

# Aplicação de métodos bioinformáticos na análise de séries de temperaturas

João Miguel Afonso  
David A. Cardoso Moreira

# Resumo

- I. Motivação
- II. Caso de Estudo
- III. Análise Estatística
- IV. Discretização
- V. Relação entre Cidades
- VI. Clustering
- VII. Procura de Padrões
- VIII. Conclusões
- IX. Trabalho Futuro

# I - Motivação

- Perceber a adaptabilidade dos **algoritmos bioinformáticos** a séries de dados **contínuos**
- Algoritmos desenvolvidos para ser usados com dados discretos, como as sequências de ADN/proteicas.
  - Ex: procura de motifs

“Global warming is real - it is man-made and it is an important problem. But it is not the end of the world.”

Bjorn Lomborg

## II - Caso de Estudo

The Kaggle logo, consisting of the word "kaggle" in a blue, lowercase, sans-serif font, is centered within a horizontal band. This band has a dark, moody background featuring a sky with white and grey clouds, and a portion of a globe is visible behind the text.

kaggle

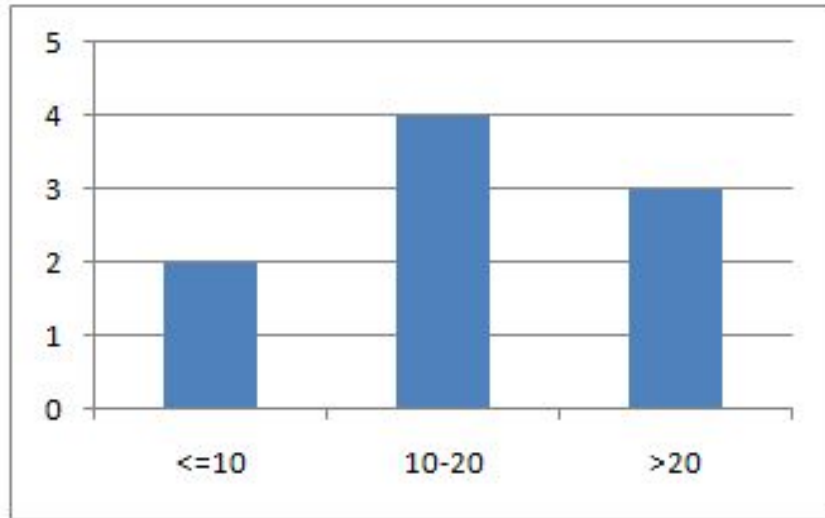
**Climate Change: Earth Surface Temperature Data**

### III - Análise Estatística

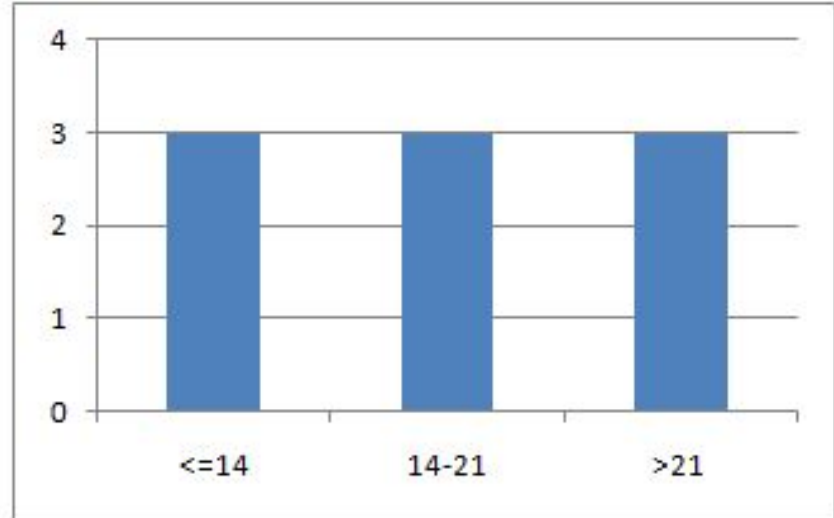
- 15 cidades
- 1521 a 3166 registros por cidade
- Temperatura mínima: -19.37 °C
- Temperatura máxima: 36.34 °C
- Temperatura média: 13.54 °C

## IV - Discretização (EW & EF)

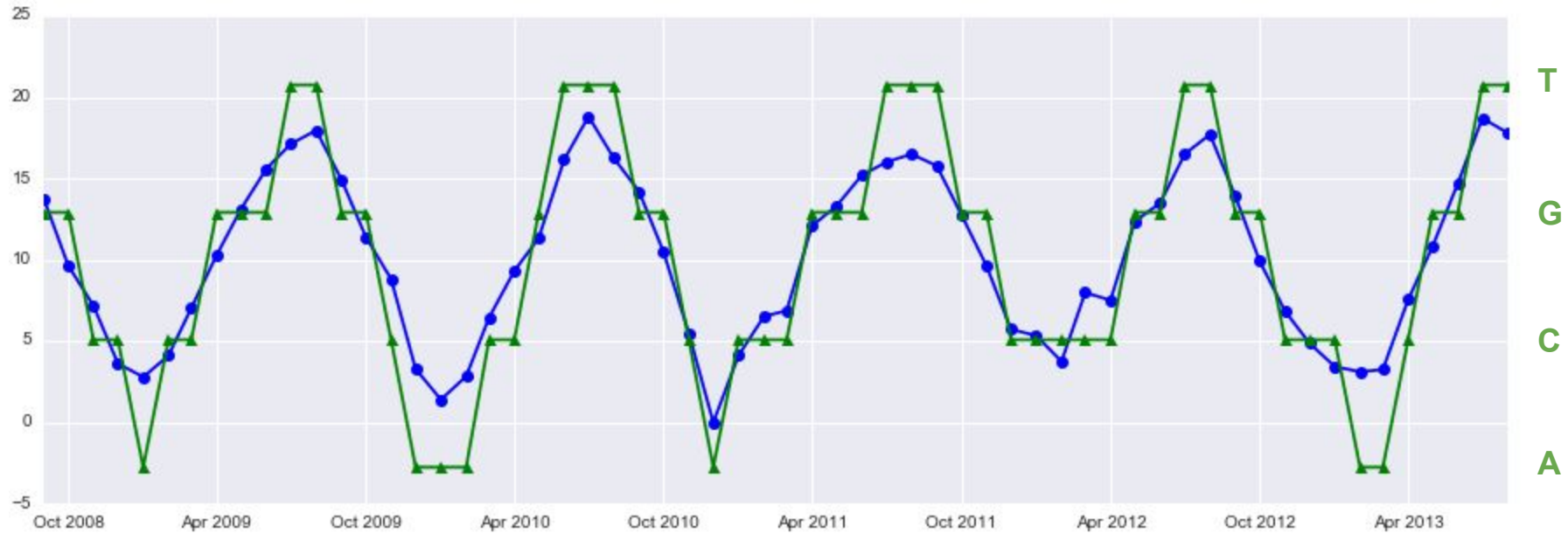
**Equal width**



**Equal frequency**

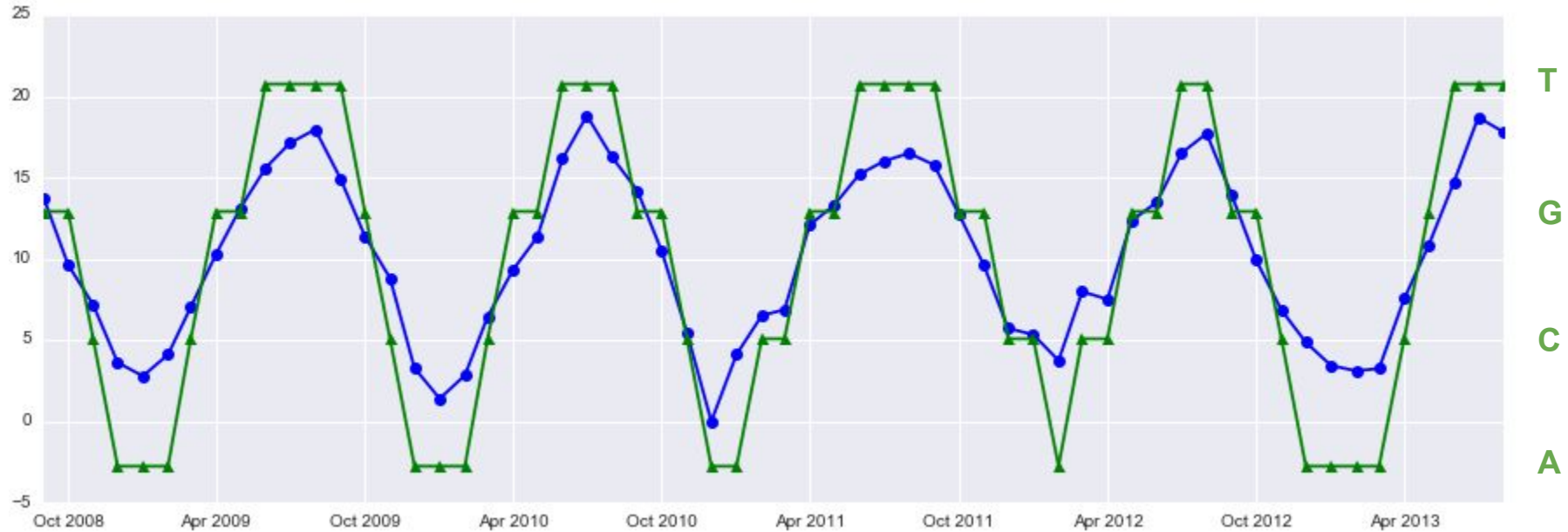


## IV - Discretização (EW)

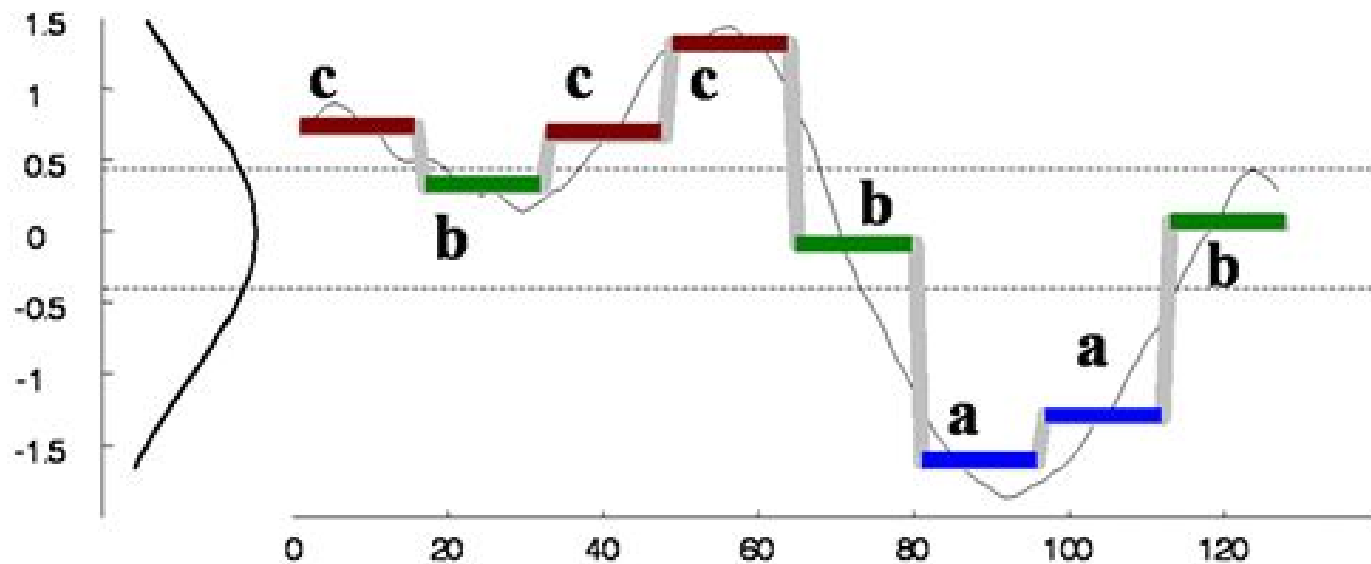




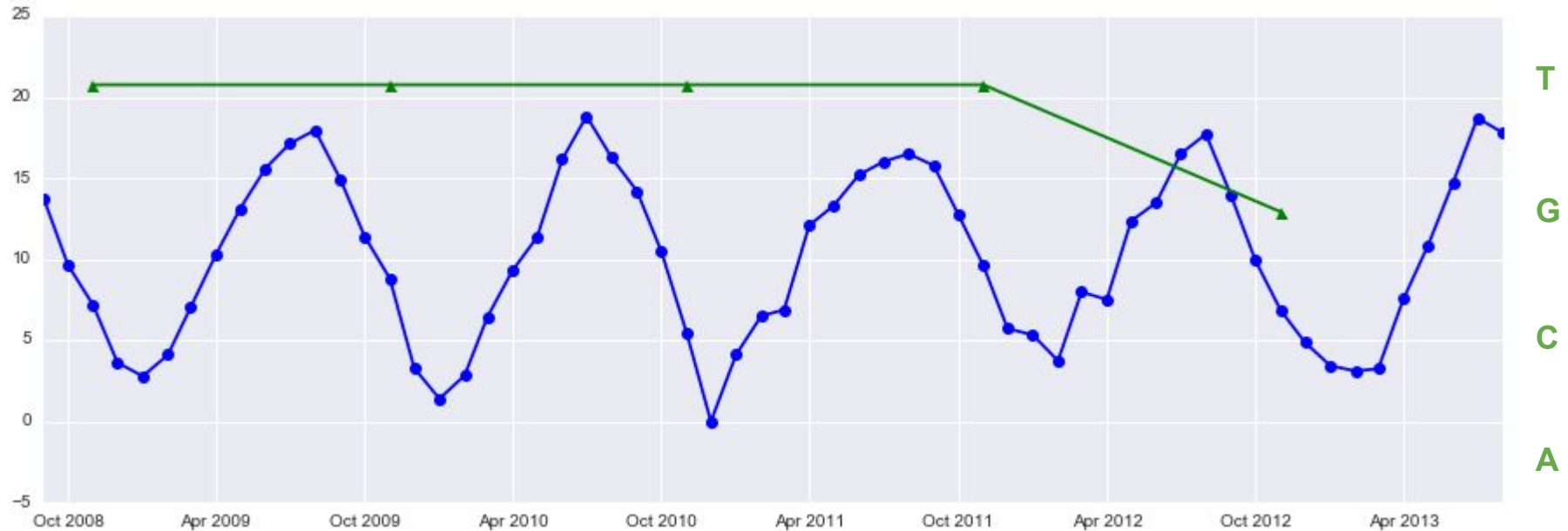
## IV - Discretização (EF)



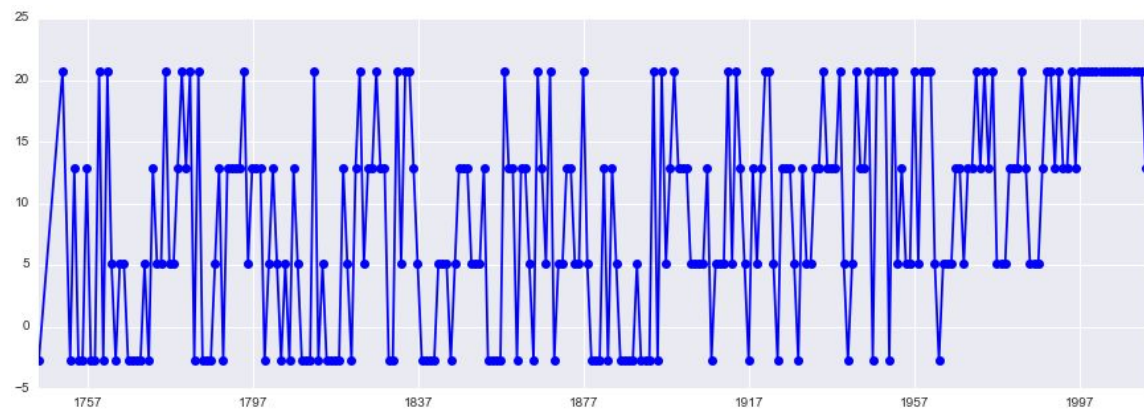
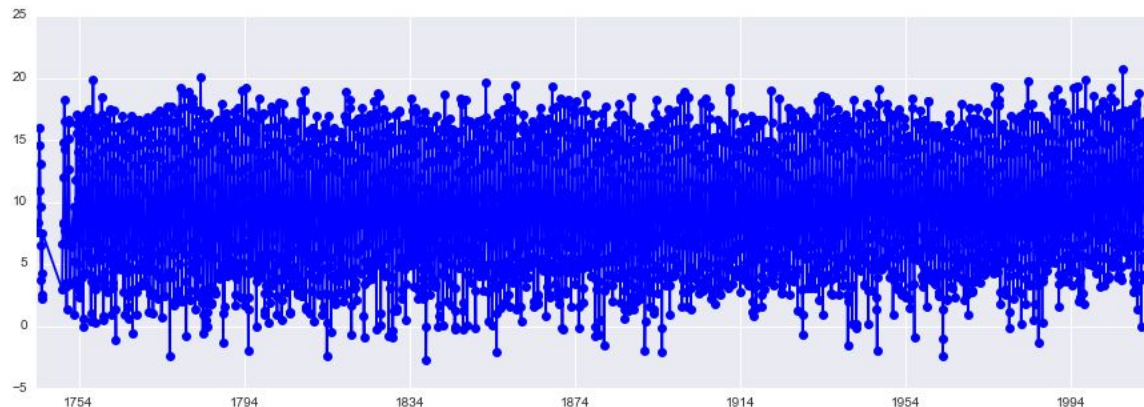
## IV - Discretização (SAX)



## IV - Discretização (SAX)

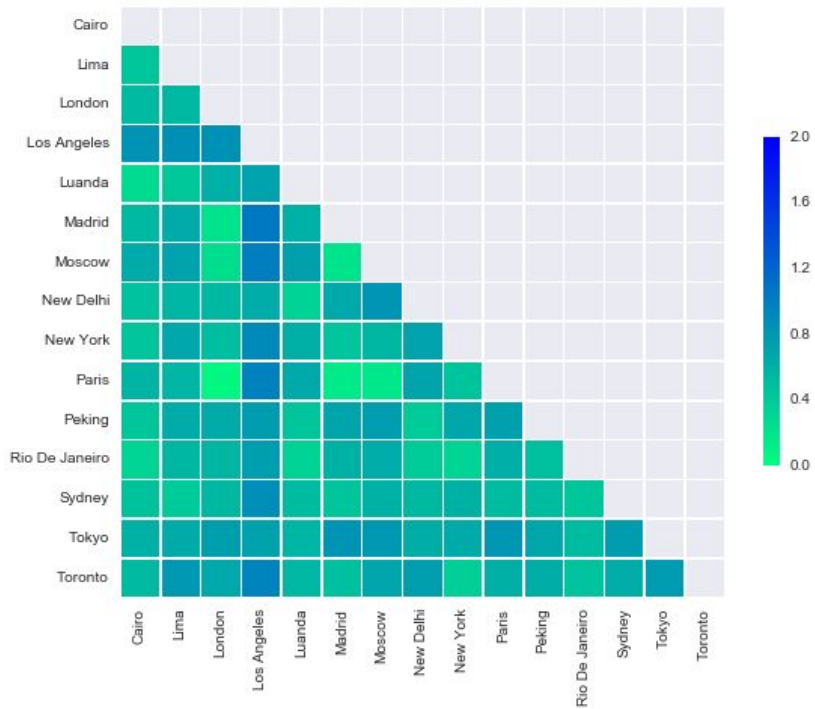
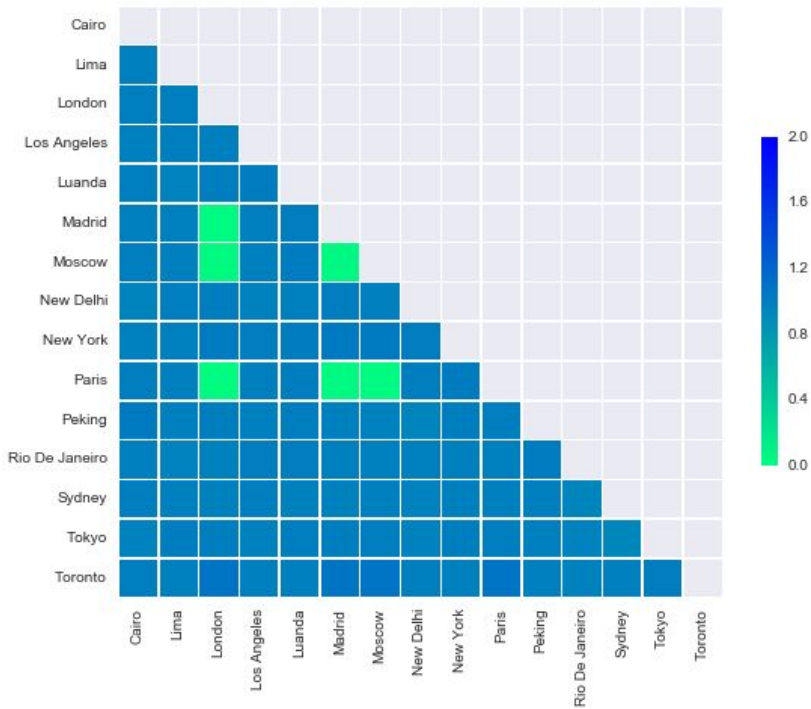


## IV - Discretização (SAX)

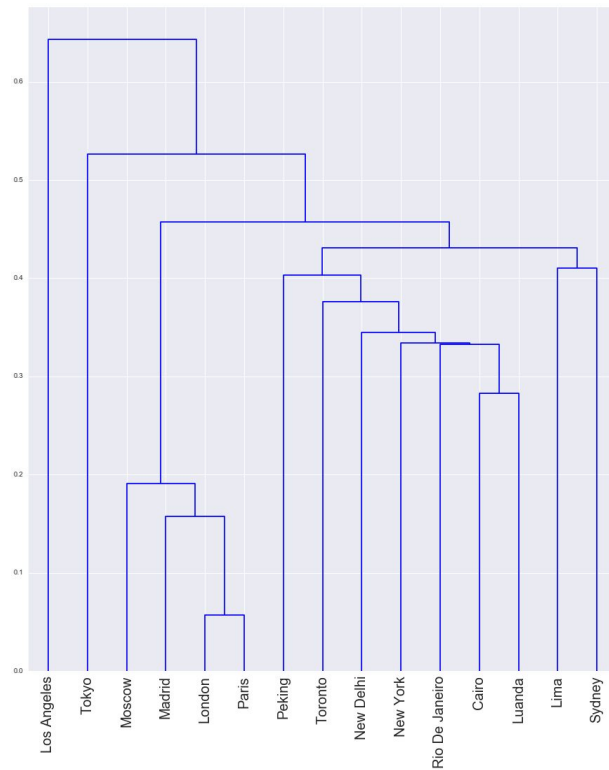
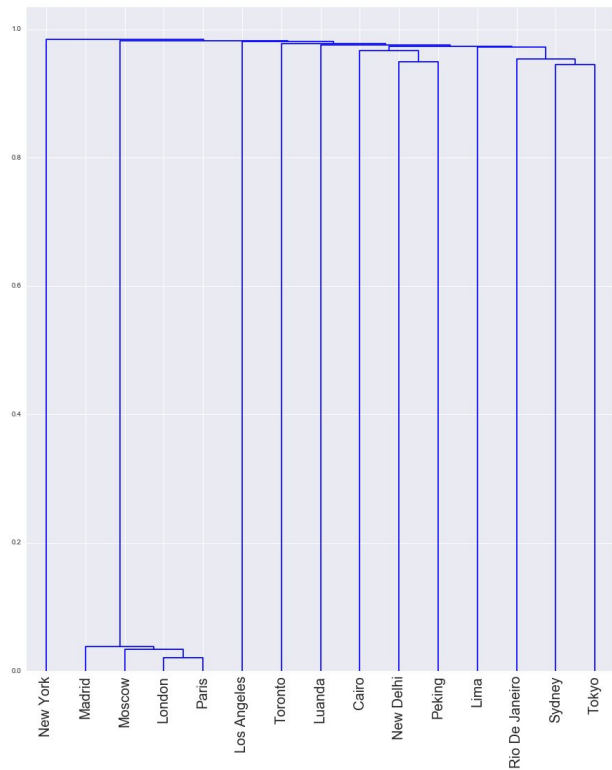


T  
G  
C  
A

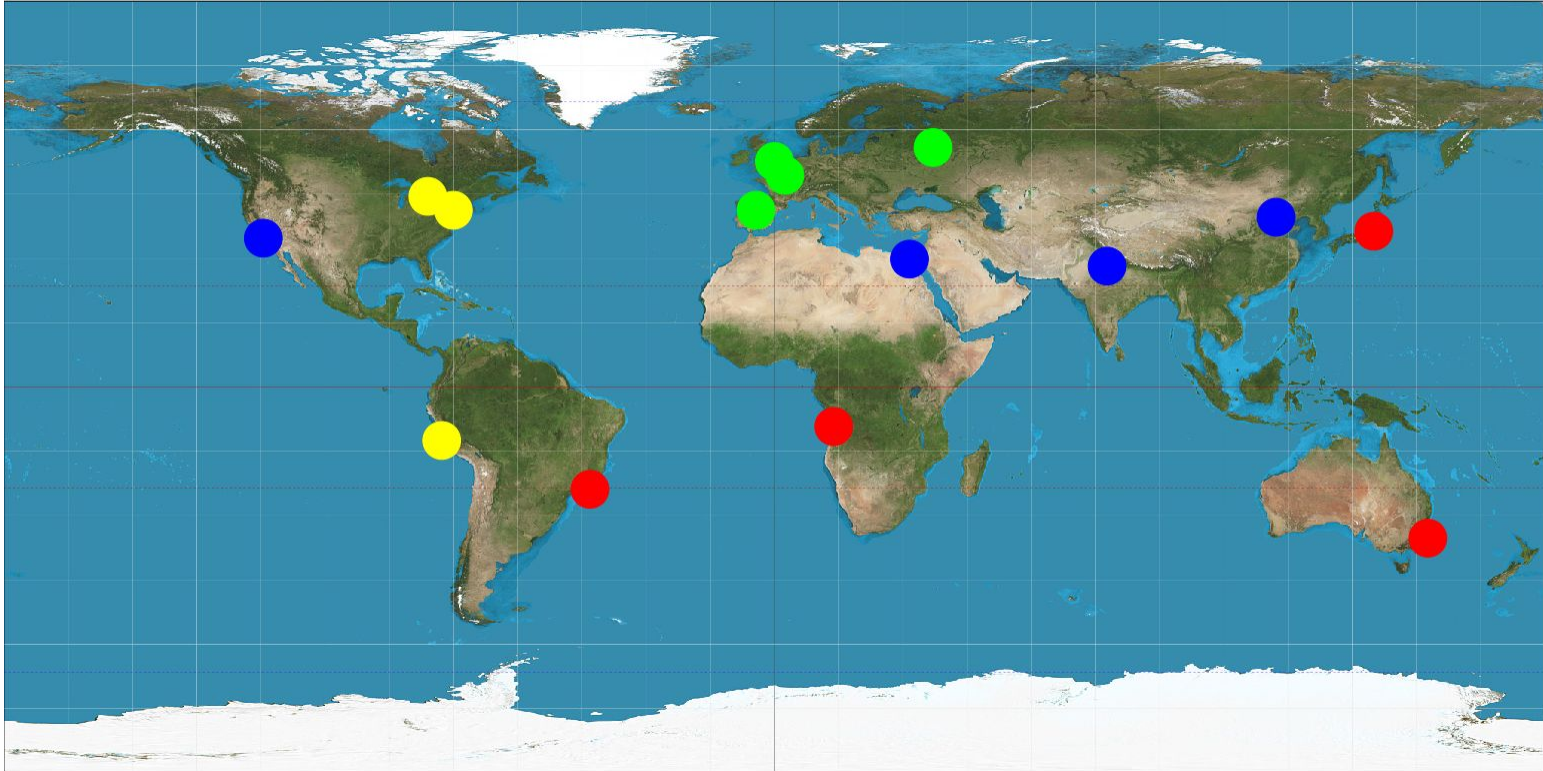
## V - Relação entre Cidades



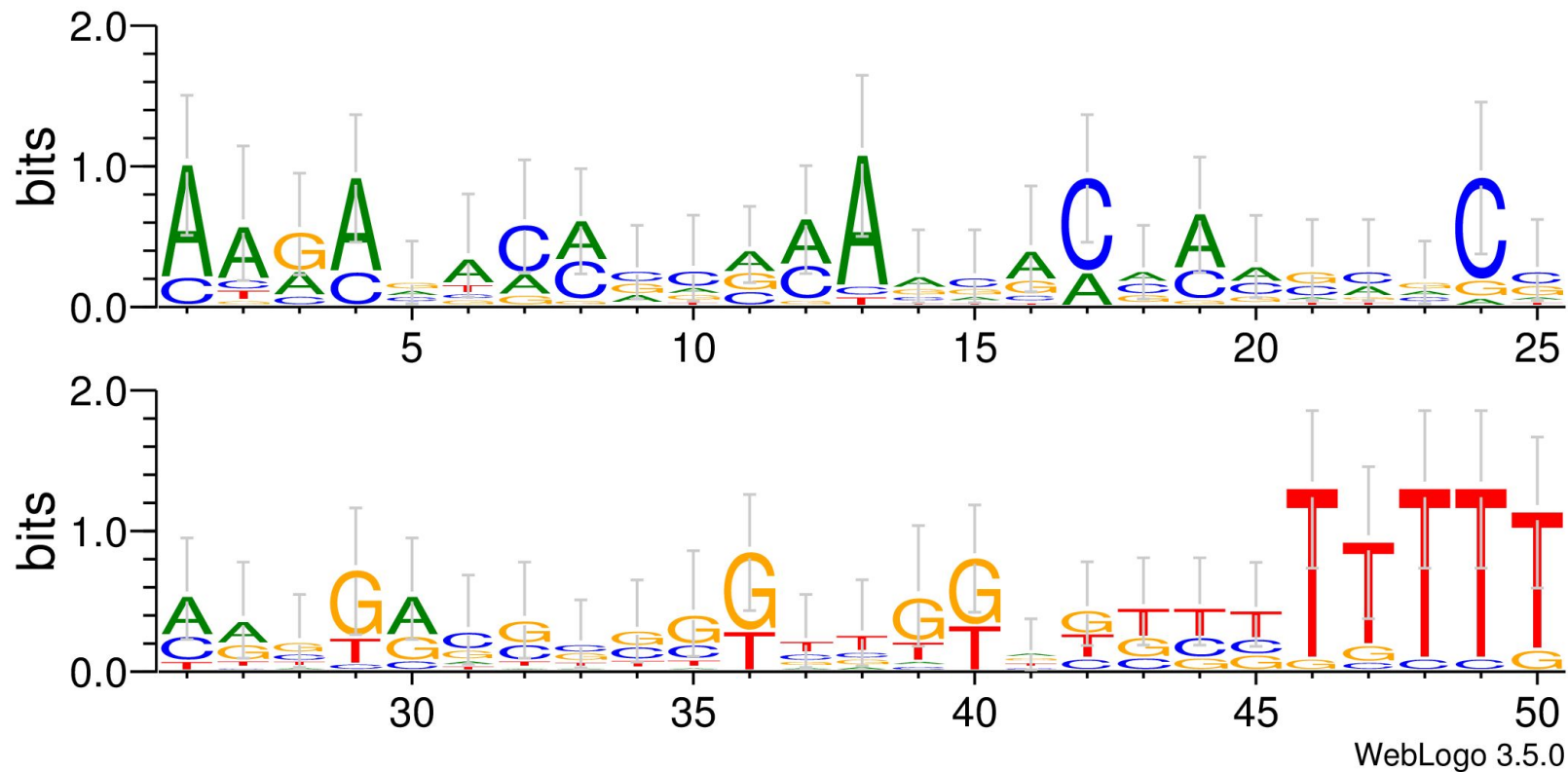
# VI - Clustering (Hierárquico)



## VI - Clustering (K-medoids)

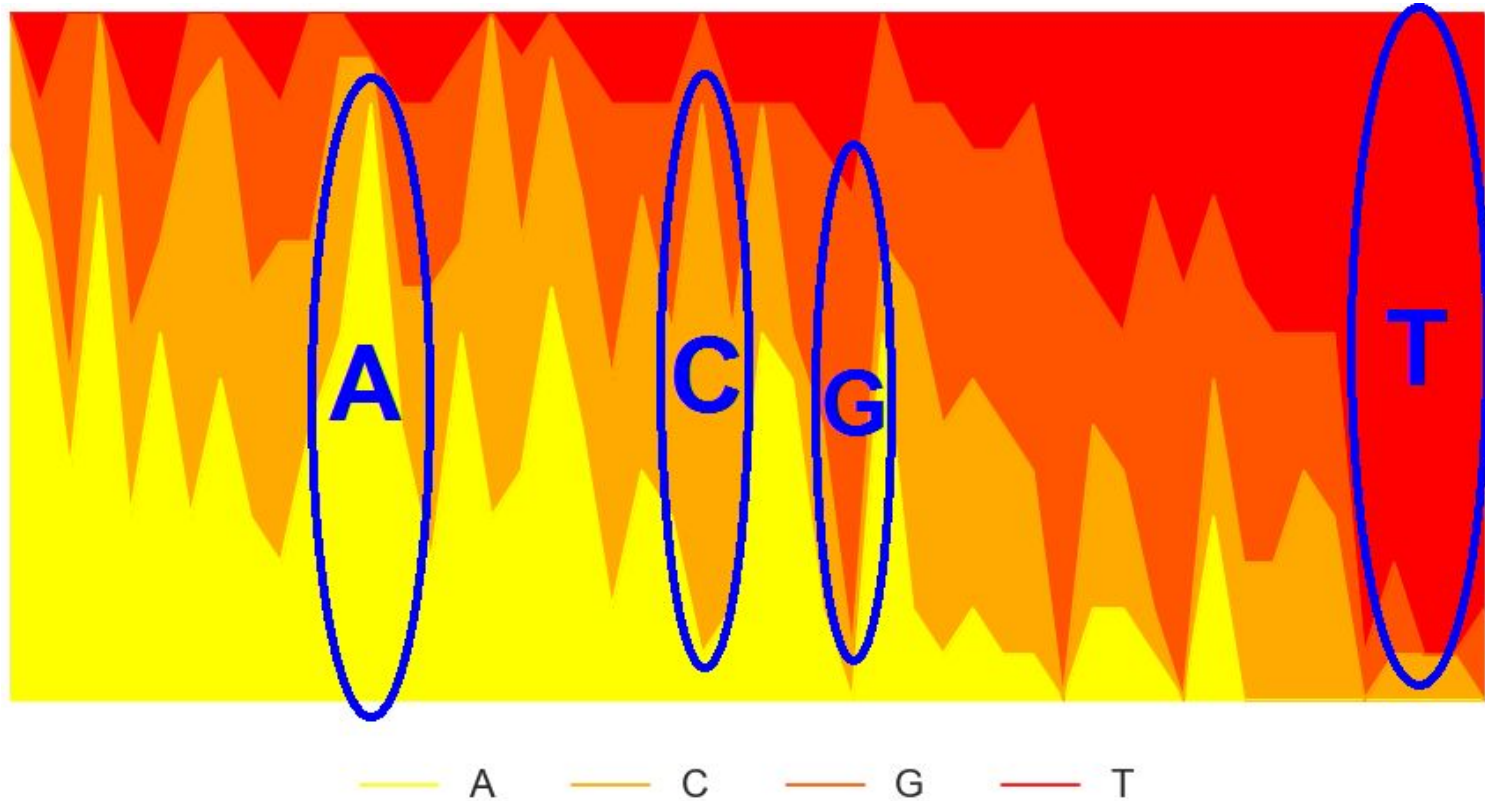


## VII - Procura de Padrões





## VII - Procura de Padrões

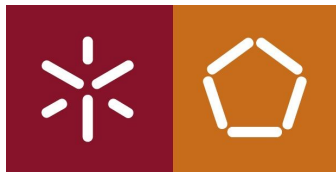


## VIII - Conclusões

- Todas as fases da análise de dados são cruciais para obter bons resultados
- A implementação de vários algoritmos e métodos de visualização não permitiu a exploração exaustiva de cada um deles
- As temperaturas estão a subir!

# IX - Trabalho Futuro

- Limpeza dos dados (previsão valores em falta)
  - Regressão Linear
  - Média aritmética
- Estudo aprofundado de alguns algoritmos
- Procura de conclusões relevantes



# Aplicação de métodos bioinformáticos na análise de séries de temperaturas

João Miguel Afonso  
David A. Cardoso Moreira