

# Predicting Real Estate Market

Team 41: Guangpeng Liu, Jiawei Zhu, Manikanta Reddy, Matt Zhou

We are building a web application to predict the rising or falling trends of future housing prices via statistically significant determinants. This builds upon existing online “housing marketplaces” data and provides users with a better understanding of a neighborhood or county’s future valuation.

## Motivation

Housing market is very sensitive. Our app aids consumers looking to invest in Real Estate that can potentially provide them high yields in long term and assist them in property value price comparisons as per market using predictive models.

## How will someone use this?

You will be able to look into the future prices and compare across various neighborhoods.

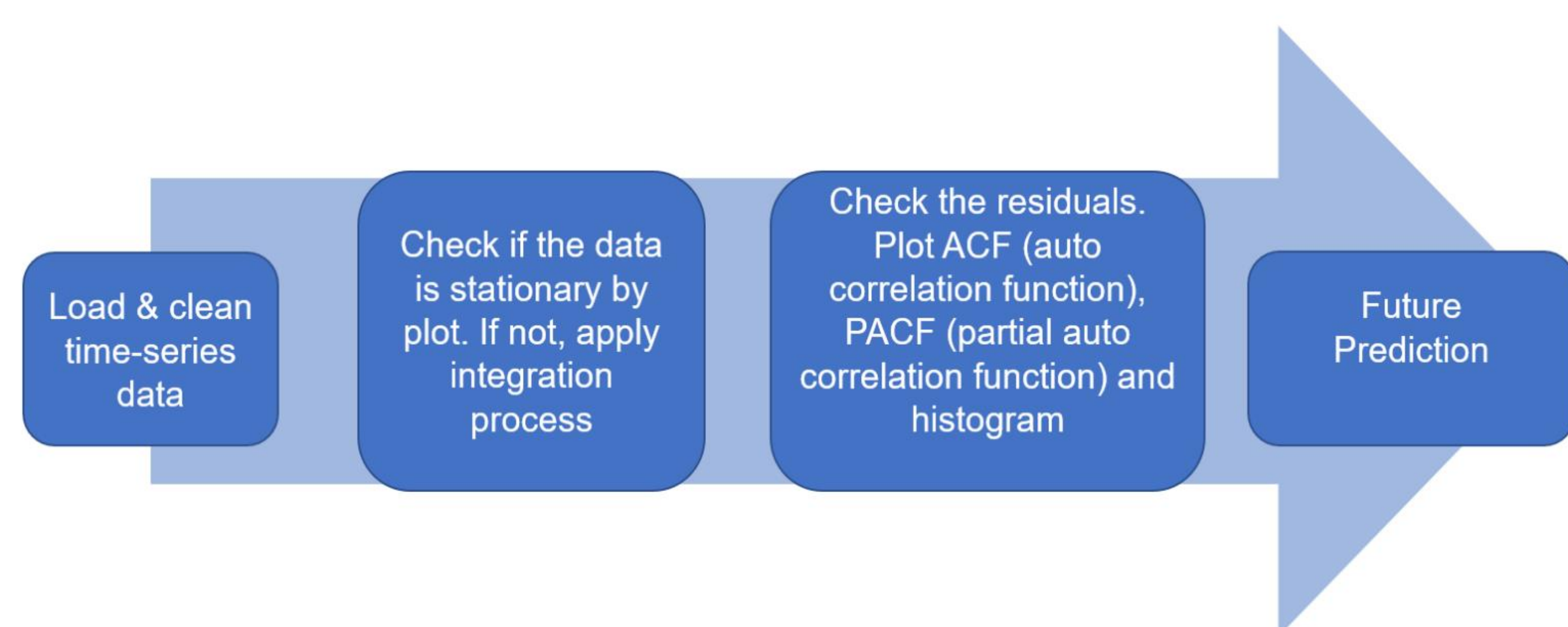
## How are we doing this?

Our algorithms model the past data and make confident estimates for future. Our visualizer helps you understand these price variations.

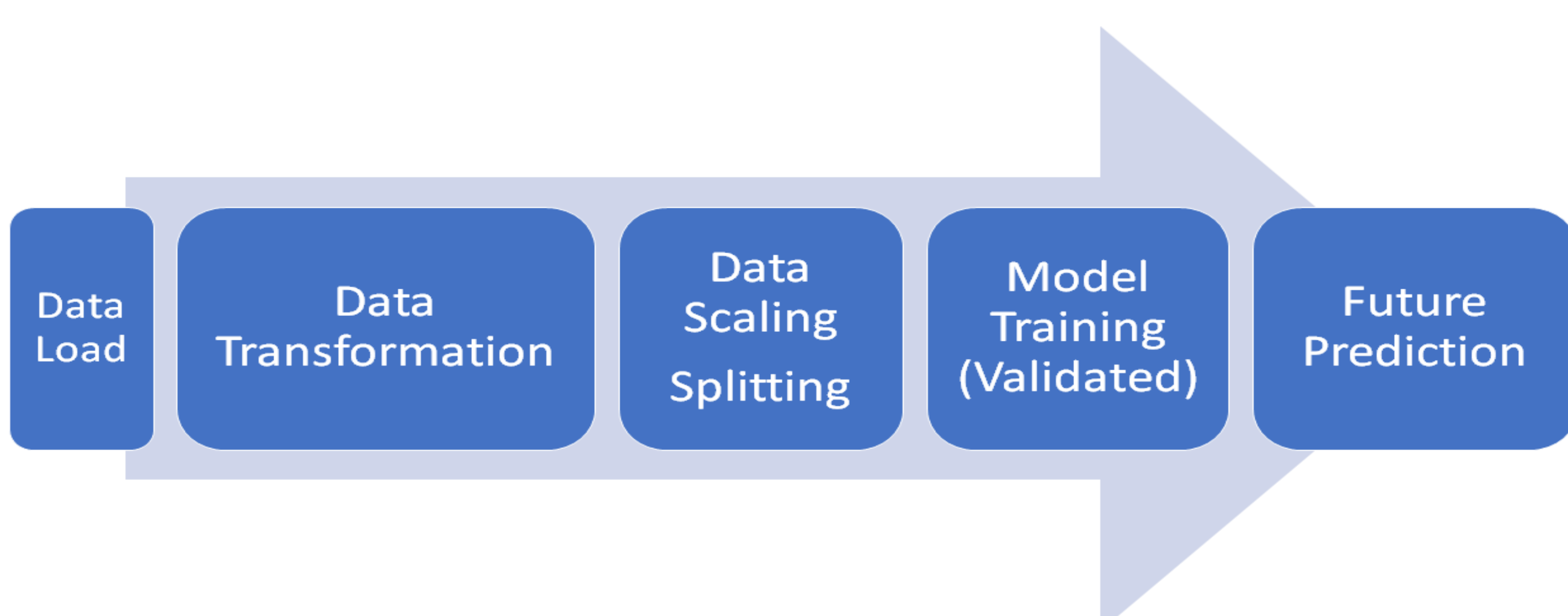
## Approaches

Two methods for the prediction are utilized including a linear regression model (autoregressive integrated moving average model, **ARIMA**) and non-Linear Long short-term memory (**LSTM**) network model.

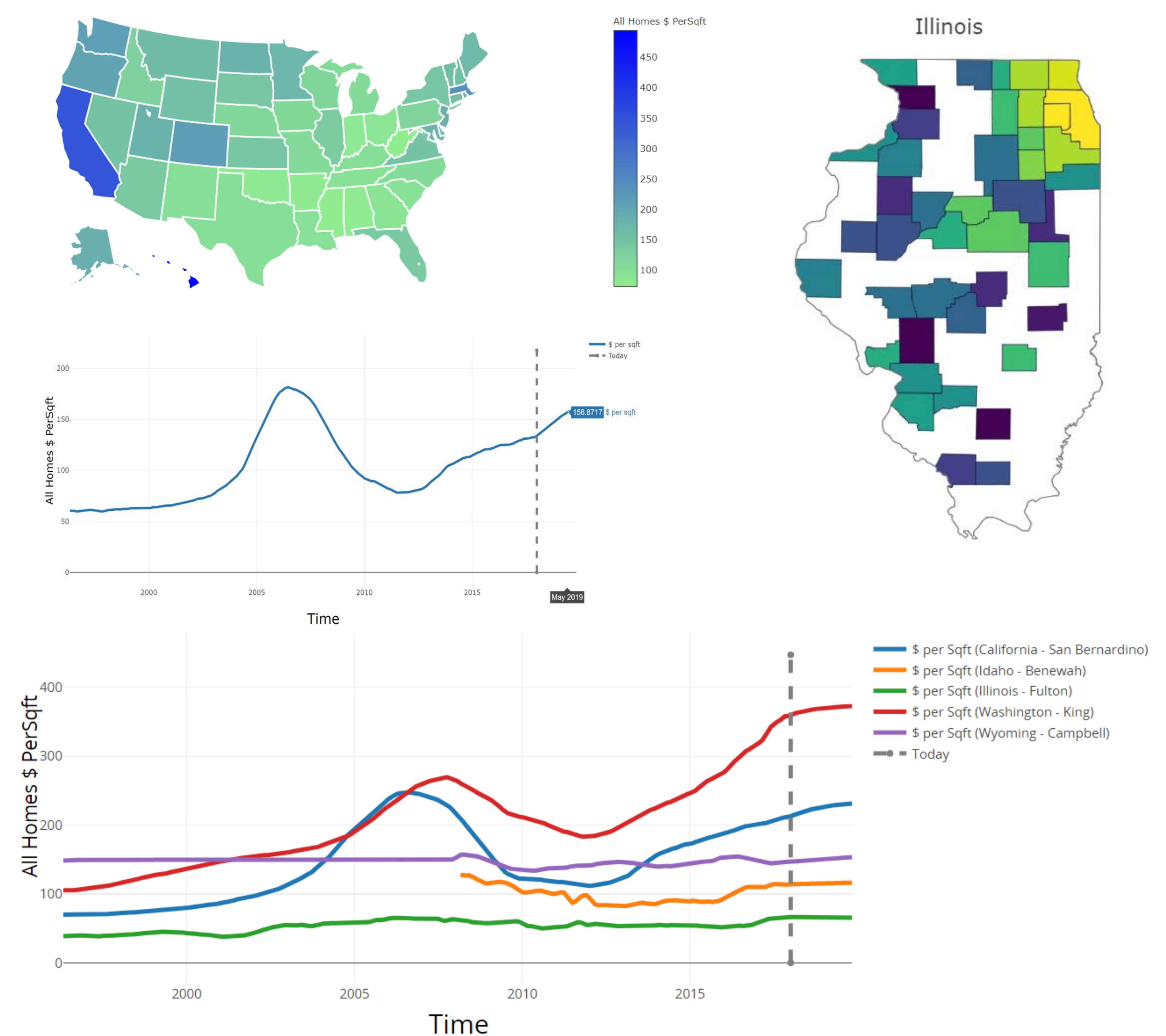
**ARIMA** model is a generalization of an autoregressive moving average model which is fitted to time series data either to better understand the data or to predict future points. Steps are as follows:



**LSTM** network model is a special kind of recurrent neural networks and is well-suited to classifying, processing and making predictions based on time series data. Steps are as follows:



## Interactive Visualization

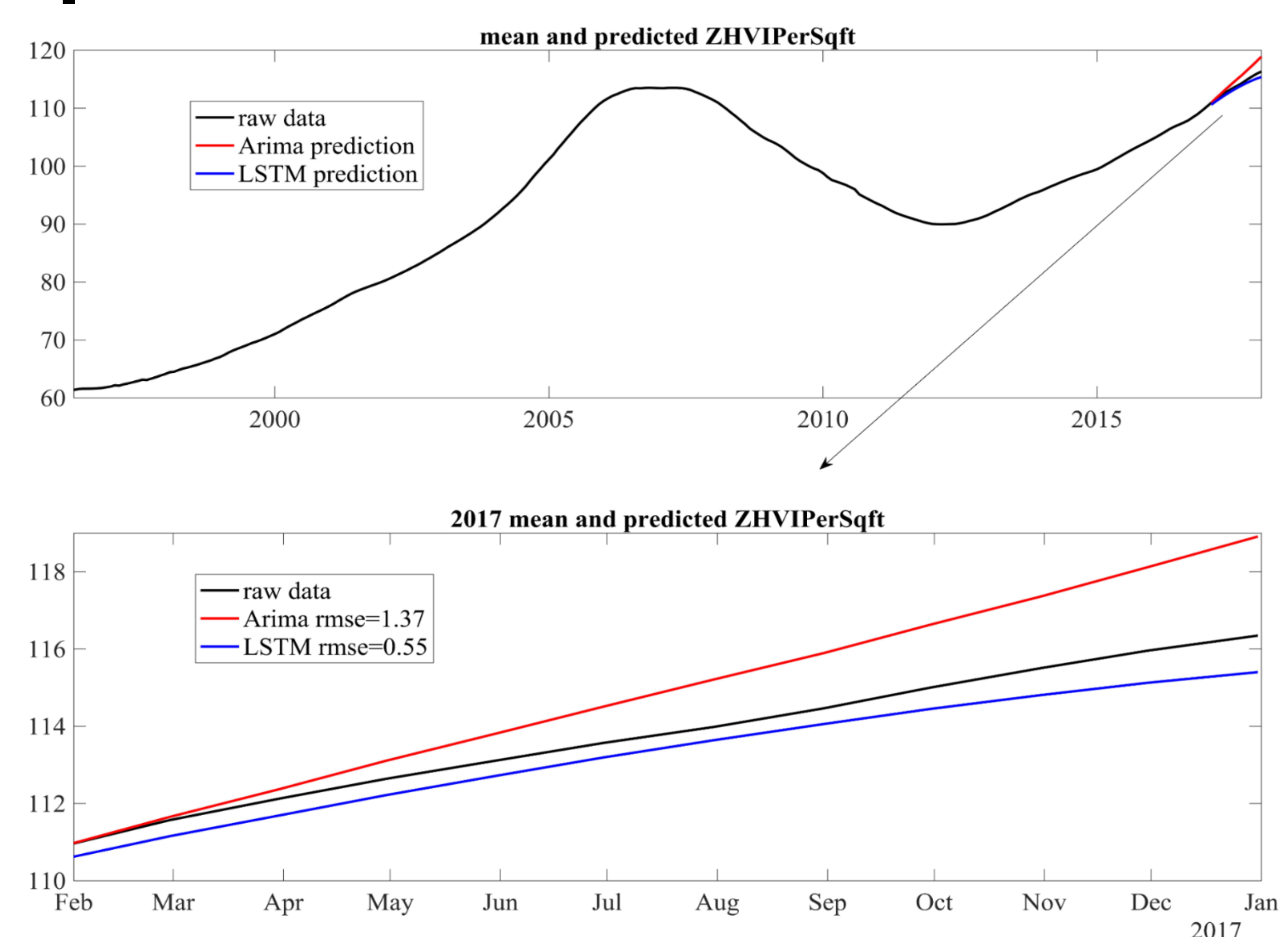


## Data

Zillow provides an API for real time polling of the latest pricing data. However the API clause prevents caching or storing of the data over long durations of time. Hence for our prediction and exploratory tasks we have decided to use the Zillow Economics Dataset already consolidated and licensed for distribution on Kaggle. Here are a few tidbits about the available data:

- Size: 511MB
- Records: 554776
- Available Features: 90

## Experiments and Evaluation



The above plot presents the difference between the mean housing values (values averaged over all counties) predicted by both ARIMA and LSTM and the existing housing data using **cross validation**.