



Da Astrofísica à Saúde: a questão interdisciplinar da Ciência de Dados



O que temos para **hoje**?

- ★ Quem sou eu
- ★ Minha trajetória acadêmica e o que aprendi nesses 15 anos
- ★ Como me tornei cientista de dados
- ★ O que a Astrofísica tem em comum com a saúde?
- ★ Algumas considerações sobre o 'ser cientista'

Patricia Novais

1

- ★ Bacharela em Física (USP)
- ★ Mestra em Astrofísica (USP)
- ★ Doutora em Astrofísica (USP)
- ★ Cientista de Dados (SulAmérica Saúde)
- ★ Professora (física, matemática, astrofísica)
- ★ Divulgadora de ciência
- ★ PyLady
- ★ Advisor do Overleaf
- ★ (Tentativa de) Bailarina

Email: patynovais@gmail.com

Minha trajetória acadêmica

E algumas coisinhas que aprendi...



Pares de Galáxias em Fusão

Iniciação Científica

Utilizando critérios matemáticos, seleção de pares de galáxias dentro de um dos maiores levantamentos de dados da Astronomia. Estudos ainda de ambiente, cores, morfologia.

O que aprendi?

- ★ Primeiro contato com Big Data
- ★ Uso de SQL para manejo de dados (MySQL)
- ★ Teste Kolmogorov-Smirnov
- ★ Fortran go





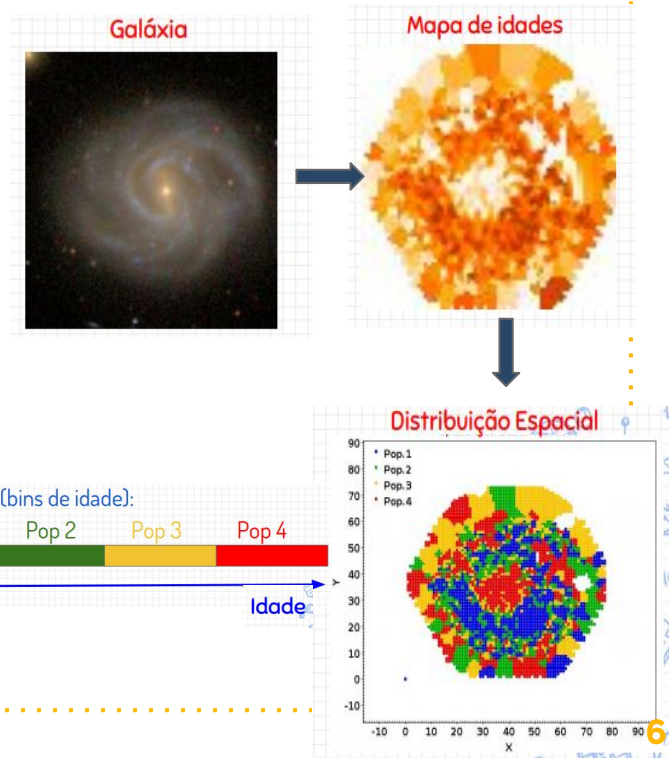
A Distribuição de Populações Estelares em Galáxias

Mestrado

Criação de método de uma nova abordagem para estudar espacialmente as populações estelares em galáxias.

O que aprendi?

- ★ Mais Fortran 90 (rs)
- ★ R básico
- ★ PCA
- ★ Processamento de Imagens
- ★ Coeficiente de Gini
- ★ Funcionais de Minkowski (Característica de Euler Poincaré)





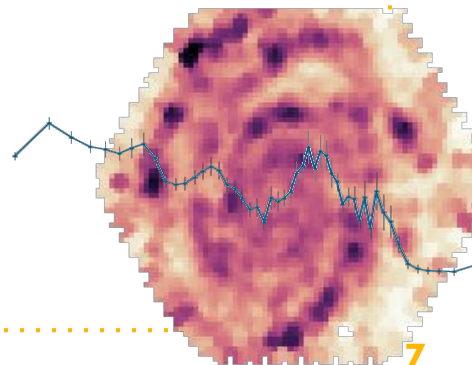
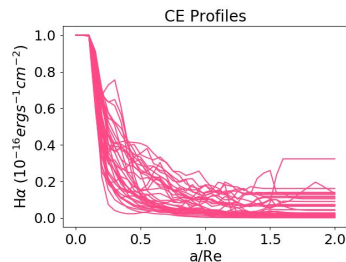
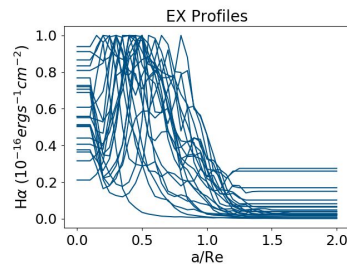
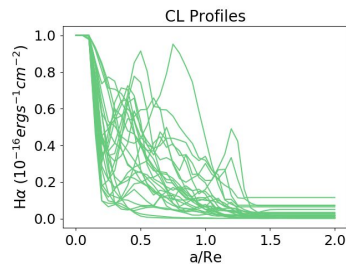
Resolvendo Galáxias: A morfologia das emissões em H α nas galáxias do projeto CALIFA

Doutorado

Estudo das emissões em H α em galáxias do projeto CALIFA, propondo um novo sistema de classificação baseado nessa medida.

O que aprendi?

- ★ Python (yes!!)
- ★ Processamento de Datacubes
- ★ Voronoi Tessellation e afins
- ★ Momentos da Imagem
- ★ Momentos invariantes de Hu
- ★ Clusterização (K-means e Mclust)



Como me tornei cientista de dados

Um pouco de *serendipity*...

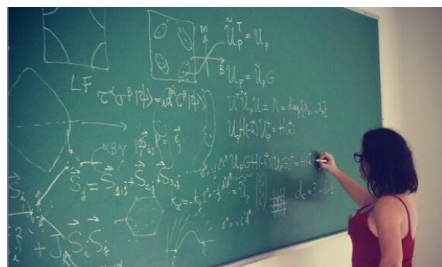
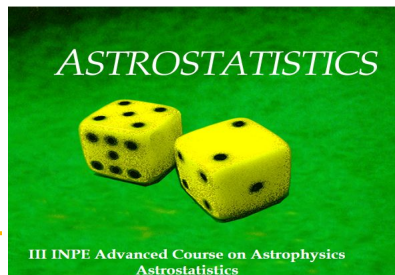
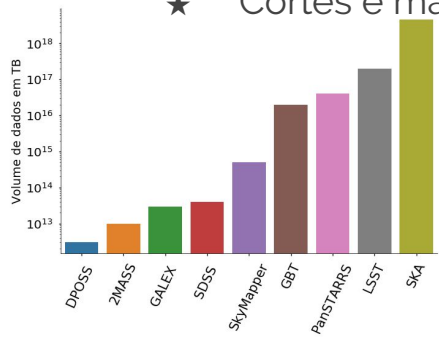


Um pouco de serendipidade ou um caminho natural?

Além da própria experiência dentro da universidade, alguns eventos e acontecimentos acabaram me direcionando para a Ciência de Dados...

Alguns fatos...

- ★ Trabalhar com dados de surveys e não com astronomia observacional
- ★ III INPE Advanced Course on Astrophysics Astrostatistics (2009)
- ★ Workshop de Ciência de Dados do IAG (2015 e 2016)
- ★ Entrar para as PyLadies e me envolver mais com tecnologia
- ★ Ter sido professora em diversos níveis educacionais
- ★ Ter feito parte de uma campanha de empoderamento feminino da Samsung
- ★ Cortes e mais cortes nos recursos destinado à pesquisa



4

Ok. Tudo muito legal.

*Mas... E o que a Astrofísica tem
em comum com a
saúde/operadoras de saúde?*



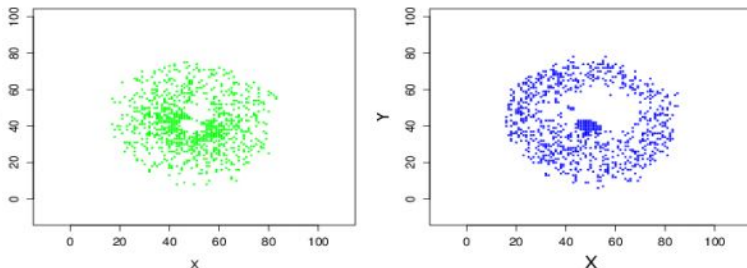


Alguns exemplos...

Característica de Euler-Poincaré

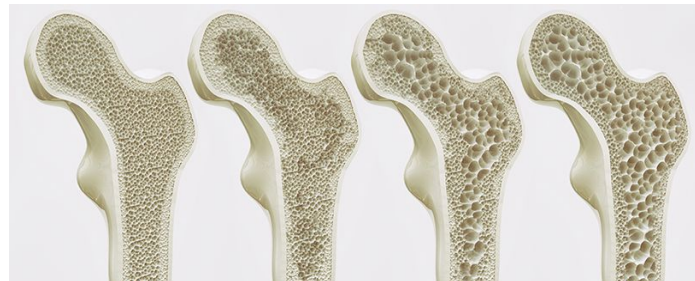
Na **Astronomia**...

Útil para determinar se uma população estelar tem distribuição espacial mais 'porosa' ou mais homogênea.



Na **Saúde**...

Usado para determinar o nível de porosidade nos ossos, útil para estudos de Osteoporose.



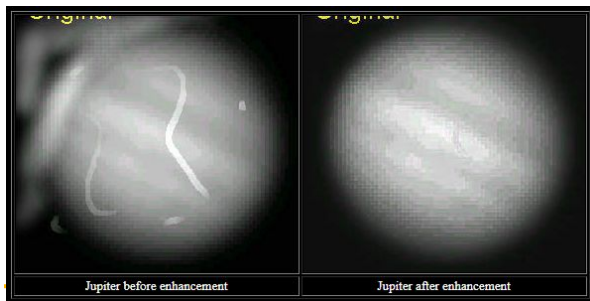


Alguns exemplos...

Processamento de Imagens

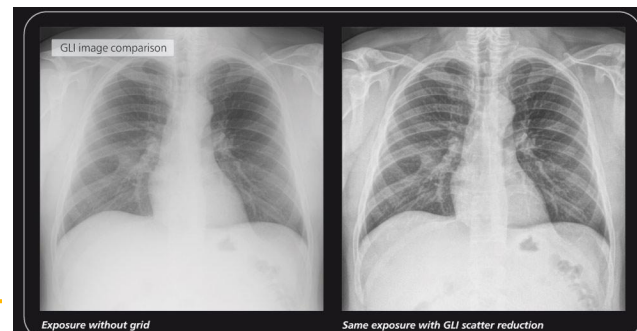
Na **Astronomia**...

Necessário para tornar as imagens astronômicas mais assertivas e obter medidas das imagens (fotometria)



Na **Saúde**...

Utilizada para melhorar as imagens médicas, favorecendo melhores diagnósticos e até automatização.

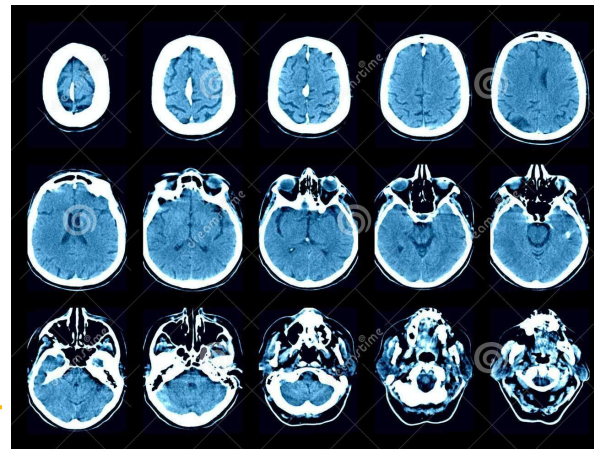
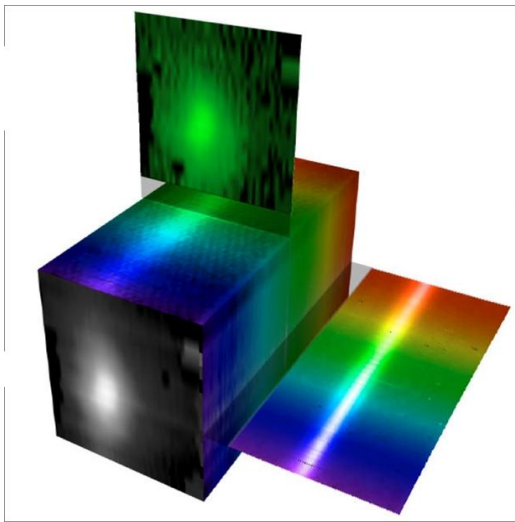




Alguns exemplos...

Análise de Datacubes

Assim como na **Saúde**, onde a tomografia nos permite estudar as estruturas através de cortes, na **Astronomia** os cubos de dados mostram 'seções' diferentes de uma galáxia, de acordo com o comprimento de onda observado.





Alguns exemplos em outras áreas...

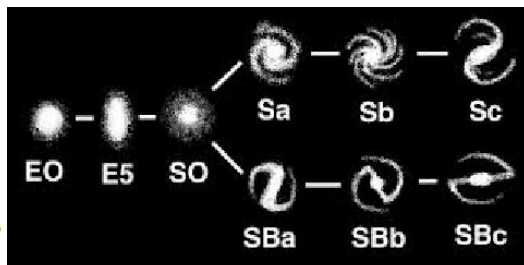
Métodos de Classificação

Na **Astronomia**...

Métodos de classificação e outros métodos são utilizados para encontrar padrões entre as galáxias.

Nas **Operadoras de Saúde**...

Podem ser úteis, por exemplo, para classificar os perfis de usuários mais propensos ao uso de planos de saúde...





Alguns exemplos em outras áreas...

- ★ **Isolation forest** e outros métodos são utilizados para tanto para encontrar 'weird galaxies' quanto fraudes bancárias.
- ★ Separação estrela-galáxia podem utilizar **regressão logística**, assim como encontrar se um dado consumidor tem mais ou menos chances de comprar um produto
- ★ **Coeficiente de Gini**, aplicado em economia para medir desigualdades, também é útil para estudar morfologias em galáxias.

5

Algumas considerações finais....

Afinal, o que é ser um cientista?

(de dados ou não)



O que é um cientista, afinal?

- ★ Ser cientista é mais do que saber rodar modelos



O que é um cientista, afinal?

- ★ Ser cientista é mais do que saber rodar modelos
- ★ É ter a capacidade de criar ferramentas de acordo com o problema



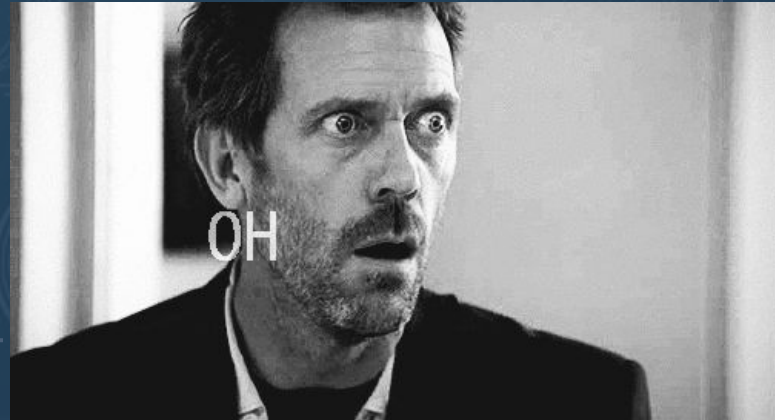
O que é um cientista, afinal?

- ★ Ser cientista é mais do que saber rodar modelos
- ★ É ter a capacidade de criar ferramentas de acordo com o problema
- ★ É conseguir olhar o todo e a parte



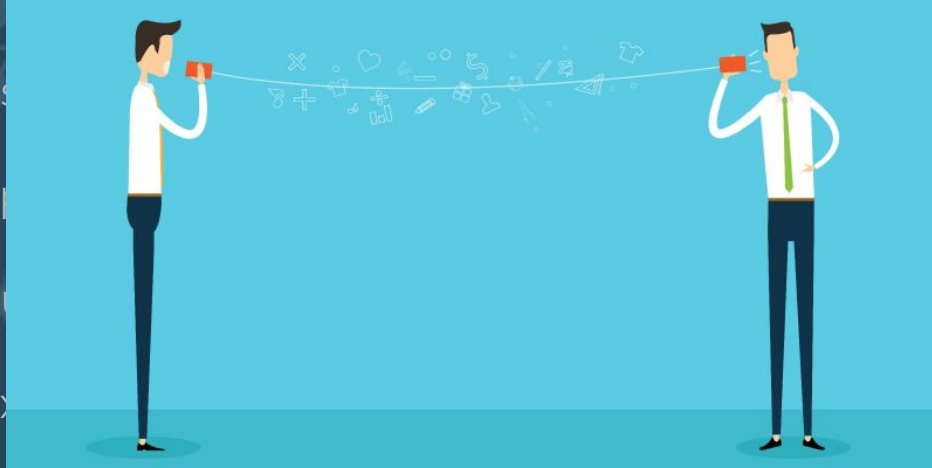
O que é um cientista, afinal?

- ★ Ser cientista é mais do que saber rodar modelos
- ★ É ter a capacidade de criar ferramentas de acordo com o problema
- ★ É conseguir olhar o todo e a parte
- ★ **É saber extrair insights, ter aqueles momentos “Dr. House”**



O que é um cientista, afinal?

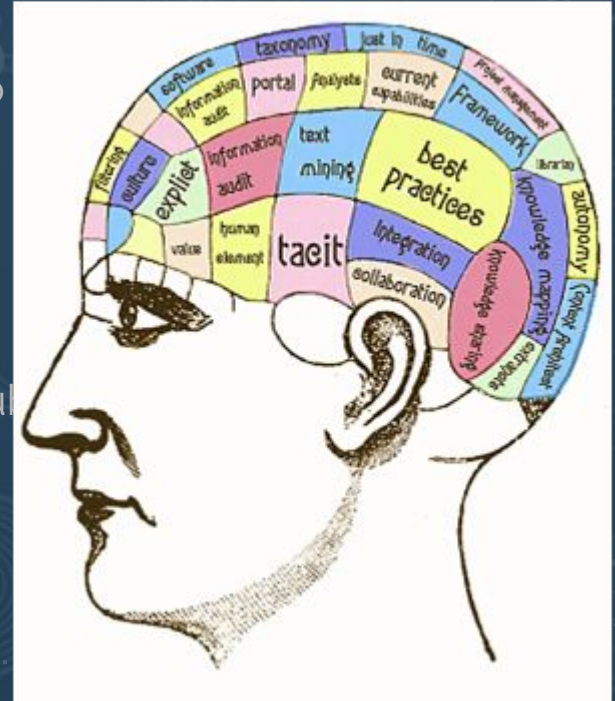
- ★ Ser cientista
- ★ É ter a capacidade de resolver um problema
- ★ É conseguir transmitir conhecimento para qualquer público
- ★ É saber explicar
- ★ É saber transmitir conhecimento para qualquer público



problema

O que é um cientista, afinal?

- ★ Ser cientista é mais do que saber rodar modelos
- ★ É ter a capacidade de criar ferramentas de acordo
- ★ É conseguir olhar o todo e a parte
- ★ É saber extrair insights, ter aqueles momentos “Dr.
- ★ É saber transmitir conhecimento para qualquer públ
- ★ **É ter o maior domínio possível do problema**



O que é um cientista, afinal?

- ★ Ser capaz de lidar com o problema
- ★ É ter a capacidade de transitar entre diversas áreas

O que é um cientista, afinal?

- ★ Ser cientista é mais do que saber rodar modelos
- ★ É ter a capacidade de criar ferramentas de acordo com o problema
- ★ É conseguir olhar o todo e a parte
- ★ É saber extrair insights, ter aqueles momentos “Dr. House”
- ★ É saber transmitir conhecimento para qualquer público
- ★ É ter domínio do problema
- ★ É ter a capacidade de transitar entre diversas áreas



OBRIGADA!

