



THE
DEVELOPER'S
CONFERENCE



THE
DEVELOPER'S
CONFERENCE

Data Science e Machine Learning

Um dia na vida de um cientista de dados

Aishameriane Schmidt

Estatista - UFRGS
Mestranda em Economia - UFSC

Matheus Magrin

Cientista da Computação - UFSC
Cientista de dados - Bravi

Agenda



THE
DEVELOPER'S
CONFERENCE

O que significa ser “cientista de dados”?

- O mito do datacórneo
- Tipos de cientistas de dados
- Conhecimentos úteis e onde obtê-los

Guia prático para análise de dados

- Entendendo a dor do cliente
- Enfrentando a horda de dados
- Transformando hipóteses em análises
- Interpretando resultados

Agenda



THE
DEVELOPER'S
CONFERENCE

O que significa ser “cientista de dados”?

- O mito do datacórneo
- Tipos de cientistas de dados
- Conhecimentos úteis e onde obtê-lo

Guia prático para análise de dados

- Entendendo a dor do cliente
- Enfrentando a horda de dados
- Transformando hipóteses em análises
- Interpretando resultados

Quem fala



THE
DEVELOPER'S
CONFERENCE

Aishameriane

- Estatista - UFRGS (2010)
- RBS, Cassol, Chaordic, Meritt, Bravi
- Economista to be - UDESC (201?)
- Mestranda em Economia com Ênfase em Finanças e Mercado de Capitais - UFSC (2019)
- Não poderia viver de programação

Matheus

- Cientista da Computação - UFSC (2013)
- Laboratório de Sistemas de Conhecimento/UFSC, Bravi
- Desenvolvedor - Bravi (2013~14)
- Cientista de dados - Bravi (2014~hoje)

Quem fala



THE
DEVELOPER'S
CONFERENCE

Aishameriane

- Estatista - UFRGS (2010)
- RBS, Cassol, Chaordic, Meritt, Bravi
- Economista to be - UDESC (201?)
- Mestranda em Economia com Ênfase em Finanças e Mercado de Capitais - UFSC (2019)
- Não poderia viver de programação

Matheus

- Cientista da Computação - UFSC (2013)
- Laboratório de Sistemas de Conhecimento/UFSC, Bravi
- Desenvolvedor - Bravi (2013~14)
- Cientista de dados - Bravi (2014~hoje)

Quem fala



THE
DEVELOPER'S
CONFERENCE

Aishameriane

- Estatista - UFRGS (2010)
- RBS, Cassol, Chaordic, Meritt, Bravi
- Economista to b UDESC (2011)
- Mestranda em Ec ... Mercado de Capitais - UFSC (2019)
- Não pod

Matheus

- Ciência
- Laboratório
- Desenvol
- Cientista de dados ... (~hoje)

**Ambos já
desenvolveram
atividades na área de
data science**

Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

O que são? - O que comem? - Onde vivem?





THE
DEVELOPER'S
CONFERENCE

One problem with job post data is that "data scientist" is a
loosey-goosey term.

- *Barb Darrow, Fortune Tech, 2015.*



One problem with job post data is that "data scientist" is a
loosey-goosey term.

Generally speaking, practitioners are expected to know
statistical analysis, **predictive modeling** and
programming.

Oh, and having a certain **artistic flair** to guide how results
are visualized is a definite plus.

- *Barb Darrow, Fortune Tech, 2015.*



THE
DEVELOPER'S
CONFERENCE

One problem with job post data is that "data scientist" is a
loosey-goosey term.

Generally speaking, practitioners are expected to know
**statistical analysis, predictive modeling and
programming.**

Oh, and having a certain **artistic flair** to guide how results
are visualized is a definite plus.

**But ask a dozen hiring managers and you may get a
dozen variations on that theme.**

- *Barb Darrow, Fortune Tech, 2015.*

Disponível em: <http://fortune.com/2015/05/21/data-science-white-hot/>

Recomendamos também:

<http://radar.oreilly.com/2013/06/theres-more-than-one-kind-of-data-scientist.html>

<https://www.datacamp.com/community/tutorials/data-science-industry-infographic#gs.8bsW=Eo>



Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

O mito do datacórnio



Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

O mito do datacórnio

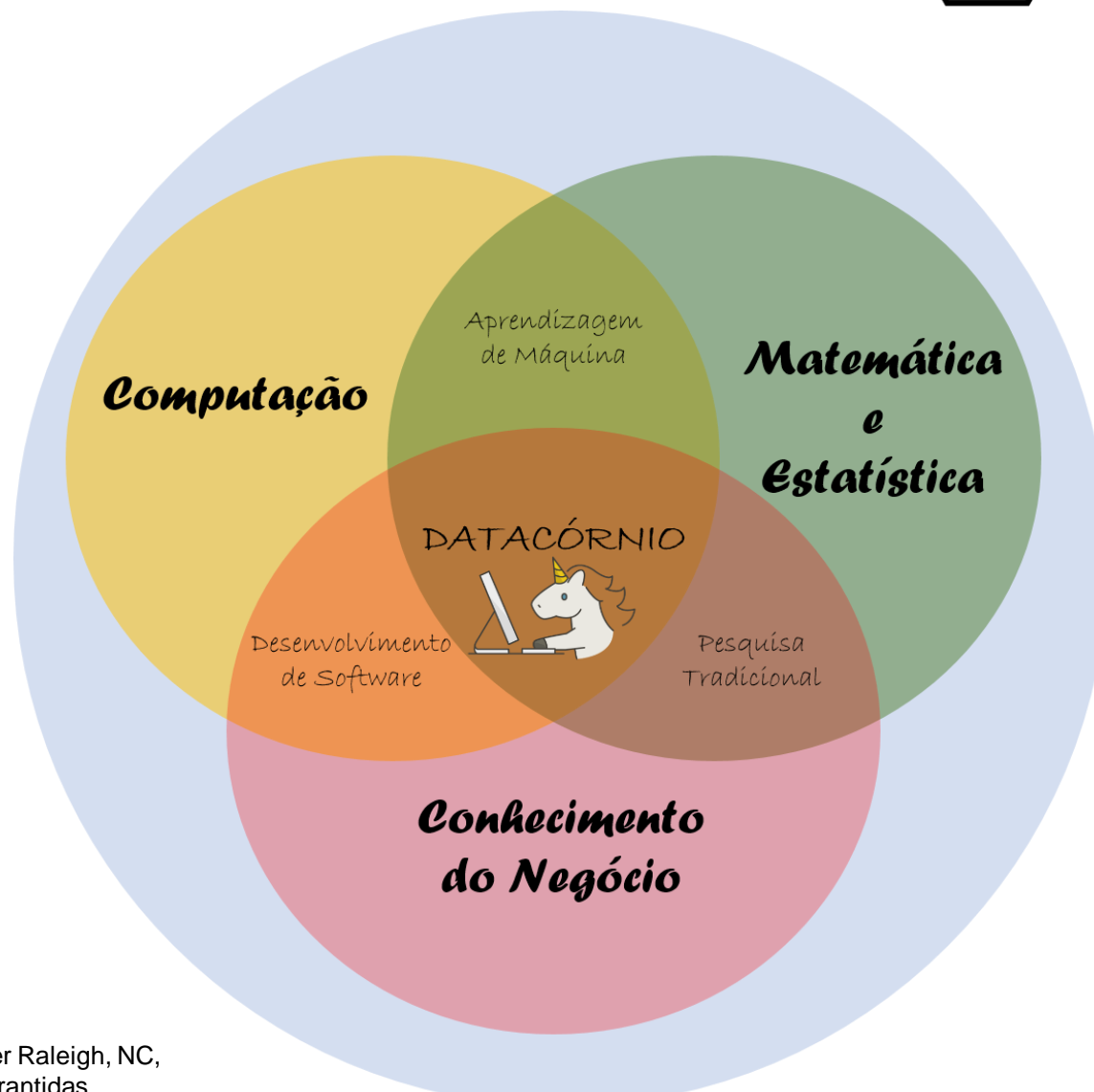




Cientistas de dados



THE
DEVELOPER'S
CONFERENCE



Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

Os tipos de cientistas de dados



Datacornius Statisticus

- Conhecimentos de métodos estatísticos
- Algum conhecimento em programação
- Conhecimentos de álgebra linear e cálculo
- Foco em conjuntos menores de dados

Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

Os tipos de cientistas de dados



Datacornius Machinus Learningus

- Conhecimentos de programação
- Conhecimentos em técnicas de aprendizagem de máquina
- Foco em técnicas de análise para grandes volumes de dados
- Algum conhecimento em matemática e estatística

Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

Os tipos de cientistas de dados



Datacornius Developeurs

- Conhecimentos de programação, estruturas de dados, importação e exportação de bancos de dados
- Conhecimentos em Big Data
- Algum conhecimento de matemática e/ou estatística

Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

Os tipos de cientistas de dados



Datacornius Gerentus

- Conhecimento do negócio
- Conhecimento em gerenciamento de projetos e organização de equipes
- Meio de campo entre o time técnico e cliente final

Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

Os tipos de cientistas de dados



Datacornius spp.

- Os papéis não são fechados e nem não-intercambiáveis
 - Podem haver pôneis com mistura de mais de uma espécie de datacórnio!

Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

Como buscar conhecimento?

coursera



 UDACITY



edx



Cientistas de dados



THE
DEVELOPER'S
CONFERENCE

Como buscar conhecimento?



Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está pensando em cancelar a assinatura?”

“Quem é mais propenso a fazer uma assinatura?”

“Quanto tempo o cliente precisa ser pagante para começar a dar retorno sobre o gasto de aquisição?”

“Quanto do aumento de clientes desde a semana passada pode ser atribuído à nova campanha de marketing?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está pensando em cancelar a assinatura?”

“Quem é mais propenso a fazer uma assinatura?”

“Quanto tempo o cliente precisa ser pagante para começar a dar retorno sobre o gasto de aquisição?”

“Quanto do aumento de clientes desde a semana passada pode ser atribuído à nova campanha de marketing?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está pensando em cancelar a assinatura?”

“Quem é mais propenso a fazer uma assinatura?”

“Quanto tempo o cliente precisa ser pagante para começar a dar retorno sobre o gasto de aquisição?”

“Quanto do aumento de clientes desde a semana passada pode ser atribuído à nova campanha de marketing?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está pensando em cancelar a assinatura?”

“Quem é mais propenso a fazer uma assinatura?”

“Quanto tempo o cliente precisa ser pagante para começar a dar retorno sobre o gasto de aquisição?”

“Quanto do aumento de clientes desde a semana passada pode ser atribuído à nova campanha de marketing?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está propenso a se tornar inadimplente?”

“Qual a faixa de crédito que pode ser concedida para o cliente?”

“O que este cliente pode gostar, sendo que nem conheço ele?”

“Em qual momento preciso interagir com o cliente para fidelizá-lo?”

“Quais clientes respondem melhor a cada um dos meus canais de comunicação?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está propenso a se tornar inadimplente?”

“Qual a faixa de crédito que pode ser concedida para o cliente?”

“O que este cliente pode gostar, sendo que nem conheço ele?”

“Em qual momento preciso interagir com o cliente para fidelizá-lo?”

“Quais clientes respondem melhor a cada um dos meus canais de comunicação?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está propenso a se tornar inadimplente?”

“Qual a faixa de crédito que pode ser concedida para o cliente?”

“O que este cliente pode gostar, sendo que nem conheço ele?”

“Em qual momento preciso interagir com o cliente para fidelizá-lo?”

“Quais clientes respondem melhor a cada um dos meus canais de comunicação?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está propenso a se tornar inadimplente?”

“Qual a faixa de crédito que pode ser concedida para o cliente?”

“O que este cliente pode gostar, sendo que nem conheço ele?”

“Em qual momento preciso interagir com o cliente para fidelizá-lo?”

“Quais clientes respondem melhor a cada um dos meus canais de comunicação?”

Situações práticas/comuns



THE
DEVELOPER'S
CONFERENCE

“Quem está propenso a se tornar inadimplente?”

“Qual a faixa de crédito que pode ser concedida para o cliente?”

“O que este cliente pode gostar, sendo que nem conheço ele?”

“Em qual momento preciso interagir com o cliente para fidelizá-lo?”

“Quais clientes respondem melhor a cada um dos meus canais de comunicação?”

Guia prático de análise de dados



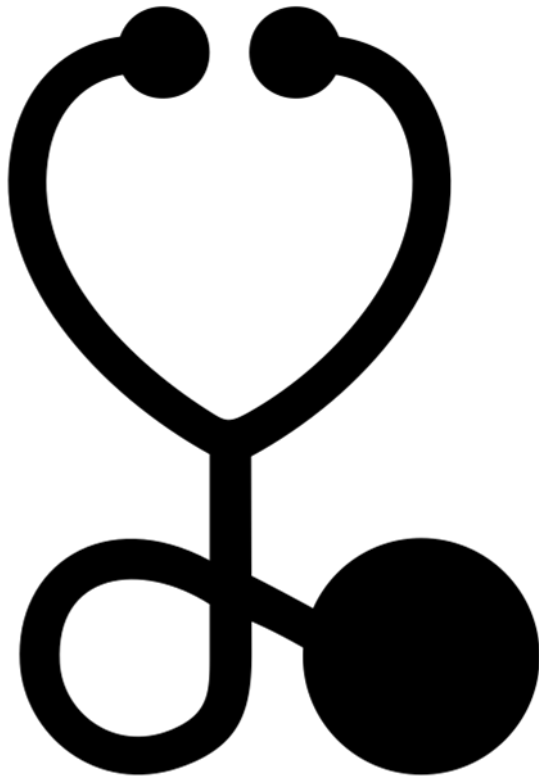
THE
DEVELOPER'S
CONFERENCE

- Delimite o problema
- Explore os dados (estatísticas e visualizações)
- Faça a análise inferencial
- Como interpretar os resultados

Delimite o problema



THE
DEVELOPER'S
CONFERENCE



1º Passo:
Entenda a dor da pessoa que
é sua cliente

Delimite o problema



THE
DEVELOPER'S
CONFERENCE

EXPECTATIVA:

- **Objetivo:**

Saber quais clientes da loja X tem risco de não pagar uma ou mais parcelas no prazo de vencimento;

- **Técnica:**

Modelo de regressão logística;

- **Variável no banco de dados:**

PARCELA_ATRASO == 1

Delimite o problema



THE
DEVELOPER'S
CONFERENCE

EXPECTATIVA:

REALIDADE:

- **Objetivo:**

Saber quais clientes da loja X tem risco de não pagar uma ou mais parcelas no prazo de vencimento;

- **Técnica:**

Modelo de regressão logística;

- **Variável no banco de dados:**

PARCELA_ATRASO == 1

DÓI
TUDO!

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Limpeza do banco de dados



Explore os dados



THE
DEVELOPER'S
CONFERENCE

Limpeza do banco de dados

- Os dados dificilmente estarão prontos para análise;
- Muitas variáveis precisarão ser criadas:
 - Número de faltas no semestre - é preciso somar as faltas de todos os dias letivos;
 - Acessou pelo menos uma vez o material virtual - é necessário fazer checagem dos logins no sistema;
 - Reprovação anterior na mesma matéria - é preciso verificar se já houve matrícula na matéria e se houve reprovação (além disso verificar se a reprovação foi por nota ou por faltas);
 - etc

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Limpeza do banco de dados

- Os dados dificilmente estarão prontos para análise;
- Muitas variáveis precisarão ser criadas:
 - Número de faltas no semestre
 - Acessou pelo menos uma vez o material virtual
 - Reprovação anterior na mesma matéria

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Limpeza do banco de dados

- Os dados dificilmente estarão prontos para análise;
- Muitas variáveis precisarão ser criadas:

- **Número de faltas no semestre**

É preciso somar as faltas de todos os dias letivos

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Limpeza do banco de dados

- Os dados dificilmente estarão prontos para análise;
 - Muitas variáveis precisarão ser criadas:
 - **Número de faltas no semestre** -
É preciso somar as faltas de todos os dias letivos;
 - **Acessou pelo menos uma vez o material virtual**
- Necessário fazer checagem dos logins no sistema

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Limpeza do banco de dados

- Os dados dificilmente estarão prontos para análise;
- Muitas variáveis precisarão ser criadas:

- **Número de faltas no semestre** -

É preciso somar as faltas de todos os dias letivos;

- **Acessou pelo menos uma vez o material virtual** - é

Necessário fazer checagem dos logins no sistema;

- **Reprovação anterior na mesma matéria**

É preciso verificar se já houve matrícula na matéria e se houve reprovação

Além disso, verificar se a reprovação foi por nota ou por faltas

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Entendimento das variáveis (dicas)

- Cuidado com as variáveis que se sobrescrevem:
logs de sistema, acessos, etc.

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Entendimento das variáveis (dicas)

- Cuidado com as variáveis que se sobrescrevem:
logs de sistema, acessos, etc.
- Cuidado com variáveis que, por mudanças de versão no sistema, deixaram de ser utilizadas;

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Entendimento das variáveis (dicas)

- Cuidado com as variáveis que se sobrescrevem:
logs de sistema, acessos, etc.
- Cuidado com variáveis que, por mudanças de versão no sistema, deixaram de ser utilizadas;
- Pegue um caso “conhecido” e verifique se as informações das variáveis são consistentes com a realidade.

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Entendimento das variáveis (dicas)

- Cuidado com as variáveis que se sobrescrevem:
logs de sistema, acessos, etc.
- Cuidado com variáveis que, por mudanças de versão no sistema, deixaram de ser utilizadas;
- Pegue um caso “conhecido” e verifique se as informações das variáveis são consistentes com a realidade.
- Cuidado com dados que estão localizados no tempo

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Entendimento das variáveis (dicas)

- Certifique-se de que haverá um prazo máximo para que suas dúvidas sejam resolvidas pelo cliente;

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Entendimento das variáveis (dicas)

- Certifique-se de que haverá um prazo máximo para que suas dúvidas sejam resolvidas pelo cliente;
- Faça um registro das dúvidas e de quais foram as respostas dadas (e por quem foram dadas);

Explore os dados



THE
DEVELOPER'S
CONFERENCE

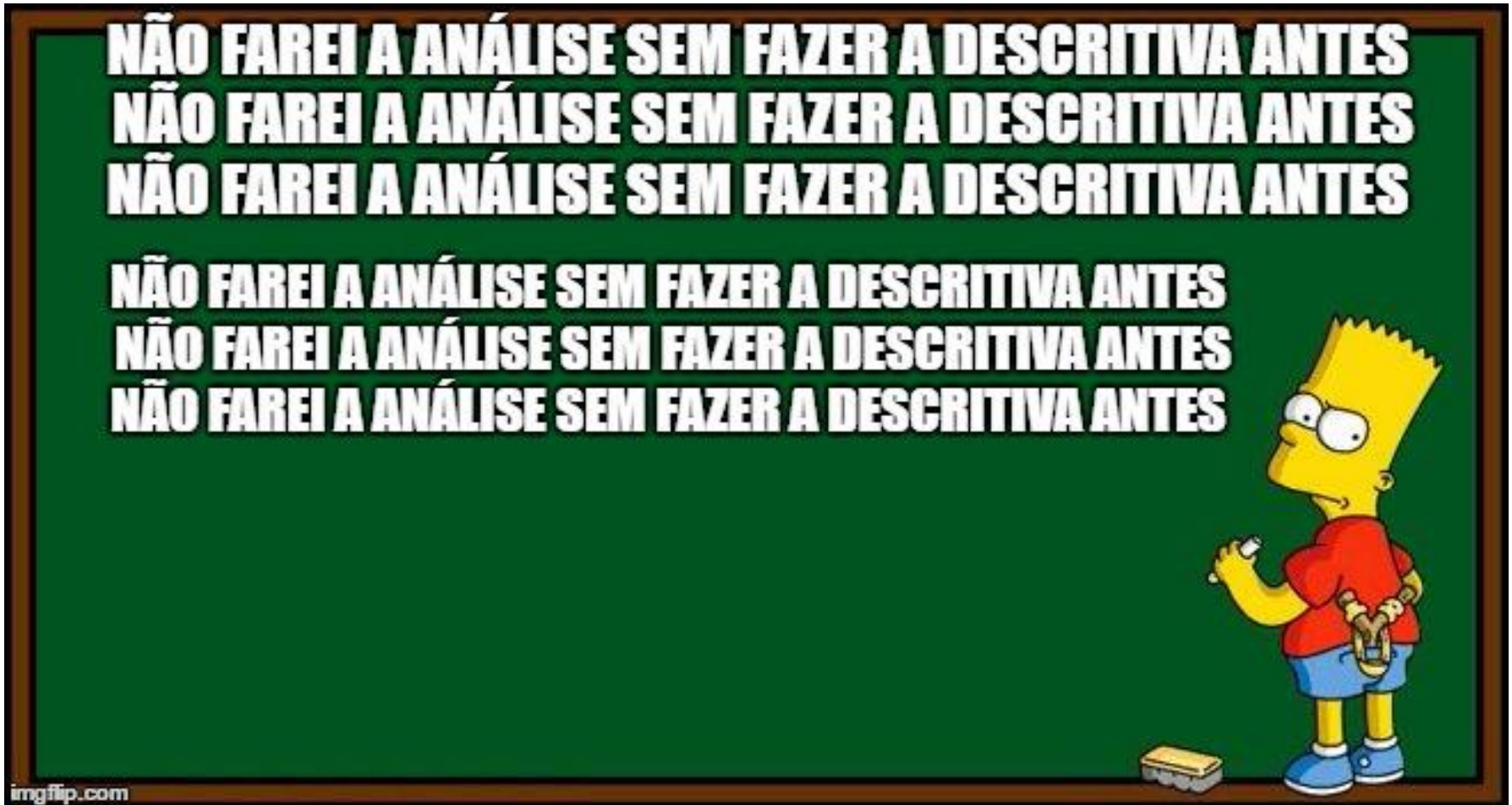
Entendimento das variáveis (dicas)

- Certifique-se de que haverá um prazo máximo para que suas dúvidas sejam resolvidas pelo cliente;
- Faça um registro das dúvidas e de quais foram as respostas dadas (e por quem foram dadas);
- Faça testes para verificar se a resposta dada é consistente com o que os dados falam;

Explore os dados



THE
DEVELOPER'S
CONFERENCE

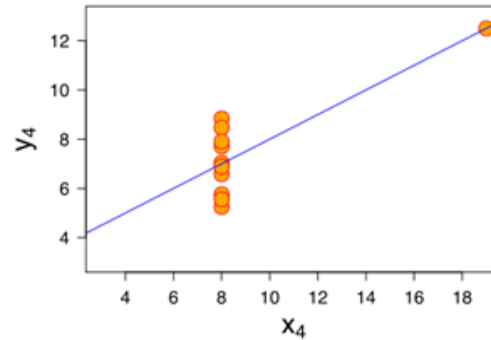
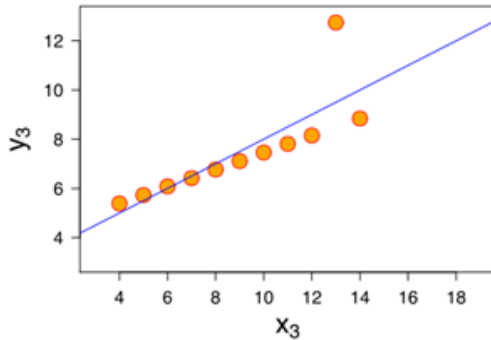
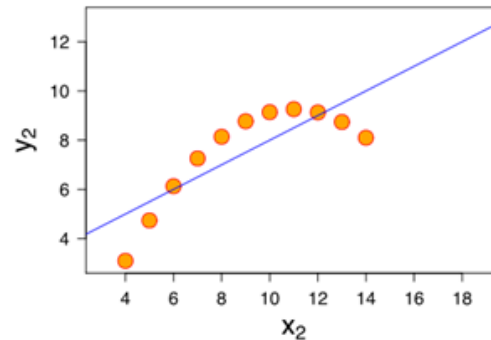
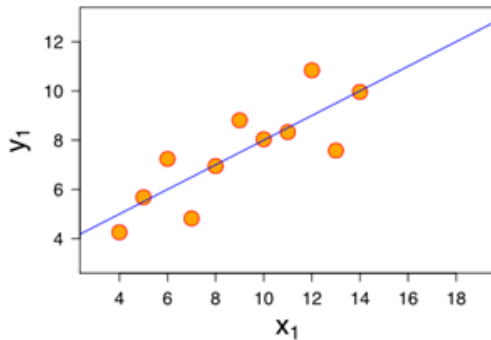


Explore os dados



THE
DEVELOPER'S
CONFERENCE

Quarteto de Anscombe

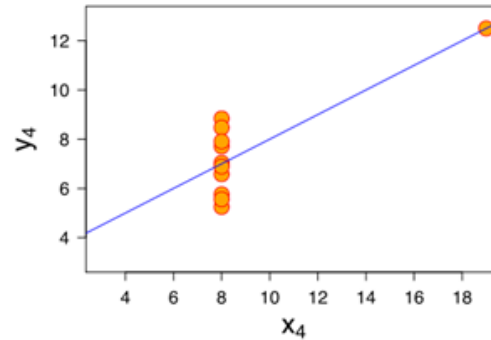
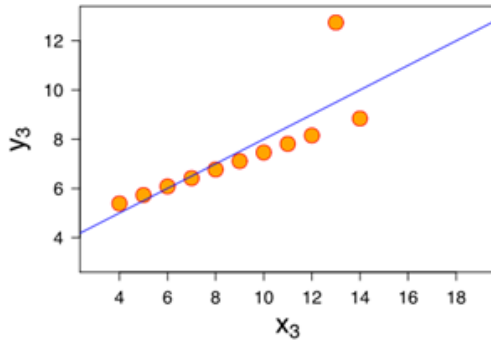
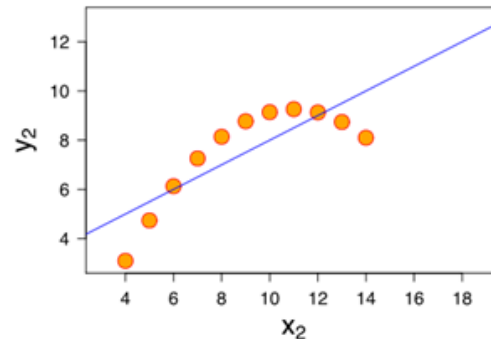
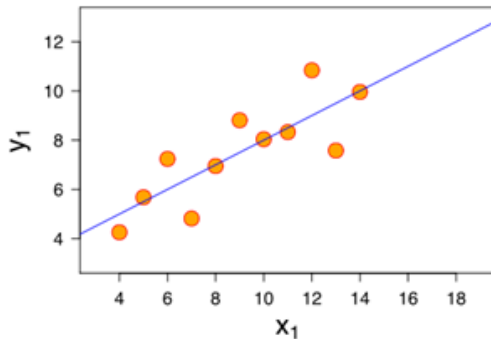


Explore os dados



THE
DEVELOPER'S
CONFERENCE

Quarteto de Anscombe



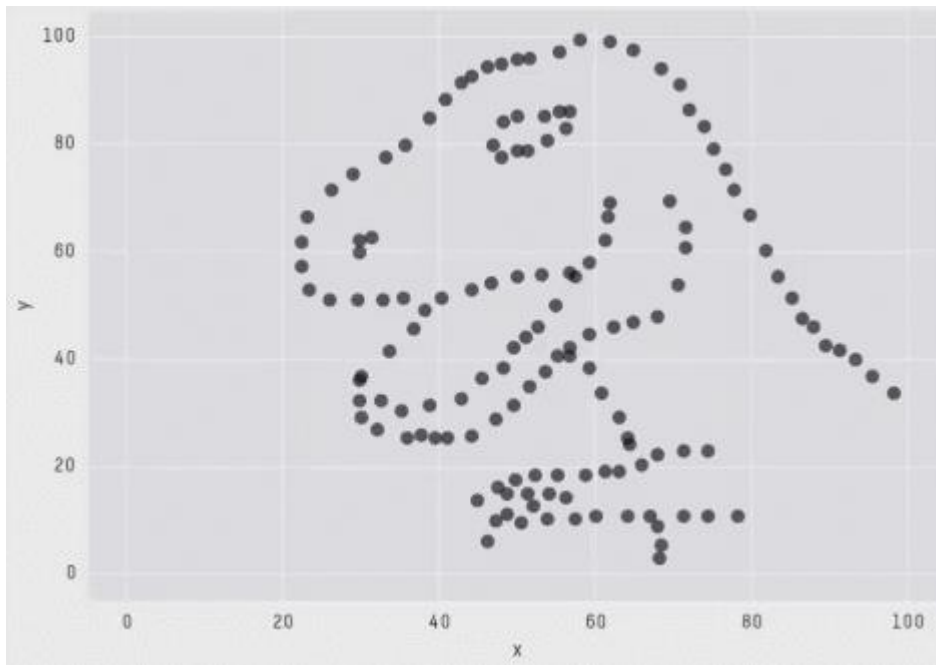
Propriedade	Valor
Média de x	9
Variância de x	11
Média de y	7,50
Variância de y	4,125
Correlação	0,816
Regressão linear	$y = 3,0 + 0,50x$

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Meet the Datasaurus Dozen!



X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526

Matejka, J.; Fitzmaurice, G.

Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics Through Simulated Annealing.

ACM SIGCHI on Human Factors in Computing Systems.

<https://www.autodeskresearch.com/publications/samestats>

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Cuidado com os zeros!

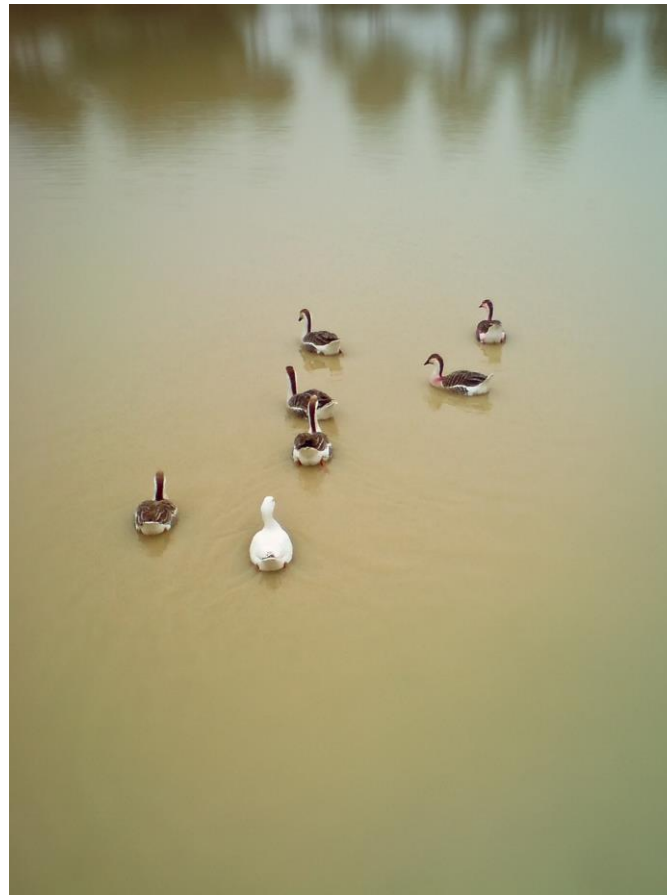


Explore os dados



THE
DEVELOPER'S
CONFERENCE

Cuidado com os valores discrepantes!



Explore os dados



THE
DEVELOPER'S
CONFERENCE

Técnicas de Análise Descritiva

- Medidas descritivas:
Média, moda, mediana, desvio padrão, correlações

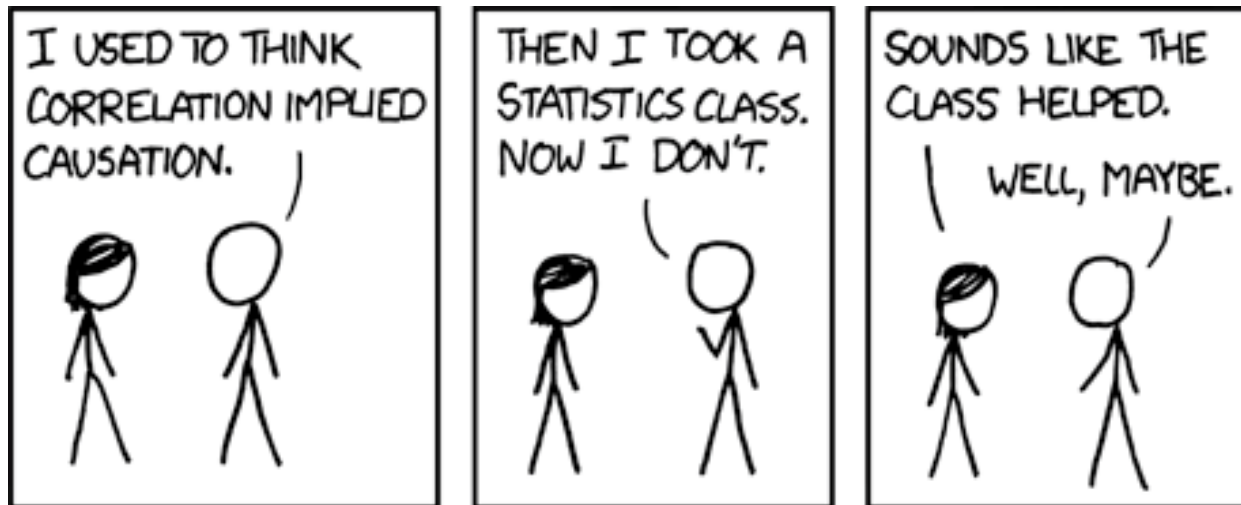
Explore os dados



THE
DEVELOPER'S
CONFERENCE

Técnicas de Análise Descritiva

- Medidas descritivas:
Média, moda, mediana, desvio padrão, correlações



Fonte: XKCD
<https://xkcd.com/552/>

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Técnicas de Análise Descritiva

- **Medidas descritivas:**
Média, moda, mediana, desvio padrão, correlações
- **Gráficos:**
Boxplot, histogramas (variáveis quantitativas)
Gráficos de barra e pizzas (variáveis qualitativas)

Explore os dados



THE
DEVELOPER'S
CONFERENCE

Técnicas de Análise Descritiva

- Medidas descritivas:
Média, moda, mediana, desvio padrão, correlações
- Gráficos:
Boxplot, histogramas (variáveis quantitativas)
Gráficos de barra e pizzas (variáveis qualitativas)
- Tabelas de frequências
- Tabelas cruzadas

Referência adicional:

Schmidt, A.; Boito, F. Z.; Pilla, L. L..

Fundamentos de Estatística para Análise de Desempenho

Mini curso da 17ª Escola Regional de Alto Desempenho - ERAD-RS 2017

Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/erad/2017/002.pdf>

Análise inferencial



THE
DEVELOPER'S
CONFERENCE

O problema fundamental da inferência estatística é inferir, através das estimativas obtidas com base em uma amostra, características dos parâmetros populacionais.

Análise inferencial



THE
DEVELOPER'S
CONFERENCE

Técnicas Inferenciais

Testes de hipóteses
Análise de Regressão
Séries Temporais
Métodos Causais
Análise de Experimentos
Etc

Análise inferencial



THE
DEVELOPER'S
CONFERENCE

Técnicas Inferenciais

Testes de hipóteses
Análise de Regressão
Séries Temporais
Métodos Causais
Análise de Experimentos
Etc

Abordagem
Bayesiana x Clássica

Análise inferencial



THE
DEVELOPER'S
CONFERENCE

Técnicas Inferenciais

Testes de hipóteses
Análise de Regressão
Séries Temporais
Métodos Causais
Análise de Experimentos
Etc

Abordagem
Bayesiana x Clássica

Métodos Paramétricos
e Não-paramétricos

Algoritmos de ML



THE
DEVELOPER'S
CONFERENCE

- Supervisionados
 - Classificação
 - Regressão
- Não supervisionados
 - Recomendação
 - Clustering
 - Redução de dimensionalidade

Algoritmos de ML



THE
DEVELOPER'S
CONFERENCE

- Supervisionados
 - Classificação
 - Regressão
- Não supervisionados
 - Recomendação
 - Clustering
 - Redução de dimensionalidade

Entenda seu problema e escolha as suas armas



THE
DEVELOPER'S
CONFERENCE

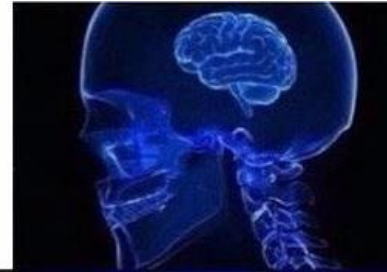


A sua arma mais importante...



THE
DEVELOPER'S
CONFERENCE

Im()



KNN,
Random Forests



Algoritmos genéticos,
Estatística Avançada
Programação



Combinar
ferramentas de
acordo com seu
problema



imgflip.com

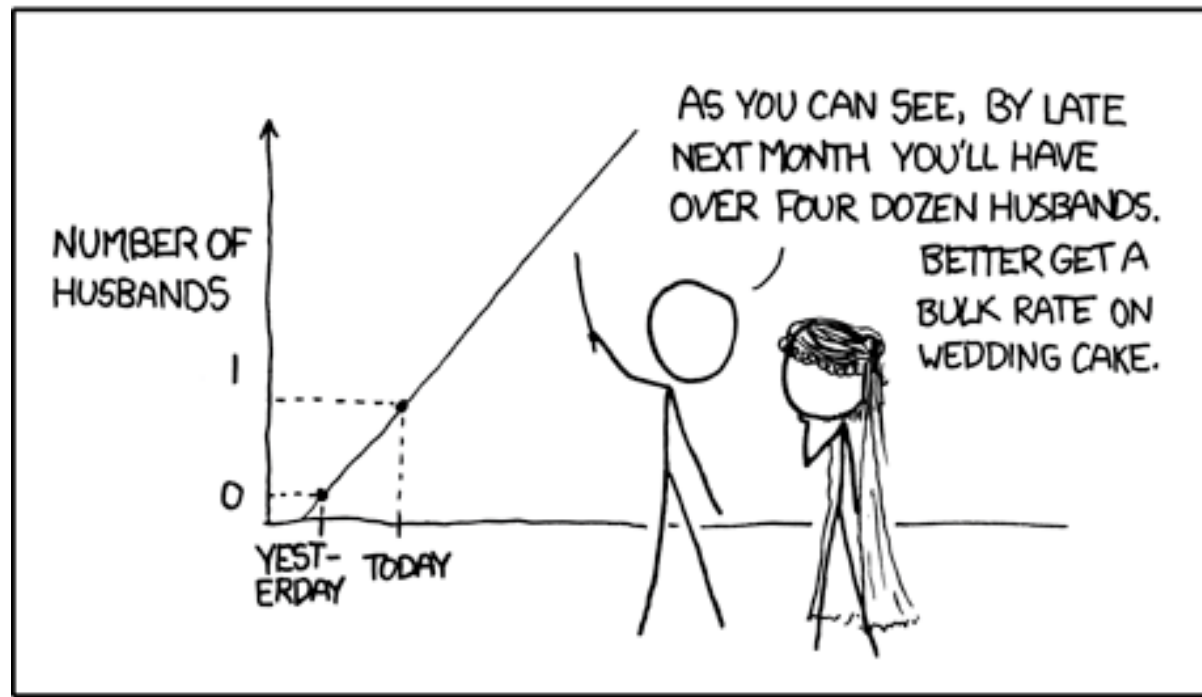
Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

Suas conclusões devem ser consistentes de acordo com o que os dados permitem concluir e as técnicas informam

MY HOBBY: EXTRAPOLATING



Fonte: XKCD
<https://xkcd.com/605/>

Referência adicional: <http://www.explainxkcd.com/wiki/index.php/Extrapolating>

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

Faça um documento técnico (relatório) detalhado com todos os passos, incluindo:

- A etapa de limpeza dos dados com as decisões tomadas;
- Etapa de análise descritiva, novamente com as decisões tomadas e achados relevantes (e aqueles que não pareceram relevantes também!);
- A construção do modelo, com as principais tentativas e a versão final;
 - Inclua a interpretação, quando houver;
 - Se for um modelo preditivo/classificatório, inclua a etapa de validação e resultados relacionados;
- Faça considerações a respeito dos possíveis impactos que as decisões tomadas (de inclusão/exclusão de variáveis, tipo de modelo, limitações nos dados) tem nos resultados obtidos e nas extrapolações.

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

Faça um documento técnico (relatório) detalhado com todos os passos, incluindo:

- A etapa de limpeza dos dados com as decisões tomadas;
- Etapa de análise descritiva, novamente com as decisões tomadas e achados relevantes (e aqueles que não pareceram relevantes também!);
- A construção do modelo, com as principais tentativas e a versão final;
 - Inclua a interpretação, quando houver;
 - Se for um modelo preditivo/classificatório, inclua a etapa de validação e resultados relacionados;
- Faça considerações a respeito dos possíveis impactos que as decisões tomadas (de inclusão/exclusão de variáveis, tipo de modelo, limitações nos dados) tem nos resultados obtidos e nas extrapolações.

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

Faça um documento técnico (relatório) detalhado com todos os passos, incluindo:

- A etapa de limpeza dos dados com as decisões tomadas;
- Etapa de análise descritiva, novamente com as decisões tomadas e achados relevantes (e aqueles que não pareceram relevantes também!);
- A construção do modelo, com as principais tentativas e a versão final:
 - Inclua a interpretação, quando houver;
 - Se for um modelo preditivo/classificador, inclua a etapa de validação e resultados relacionados;

Interpretando e comunicando achados



Faça um documento técnico (relatório) detalhado com todos os passos, incluindo:

- A etapa de limpeza dos dados com as decisões tomadas;
- Etapa de análise descritiva, novamente com as decisões tomadas e achados relevantes (e aqueles que não pareceram relevantes também!);
- A construção do modelo, com as principais tentativas e a versão final;
 - Inclua a interpretação, quando houver;
 - Se for um modelo preditivo/classificatório, inclua a etapa de validação e resultados relacionados;
- Faça considerações a respeito dos possíveis impactos que as decisões tomadas a respeito dos dados tenham sobre os resultados obtidos e nas extrapolações.

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

- Faça uma pré-apresentação mais detalhada para a equipe técnica do cliente, para alinhar resultados e corrigir pequenos problemas;

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

- Faça uma pré-apresentação mais detalhada para a equipe técnica do cliente, para alinhar resultados e corrigir pequenos problemas;
 - Se necessário, refaça essa pré-apresentação com pontos que não ficaram bons da primeira vez;

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

- Faça uma pré-apresentação mais detalhada para a equipe técnica do cliente, para alinhar resultados e corrigir pequenos problemas;
 - Se necessário, refaça essa pré-apresentação com pontos que não ficaram bons da primeira vez;
- Prepare uma apresentação final, ressaltando os pontos principais e contando a história para o cliente

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

- Faça uma pré-apresentação mais detalhada para a equipe técnica do cliente, para alinhar resultados e corrigir pequenos problemas;
 - Se necessário, refaça essa pré-apresentação com pontos que não ficaram bons da primeira vez;
- Prepare uma apresentação final, ressaltando os pontos principais e contando a história para o cliente - **se possível envolva a equipe técnica da empresa cliente junto;**

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

- Faça uma pré-apresentação mais detalhada para a equipe técnica do cliente, para alinhar resultados e corrigir pequenos problemas;
 - Se necessário, refaça essa pré-apresentação com pontos que não ficaram bons da primeira vez;
- Prepare uma apresentação final, ressaltando os pontos principais e contando a história para o cliente - se possível envolva a equipe técnica da empresa cliente junto;
 - Deixe os detalhes sórdidos para seu relatório técnico

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

- Dependendo do contrato, entregue as bases de dados transformadas e códigos finais;

Interpretando e comunicando achados



THE
DEVELOPER'S
CONFERENCE

- Dependendo do contrato, entregue as bases de dados transformadas e códigos finais;
- Lembre-se de especificar no contrato inicial a periodicidade de avaliação de performance do modelo além dos custos para atualização do modelo.

Wrapping up



THE
DEVELOPER'S
CONFERENCE

- Data Science não é: matemática ou estatística ou programação ou machine learning;

Wrapping up



THE
DEVELOPER'S
CONFERENCE

- Data Science não é: matemática ou estatística ou programação ou machine learning;
 - É um conjunto de técnicas de diversas áreas que possibilita trazer luz sobre situações reais quando temos dados disponíveis;
 - Você não precisa dominar tudo!

Wrapping up



THE
DEVELOPER'S
CONFERENCE

- Data Science não é: matemática ou estatística ou programação ou machine learning;
 - É um conjunto de técnicas de diversas áreas que possibilita trazer luz sobre situações reais quando temos dados disponíveis;
 - Você não precisa dominar tudo!
- Não existe receita de bolo (sry about that)



Wrapping up



THE
DEVELOPER'S
CONFERENCE

- Registre tudo em um daily-log da equipe;

Wrapping up



THE
DEVELOPER'S
CONFERENCE

- Registre tudo em um daily-log da equipe;
- A skill mais importante é a capacidade de **transformar um problema do cotidiano em uma pergunta de data science**;

Wrapping up



THE
DEVELOPER'S
CONFERENCE

- Registre tudo em um daily-log da equipe;
- A skill mais importante é a capacidade de transformar um problema do cotidiano em uma pergunta de data science;
- Valorize o trabalho em equipe e discuta os resultados constantemente.



THE
DEVELOPER'S
CONFERENCE



THE
DEVELOPER'S
CONFERENCE

MUITO OBRIGADO!



aishameriane@gmail.com



braunmagrin@gmail.com

