BARCELONA GRADUATE SCHOOL OF ECONOMICS MASTER'S DEGREE IN DATA SCIENCE

Data Visualization − Project: City Weather Radials

MIQUEL TORRENS (NIS: 55550)

April 11th, 2016

Framework

The project conducted is devoted to analysing some notorious weather conditions in different cities of the world. The technique used throughout the project is the weather radial, which analyses these weather conditions using a 360° chart that plots all observations in a circular form. The results of these analyses can be explored through the following web link:

- `https://mtorrens.shinyapps.io/dvis/`

In this report we summarise the main points of the project.

The aim of the project is to deepen in the knowledge of the climatic conditions of the included cities with this powerful tool. The weather radial is capable of showing the polyhedral behaviour of the data, as it can show time, average, variability and consistency of the data.

Data

The data used has been retrieved from the Weather Uunderground website:

- `https://www.wunderground.com/`

Weather underground is a commercial weather service property of The Weather Channel, that offers free worlwide weather data online. It collects informations of most weather stations, and gathers all important cities in the world.

Specifically, the data obtained belongs mostly to the airports of the selected cities. The cities for which we perform the analysis were filtered manually and the data was scrapped directly from the website. The code to do so can be found on the Github repository of the project:

- `https://github.com/mtorrens/weather/`

The data obtained contains information referring to:

1. Temperature
2. Precipitation
3. Relative humidity
4. Sea-level pressure
5. Wind speed
6. Dew point

7. Visibility
8. Cloud coverage.

The data was scrapped on a yearly basis for a period of 16 years, comprising the 2000-2015 time span. We recorded these variables for all the cities during this period, except for those cities which were not available for all the years. Each record of the dataset contains daily information regarding these variables, including minimum, maximum and mean values of each day.

Display

We analyse all the variables for each city using the weather radials. Each of this radials takes a city, a year and a variable as an input. Then it grabs the corresponding data and plots a bar per observation (a total of 365, one per day) in a radial fashion, according to each days' maximum, minimum and mean. It plots the oscillation in temperatures of the city across an entire year. The color of the bar shows the average temperature with respect to the the rest of the year.

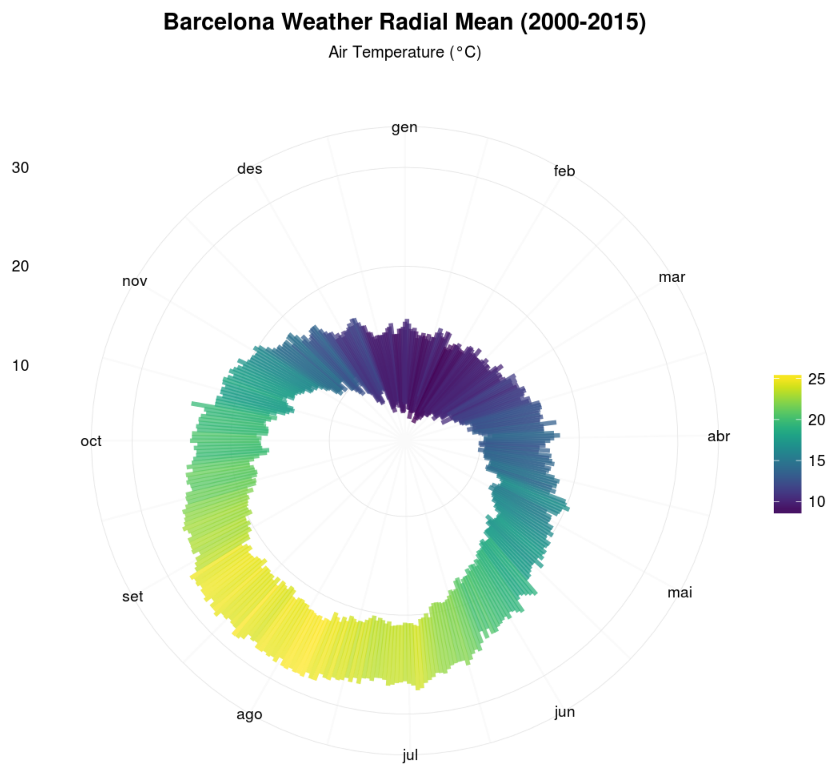
The main tab of the app cointains this analyses, that can be explored using the dropdown menus. One can choose the desired city and then select the variable of interest, and a time span. We include radials for all years in the 2000-2015 span, as well as a couple of aggregates: the mean of the period (each bar shows the average values of that day, say, the mean value of April 10 during the 16 years in the April 10 bar), as well as the median. The median is introduced in case there were some exceptional values in some year that could unbalance the arithmetic mean.

Additionally, we have built more ellaborate plots for a specific set of cities in the rest of tabs. These plots take advantage of the D3 opportunities to customize the plot and make it clearer and more complete. For a narrow set of interesting cities, they plot the temperature data and additionally introduce simultaneously the precipitation data. The blue balls illustrate (when available) if there has been significant precipitation during that day, whose size indicates the amount of precipitation.

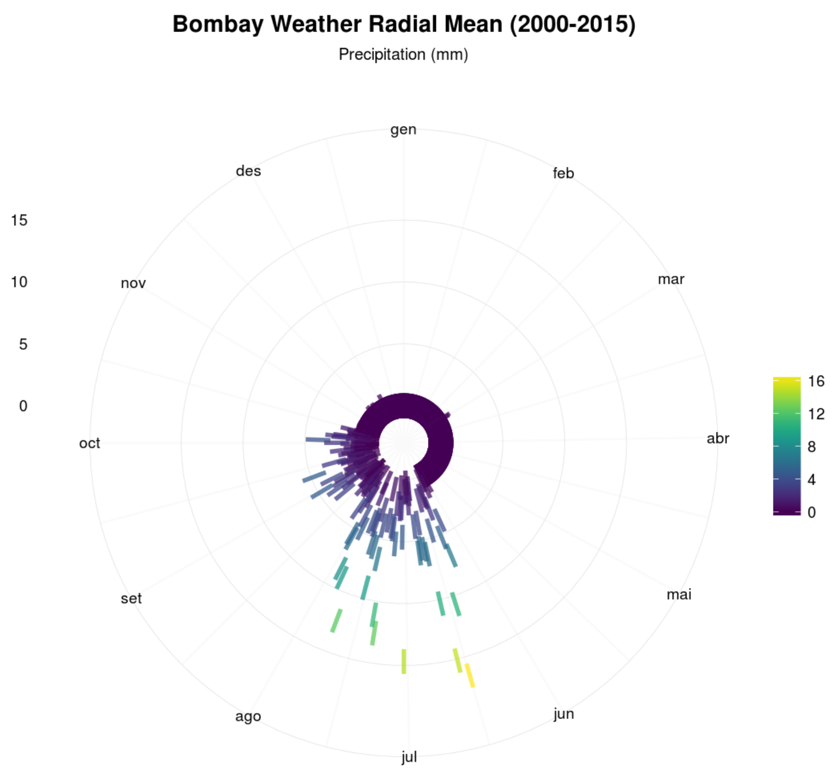
Results

We present some of the features that can be easily spotted from the data using the weather radial.

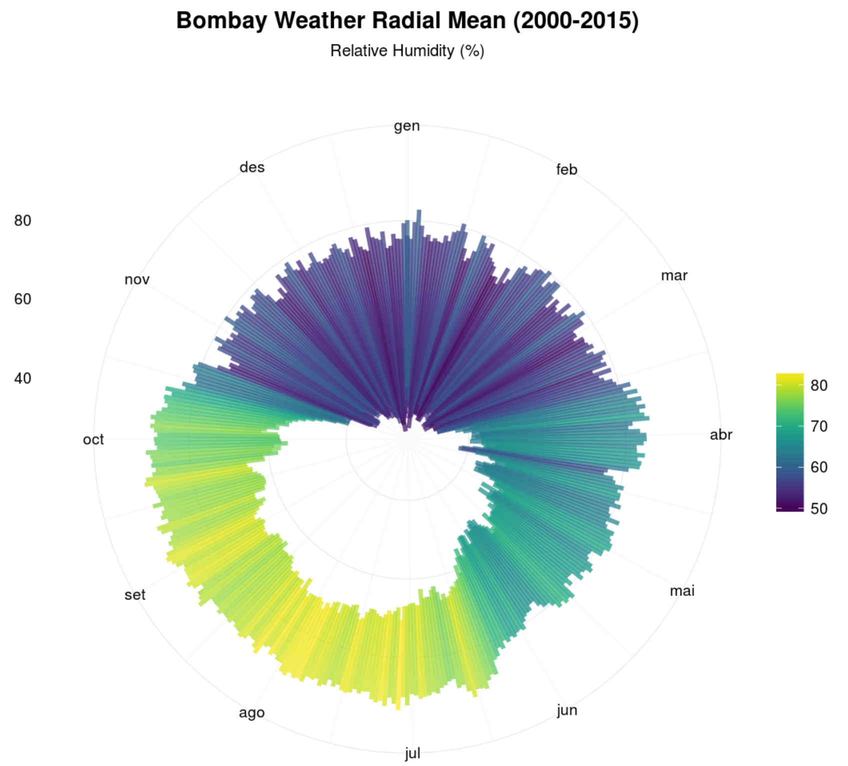
- Example 1. Spotting seasons through temperature, useful in temperate climates:



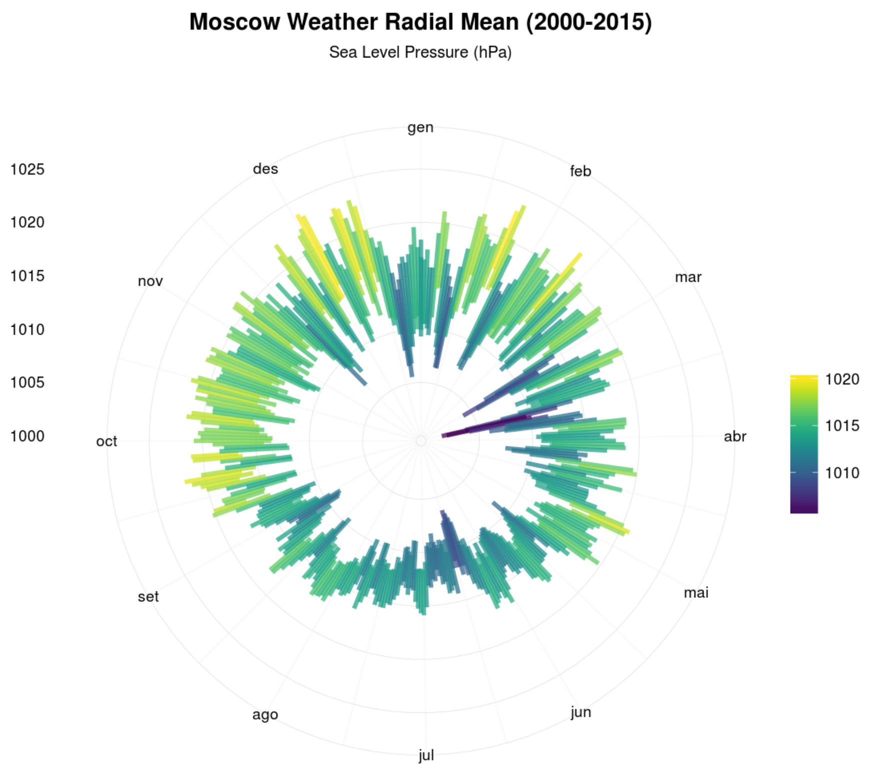
- Example 2. Spotting rainy seasons, helpful in tropical climates:



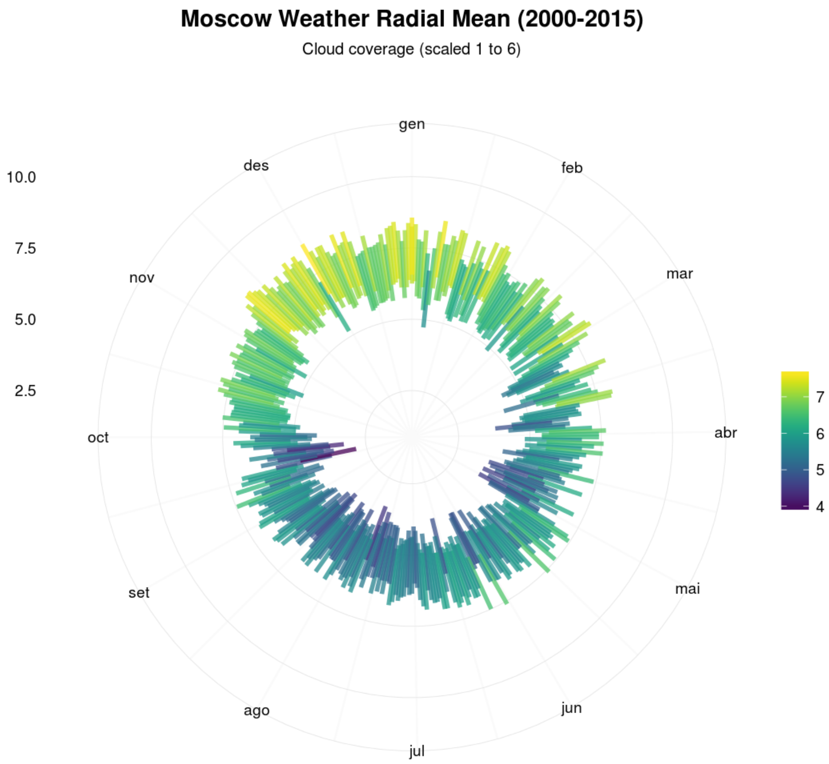
- Example 3. Humidity generally goes along with rain, and we can also spot its seasonality:



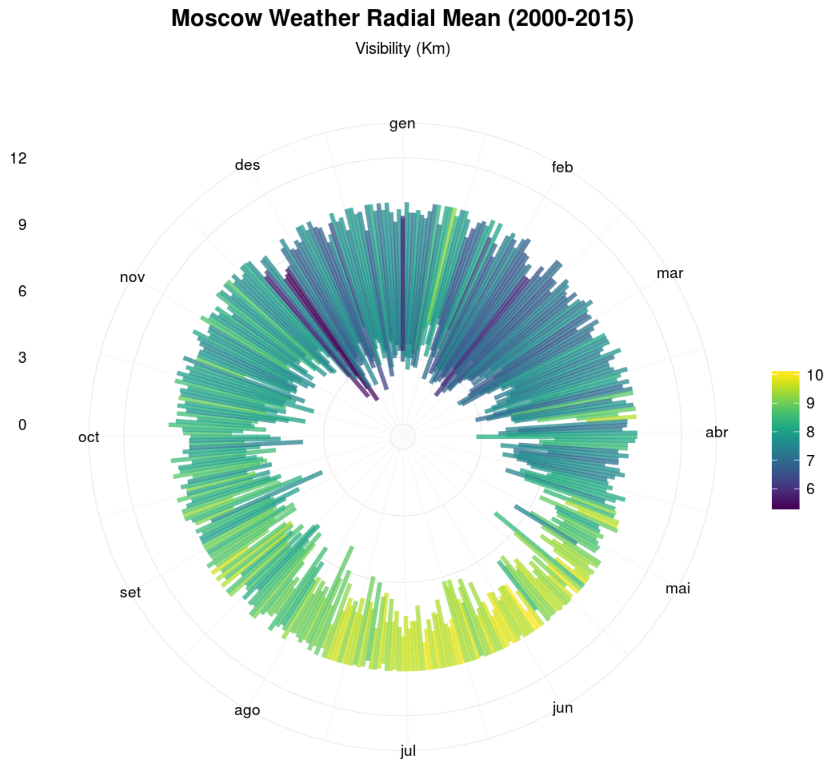
- Example 4. Not only temperatures or rain are seasonal, also air pressure can vary during the seasons:



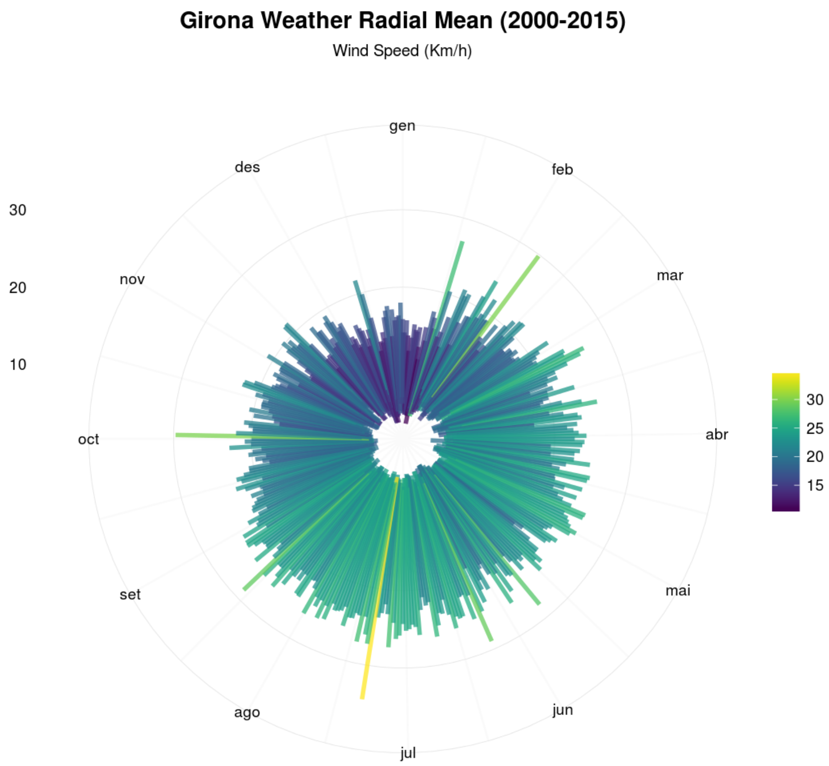
- Example 5. And with it, also cloud coverage and visibility are seasonalised:



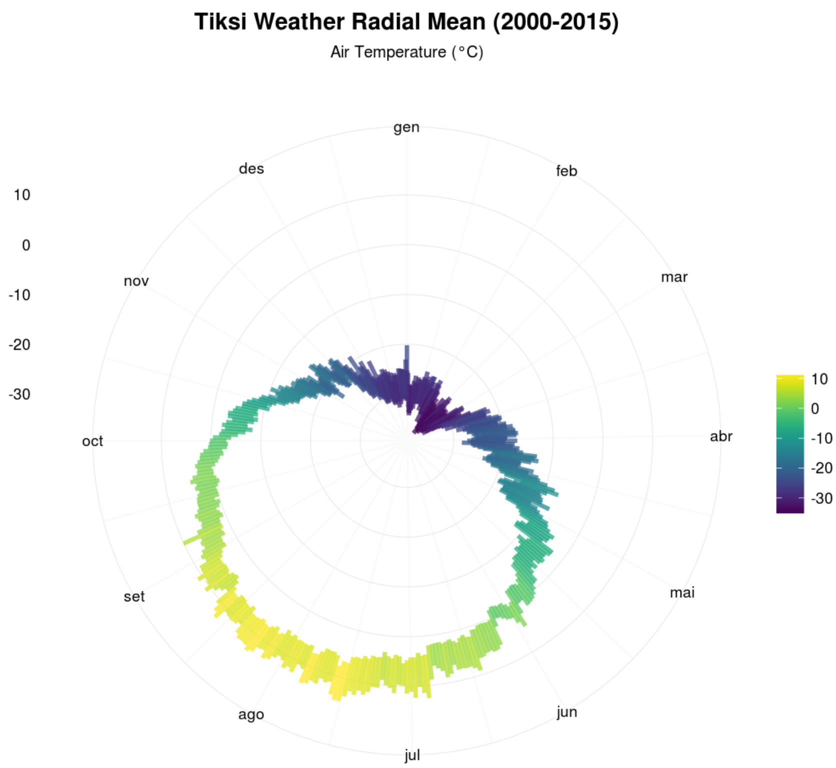
- Example 6. In the case of visibility, there is also intra-daily variability, as in winter it seems to vary more than in summer:

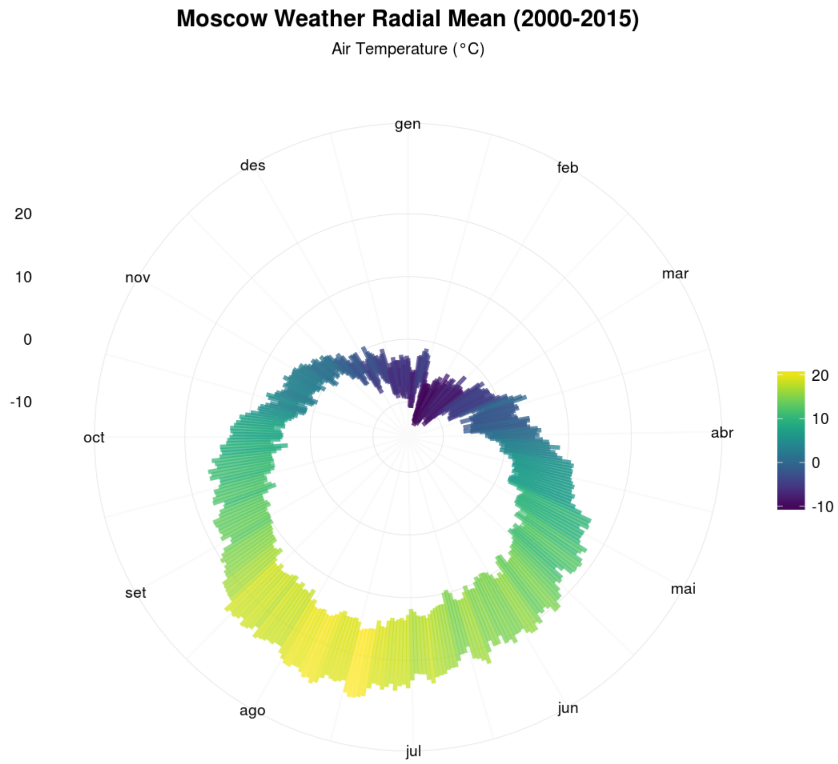


- Example 7. Finally, wind can also has its season in the year.



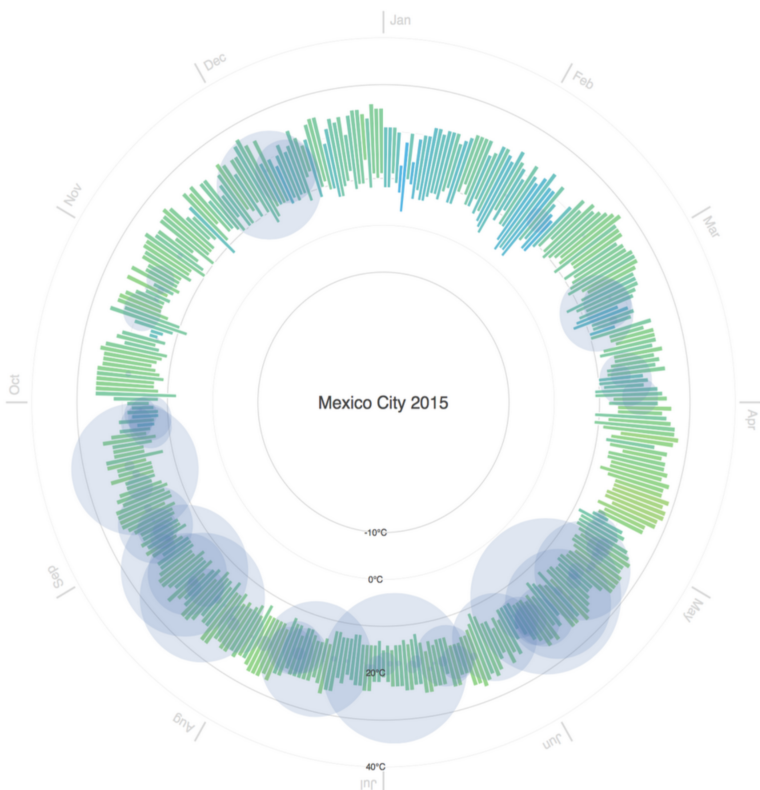
- Example 8. Equally important to season, there are cities that with a similar year-round weather can exhibit different stability during the day (wider intra-day range):



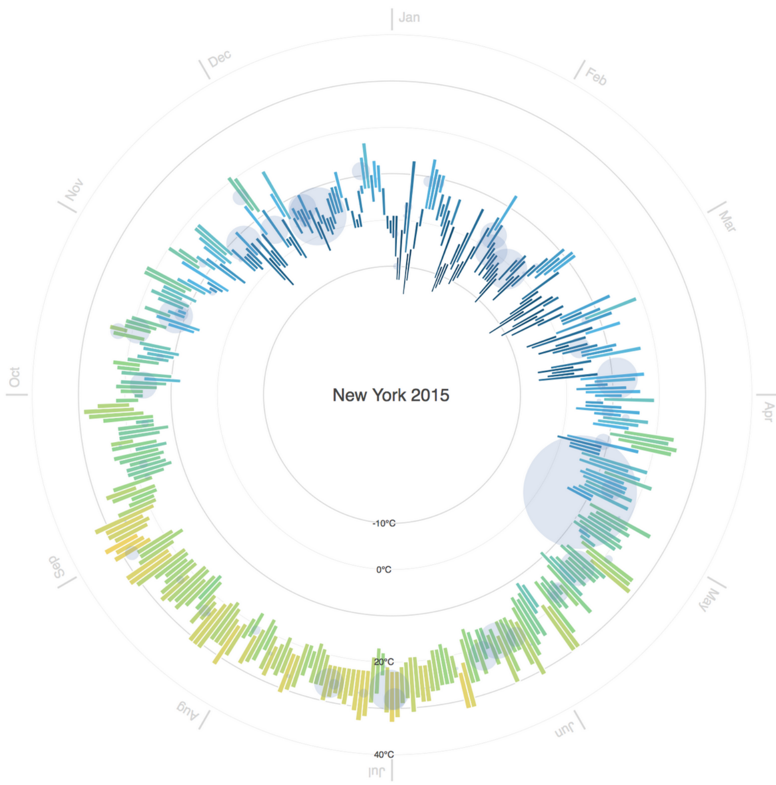


Using the possibilities of D3, we can observe the data with more detail.

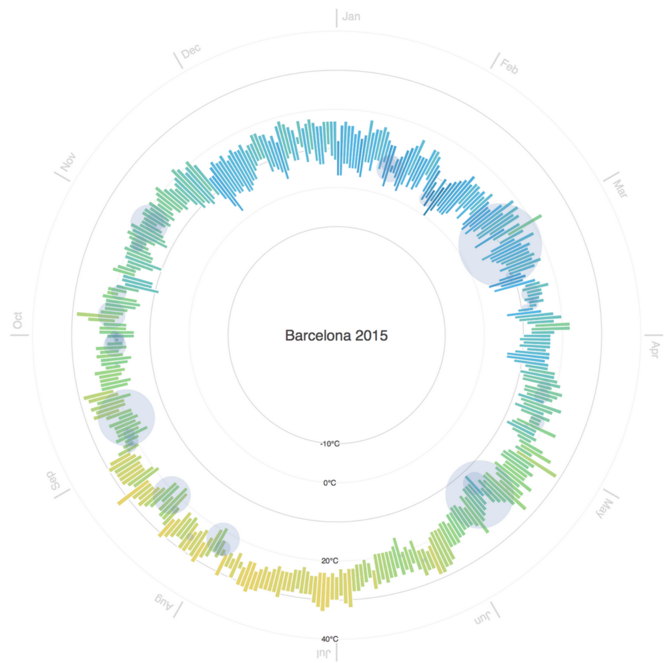
- Example 9. Spotting rainy seasons is clearer, and observing its correlation with temperature (more rain, less intra-day variability, as the effect of the sun is neutralised):



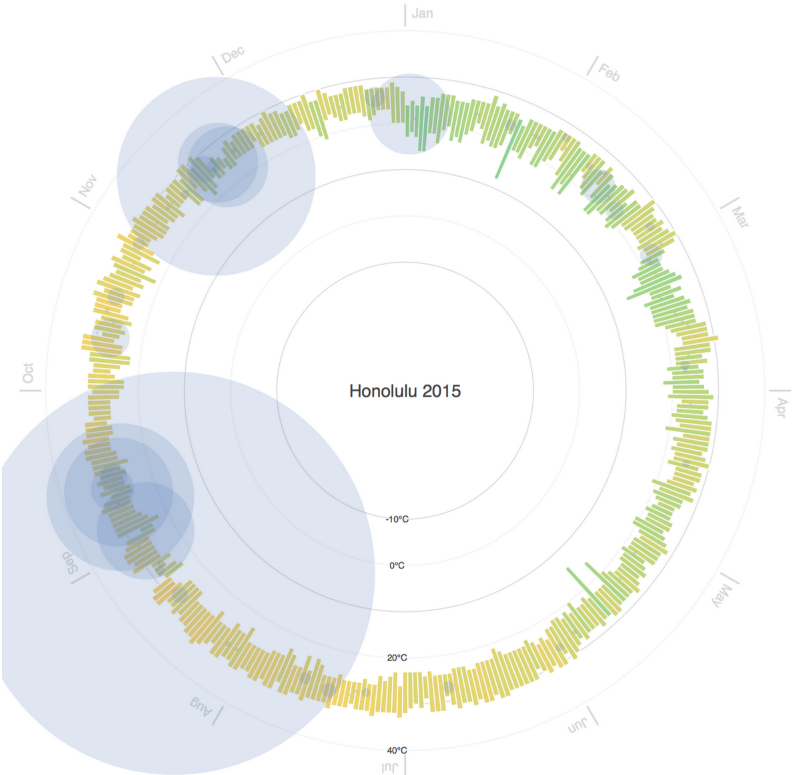
- Example 10. We can observe year variability, that is how different the seasons are and how the variability in consecutive days evolves in each season:

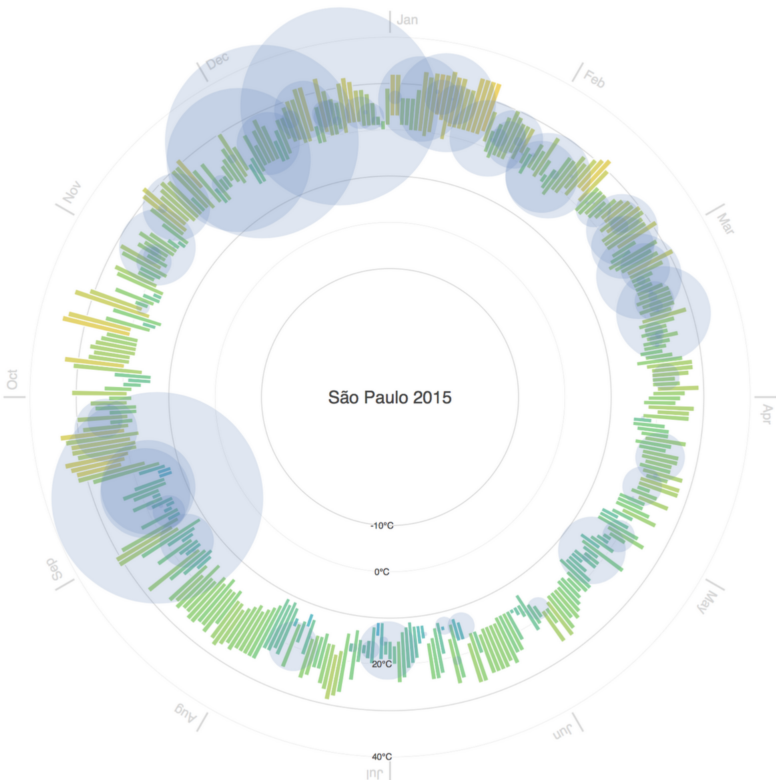


- Example 11. Season stability is also a factor: as intra-day variability changes across the seasons (in this case, more variability in winter, less variability in summer, usual in hotter climates as there is less sun time in the winter):



- Example 12. Some climates have higher thermal stability in the day and between days despite having global stability in the year-round temperatures:





This are some examples of all the things that can be observed using radial weather charts.

Conclusions

This project shows that the radial analysis can be really helpful when analysing weather data. It can capture with ease many features of the data simultaneously and the interpretation of the graph is intuitive and rich, despite requiring very few prior knowledge. Here we have explored some of the power it can have for generic cases, but surely the user can expand its uses and interpretations for particular cases and specific climates. For this reason we have built the aforementioned Shiny app associated to this report.

It would be interesting in future implementations of this analysis to extend it to multivariate data. Including the precipitation circles is a first approach to that, which already makes the chart naturally richer. A next step would be to include multiple variables and clearly observe correlations at a glance, so as to make deeper analyses.

All in all, the project has been useful to shed some light on innovative methodologies to display weather data, far from classical line charts and thermal maps, but maybe with even more power and versatile after all.