PageRank
A geospatial use case

Carlos David García Hernández
Data Scientist

PageRank

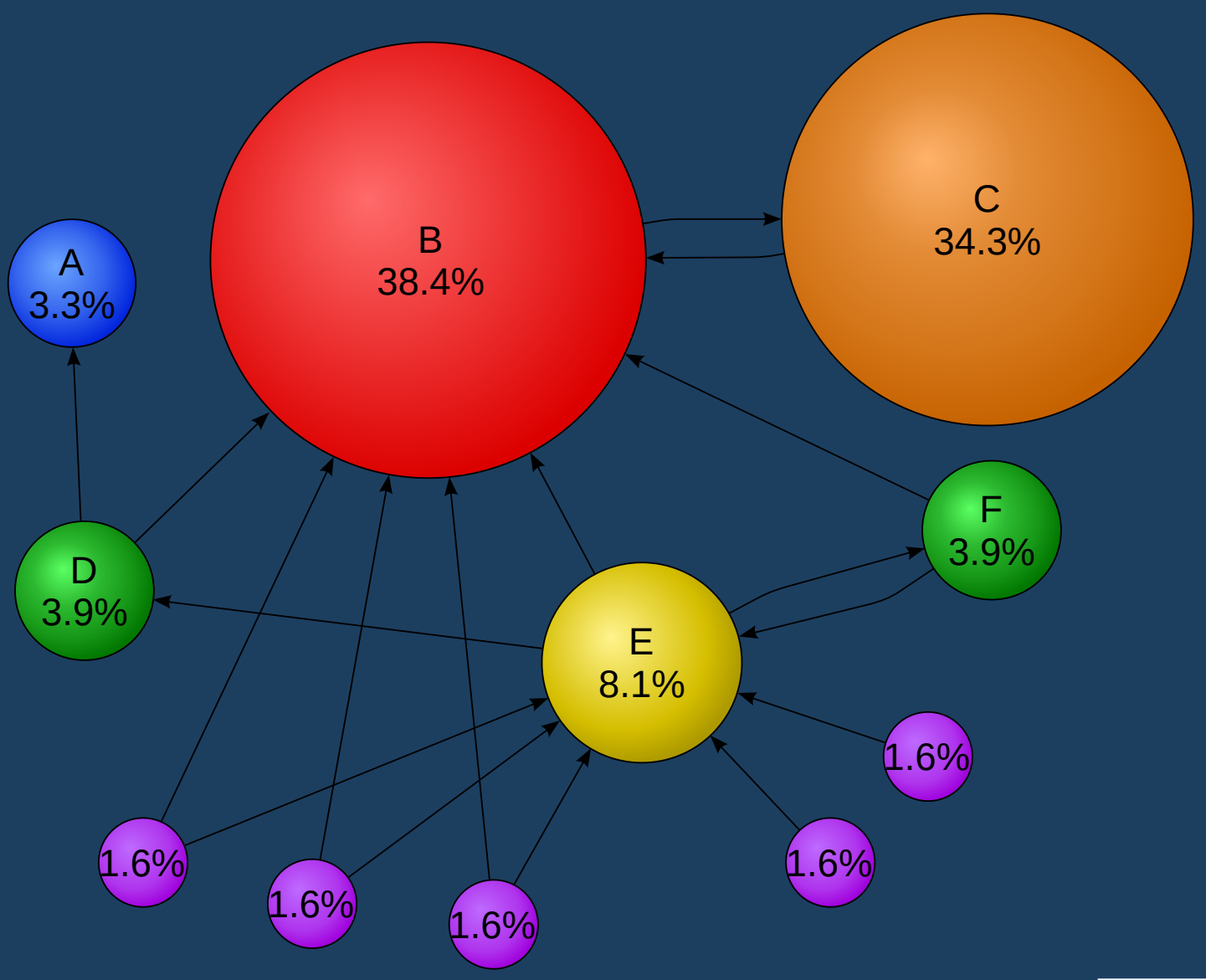
It’s the classic algorithm that Google developed to order by importance the web pages all around the internet.

PageRank was the most useful metric to position yourself on the web (Hello SEO analysts)

As a structure

You can see it as a network with nodes and edges.

If your node is big, then you are the coolest guy in the network.



PageRank math

To be that cool node you need to have connections but watch out, quality “PR(i)” over quantity “C(i)” it’s the way to go.

Damping factor “d” is just that randomness in life that you just can’t control.

$$\mathrm{PR}(A) = (1-d) + d\sum_{i=1}^{n}\frac{\mathrm{PR}(i)}{C(i)}$$

GeoSpatial example

Think about Mexico City and the Ecobici service, the bicycle rental service.

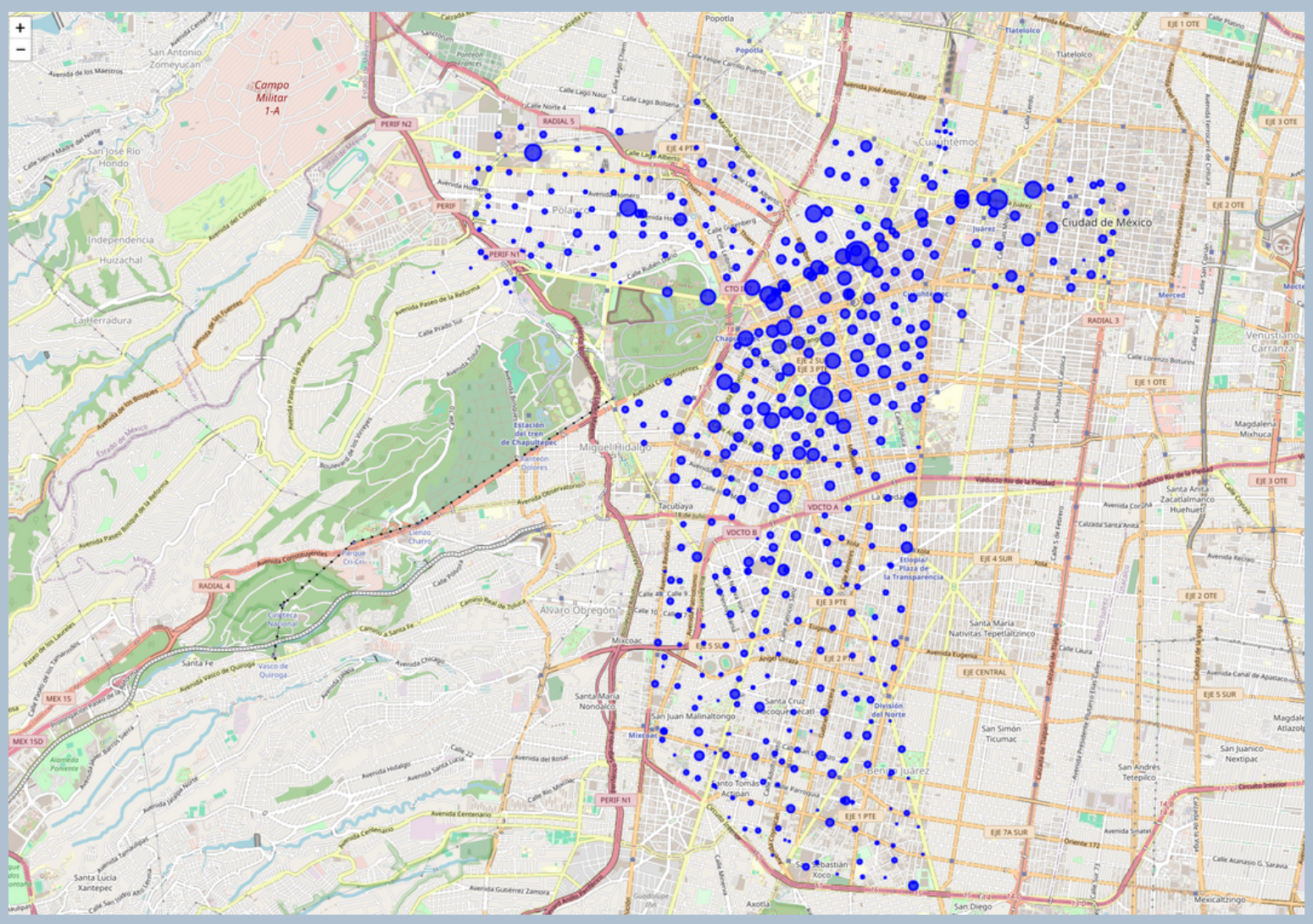
It has many stations, and people go from one to another, sometimes even going to the same station.

The most popular season is relevant as it is more likely that there are more places to visit.

GeoSpatial example

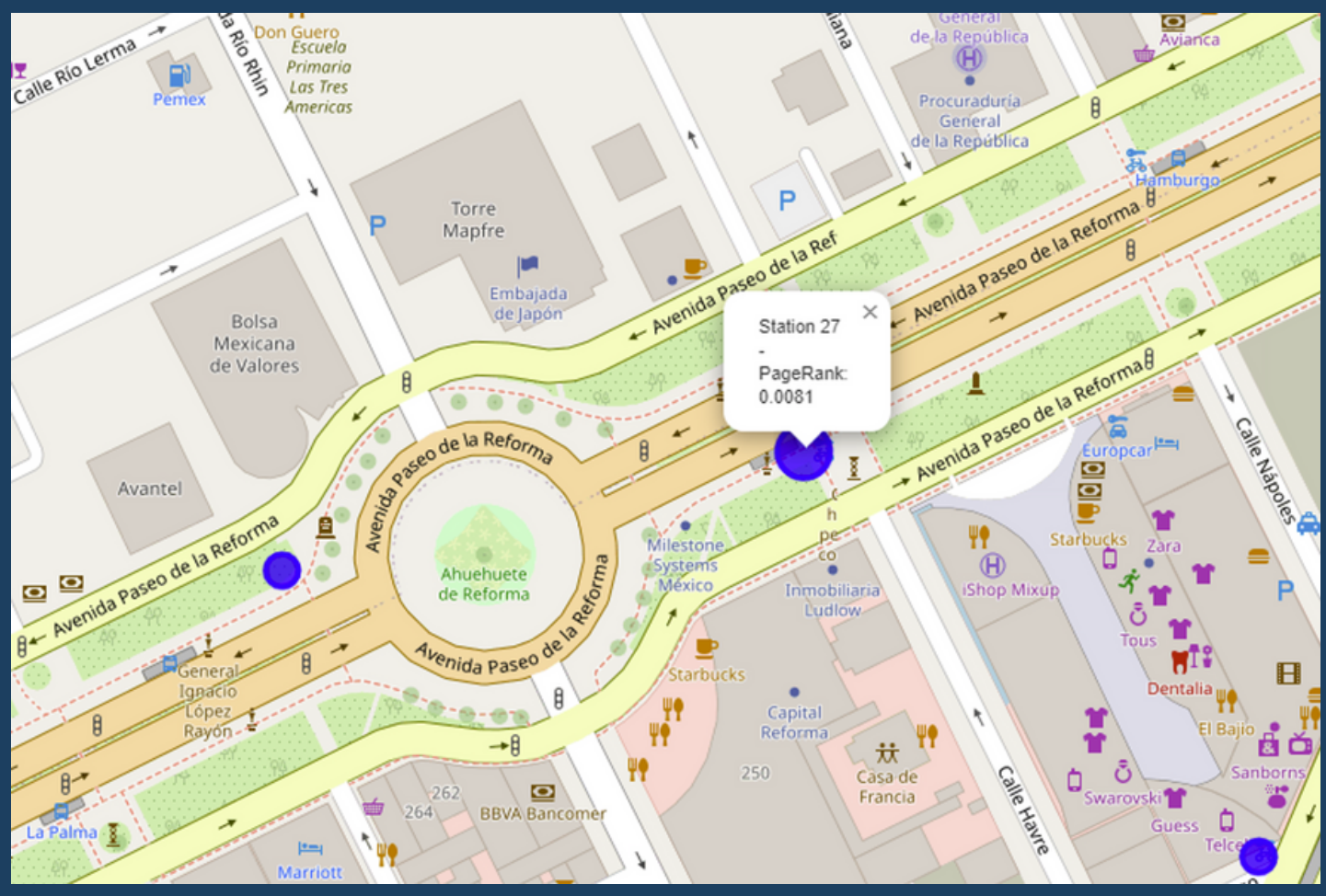
I introduce to you the Ecobici stations in Mexico City.

We can see some bigger circles, but lets make a little zoom.



GeoSpatial example

Station number 27 is the most popular station, probably the most traveled by workers and people who wants to enjoy Paseo de la Reforma Avenue.



References

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems, 30(1–7), 107–117. https://doi.org/10.1016/s0169-7552(98)00110-x

PageRank. (2013, December 11). In Wikipedia. https://es.wikipedia.org/wiki/PageRank

Thank you for your time

I will read and answer you in the comment section.