

---

---

# Python & Data Science

— Mariana Lopes —  
mariana.lobes@bravi.com.br  
2105.mari@gmail.com

---

---

Eu

Mariana Vieira Ribeiro Lopes

Bacharel em Ciência da Computação (Turma 2008) - UFSCar

Mestre em Ciência da Computação (Inteligência Artificial) - UFSCar

Data Scientist - Bravi (Desde agosto/2015)

Tutora do curso de BSI da EaD - UFSCar

# DATA SCIENTIST

Sexiest job of the 21<sup>st</sup> century

Harvard Business Review



**IDF15**  
INTEL DEVELOPER FORUM



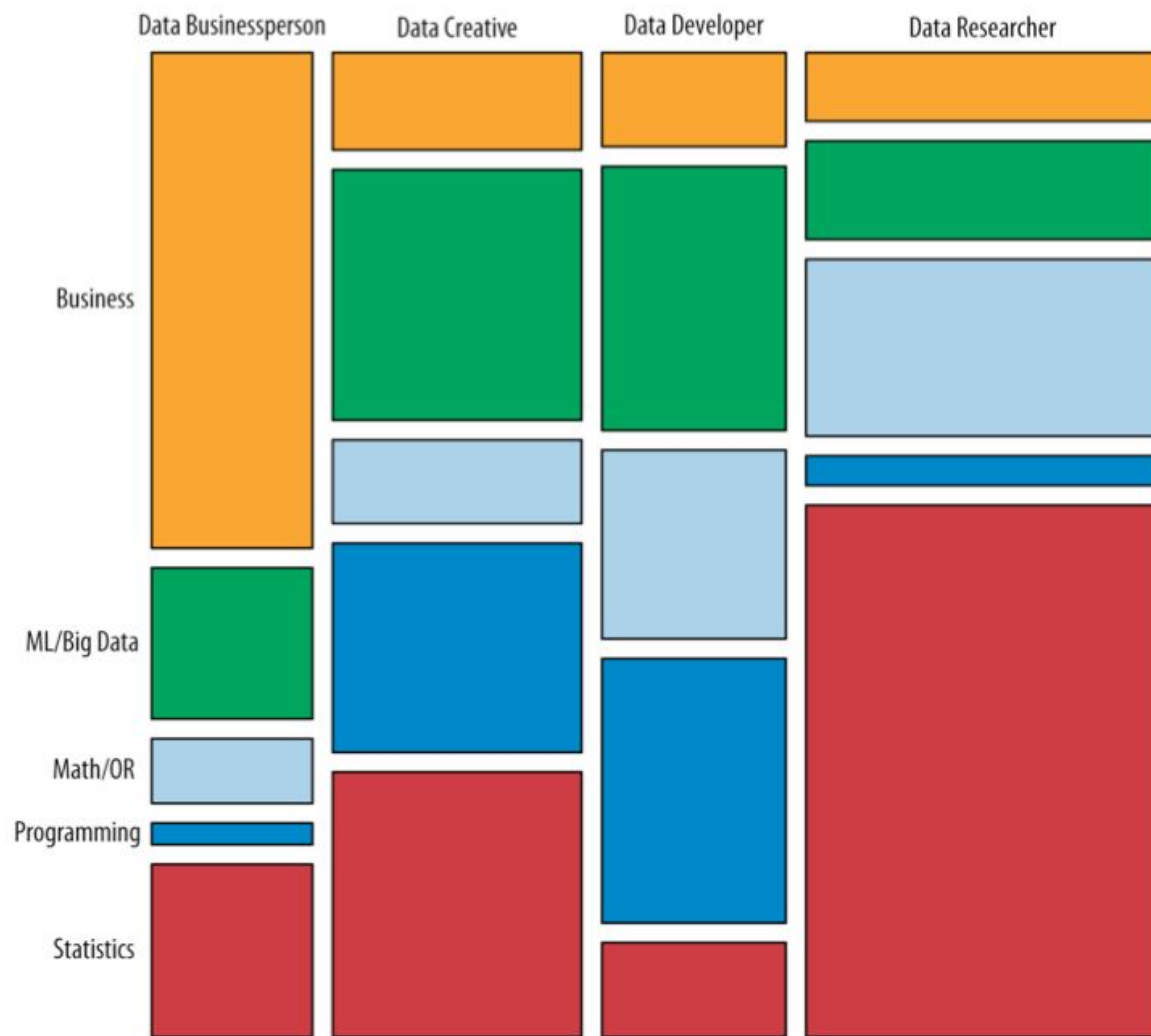
# Data Scientists

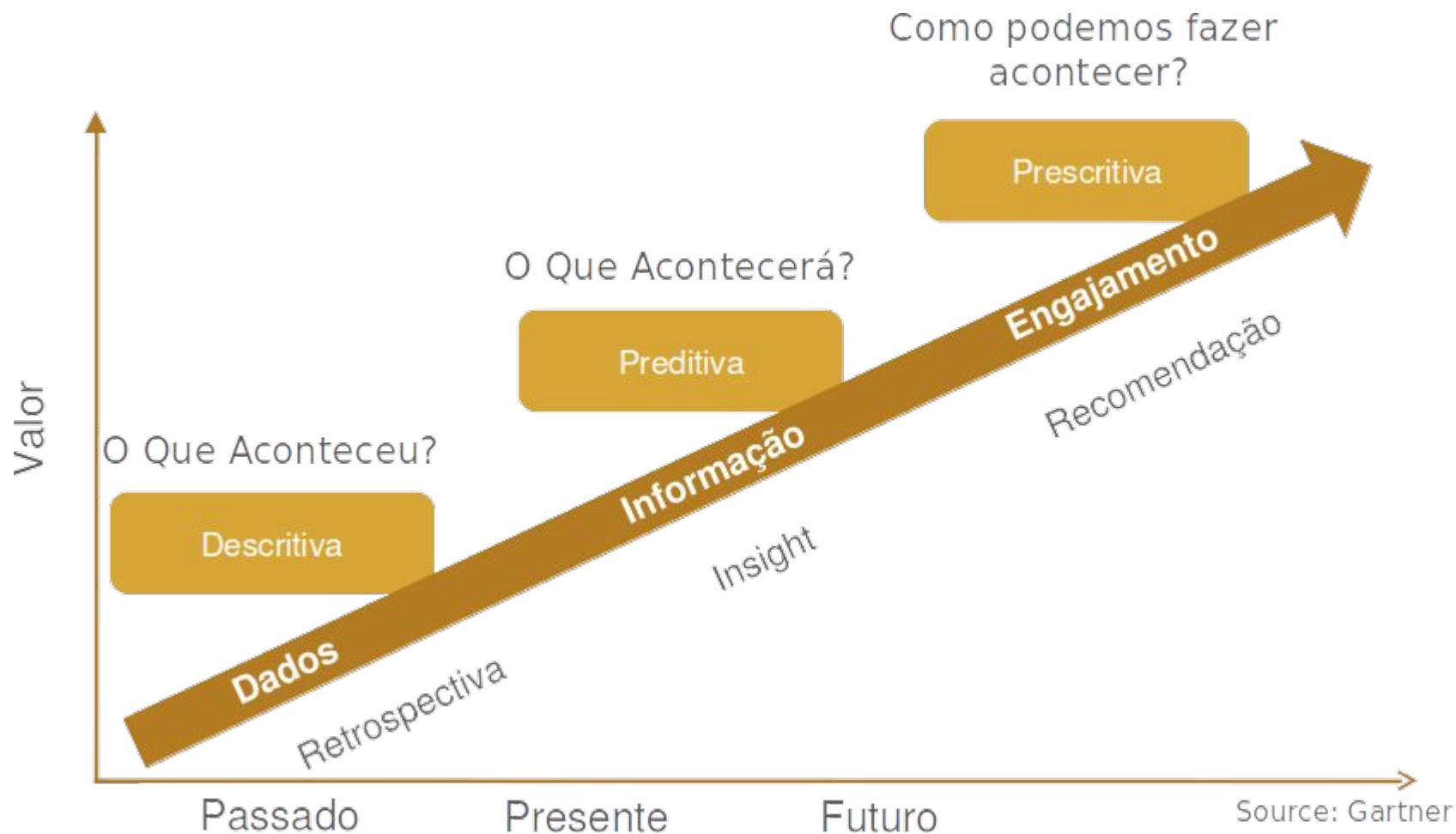
Quem são?  
O que fazem?  
Onde vivem?



Data Literacy:  
“Analfabetismo de dados”

“Cientistas de Dados – profissionais  
**criativos** com um mix de habilidades em  
**codar, estatística e negócios**”  
[NESTA UK].







**O que eu faço hoje?**

# Learning Analytics

- “ Descobrir padrões escondidos nos dados educacionais e utilizá-los a fim de entender melhor o processo educacional, o aprendizado e perfil dos estudantes.”
- Beneficia estudantes e instituições
- Modelos Preditivos

V1	V2	V3	Evadiu
...	...	...	Sim
...	...	...	Não

Dados Históricos  
Crus

Limpeza  
Sanitização  
Engenharia de Features

F1	F2	F3	Evadiu
...	...	...	Sim
...	...	...	Não

Dados Históricos  
Processados

Treinamento

Modelo  
Preditivo

Dados Correntes  
Crus

V1	V2	V3
...	...	...
...	...	...



Limpeza  
Sanitização

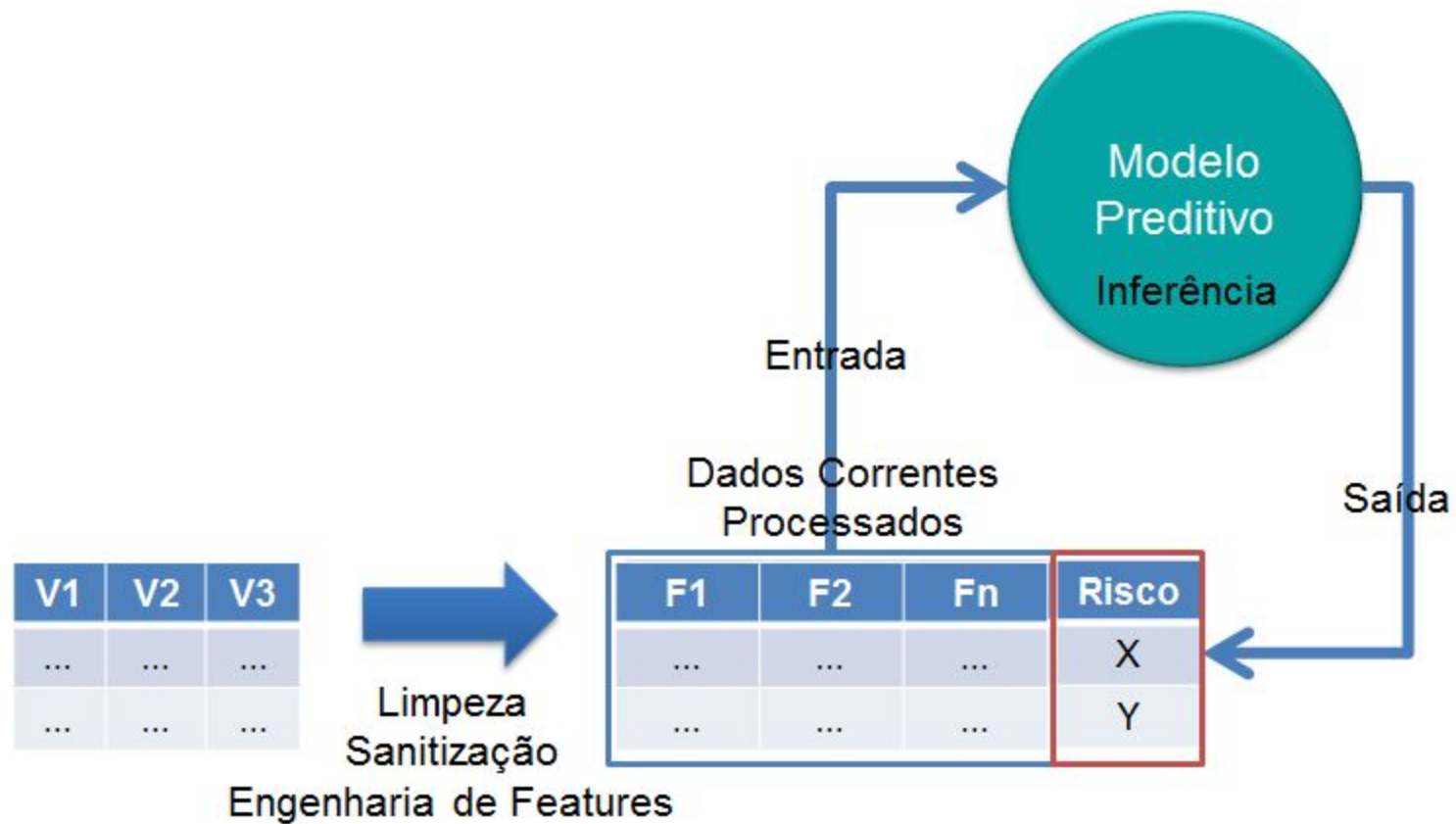
Engenharia de Features

Dados Correntes  
Processados

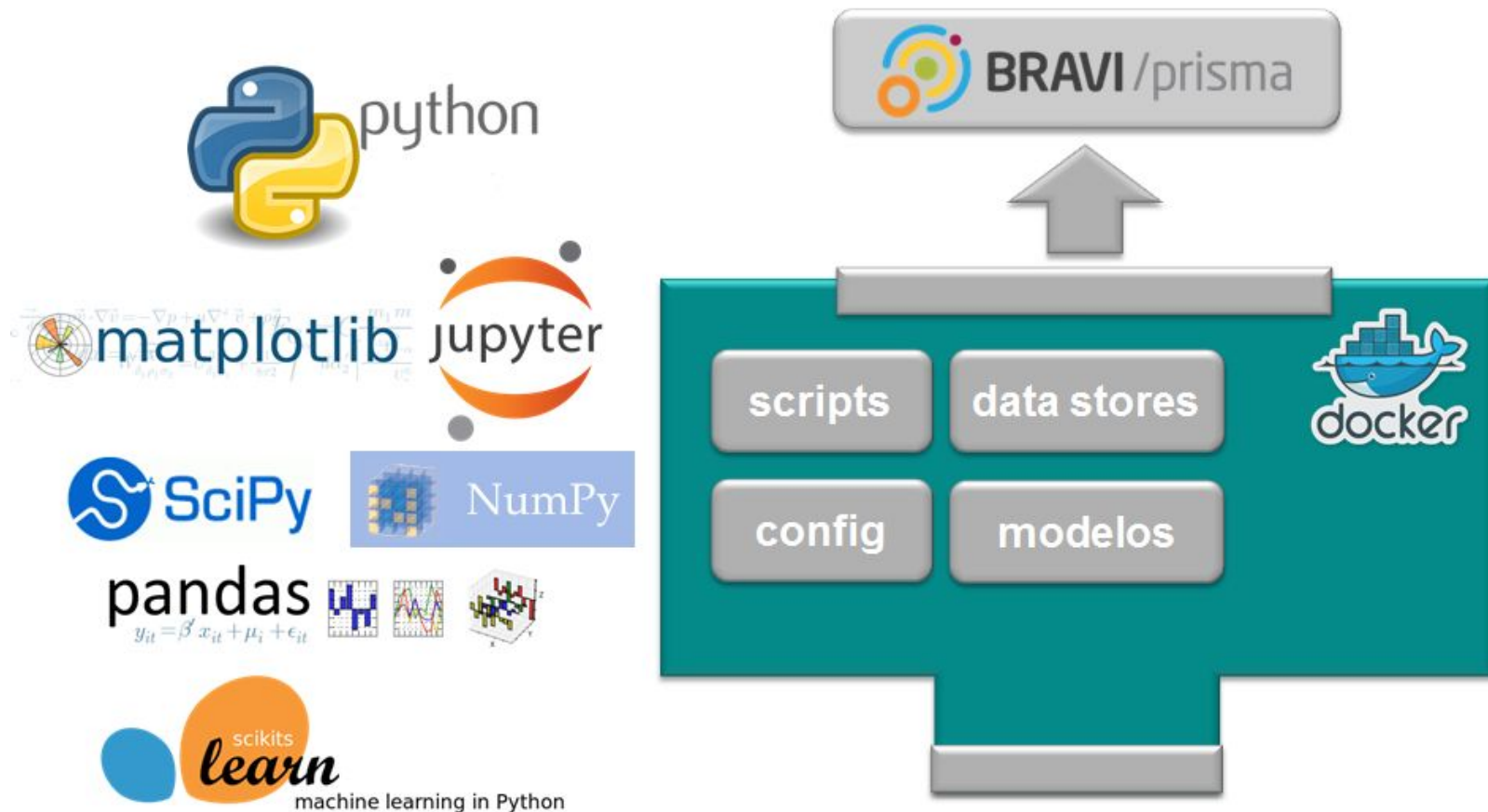
F1	F2	F <sub>n</sub>
...	...	...
...	...	...

Entrada





# Data Product Customizado



Spark



cassandra



BRAVI/prisma

API Dados

Data Product

API Dados

Repositório de Dados

Formato  
Template Bravi

API Dados

Cliente 1

Cliente 2

Cliente 3

...

Cliente n

# Chega de blá blá blá e vamos ver um pouco na prática...

- Demo Prisma
- Construir dados
  - `python build/BUILD.py`



# Links

- Github do Brandon Rhodes com o material que usei de base para o Pandas:
  - <https://github.com/brandon-rhodes/pycon-pandas-tutorial>
- Links para downloads dos conjuntos de dados:
  - <ftp://ftp.fu-berlin.de/pub/misc/movies/database/actors.list.gz>
  - <ftp://ftp.fu-berlin.de/pub/misc/movies/database/actresses.list.gz>
  - <ftp://ftp.fu-berlin.de/pub/misc/movies/database/genres.list.gz>
  - <ftp://ftp.fu-berlin.de/pub/misc/movies/database/release-dates.list.gz>
  - <https://raw.githubusercontent.com/uiuc-cse/data-fa14/gh-pages/data/iris.csv>
- Link para vídeo da apresentação do Brandon Rhodes no PyCon 2015:
  - <https://www.youtube.com/watch?v=5JnMutdy6Fw>