Attention: staggering demand for house rental in Boston!

In this article I want to share with you results of my analysis that I perform on the Boston Airbnb open dataset. The dataset is provided by kaggle.com and recommended by Udacity during the Nanodegree Data Scientist course.

The following Airbnb activity is included in this Boston dataset:

- Listings, including full descriptions and average review score
- Reviews, including unique id for each reviewer and detailed comments
- Calendar, including listing id and the price and availability for that day

Throughout the analysis I want to follow CRISP-DM process, actually with a few modifications. The original CRISP-DM process:
* business understanding
* data understanding
* data preparation
* data modeling
* results
* deployment

Here, I do not perform modelling as I do statistical analysis and do not train any machine learning model. I also do not deploy the results but present them in the article.

To understand the domain, I had a brief look at the data. I was thinking what would be interesting to know, from my perspective, if I would like to start renting my apartment. I came with the following questions that I want to get answered after my analysis of the data:

* What is the price distribution for each room type?
* How is the price changing over time (count in months) and what are the top 5 average most expensive zipcodes all time?
* What is the distribution of demand over the entire time, per month?

Let's get started.

. . .

I obtained a lot of information in the datasets. I could have a look at the reviews content, who gave them and when it happened. Further, I could see the availability of each airbnb apartment on various dates, and the price connected with the given apartment on the given date.

A special abundance of features was available in the data containing information about the listings. To name just a few:

- posting URL, its description, picture, …
- review summary, scores, accuracy, cleaniness, …

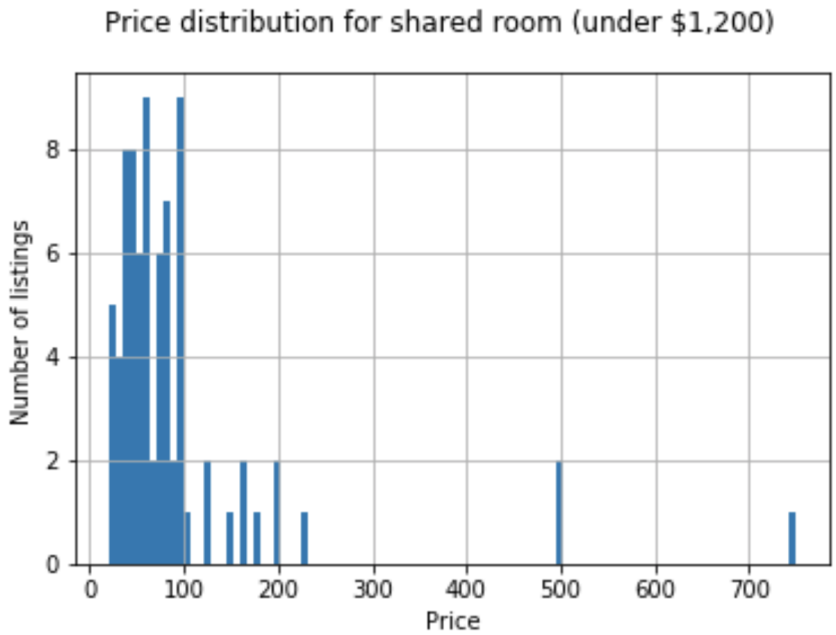
- neighbourhood overview, zipcode, …
- price per night, reviews per month, …
- property type, room type, number of bedrooms, rooms, …

The crucial step was then to identify, what I actually need to know, to answer my questions. I decided to focus on a few aspects only: price per night, zipcode, type of the property and of the room.

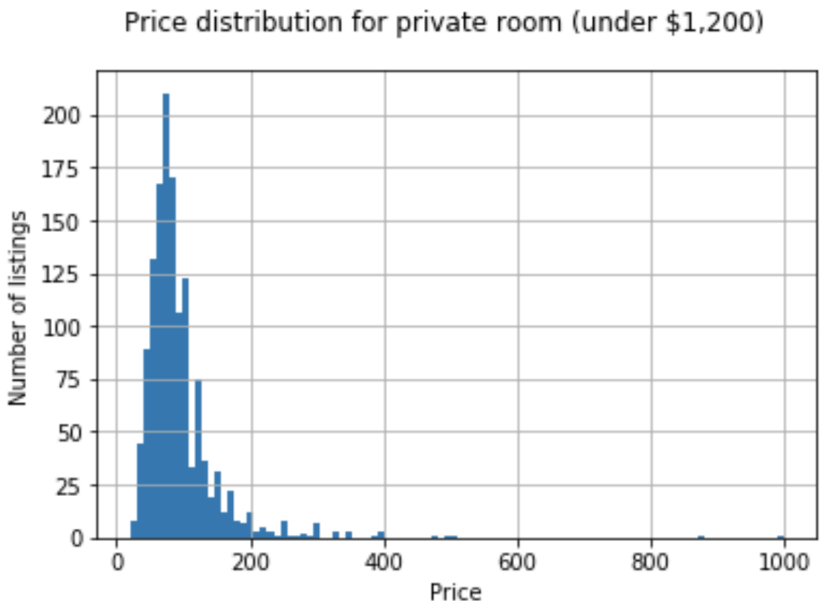
Let's have a look at what the data tells us!

. . .

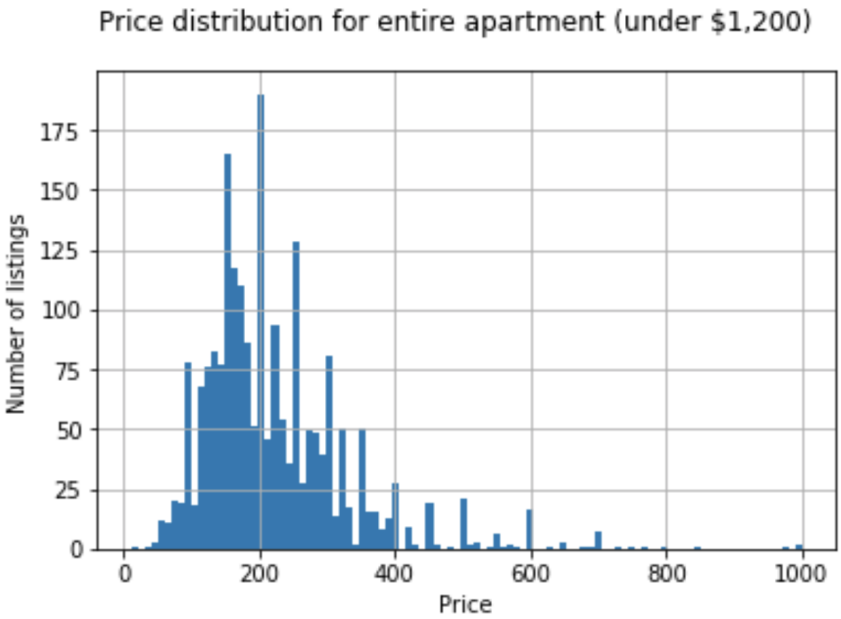
The distribution of the prices varied between 20 USD and 4000 USD. However the great majority I could observe between 20 USD and 1000 USD so I present you the distribution for that range, for each particular type of the room. There were three of them:



Distribution for the private rooms: there were in total 79 offerings of private rooms. The cheapest one 20$, on average 93$ and 75% of all the offerings were cheaper than 99$.

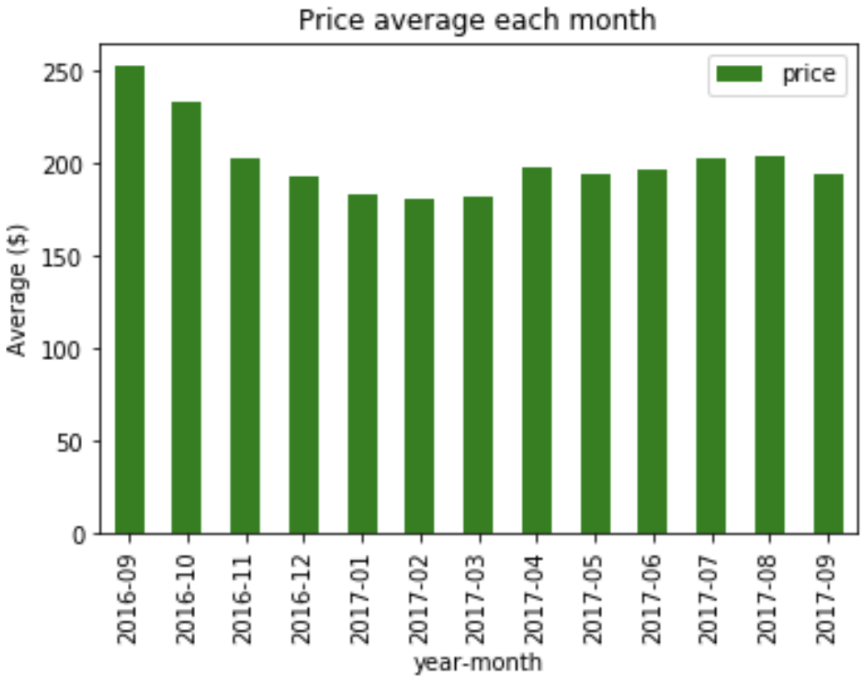


Distribution for the private rooms: there were in total 1354 offerings of private rooms. The cheapest one 20$, on average 95$ and 75% of all the offerings were cheaper than 100$.



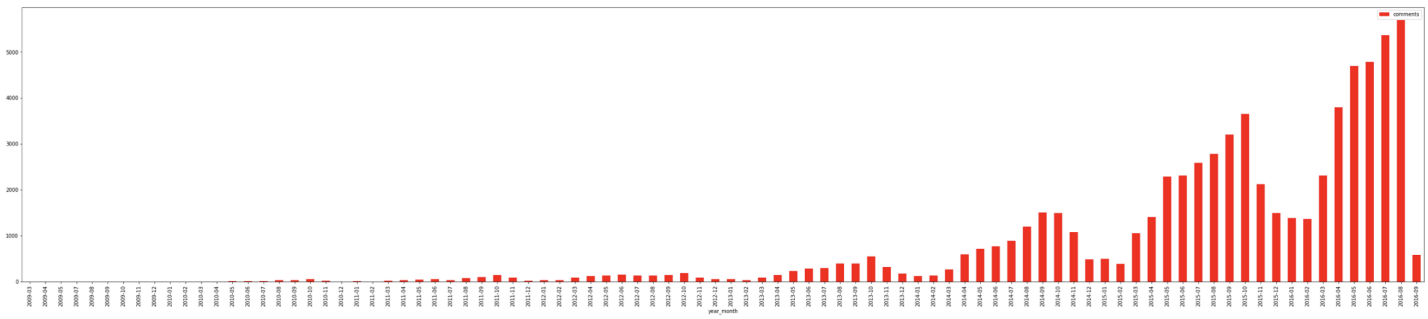
Distribution for the private rooms: there were in total 2111 offerings of private rooms, the biggest group of all. The cheapest one 11$, on average 228$ and 75% of all the offerings were cheaper than 270$.

It turned out that the price between Sep. 2016 and Sep. 2017, for the data available for me, did not change practically at all on average. This does not consider how prices of individual areas fluctuated and what was the spread and if that changed, but on average this is the result. I would assume that exactly the fact of some areas becoming more and some less expensive could have averaged out.



Average rental prices in the whole area of Boston over one year period.

In the next plot, clearly we can see a pattern that demand is increasing, even drastically over the last years. It might be a bit hard to read directly from the plot so let me stress that the data dates back to 2009 the earliest. The drastic increase is observed in the last 4 years, more and more. This is also seasonal pattern, usually increasing up to September—October, peaking and then dropping in winter. Remember that if you want to rent an apartment in the Boston area!



Clearly observed pattern of increasing demand over the last few years for rentals in the Boston area.