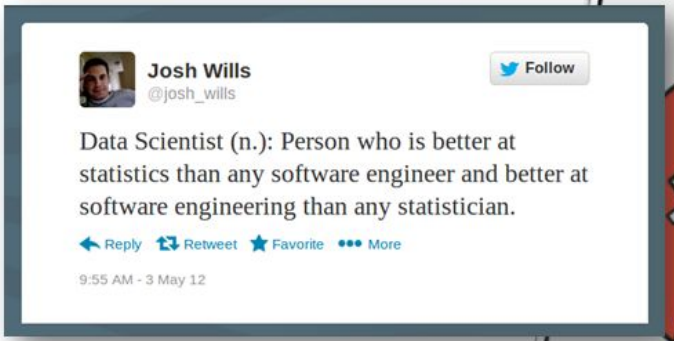
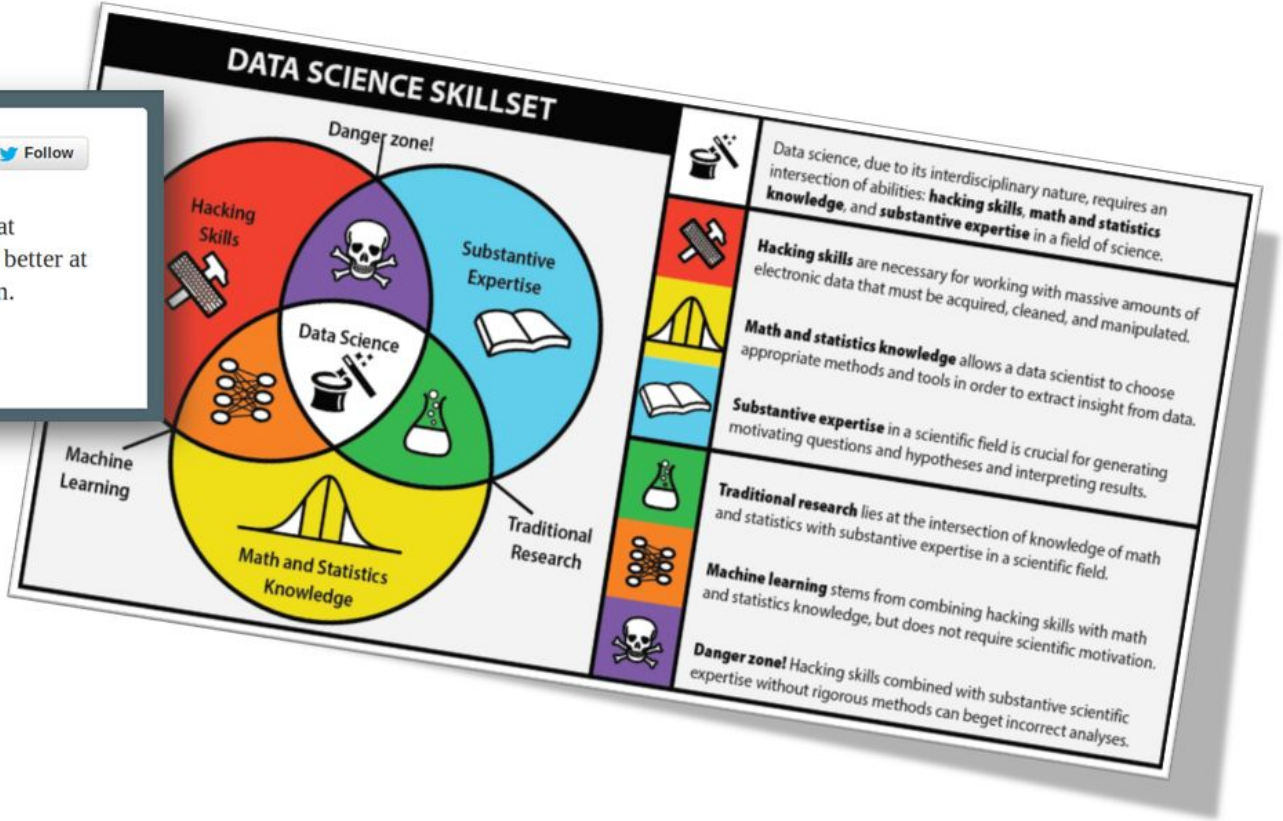
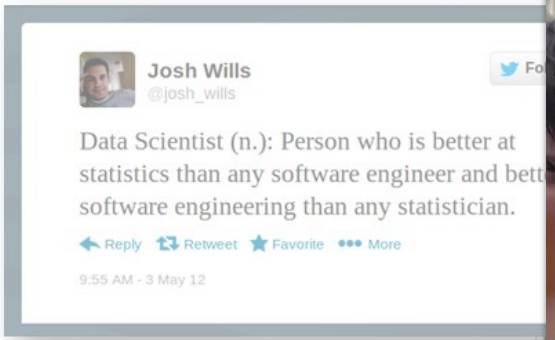
Como ingressar na área de Ciência de Dados

Júlia Tessler - Data Scientist

O que é ciência de dados?


Josh Wills
@josh_wills
Follow
Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.
Reply
Retweet
Favorite
More
9:55 AM - 3 May 12


DATA SCIENCE SKILLSET
Danger zone!
Hacking Skills
Substantive Expertise
Data Science
Machine Learning
Traditional Research
Math and Statistics Knowledge
Data science, due to its interdisciplinary nature, requires an intersection of abilities: **hacking skills**, **math and statistics knowledge**, and **substantive expertise** in a field of science.
Hacking skills are necessary for working with massive amounts of electronic data that must be acquired, cleaned, and manipulated.
Math and statistics knowledge allows a data scientist to choose appropriate methods and tools in order to extract insight from data.
Substantive expertise in a scientific field is crucial for generating motivating questions and hypotheses and interpreting results.
Traditional research lies at the intersection of knowledge of math and statistics with substantive expertise in a scientific field.
Machine learning stems from combining hacking skills with math and statistics knowledge, but does not require scientific motivation.
Danger zone! Hacking skills combined with substantive scientific expertise without rigorous methods can beget incorrect analyses.



ata science, due to its interdisciplinary nature, requires an
tersection of abilities: **hacking skills**, **math and statistics**
owledge, and **substantive expertise** in a field of science.

cking skills are necessary for working with massive amounts of
tronic data that must be acquired, cleaned, and manipulated.

and statistics knowledge allows a data scientist to choose
priate methods and tools in order to extract insight from data.

ntive expertise in a scientific field is crucial for generating
ing questions and hypotheses and interpreting results.

al research lies at the intersection of knowledge of math
tics with substantive expertise in a scientific field.

arning stems from combining hacking skills with math
s knowledge, but does not require scientific motivation.

Hacking skills combined with substantive scientific
bout rigorous methods can beget incorrect analyses.

Vaga

- Conhecimento de TI em desenvolvimento e infraestrutura para solução integrada de problemas.
- Perfil consultivo nas áreas de BI, Data Warehousing, Integração de aplicações, infraestrutura Cloud e plataformas de Big Data.
- Sólidos conhecimentos em padrões de arquitetura e facilidade para absorver novas tecnologias.

Vaga

- Conhecimento de TI em desenvolvimento e infraestrutura para solução integrada de problemas.
- Perfil consultivo nas áreas de BI, Data Warehousing, Integração de aplicações, infraestrutura Cloud e plataformas de Big Data.
- Sólidos conhecimentos em padrões de arquitetura e facilidade para absorver novas tecnologias.

Sólidos conhecimentos em:

- Hadoop, Spark, Mapreduce, Hive e HBASE.
- Sistemas operacionais Unix/Linux.
- Java, Python, ou Scala (mais de uma é desejável);

Vaga

- Conhecimento de TI em desenvolvimento e infraestrutura para solução integrada de problemas.
- Perfil consultivo nas áreas de BI, Data Warehousing, Integração de aplicações, infraestrutura Cloud e plataformas de Big Data.
- Sólidos conhecimentos em padrões de arquitetura e facilidade para absorver novas tecnologias.

Sólidos conhecimentos em:

- Hadoop, Spark, Mapreduce, Hive e HBASE.
- Sistemas operacionais Unix/Linux.
- Java, Python, ou Scala (mais de uma é desejável).
- Modelagem de dados, Data Warehousing e bancos de dados MPP.
- NoSQL e práticos em alguma solução (Hbase, Cassandra, Riak, Redis, MongoDB, etc).

Vaga

- Conhecimento de TI em desenvolvimento e infraestrutura para solução integrada de problemas.
- Perfil consultivo nas áreas de BI, Data Warehousing, Integração de aplicações, infraestrutura Cloud e plataformas de Big Data.
- Sólidos conhecimentos em padrões de arquitetura e facilidade para absorver novas tecnologias.

Sólidos conhecimentos em:

- Hadoop, Spark, Mapreduce, Hive e HBASE.
- Sistemas operacionais Unix/Linux.
- Java, Python, ou Scala (mais de uma é desejável).
- Modelagem de dados, Data Warehousing e bancos de dados MPP.
- NoSQL e práticos em alguma solução (Hbase, Cassandra, Riak, Redis, MongoDB, etc).
- Dimensionamento de hardware, arquiteturas de rede, TCP/IP, DNS, Firewalls.
- Linguagens Shell e SQL.

Vaga

- Conhecimento de TI em desenvolvimento e infraestrutura para solução integrada de problemas.
- Perfil consultivo nas áreas de BI, Data Warehousing, Integração de apl infraestrutura Cloud e plat e Big Data.
- Sólidos conh em padrões de arquitet dade para absorver n logias.

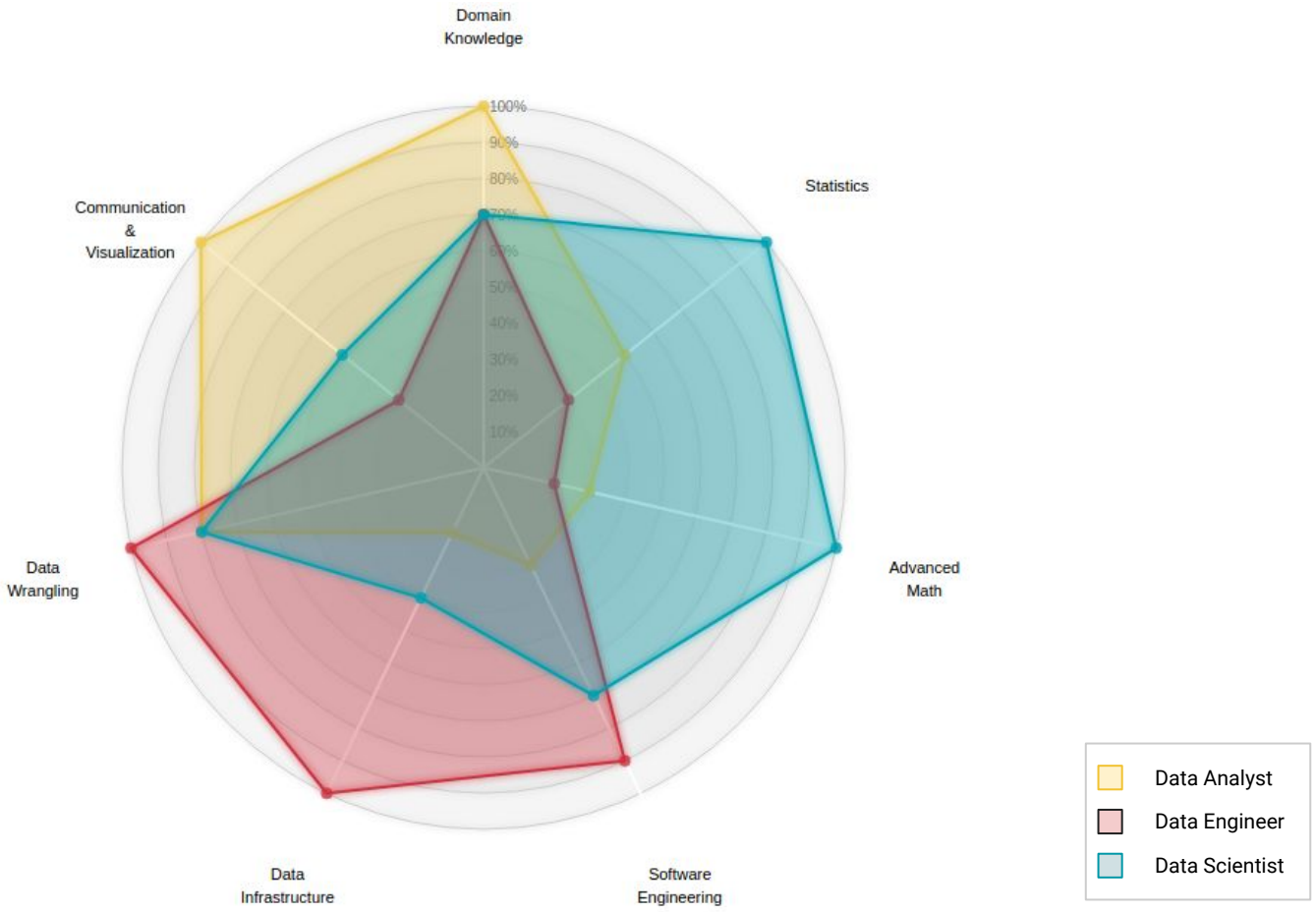
Sólidos conhecimentos em:

- Hadoop, Spark, Mapred HBASE.
- Sistemas operac Linux.
- Java, Pyth a (mais de uma é des
- gem de dados, Data Warehousing bancos de dados MPP.
- NoSQL e práticos em alguma solução (Hbase, Cassandra, Riak, Redis, MongoDB, etc).
- Dimensionamento de hardware, arquiteturas de rede, TCP/IP, DNS, Firewalls.
- Linguagens Shell e SQL.

[E sabendo conectar tudo isso com as metas e estratégias da empresa]

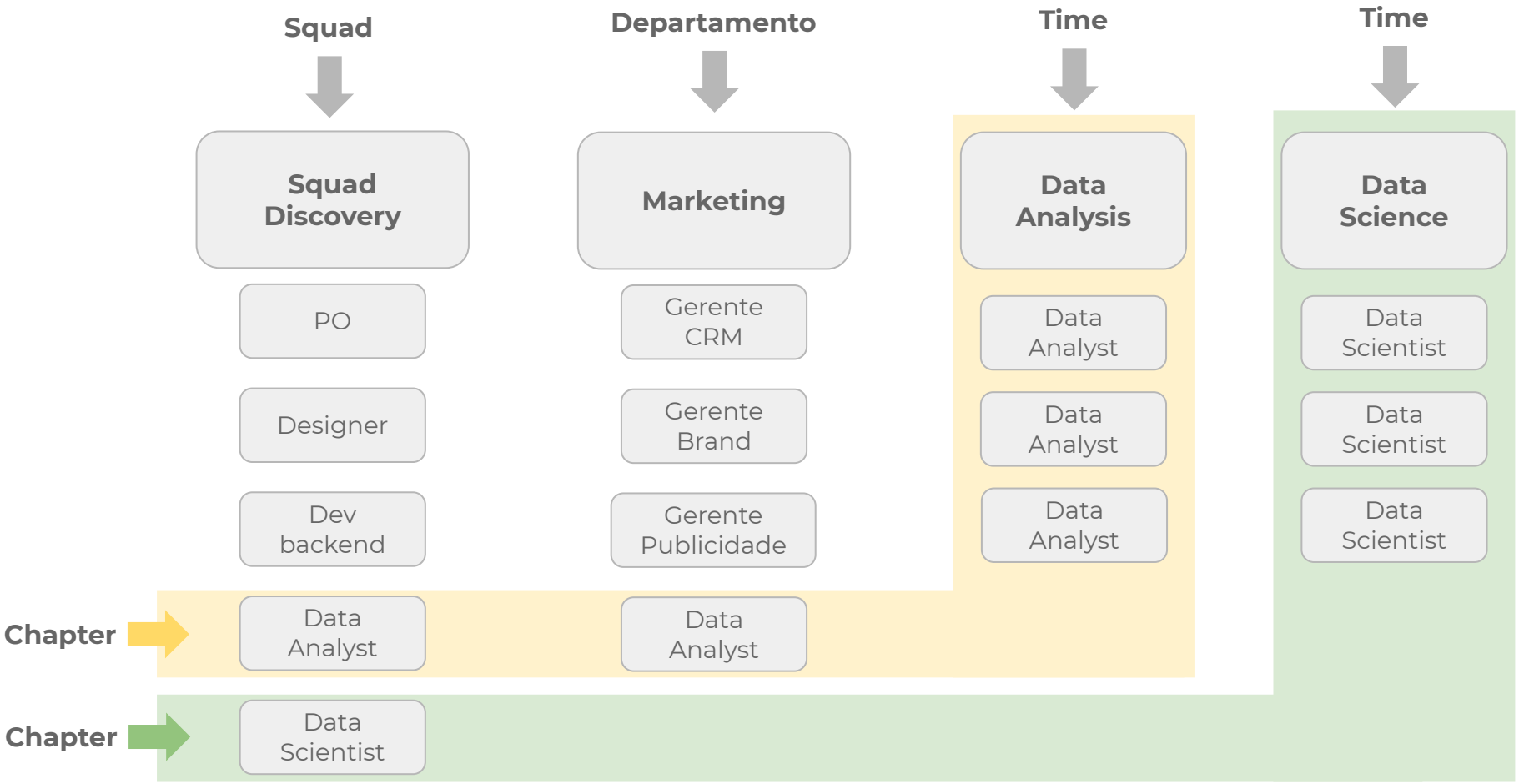
... e como eu vim parar aqui?

tudo deu errado, mas deu certo no final (:


Domain Knowledge
Statistics
Advanced Math
Software Engineering
Data Infrastructure
Data Wrangling
Communication & Visualization
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
Data Analyst
Data Engineer
Data Scientist

DATA PLATFORM	**Garantir a confiabilidade e disponibilidade dos dados da fonte única da verdade e suas plataformas de aplicação.**
DATA ANALYSIS	**Interpretar os dados e identificar insights para suportar as tomadas de decisão em todos os níveis da empresa.**
DATA SCIENCE	**Desenvolver aplicações e modelos estatísticos, matemáticos e de inteligência artificial que otimizem processos e produtos.**

Time & Business Partners



feed the data monster

http://feedthedatamonster.com/

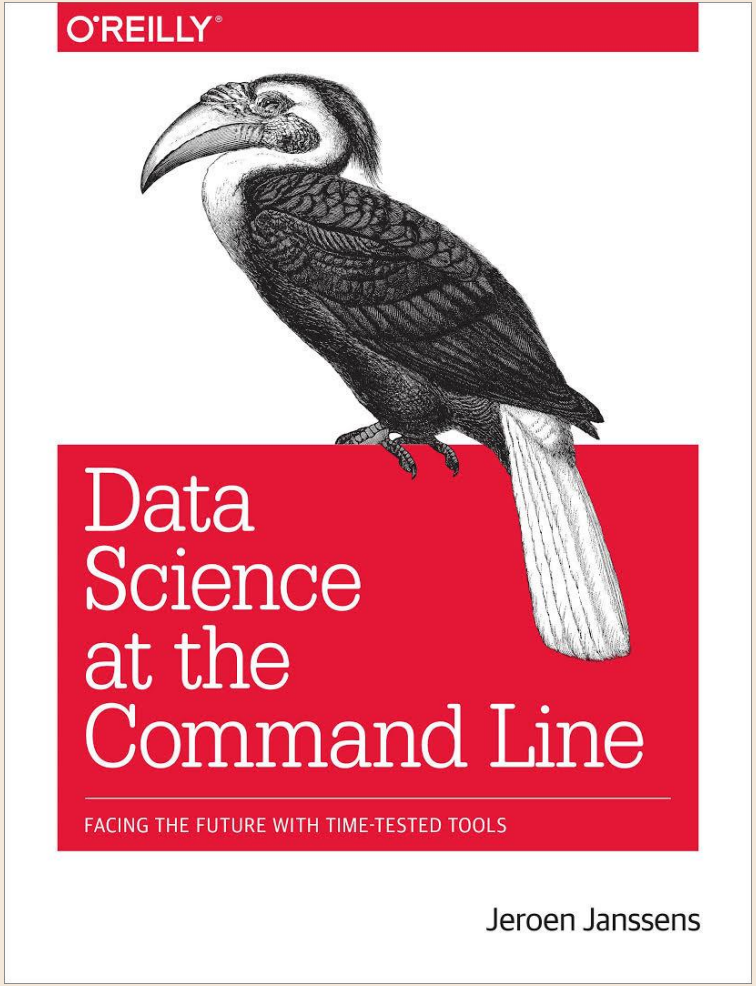
Identificação do problema

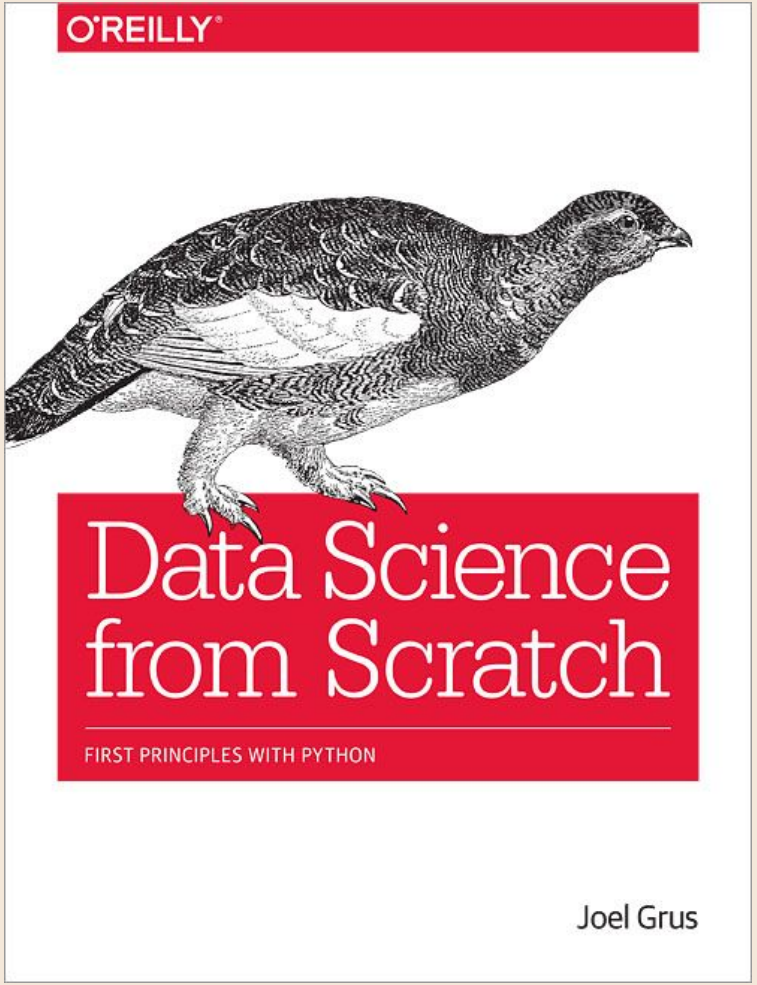
Pré-processamento

Treinamento e predição

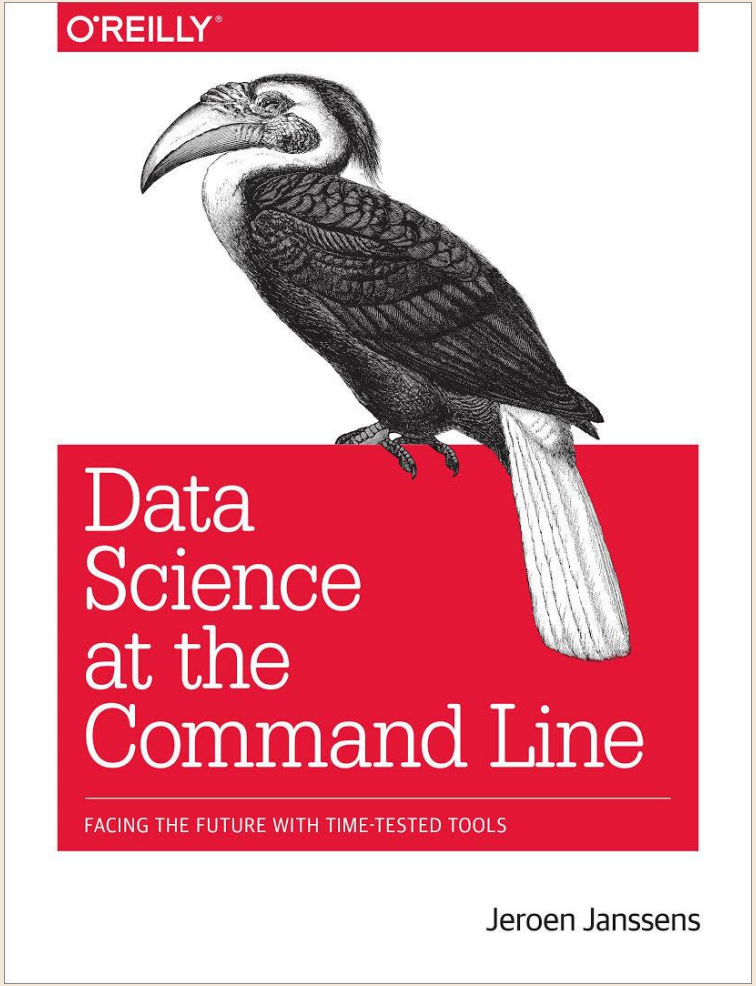
Avaliação dos resultados

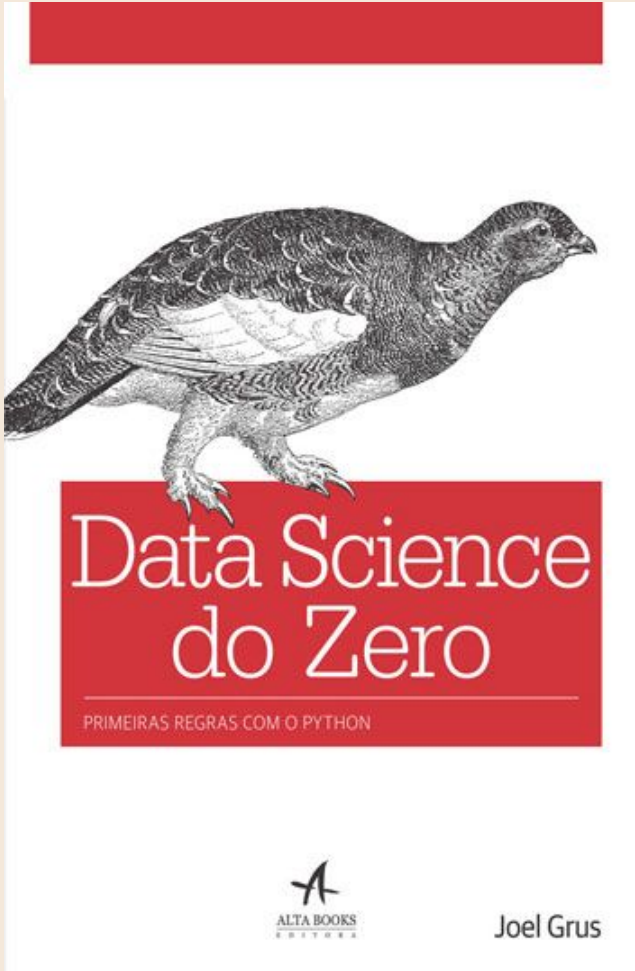
Uso do conhecimento



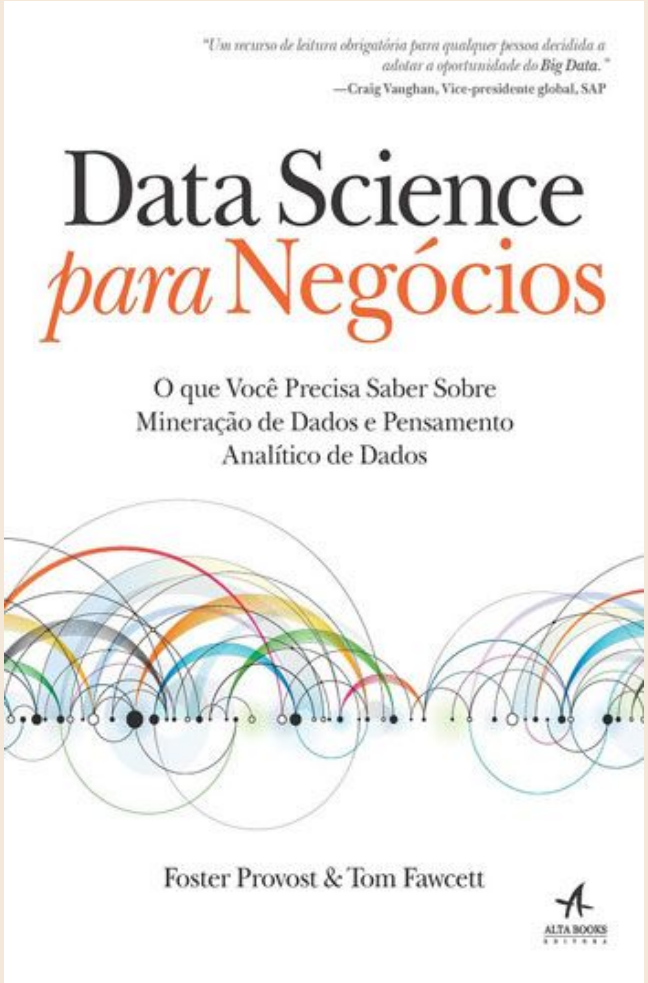


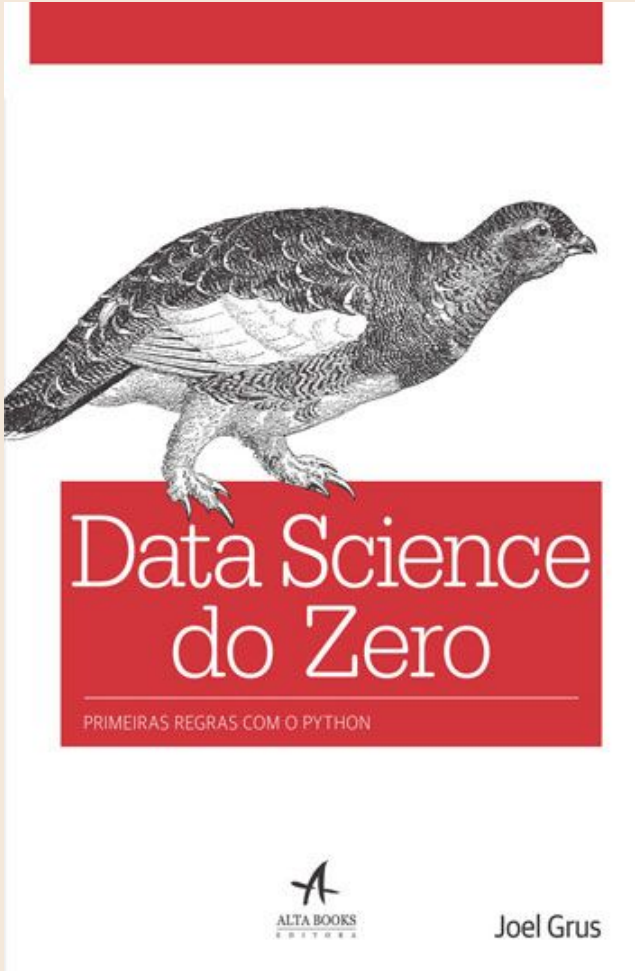
https://www.datascienceatthecommandline.com/





https://www.datascienceatthecommandline.com/


"Um recurso de leitura obrigatória para qualquer pessoa decidida a adotar a oportunidade do Big Data."
—Craig Vaughan, Vice-presidente global, SAP
Data Science para Negócios
O que Você Precisa Saber Sobre Mineração de Dados e Pensamento Analítico de Dados
Foster Provost & Tom Fawcett
ALTA BOOKS


Data Science do Zero
PRIMEIRAS REGRAS COM O PYTHON
ALTA BOOKS
Joel Grus

Obrigada!

Júlia Tessler
julia.tessler@ifood.com.br

http://ifood.delivery
https://jobs.kenoby.com/ifood

ifood