



HOSDA CAI - Data Visualisation, Sean Carroll

I8187226 Jose Ari Formiga - I8151183 Dario Ivone
I8200541 Lorenzo Di Mattia - I8151175 Monica Tedesco

Rome by numbers

The city of Rome is known to be one of the most touristic places on earth, the historical importance, and its attractions make the city a strong candidate for new and old travellers. Monuments, museums, galleries, parks, villas, squares, churches are some of the many pieces of history that trigger curiosity in many of those culture lovers.

The tourism data support this claim, for 2018 the arrival number grew 4% as well the total tourist spending that had grown of 4% for the same year. The transportation models had an increase in flights, train, and vehicles usage. [1]



General accommodation can be a representative budget allocation for any trip, so increasing variety can have beneficial impacts on lowering costs and increasing the offer of services available to the public. Global travellers are also willing to make more sustainable, or eco-friendly, decisions towards their choices in where to stay. [2]

On the 2015 report from the National Institute of Statistics, 421 million were the nights used as accommodation for the entire country of Italy, 50 cities had a total market share of 46% of those nights and Rome was the principal sink accounting for 6,4%, the second position was completed by Milan and Venice, with 2,8% each. [3]

Rome by numbers

Wikipedia reports that the two of the most visited places are the Vatican Museum (approximately 4.2 million tourists per year) and the Colosseum (approximately 4 million tourists per year). [4]



Like any other sector of the economy, the accommodation industry was also shaken by the shared economy. One of the key players in this field is Airbnb, a marketplace where customers can offer lodging and services named experiences, activities delivered by local people that goes beyond a regular tour.

This project will have all the data gathered from the “Inside Airbnb” that is an independent source of data not linked to the Airbnb company, all the data is public and no private information will be used.

This project aims to evaluate graphically Airbnb listings throughout Rome, for the years 2008 - 2019, and explore insights of this field of the accommodation since its introduction on the Italian Capital.

Rome, Lazio, Italy			
See Rome data visually here .			
Date Compiled	Country/City	File Name	Description
14 May, 2019	Rome	listings.csv.gz	Detailed Listings data for Rome
14 May, 2019	Rome	calendar.csv.gz	Detailed Calendar Data for listings in Rome

According to Ben Fry, the process of understanding data begins with a set of numbers and a question. The following steps form a path to the answer: [5]



01

Pearse & Filter

Provide some structure for the data's meaning, and order it into categories, then remove all but the data of interest.
(page 3)

02

Mine

Apply methods from statistics or data mining as a way to discern patterns or place the data in mathematical context.
(page 3)

03

Represent

We have chosen 11 basic visual model, such as a bar graph, list, or tree, after 11 analysis, in order to find out which locations are preferred and why.
(pages 3-16)

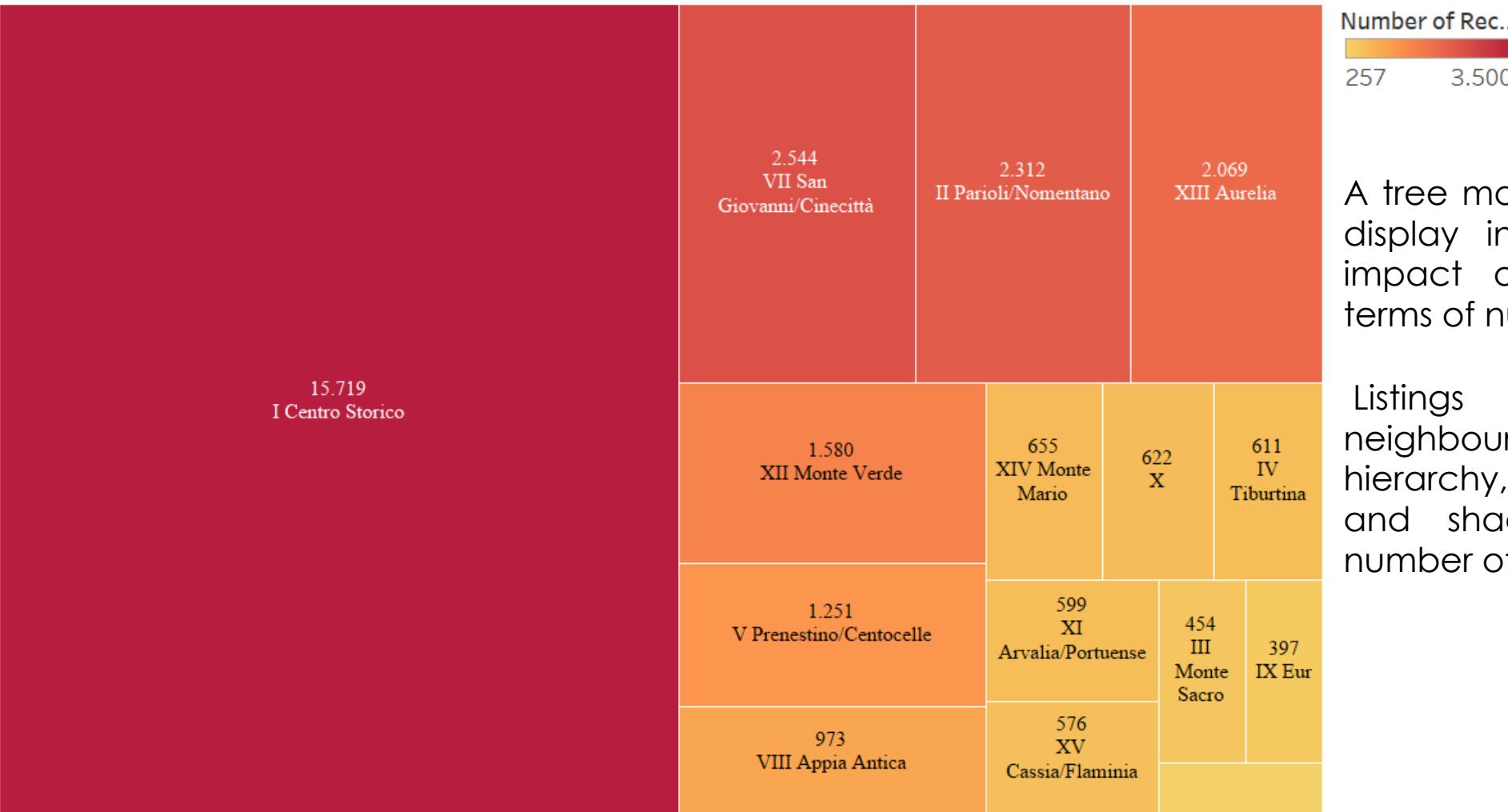
04

Refine & Improve

Add methods for manipulating the data or controlling what features are visible.
(page 17)

Graph #1

Listings per Neighbourhood



Number of Rec..
257 3.500

A tree map is one of the best ways to display in easily and efficiently the impact of each neighbourhood in terms of number of listings.

Listings were grouped by neighbourhood in order to visualize the hierarchy, and both rectangles size and shade are affected by the number of listings.

Sum of Number of Records and neighbourhood. Color shows sum of Number of Records. Size shows sum of Number of Records. The marks are labeled by sum of Number of Records and neighbourhood.

Graph #3

Annual growth by neighbourhood

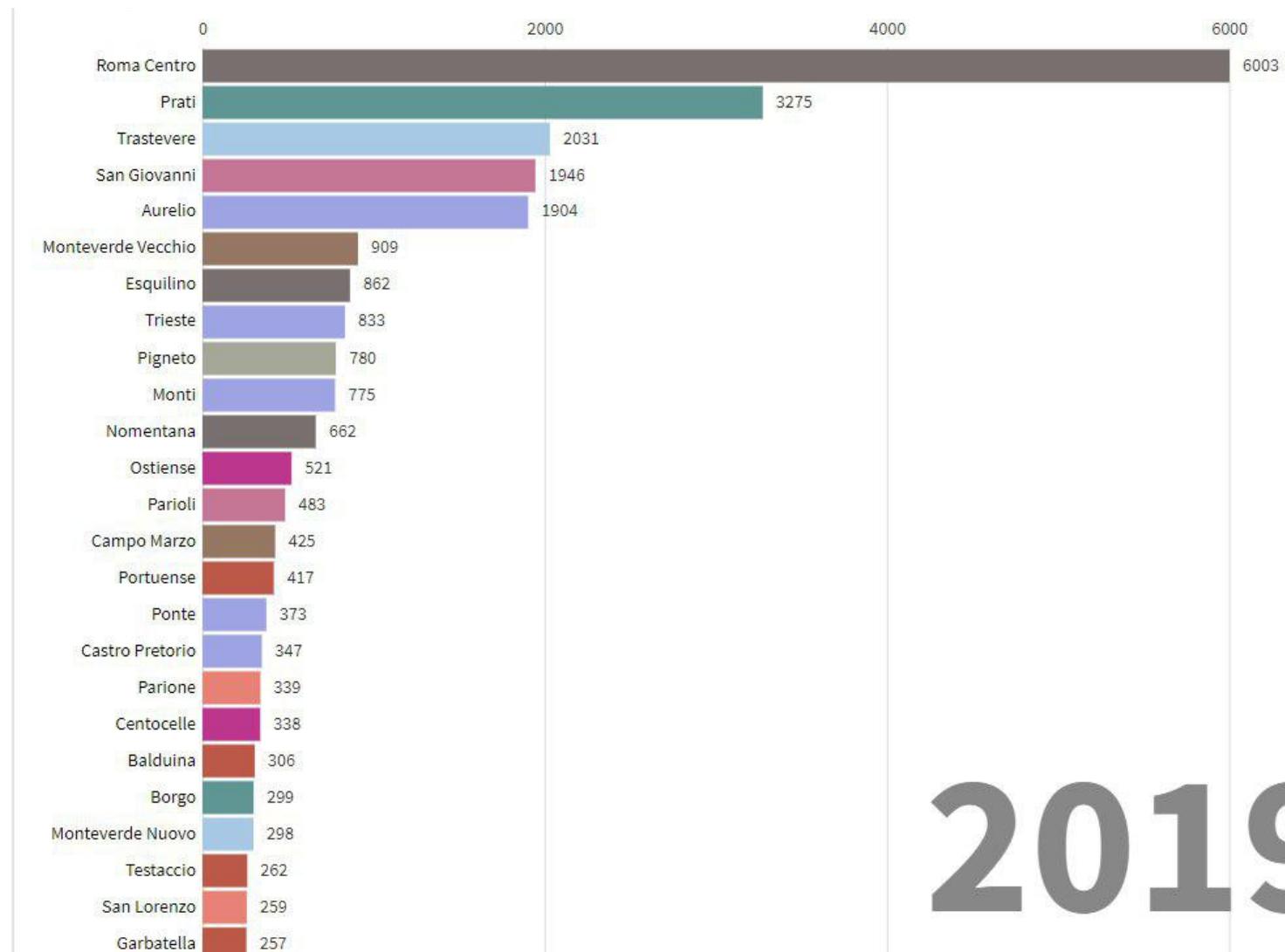
A bar chart represents categorical data with rectangular bars with heights or lengths proportional to the values that they represent [6].

And that's why it's perfect to represent our data, especially with a dynamic chart that clearly shows the rapid growth and in which location.

(...well, that's a GIF, but if you click on the link on the bottom right of this slide to see the dynamic bar chart!)

As noticeable from the chart on the right There was a rapid growth, especially at the beginning when the service since was in usage, with a big concentration in the city centre "Roma Centro", as expected since all the major tourist attractions are there.

In 2019 we have reached 30k of available accommodations.



2019

TEMPLATE CREDITS
Bar chart race by [Flourish team](#)

[On-line version](#)

Graph #4

Most Frequent Amenities

The word cloud graph is one of the best options for presenting text information originated from a no organized file as descriptions and reviews. Understanding the frequency of words can be a powerful tool for business marketing in order to increase customer conversion.



In this analysis, the objective was to find the most frequent amenities for Rome listings offered by each host and have them as “must have amenities” by the accommodation supplier. A basic guide using supplier data.

TV, Kitchen, Air conditioning, Wi-Fi, and Hairdryer were the most frequent words throughout the more than 35,000 listings evaluated.

Graph #5

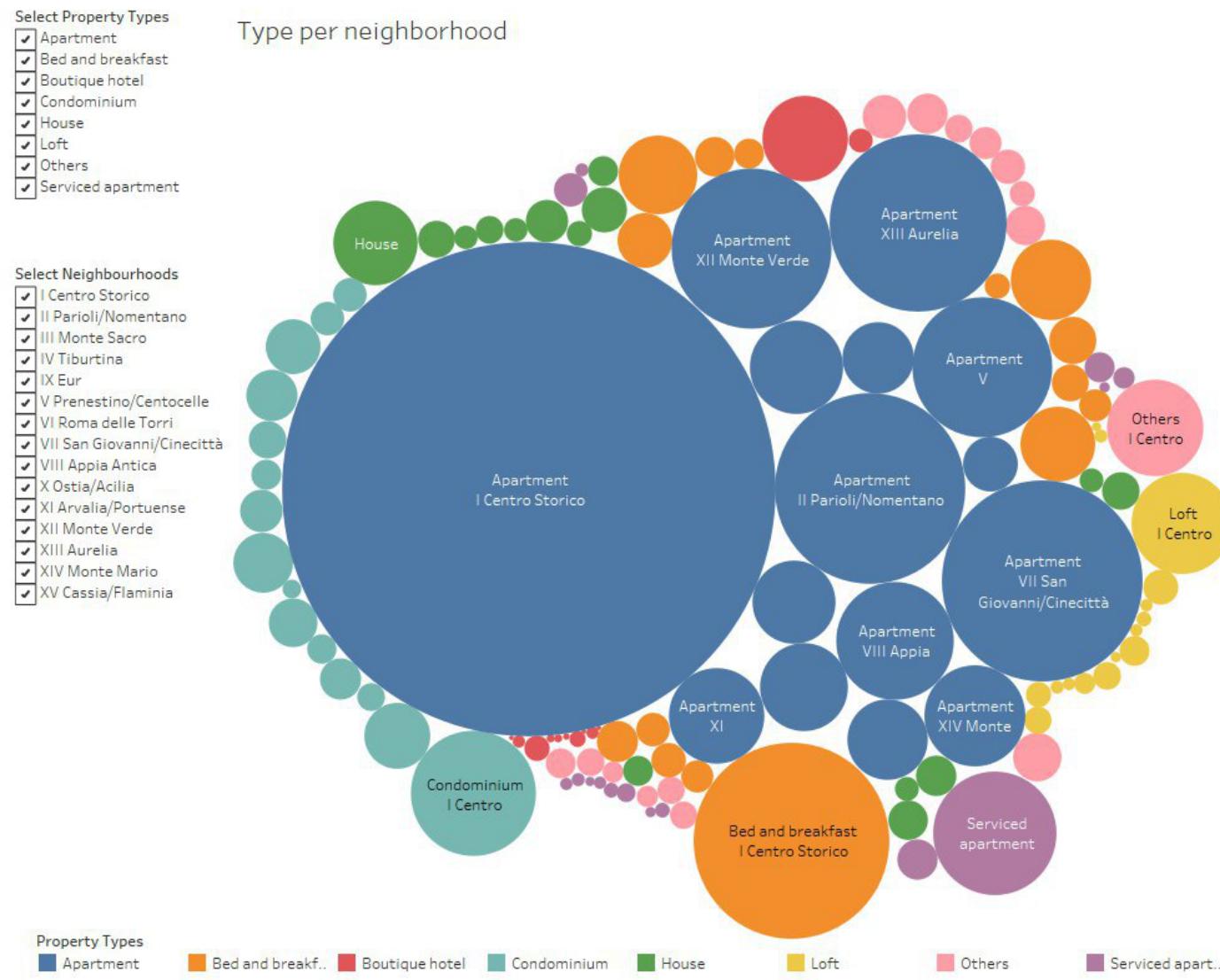
Type per Neighbourhood

Packed Bubble Chart is a cluster of circles where the dimension is proportional to the value that we are representing. Because of its simplicity is perfect for this representation in fact for any circle it is possible to read two information:

- the amount of accommodations represented by the dimension of the circle
- the type of property (only the main ones) represented by the colour.

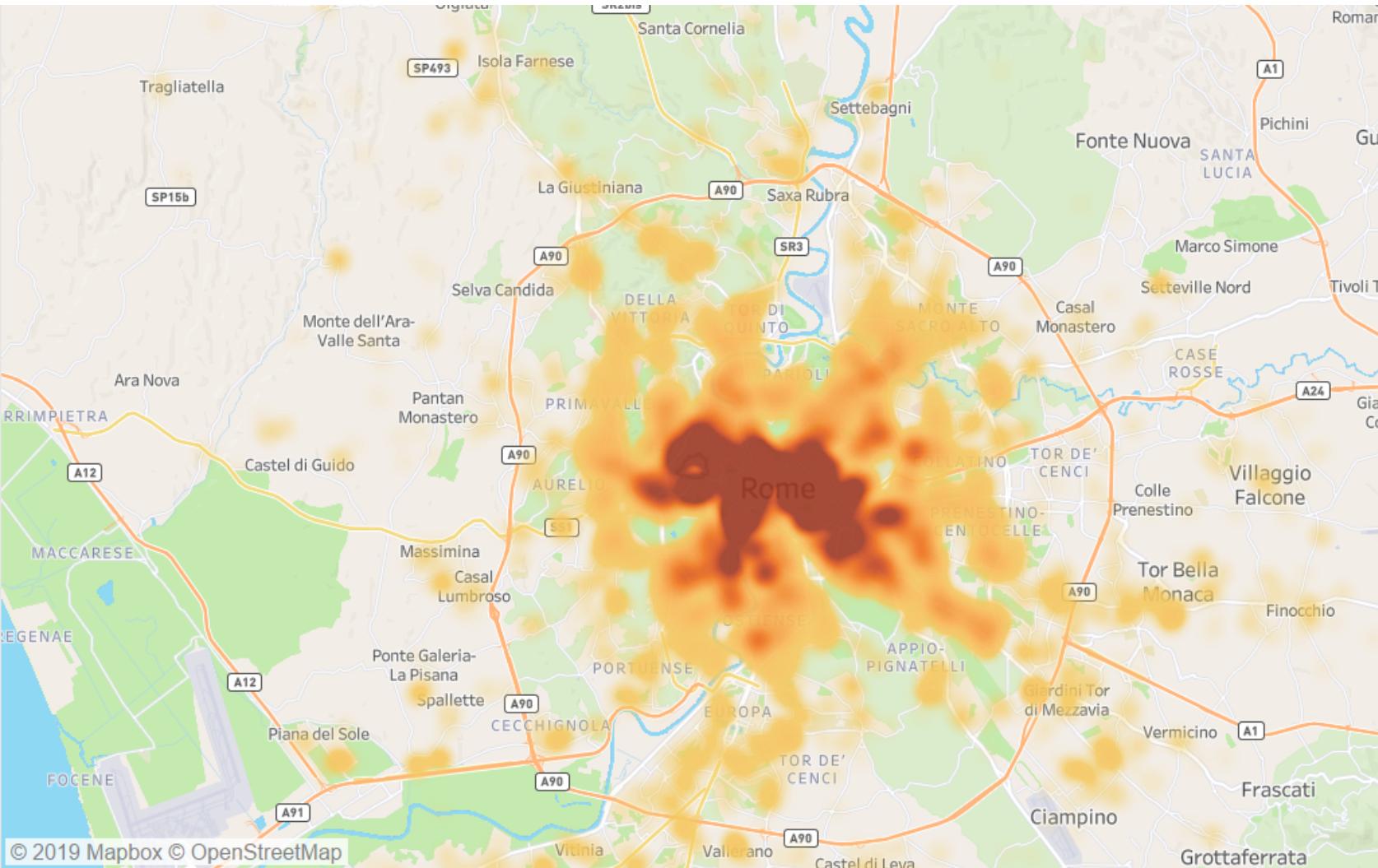
There are filters in order to select only what we are interested in, and the list seems to be a proper choice to visualise at a glance all the values.

The scope is not to give values but only to graphically represent with an easy interpretation our two parameters. In fact you can easily identify that the biggest amount of accommodations are apartment situated in Rome historical centre "Centro Storico".



Graph #6

Density of Airbnb listings in Rome



Map based on longitude (copy) and latitude (copy). Details are shown for property_type.

This map shows different property types listed on Airbnb in Rome: they can go from a Hotel to a Treehouse.

Results with less than 1 listing (like Igloos) were filtered. Showing the results on a map can also give an idea about the type of properties in the different areas of the city.

As we can clearly see from the map, the most common listing type is of course an apartment.

Graph #7

Average Review Analysis based on the Neighbourhood

In a radar chart is used polar coordinates instead of cartesian coordinates, practically three or more quantitative variables are represented on axes starting from the same point.

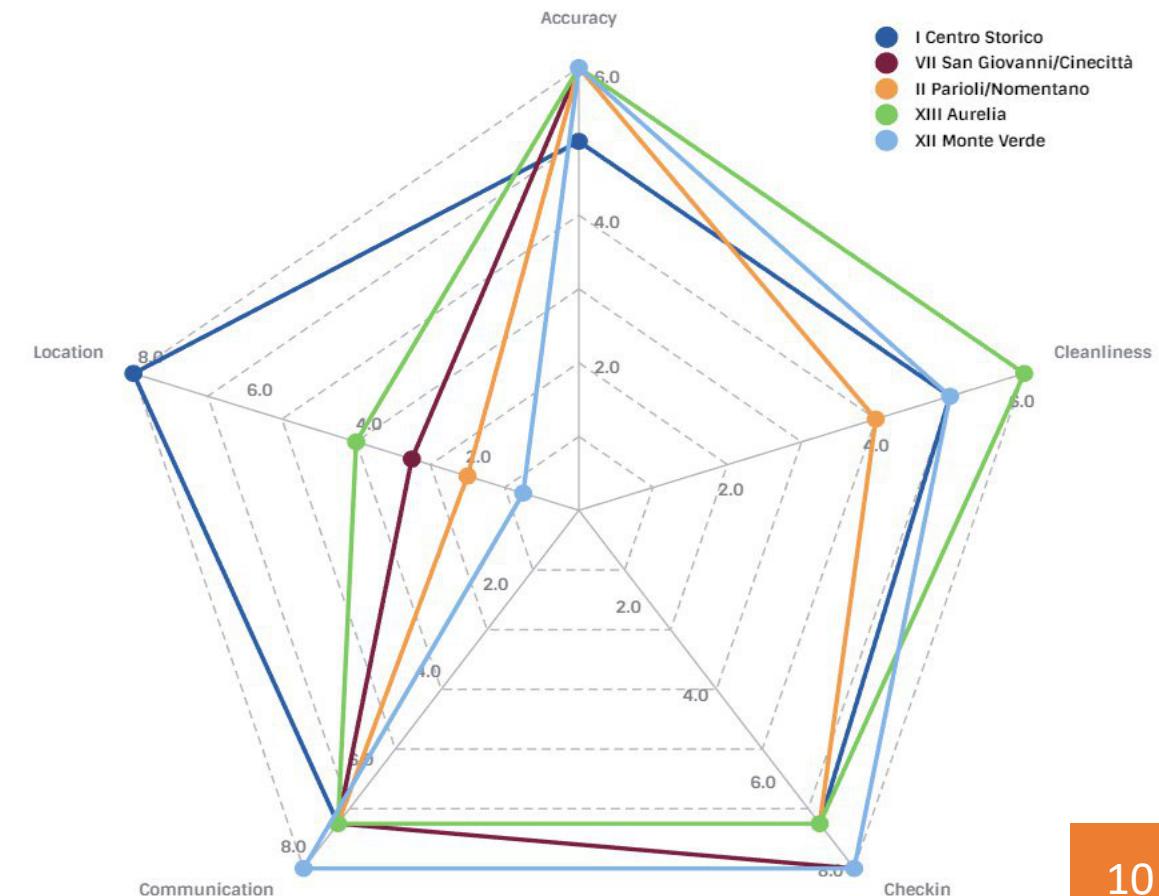
This chart is used mainly to display a certain number of observations, in our case the neighbourhoods with three or more variable (reviews based on different criteria).

In fact with this analysis we have identified if the most visited places (the main 5 neighbourhoods) have better or worst reviews than others and why.

We have created an analysis based on five categories/variables: location, accuracy, check-in, communication and cleanliness.

To emphasise the difference between the observation since the values are between 9.0 and 10.0, it is used only the decimal values so the range is 0.1 to 1.

It is clear that the main difference is on "location" that it's independent of the quality of the accommodation but intrinsic of the position of the listing. Of course also this time Rome historical centre (Centro Storico) is the winner and this justifies the big amount of accommodations that are in this location.,



Graph #8

Percentage of Professional Host vs Host and relative listings

The People chart shows in an easy way the percentage of hosts with more than 6 listings (defined Professional Hosts, or shortly Pro Host) against the percentage of regular hosts (accounts with less than 6 listings). In order to find this percentage we have subset the number of Pro Hosts and then calculated their percentage against the number of total hosts in Rome.

Once identified the population of Pro Hosts, we calculated the number of listings managed by them.

As we can see, the percentage of professional hosts is quite low (4%) but the number of listings managed by them is actually more relevant (23%).

% of Pro host



% Pro Listings



Pro Host
False
True

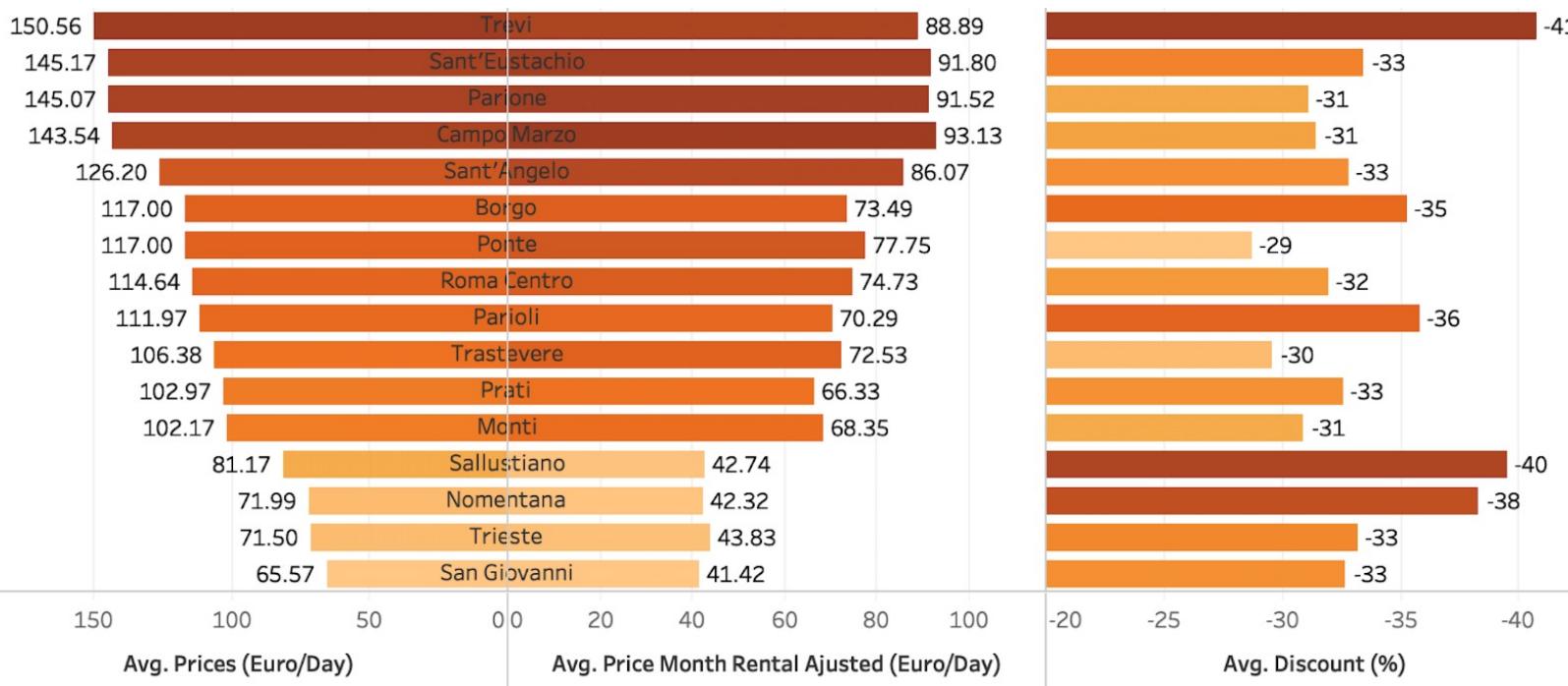
Pro Listings
False
True

Graph #9

Discount for Monthly Rental By Neighbourhood

For the discount evaluation, where each location would have two values to be shown together, the representation that fitted better was as a butterfly chart.

On this figure, we present a ranking of the higher average price per neighbourhood (Euro/Day) and also the Discount Price (Euro/Day) for who chose a long stay (month). Only 11.9% (3634) of the records had both prices available.



For the 17 records, the maximum discount was 40.71 % and lowest 28.68%, and the most part of values, 12, varied from 30 to 35%, which indicates a pattern price for long stays and also the discount that a customer can expect.

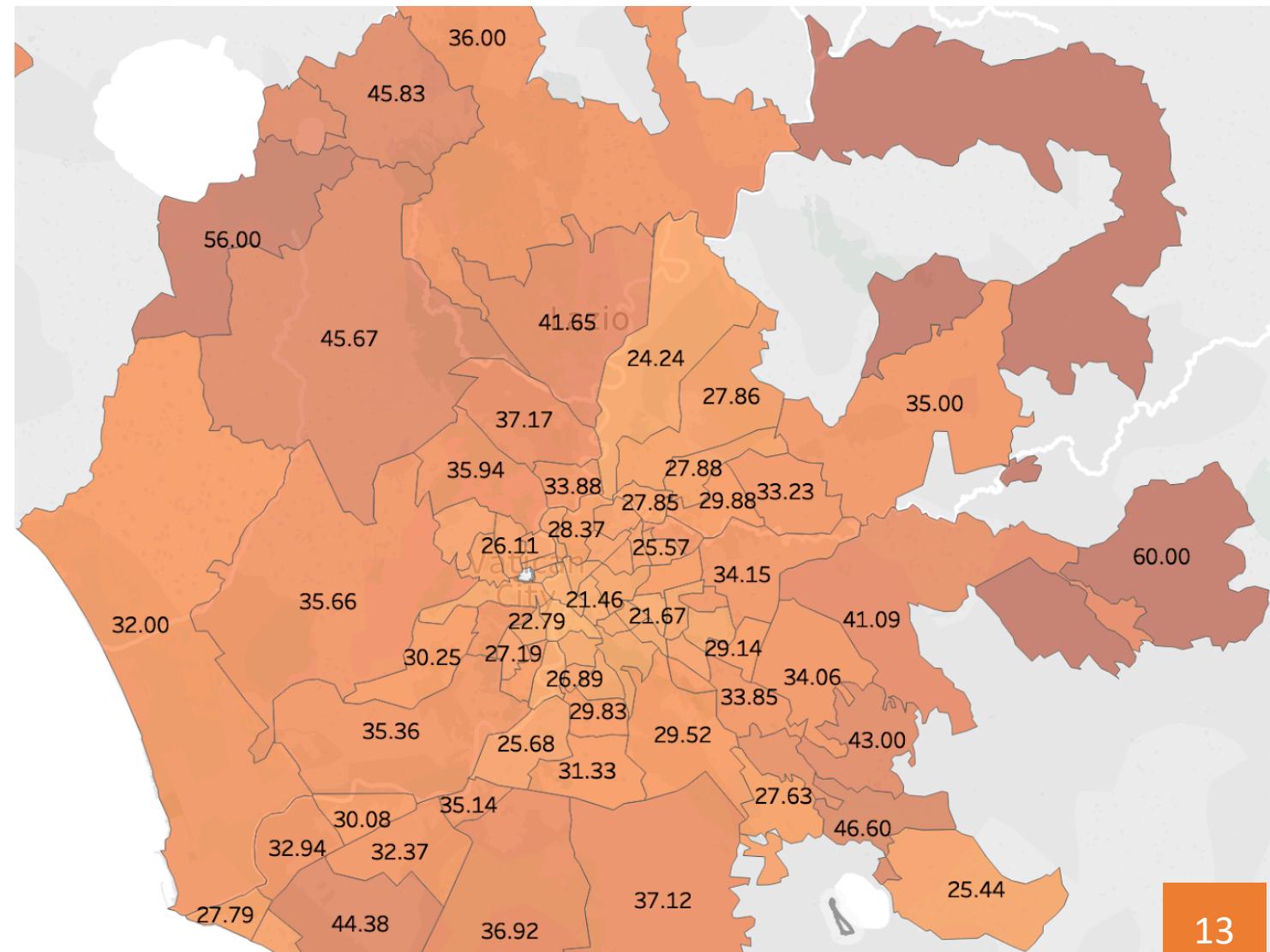
Graph #10

Room Availability per Neighbourhood (next 60 days) in Rome Region

To represent geographical data our choice was a heat map, the analysis used the neighbourhood (physical boundaries) as the basis of the calculation and used average values of availability for the next 60 days, as the darkness or lightness factor for each of them.

The graph tonality, darker orange, greater availability and, lighter orange fewer availability, clearly confirm our claim

This graph confirms that further regions from the city centre will have greater availability for the next 60 days from the evaluated date.



Graph #11

Price / Bed per Neighbourhood

With a simple table, you can clearly distinguish all the numbers you need, but it won't be immediate and easy to take out from a bunch like that the significant ones.

In order to analyse the data and to find the meaning and the story in it, we could use an highlight table.

It is a simple concept but you'll have a quick and visually clear result of what you're looking for.

In our example it is very clear how a highlight table helps, just imagine the same table without the highlights.

For some types of properties the price is rather homogeneous (i.e. apartment) and there are some peaks for specific type in some neighbourhoods.

Property Type	I Centro Storico	II Parioli/Nomentano	III Monte Sacro	IV Tiburtina	IX Eur	V Prenestino/Centro..	VI Roma delle Torri	VII San Giovanni/..	VIII Appia Antica	X Ostia/Acilia	XI Arvalia/Portuense	XII Monte Verde	XIII Aurelia	XIV Monte Mario	XV Cassia/Flaminia
Apartment	47.8	35.0	27.5	24.8	30.0	27.5	20.0	30.0	28.2	25.0	28.3	32.1	32.5	30.0	32.0
Bed and breakfast	55.0	41.7	40.0	36.3	40.0	30.0	25.0	39.0	37.5	47.0	52.0	40.0	45.0	44.0	44.0
Boutique hotel	70.0	74.0	65.0	26.3	50.0			31.3	77.5	60.0	60.0		50.0	60.0	37.5
Condominium	40.0	30.0	25.0	25.0	30.0	24.8	20.0	30.0	30.0	30.0	30.0	33.3	27.5	22.5	29.4
Guest suite	55.0	61.0		65.0	37.0	31.0	45.0	22.5	20.0	31.0	10.0	25.2	44.0	57.5	54.0
Guesthouse	55.0	77.5	32.3	33.8		35.0	16.0	39.5	20.0	30.5	8.7			34.0	45.0
House	47.0	42.9	24.5	30.0	35.0	25.0	25.0	30.0	31.5	26.8	35.0	35.0	35.0	27.5	28.0
Loft	55.0	40.0	26.3	20.8	49.8	30.0	34.0	31.3	19.5	42.0	37.3	33.8	37.5	62.5	34.5
Other	60.0	34.0	20.0	65.0	9.7	22.0	15.0	100.0	48.0	24.7	78.0	25.0	44.5		27.1
Serviced apartment	59.0	47.5	26.7	33.1	32.5	28.7	24.5	35.0	27.5	20.8	34.5	22.5	32.0	24.3	49.0
Tiny house	55.5	110.0		38.0		27.0	40.0	38.8		40.0	35.0	45.0	42.5	40.0	30.0
Townhouse	47.0	12.5	61.3	35.0	23.5	27.0	49.0	25.5		25.0	32.0	30.0	27.5	13.3	25.0
Villa	47.8	69.5	45.0	28.3	30.0	83.3	24.8	39.5	57.1	25.0	33.3	37.5	35.0	25.0	40.0

Price per Bed (\$)

8.7

Refine & Improve

We have improved our representations with this slides but we have made them even more clear and more visually engaging with an infographic!



The image shows a slide cover for an Airbnb report. The background is a black and white aerial photograph of a dense urban area, likely Rome. Overlaid on the left side is large, bold text: "MAY 2019 AIRBNB REPORT!". Above this text, in a white rectangular box with a thin black border, are the words "ITALY" in a small, sans-serif font and "ROME" in a larger, bold, sans-serif font. The overall design is clean and modern, using a high-contrast color scheme of black, white, and orange.

Sources

Dataset, Inside Airbnb Data, work published from: United States

<http://insideairbnb.com/get-the-data.html>

Icons © 2019 Piktochart. All Rights Reserved. Malaysia Incorporated Company

References

1. Ministero delle politiche agricole alimentari forestali e del turismo (2018) *Tourism Data 2018*. Available at: <https://www.turismo.politicheagricole.it/media/dati-turismo-2018/> [accessed 26/05/2019]
2. Tourism Statistic (2019) *Travel statistics to know in 2019-2020*, Available at: <https://www.trekksoft.com/en/blog/65-travel-tourism-statistics-for-2019> [accessed 26/05/2019]
3. ISTAT (2017) *Tourist flow in Italy*. Available at: <https://www.istat.it/en/archivio/224433> [accessed 26/05/2019]
4. Wikipedia (2019) *Tourism in Rome*. Available at: https://en.wikipedia.org/wiki/Tourism_in_Rome [accessed 26/05/2019]
5. Visualizing Data, by Ben Fry, 12/2007 (ISBN 9780596514556)
6. Wikipedia, 2019, Bar Chart [ONLINE] Available at: https://en.wikipedia.org/wiki/Bar_chart [Accessed 30/04/2019]