MSBA Datathon

Defining and Modeling the Housing Affordability



# Background Information

Clayton Homes strives to open doors to a better life for more people by providing affordable housing options, both through manufactured and traditional site-built homes. However, meeting that affordability goal looks different by market. There are many factors, such as income level versus typical home prices, rent prices, price per square foot, etc. Clayton would like to know what affordability means across the United States so that they can be sure to design, build, and sell products to achieve their mission.

# Project Description

In this project, you are asked to utilize the available data, along with any additional data that you may find relevant, and construct the metric of *relative housing affordability*, and identify characteristics that are related to the metric. In addition, it’s of interest to construct predictive models for the metric for the future, for new or existing zip codes.

# Data Sources

There are four data files that include:

* Zillow monthly median home value since January 2000
* Zillow monthly median rent value since March 2015
* Household count and median income from census.gov
* Annual interest rate for 30-years mortgage loans since 1974

Link to download data: <https://drive.google.com/drive/folders/1FsSEYXAeh2Xr6oEFB4mFfKdLOORm-Sbf>

Link to R notebook to load the data:

<https://colab.research.google.com/drive/1PCs8ZJuVcgTC010ANLLOS1jRo_RHUTgo>

Link to Python notebook to load the data:

<https://colab.research.google.com/drive/1Z882xsSCXh4BGsousQ1beejuMUXMF706>

Link to R Markdown to load the data:

<https://drive.google.com/file/d/1MWbzfrFTUydM9YALBBfIo2a4dq4aWhy6>

# Submission

**Please upload your presentation slides for both the courses, due at 9:30am, Friday, Nov 18.**

# Presentation

Location and Schedule

* **Location: Student Union (SU) 262C**
* **Date: Friday, November 18**
* Coffee, tea, and water available from 8am to 2:30pm.
* **09:30: Check in**
* 10:00: Presentation starts
* 12:10: Presentation ends
* **12:15: Lunch (sandwich and salad options)**

Instruction

**Each team will have 10 minutes to present their innovative approach and findings.** Your presentation should focus on:

1. Problem statement.
2. Description of the data, and any data pre-processing steps.
3. Methods considered, and the best method(s).
4. Summary of (interesting) results.

In your presentation please use as many graphics as necessary that are clear. **The presentations should be delivered as though you are talking to a business audience with not a lot of mathematics, statistics, or data analytics knowledge.**

Grading Rubric

Please grade the student group on the following (5 being the best score).

**Group #:**

**Results:**

1. Student group obtained results that answer the research problem.

1 2 3 4 5

1. Student group examined their results for logical and data consistency.

1 2 3 4 5

**Methodology:**

1. Student group used appropriate methodologies/methodological approaches.

1 2 3 4 5

1. Student group discussed alternative methodologies, or assumptions underlying their methodologies.

1 2 3 4 5

**Presentation:**

1. Student group clearly stated the problem, methodology, results, and conclusions.

1 2 3 4 5

1. Student group used legible slides, and clear, appropriate, appealing visuals.

1 2 3 4 5