

10/18 DATA ANALYTICS - IRONHACK BOOTCAMP

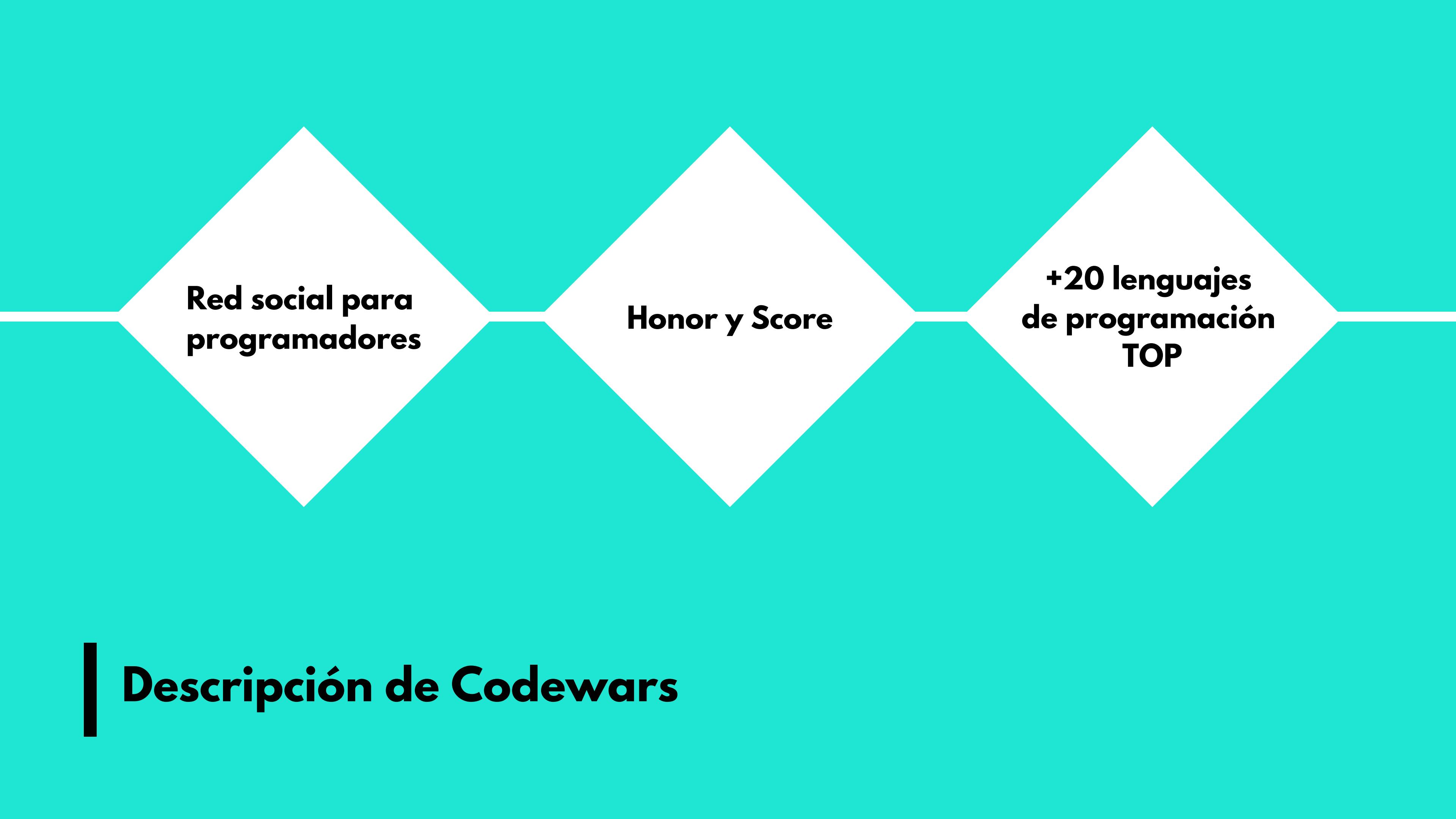
# Sistema de Recomendación: Codewars

**Alberto García Cobo**

# CONTENIDO



Descripción de Codewars  
Descripción del Dataset  
Data cleaning and wrangling  
Mayores retos  
Machine Learning  
Problemas  
Resultados



**Red social para programadores**

**Honor y Score**

**+20 lenguajes  
de programación  
TOP**

**Descripción de Codewars**

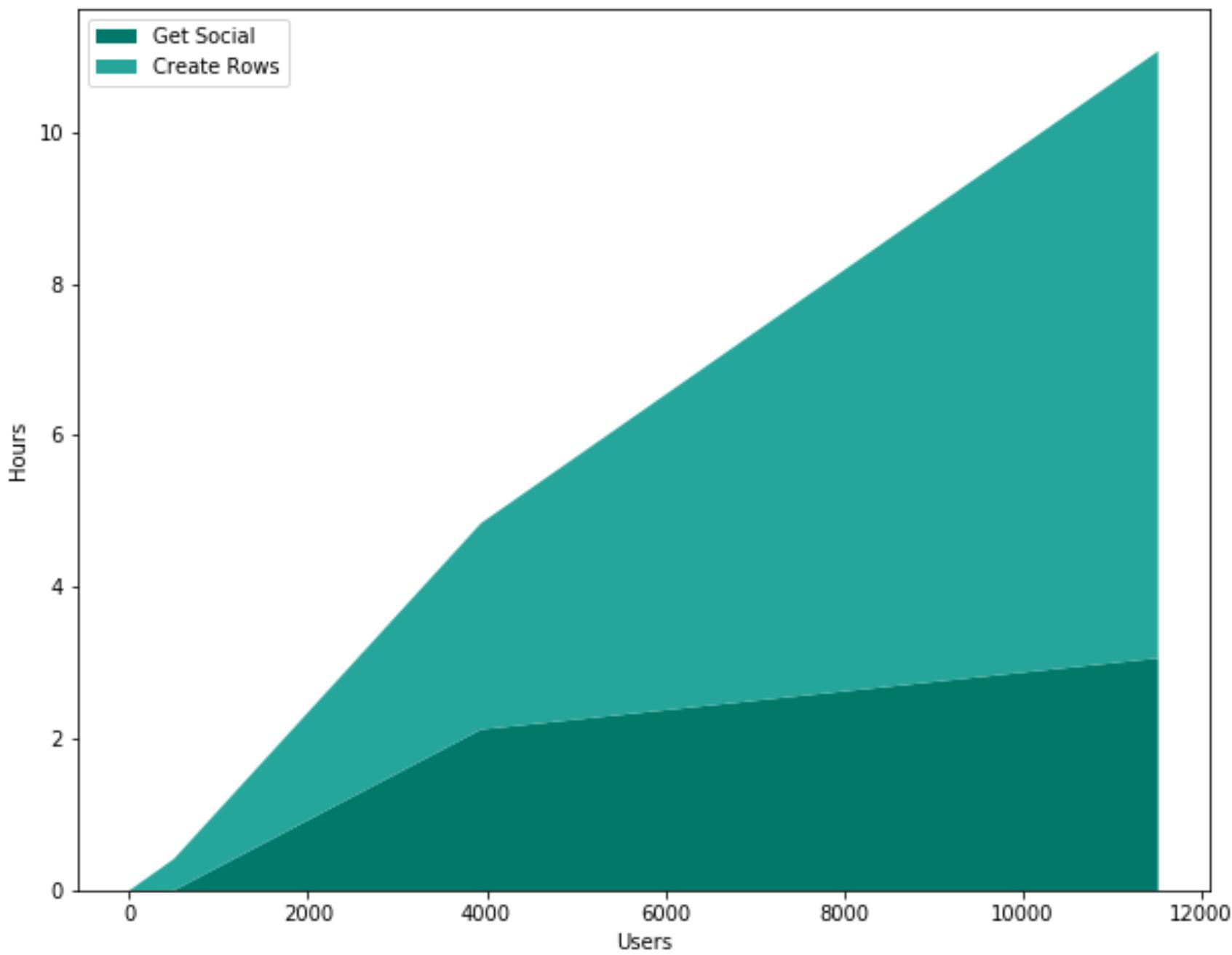
**Usuarios de  
Codewars**

**Web Scraping  
+ API requests**

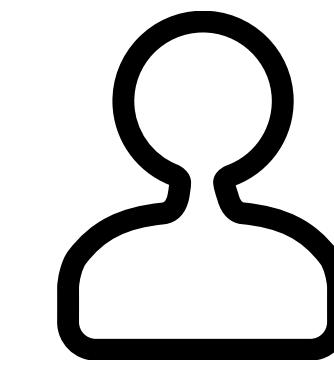
**11500 Registros  
3 Iteraciones**

## | Descripción del Dataset

Coste en Tiempo en cada diferente iteración

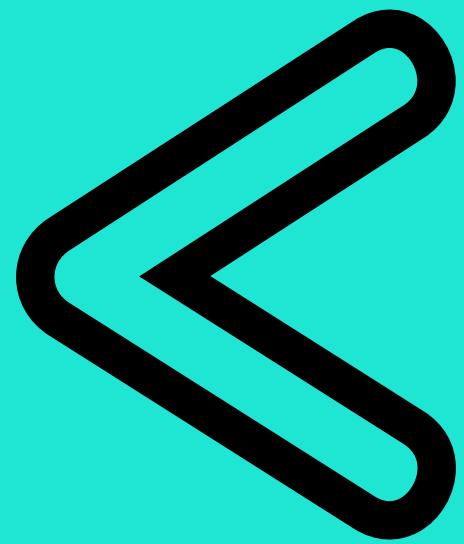
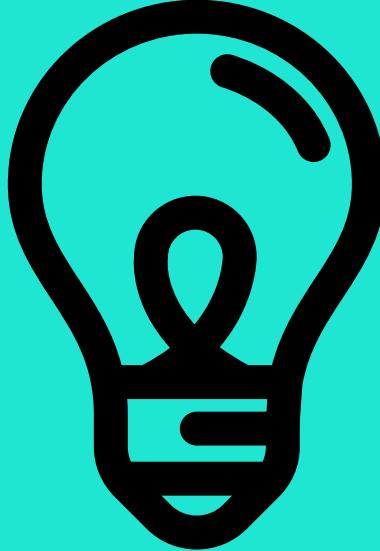


## Descripción del Dataset



0.0+0.4 horas  
2.1+2.7 horas  
3.0+8.0 horas

500  
3939  
11500



"SEP 2018"

"3,101"

"JAVASCRIPT (5 KYU)"

"3 (1 REPLIES)"

# | Data cleaning and wrangling

**Combinar Datos  
de la API y Web  
Scraping**

**Treemap:  
javascript +  
json ad hoc**

**Sistema de  
Recomendación  
Vectorial**

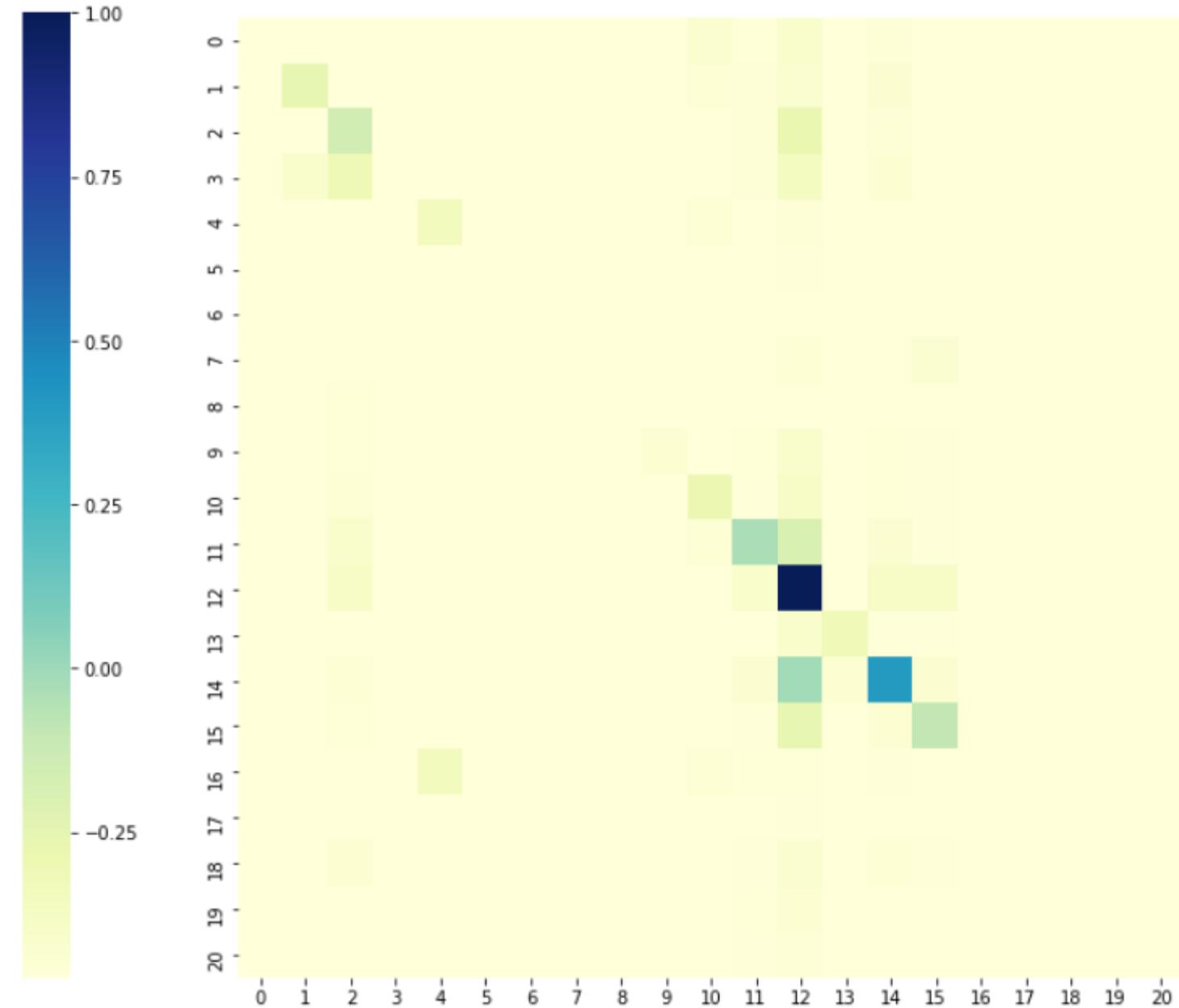
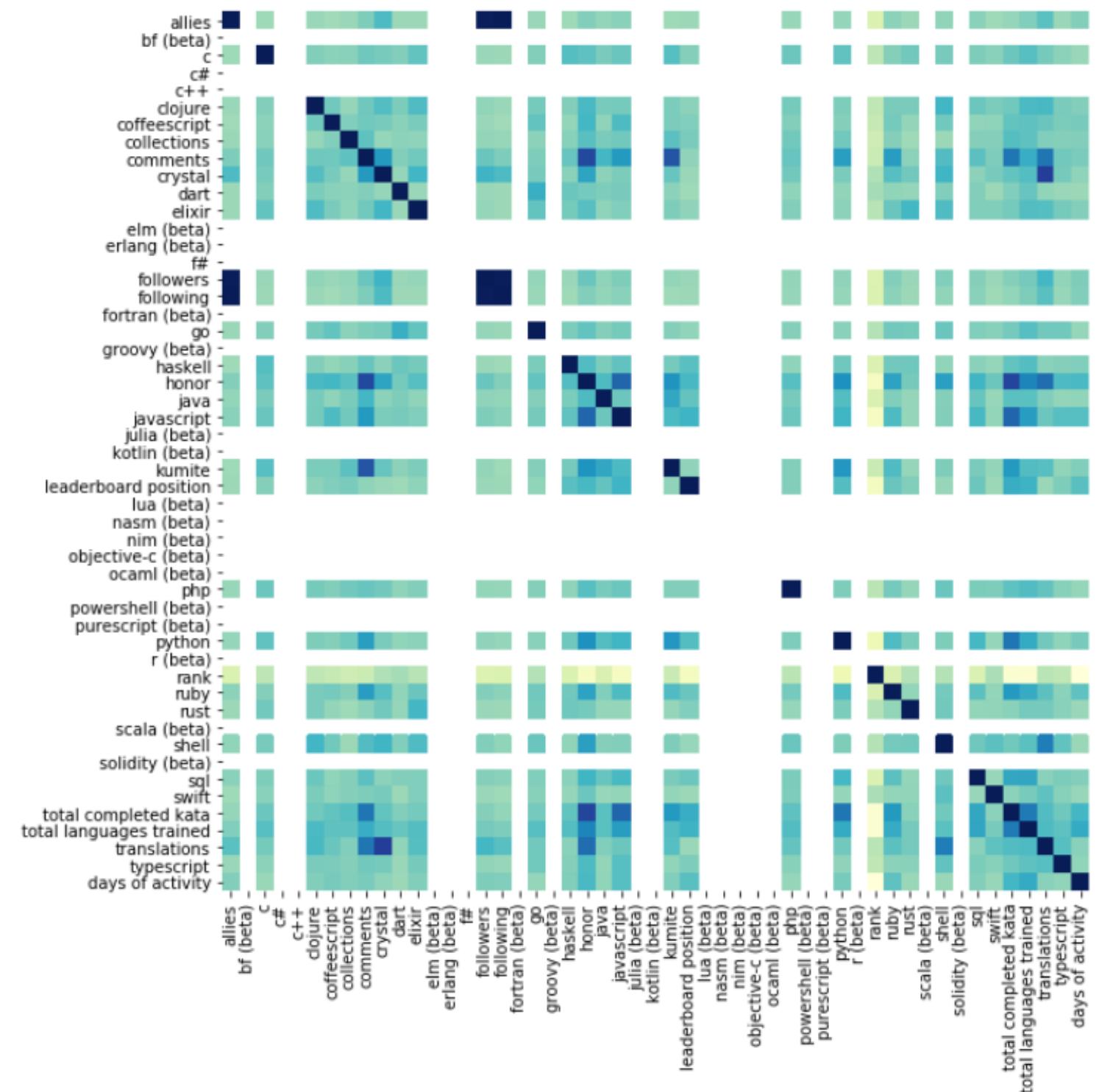
**Mayores retos**

# | Machine Learning

**Correlación  
heatmap**

**Supervised  
Learning:  
score = 0.51**

**Unsupervised  
Learning**



# Correlation y confussion heatmap

# Problemas

**Tiempos de  
ETL**

**Web Scraping  
de una Web  
Dinámica**

**Creación  
de Grafos**

## **Javascript, la gran ganadora**

## **Sistema de recomendación por distancia euclídea**

## **Grafos de relaciones**

# **Resultados**



| d3 Treemap

# Buscamos expertos en: python, java y javascript



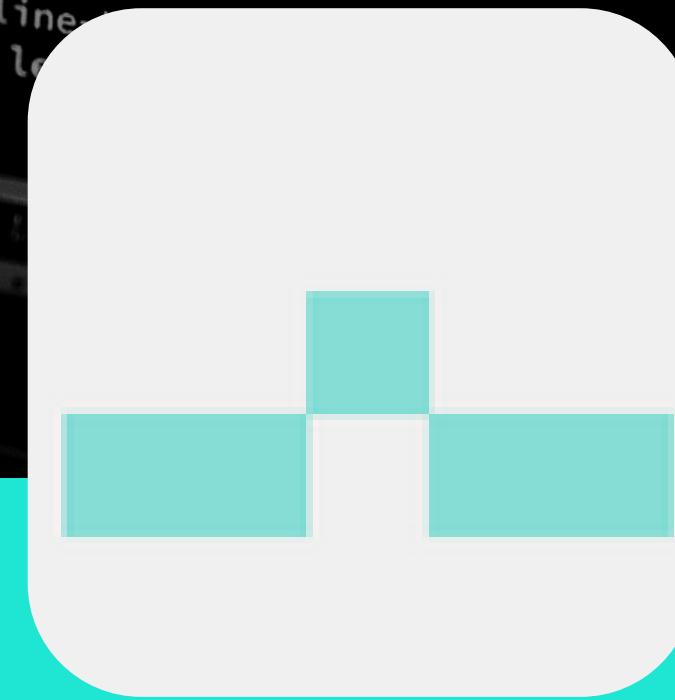
**Blind4Basics**

<https://github.com/Blind4Basics>



**Bubbler**

<https://github.com/Bubbler-4>



**Lyapunov**

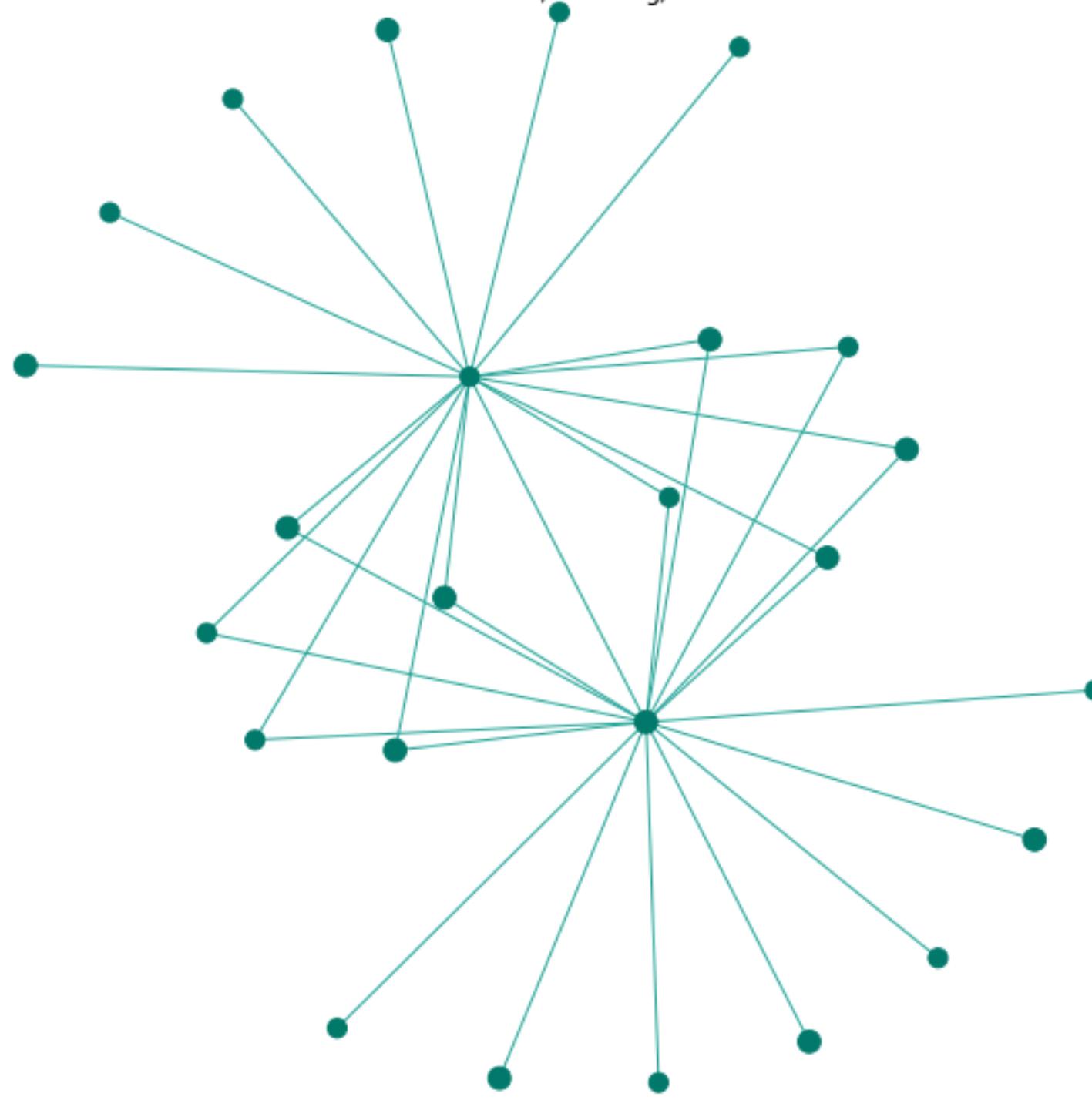
<https://github.com/Lyapunov>

| **Sistema de recomendación**

C	Javascript	Python	Java	R	Ruby
0	1	1	1	0	0
0.1	0.5	0.7	0.9	0	0.1
0.8	0.1	0.1	0.1	0.8	0.9

# | Sistema de recomendación

Grafo de aliados/following/followers

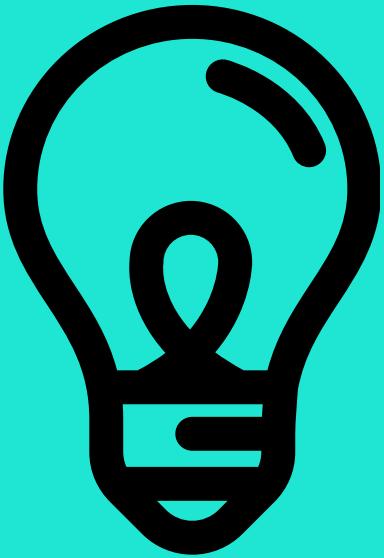


## | Grafos de networkx

# Grafos de networkx

Honor de los aliados





```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import networkx as nx
import requests
import json
import BeautifulSoup
```

| Fin