106638, 107145, 10740,
14 13.32149, Indianapolis, IN, Marion County, 1, 63326, 64575, 60216, 60767, 68942, 70044, 70751, 7346, 72127, 7435, 74834, 75589, 76288, 7681, 76558, 7649, 76480, 76695, 7377, 76810, 76882,
15 14.20330, San Francisco, CA, San Francisco County, 1, 89618, 89618, 910234, 914363, 914361, 926136, 927137, 925700, 928170, 928170, 928170, 928170, 928170, 928170, 928170, 928170, 928170, 928170, 928170, 928170,
16 15.28043, Charlotte, NC, Mecklenburg County, 1, 1595123, 159324, 159714, 160585, 162044, 163599, 165677, 16440, 165798, 165798, 16440, 167453, 167788, 166788, 166788, 166788, 166788, 166788, 166788, 166788,
17 16.14972, San Jose, CA, Santa Clara County, 1, 118886, 119461, 120851, 122219, 13372, 128518, 123526, 125457, 12640, 12671, 12798, 128252, 129311, 129747, 129810,
18 17.7481, Tucson, AZ, Pima County, 1, 118886, 119461, 120851, 122219, 13372, 128518, 123526, 125457, 12640, 12671, 12798, 128252, 129311, 129747, 129810,
19 18.10920, Columbus, OH, Franklin County, 1, 91910, 92201, 94517, 95040, 96279, 97194, 97844, 98317, 98719, 99197, 99197, 99197, 99197, 99197, 99197, 99197, 99197, 99197, 99197, 99197,
20 19.12455, Louisville, KY, Jefferson County, 1, 97444, 97781, 97817, 97836, 98208, 98888, 99224, 99633, 99430, 99545, 99590, 100174, 100751, 100852, 102450, 1,
21 20.13121, Orlando, FL, Orange County, 1, 94546, 96118, 97392, 98210, 99077, 99872, 100863, 100971, 103101, 104580, 105783, 106357, 107054, 106947, 107355, 10,
22 21.17933, El Paso, TX, El Paso County, 1, 82408, 81933, 81938, 82547, 83193, 83922, 84583, 85207, 85368, 85912, 86321, 86986, 87044, 88153, 88747, 89064, 8941,
23 22.11692, Detroit, MI, Wayne County, 1, 118886, 119461, 120851, 122219, 13372, 128518, 123526, 125457, 12640, 12671, 12798, 128252, 129311, 129747, 129810,
24 23.11893, Denver, CO, Denver County, 1, 273616, 274328, 274321, 280518, 281525, 288258, 288315, 288258, 288315, 288258, 288315, 288258, 288315, 288258, 288315, 288258, 288315, 288258, 288315, 288258, 288315,
25 24.16037, Seattle, WA, King County, 1, 590438, 59328, 518803, 520091, 518428, 51773, 5192
```

Days_On_Market.csv X

C:\Users\J School\Desktop> mkdir hmkaramam > Projects > NU-Project-1 > Output > Days_On_Market.csv

```

1 StateID,StateName,Feb2018,Mar2018,Apr2018,May2018,Jun2018,Jul2018,Aug2018,Sep2018,Oct2018,Nov2018,Dec2018,Jan2019,Feb2019,Mar2019,Apr2019,May2019,Jun2019,Jul2019,Aug2019
2 45,Texas,82,64,59,67,60,61,63,69,75,79,83,88,88,74,64,64,64,74,79,82,92,91
3 43,Nevada,147,148,112,106,108,112,118,122,127,128,140,146,156,145,119,112,111,113,120,125,130,132,140,149
4 41,Florida,92,92,92,93,95,96,96,96,96,96,96,96,96,96,96,96,96,96,96,96,96,96,96,96
5 21,Illinois,126,109,77,77,78,79,80,85,91,95,104,116,125,120,82,74,76,78,81,86,92,95,100,116,125
6 47,Pennsylvania,113,110,86,78,74,74,77,81,84,87,89,99,107,105,83,74,72,74,76,79,82,80,86,98,99
7 44,Ohio,94,80,65,62,60,61,63,67,71,73,78,89,86,82,66,60,59,61,63,65,68,70,75,86,86
8 30,Michigan,96,89,68,61,60,61,64,68,72,76,78,81,91,97,91,73,62,61,63,65,71,76,78,82,94,95
9 16,Georgia,104,74,65,62,63,62,63,70,74,84,90,90,69,64,64,65,67,70,73,78,80,83,90,93,92
10 35,North Carolina,71,61,63,63,64,64,67,80,84,83,84,93,83,83,84,93,63,63,65,68,71,74,79,81,89
11 40,North Jersey,137,120,103,100,100,100,104,114,117,117,124,132,130,113,103,100,104,106,114,118,118,115,125,126
12 56,Virginia,100,73,61,62,65,66,72,76,79,83,94,95,70,60,58,61,63,65,68,69,73,75,85,79
13 59,Washington,62,48,43,42,45,48,52,56,61,66,75,84,80,59,47,47,47,51,55,57,60,63,69,76,79
14 26,Massachusetts,97,81,65,65,66,70,74,73,73,73,93,105,84,70,70,70,66,69,72,72,72,72,80,95,105
15 22,Indiana,89,77,66,62,58,59,61,63,64,67,74,81,88,80,66,59,58,59,60,63,65,67,72,81,84
16 9,Arizona,74,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62,62
17 51,Tennessee,92,76,85,62,62,63,64,66,71,76,78,78,78,78,78,78,78,78,78,78,78,78,78,78
18 32,Missouri,88,70,59,58,57,67,60,64,69,71,76,90,93,79,62,58,57,60,60,63,65,66,71,81,84
19 27,Maryland,111,107,83,78,77,81,86,90,90,91,96,106,111,105,83,73,75,81,85,91,92,91,94,104,102
20 16,Miscosin,106,87,70,66,66,63,64,66,71,74,79,94,101,84,71,65,63,63,64,67,78,78,92,97
21 31,Minnesota,91,66,61,61,61,61,61,67,73,77,99,91,74,62,68,59,63,63,66,69,72,78,91,89
22 10,Colorado,54,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46,46
23 4,Alabama,116,100,77,73,76,76,84,84,90,95,105,102,76,72,71,72,79,80,82,88,100,94
24 51,South Carolina,96,87,73,71,72,74,76,79,83,85,89,97,98,87,75,72,74,75,77,80,84,88,98,99
25 25,Louisiana,108,101,85,82,82,82,82,92,99,96,105,109,117,107,91,88,83,84,86,91,99,96,102,109,111
26 24,Kentucky,93,77,65,63,61,61,64,67,70,76,84,90,88,76,64,64,64,64,64,64,64,64,64,64
27 46,Oregon,81,60,63,58,58,60,63,63,63,63,63,63,63,63,63,63,63,63,63,63,63,63,63,63
28 40,Oklahoma,109,82,75,73,75,78,83,91,91,91,91,91,91,91,91,91,91,91,91,91,91,91,91,91
29 11,Connecticut,122,129,87,78,80,85,88,95,97,102,104,110,122,125,97,87,81,82,88,94,97,101,103,118,121
30 19,Iowa,182,94,73,69,69,68,70,74,76,78,81,84,86,106,108,81,78,69,69,73,74,77,86,99,107
31 34,Mississippi,102,90,83,79,74,76,76,81,89,89,93,109,105,93,80,76,75,73,80,85,82,90,98,94
32 6,Arkansas,102,100,79,71,71,74,78,81,85,89,98,101,108,89,84,77,75,71,74,77,80,83,91,98
33 2,Kansas,85,68,62,75,75,75,75,75,75,75,75,75,75,75,75,75,75,75,75,75,75,75,75
34 53,Utah,60,47,45,44,45,47,50,49,56,58,63,72,71,59,62,49,51,54,56,58,59,66,75,83
35 42,Nevada,61,50,53,53,53,54,56,58,61,65,71,82,85,79,69,68,64,68,72,75,76,77,83,89,85

```

Days On Market.csv

Housing Prices.csv


1

PostgreSQL Database

Sample code

- 1 Initiate database and drop existing tables with desired table names
- 2 Create two tables for days_on_market CSV file and housing_prices CSV file
- 3 Import relevant datasets with headers
- 4 Perform SELECT ALL query to ensure datasets imported correctly

Troubleshooting: needed to rename all date columns, removing hyphen from column names, in order for Flask query to work.



```
National_Housing_Database/postgres@PostgreSQL 14
Query Editor  Query History
1  -- Delete pre-existing tables
2  DROP TABLE IF EXISTS days_on_market;
3  DROP TABLE IF EXISTS housing_prices;
4
5  -- Create new tables:
6  CREATE TABLE days_on_market (
7      StateID INT NOT NULL PRIMARY KEY,
8      StateName VARCHAR(50) NOT NULL,
9      "Feb2018" INT,
10     "Mar2018" INT,
11     "Apr2018" INT,
12     "May2018" INT,
13     "Jun2018" INT,
14     "Jul2018" INT,
15     "Aug2018" INT,
16     "Sep2018" INT,
17     "Oct2018" INT,
18     "Nov2018" INT,
19     "Dec2018" INT,
20     "Jan2019" INT,
21     "Feb2019" INT,
22     "Mar2019" INT,
23     "Apr2019" INT,
24     "May2019" INT,
25     "Jun2019" INT,
26     "Jul2019" INT,
27     "Aug2019" INT,
28     "Sep2019" INT,
29     "Oct2019" INT,
30     "Nov2019" INT,
31     "Dec2019" INT,
32     "Jan2020" INT,
33     "Feb2020" INT
34 );
```

PostgreSQL Database

Screenshot example

1 Create “National Housing Database”



2 Create “Housing Prices” table and import CSV file with headers

The screenshot shows the PostgreSQL Query Editor with the following query:

```
1 SELECT * FROM housing_prices;
```

Below the query editor, the 'Data Output' tab is selected, displaying the results of the query. The table has 10 columns: Index [PK] integer, cityid integer, cityname character varying (50), State character varying (50), countyname character varying (50), noofbeds integer, Mar2018 integer, Apr2018 integer, May2018 integer, and an unlabeled integer column.

	Index [PK] integer	cityid integer	cityname character varying (50)	State character varying (50)	countyname character varying (50)	noofbeds integer	Mar2018 integer	Apr2018 integer	May2018 integer	
1	0	6181	New York	NY	Queens County	1	611603	613536	614251	
2	1	12447	Los Angeles	CA	Los Angeles County	1	517180	522014	524605	
3	2	39051	Houston	TX	Harris County	1	99702	100296	100960	

3 Create “Days on Market” table and import CSV file with headers

National_Housing_Database/postgres@PostgreSQL 14

Query Editor

Query History

1

SELECT * FROM days_on_market;

Data Output

Explain

Messages

Notifications

	stateid [PK] integer	statename character varying (50)	Feb2018 integer	Mar2018 integer	Apr2018 integer	May2018 integer	Jun2018 integer	Jul2018 integer	Aug2018 integer	Sep2018 integer	Oct2018 integer
1	9	California	62	50	50	50	52	55	56	60	
2	54	Texas	82	64	59	57	60	61	63	69	
3	43	New York	148	147	140	112	106	108	112	118	

Flask-Powered API

Sample code

1

Establish connection with PostgreSQL database and assign variables for each table

```
# Database setup and table references
engine = create_engine("postgres://postgres:admin@localhost/National_Housing_Database")
Base = automap_base()
Base.prepare(engine, reflect = True)
Base.classes.keys()
Days_On_Market = Base.classes.days_on_market
Housing_Prices = Base.classes.housing_prices
```

2

Setup Flask and list all available API routes on homepage

```
# Flask routes
@app.route("/")
def Homepage():
    """List all API routes."""
    return (
        f"<h1>Welcome to the National Housing Prices API</h1>"
        f"<h2>Find information about average housing price trends for cities across the United States fr"
        f"<strong><u>Available Routes:</u></strong><br>"
        f"1. <strong>Housing Prices for 1 Bedroom Homes:</strong> /api/v1.0/1bedroom<br>"
        f"2. <strong>Housing Prices for 2 Bedroom Homes:</strong> /api/v1.0/2bedroom<br>"
        f"3. <strong>Housing Prices for 3 Bedroom Homes:</strong> /api/v1.0/3bedroom<br>"
        f"4. <strong>Average Days on Market:</strong> /api/v1.0/days_on_market<br>"
        f"<br>"
        f"Note: API references <strong><i>Zillow House Price Data</i></strong> found on Kaggle: <i>https"
    )
```

3

Create API routes:

- Create session from Python to database
- Query table for desired subset of data
- Convert query results into a dictionary
- Return JSON representation of data

```
# APP ROUTE 1 - HOUSING PRICES FOR 1 BD
@app.route("/api/v1.0/1bedroom")
def one_bedroom():
    # Create our session from Python to the DB
    session = Session(engine)
    one_bedroom_result = session.query(Housing_Prices.cityname, Housing_Prices.State, Housing_Prices.noofb
    Housing_Prices.Jul2018, Housing_Prices.Aug2018, Housing_Prices.Sep2018, Housing_Prices.Oct2018, Housin
    Housing_Prices.Mar2019, Housing_Prices.Apr2019, Housing_Prices.May2019, Housing_Prices.Jun2019, Housin
    Housing_Prices.Nov2019, Housing_Prices.Dec2019, Housing_Prices.Jan2020, Housing_Prices.Feb2020, Housin
    filter(Housing_Prices.noofbeds == 1).\
    all()
    session.close()
    # Convert the query results to a dictionary using CityName as the key
    one_bedroom = []
    one_bedroom_dict = {}
    for cityname, State, noofbeds, Mar2018, Apr2018, May2018, Jun2018, Jul2018, Aug2018, Sep2018, Oct2018,
    one_bedroom_dict["CityName"] = cityname
```

Flask-Powered API

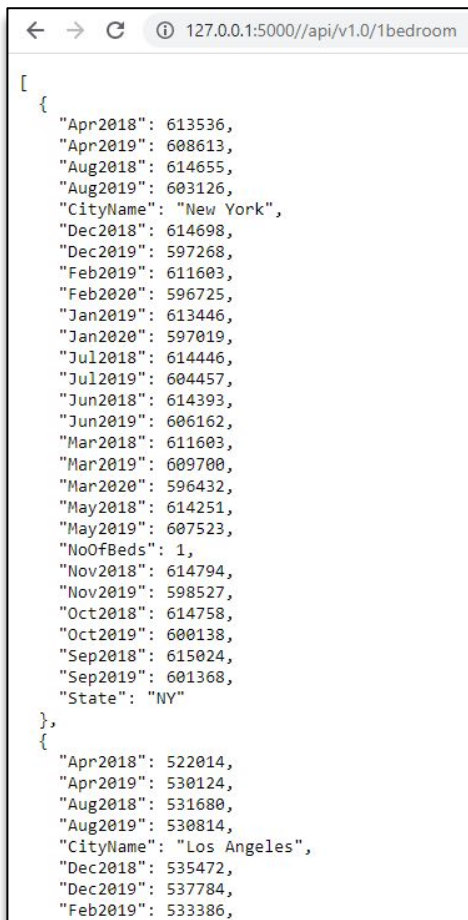
Screenshot example

1 Homepage listing four API Routes...



2

...JSONified dataset



Visualizations

Sample code

1

Create html framework and download the anime.js library

2

Set Canvas size and choose colors for the fireworks

3

Create circles for the firework effect and add an clicking option to the page

```
<script src="https://cdn.jsdelivr.net/npm/animejs@3.0.1/lib/anime.min.js"></script>
<canvas class="fireworks"></canvas>
<script src="fireworks.js"></script>
```

```
window.human = false;

var canvasEl = document.querySelector('.fireworks');
var ctx = canvasEl.getContext('2d');
var numberOfParticules = 30;
var pointerX = 0;
var pointerY = 0;
var tap = ('ontouchstart' in window || navigator.msMaxTouchPoints) ? 'touchstart' : 'mousedown';
var colors = ['#FF1461', '#FFA500', '#FBF38C'];

function setCanvasSize() {
  canvasEl.width = window.innerWidth * 2;
  canvasEl.height = window.innerHeight * 2;
  canvasEl.style.width = window.innerWidth + 'px';
  canvasEl.style.height = window.innerHeight + 'px';
  canvasEl.getContext('2d').scale(2, 2);
}
```

Visualizations

Screenshot example

