**A visualisation of suburban house prices in Cape Town (2011 – 2019)**

**Introduction:**

We’ve seen tremendous property price increases in South Africa over the past two decades, with many people in my generation feeling like it’s increasingly difficult to get onto the property ladder. However, with the continuing energy crisis, sluggish economy and political uncertainty in the country, I was intrigued to see how this might be impacting house prices.

I also wanted to find a way to create a visually impactful data science project that uses publicly available data that could have economic value to someone. So, I sat down and created a map of average house prices, coupled with the average *change* in house prices for most of the suburbs in Cape Town over the time period 2011 - 2019. The idea behind looking at the change in house prices is that one could potentially find an area that traditionally has seen strong growth in prices, but is currently undervalued; this would be seen through a historic growth in prices followed by a decline in prices over a recent period. I chose to focus on Cape Town because I wanted to use something from the City of Cape Town Open Data Portal [insert link] and also because I recently moved back to Cape Town from Joburg to pursue a Masters in Data Science at UCT.

**The process:**

When recently looking for a place to stay in Cape Town, I noticed that Property 24 had a graph on the main page for each suburb indicating suburb price trends. I created a web-scraper using BeautifulSoup in Python to scrape this data (see Github [link] for details of this web-scraper). It took a few attempts to ensure that I got the majority of suburbs in the City of Cape Town suburbs shapefile and it took a bit of searching to find the JavaScript-generated data points I was after, but the data was relatively easy to get.

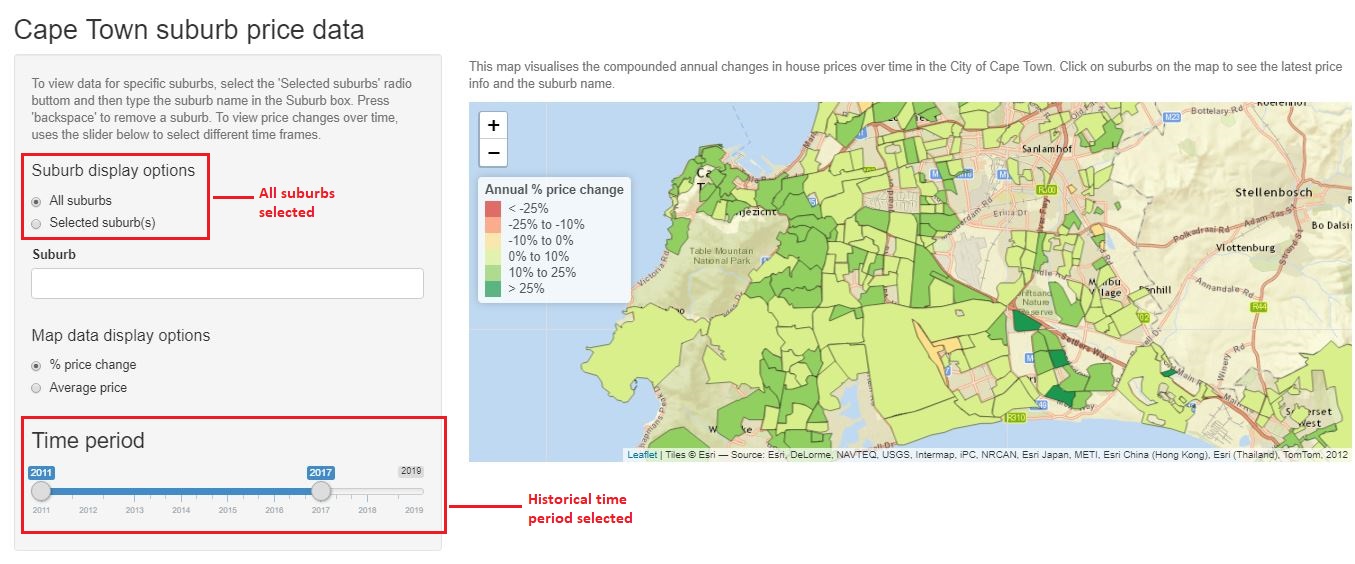
I also downloaded the shapefiles for the boundaries of suburbs in Cape Town from the City of Cape Town (CoCT) Open Data Portal which I would use to display the data.

From here, I moved over to R to take advantage of the ease of plotting spatial data with the SF and Leaflet packages and to make the data easy to navigate with a Shiny App. The scraped data was processed to remove NA values followed by calculating the annualized property price growth (or decline) per suburb for each pair of years (e.g. 2011 & 2012, 2011 & 2013… 2018 & 2019), similarly, the average house price was also calculated for these periods. The scraped data and CoCT data sets were then joined to be displayed in the Shiny App, with differences in the names of some of the suburbs adjusted on the CoCT data set to enable the joining of the datasets.

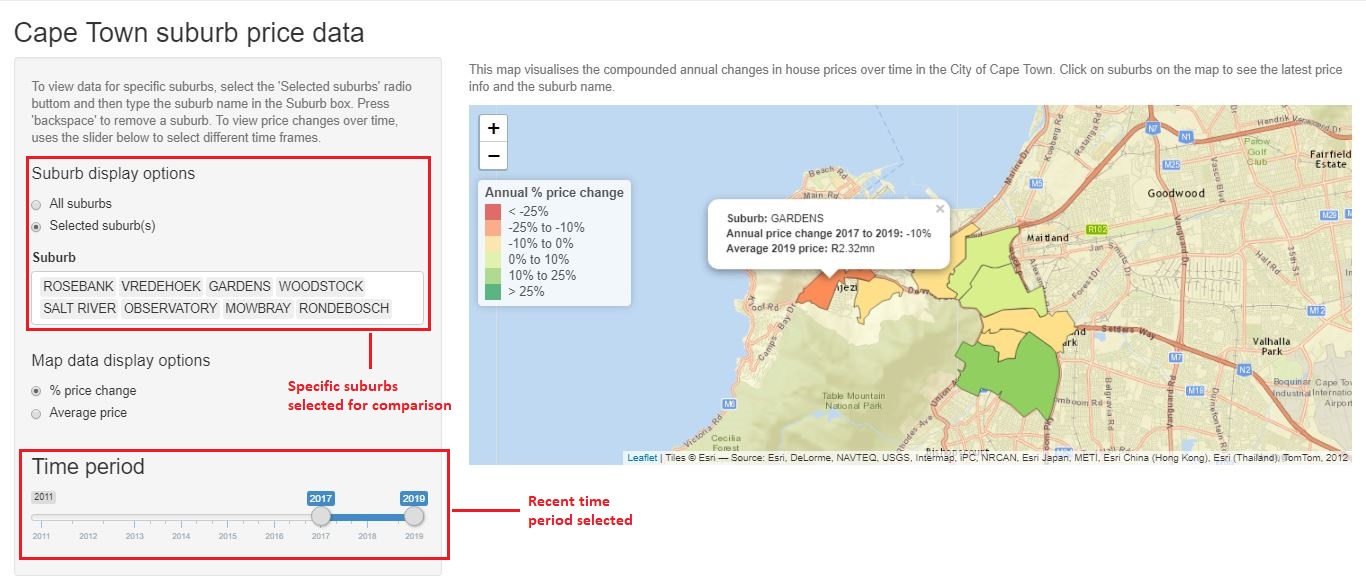
In terms of functionality, the Shiny app allows users to view the change in house price (as an annualised %) or the average house price over a time period specified by the slider. One is also able to select specific suburbs to compare.

**Extracting insights from the data:**

A useful exercise to perform using the app is to look for the suburbs that have shown the greatest historic (2011 – 2017) price increase; as there are many suburbs in Cape Town that have shown an annual price increase of between 10% to 25%, we’ll pick a subset of these here for this analysis.



For this example, I picked Vredehoek, Gardens, Woodstock, Salt River, Observatory, Mowbray, Rosebank and Rondebosch as examples of suburbs where house prices had grown on average by more than 10% per year from 2011 to 2017. The map was then adjusted to see how these house prices had changed over the past two years.



one can see that Gardens has seen an average 10% annualised reduction in price over the past two years and Vredehoek an 8% annualised reduction. Depending on your viewpoint, this either suggests that properties are currently undervalued in these suburbs and could potentially be good purchases, or that prices were historically overinflated.

If you’re looking for places that meet your budget, a similar exercise could also be carried out by looking at the average house price for the past two years on the map and seeing where you can afford to live. All else being equal, one could then look at the change in house prices to get a sense of which areas are on the up and which areas might be undervalued before looking for a property in that area.

**Challenges with the data**

The way Property24 organises suburbs and the way CoCT organises suburbs (i.e. their respective suburb hierarchies) isn’t always the same. This has led to some gaps on the map as well as some potential errors in average suburb prices. For example, Property24’s data for “Noordhoek” comprises the average of the average house price for a number of suburbs within that area, which when joined with the City of Cape Town data gives an inaccurate picture of house prices in the area; this is exacerbated if one of the areas on Property24 comprises numerous socioeconomic levels within the area.

Secondly, in a minority of cases the Property24 average house prices swing wildly from year to year. This may be due in some cases to actual wild fluctuations of house prices in an area, or perhaps more likely, to an outlier price (such as an apartment block going on sale) which skews the mean.

**In closing:**

This project showed how an exploratory data analysis approach can help to uncover high level insights in a data set that can point one towards insights in the data (for example, where have prices gone down recently). It also allowed me to use some of the great libraries across Python and R and to use publicly available data to answer a question I had for myself, namely, “Does it appear that the downturn in the economy is affecting house prices”.