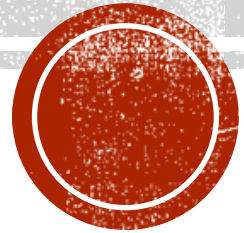
BAY-AREA

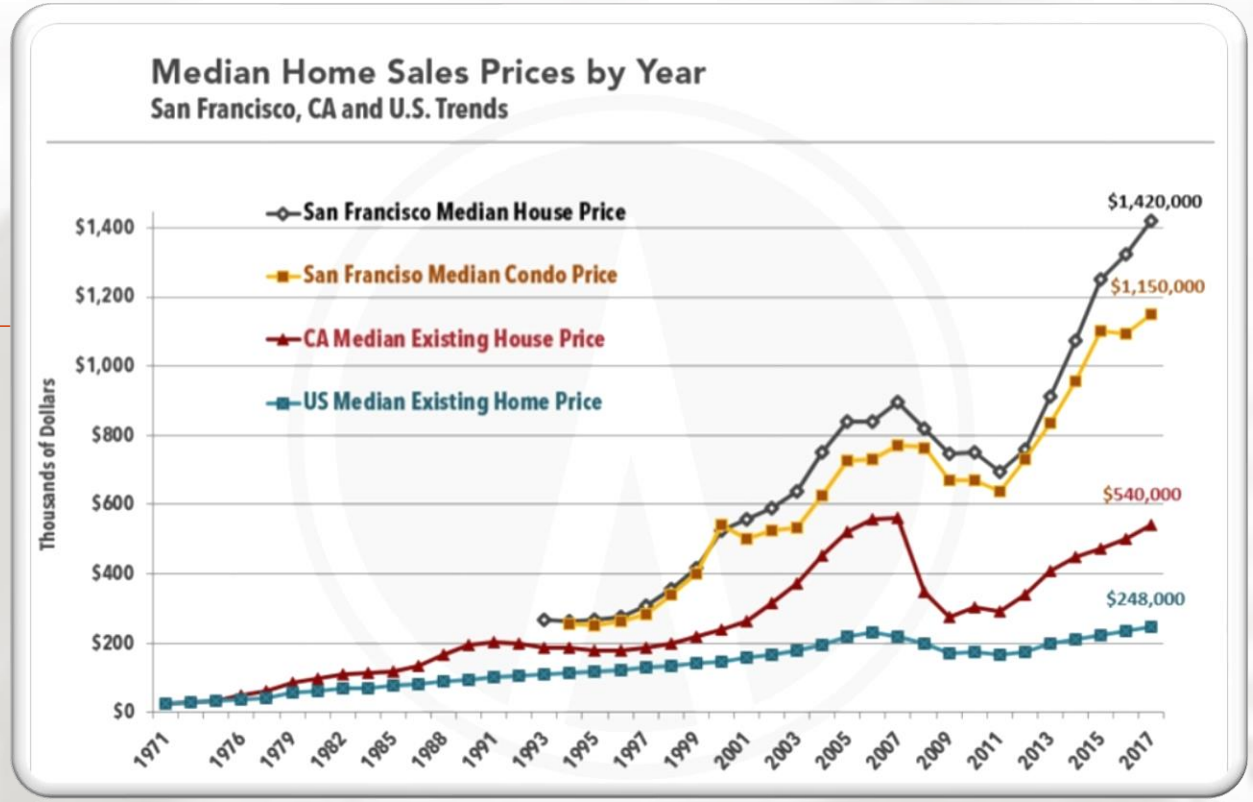
HOUSING PRICE PREDICTION

UC BERKELEY EXT. 2018 – FINAL PROJECT PRESENTATION

Indrani Kompella | Niyati Desai



WHY?



IT IS A WORN-DOWN, DECOMPOSING WOODEN SHACK THAT WAS BUILT IN 1906, AND THE INTERIOR IS UNLIVABLE IN ITS CURRENT CONDITION. THE SAN FRANCISCO HOUSE IS ALSO SELLING FOR **$350,000.00**

http://fortune.com/2015/09/25/san-francisco-cheapest-home/

COUNTIES

- ALAMEDA
- CONTRA COSTA
- MARIN
- NAPA
- SAN FRANCISCO
- SANTA CLARA
- SAN MATEO
- SOLANO
- SONOMA

DATA

THE DATA WAS COLLECTED FROM FOLLOWING SOURCES:

- CALIFORNIA ASSOCIATION OF REALTORS[CAR]
- ZILLOW - HOUSING ESTIMATES
- US CENSUS BUREAU
- BUREAU OF LABOR STATISTICS
- METROPOLITAN TRANSPORTATION COMMISSION (HTTP://WWW.VITALSIGNS.MTC.CA.GOV)

MACHINE LEARNING MODEL

HYPOTHESIS : THE REASON FOR ASTRONOMICAL RISE IN THE HOUSING PRICES IN THE BAY AREA IN THE PAST FEW YEARS, IS THE INCREASING GAP BETWEEN THE DEMAND (POPULATION AND WAGE GROWTH) AND SUPPLY (NEW HOUSING CONSTRUCTION).

MODEL: TIME SERIES ANALYSIS

STEPS :

- IDENTIFY DEPENDENT AND POTENTIAL PREDICTOR VARIABLES
- CHECK INITIAL CORRELATION BETWEEN DEPENDENT AND POTENTIAL PREDICTOR VARIABLES
- SHORT-LIST 3 MOST CORRELATED PREDICTOR VARIABLES
- RUN UNI-VARIATE LINEAR REGRESSION MODEL FOR EACH PREDICTOR VARIABLE
- RUN MULTI-VARIATE LINEAR REGRESSION MODEL FOR ALL 3 PREDICTOR VARIABLE
- USE THE COEFFICIENTS AND Y-INTERCEPT TO PREDICT FUTURE OUTCOMES

RESULT & CONCLUSION :

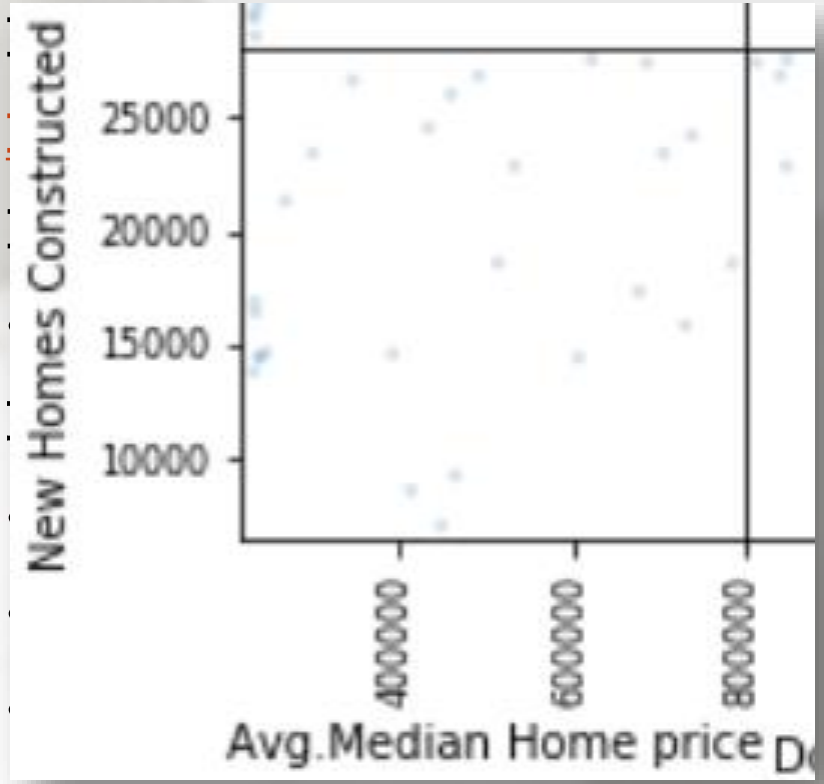
A PEEK INTO THE FUTURE

WHAT IF THE WAGES CONTINUE TO RISE AT THEIR CURRENT RATE AND THE MORTGAGE RATE (30 YEAR FIXED) HOLDS STEADY AT 4.6 %,

.....CAN YOU AFFORD TO BUY A MEDIAN PRICE HOME IN THE FUTURE???

MACHINE LEARNING MODEL

POTENTIAL PREDICTOR VARIABLES	BAY AREA LEVEL correlation	COUNTY LEVEL correlation
TOTAL POPULATION	0.799286	0.835920
HOUSEHOLD POPULATION	0.800972	-
HOUSEHOLDS	**0.814675**	**0.851533**
IMMIGRATION	0.177585	0.569903
NET MIGRATION	-0.180791	-0.137122
JOBS	**0.725683**	**0.824421**
INTEREST RATE	-0.683790	-0.711596
NEW HOMES CONSTRUCTED	0.291740	0.052245
EMPLOYERS	0.738533	0.783928
AVERAGE ANNUAL PAY	**0.829636**	**0.848847**



MACHINE LEARNING MODEL

AFTER RUNNING MACHINE LEARNING MODEL, OUR INITIAL HYPOTHESIS DID NOT GIVE EXPECTED RESULTS.

FINAL HYPOTHESIS - THE MEDIAN ANNUAL HOME PRICE IN BAY-AREA IS A FUNCTION OF AVERAGE ANNUAL PAY, NO. OF HOUSEHOLDS & JOBS.

$$y = m_o X_o + m_1 X_1 + m_2 X_2 + b$$

y = Median Annual Home Price

X_o = Average Annual Pay,

X_1 = No. of Households

X_2 = Jobs

m = Multipliers/Coefficients

b = Y-Intercept

THANKS TO

GITHUB

WORK-FLOW, PARTNER CO-ORDINATION & PROJECT COMPILATION

JUPYTER NOTEBOOK, SPYDER & VS CODE

DATA MUNGING, INITIAL MACHINE LEARNING MODEL & CODING

FLASK, HTML, CSS, JAVASCRIPT, BOOTSTRAP

BUILD THE INTERACTIVE PREDICTION APP WEBPAGE

TABLEAU

VISUALIZE PAST DATA AND ITS TRENDS & PLOT FOR USER INTERACTION ON THE WEBPAGE

PLOTLY & D3

INTERACTIVELY PLOT VISUALIZATION FOR FUTURE PREDICTIONS AS PER USER SELECTION ON THE WEBPAGE