



AUDITORIA DE DADOS

Flávio Brito



AGENDA

MÓDULO I

Introdução

Proposta do Curso e objetivos

Material do Curso e dinâmica

Recursos

MÓDULO 2

Fundamentos sobre Auditoria de Dados

Conceitos gerais e definições

Histórico e utilizações

Aplicações corporativas

O mercado

Preparação profissional

Recursos

AGENDA

MÓDULO 3

Arquivos

Metadados e Armazenagem
Integração com Excel, Python e R
Operações e Status

Listagem em Diretório

Permissionamento

Mudança de Diretório

Cópia

Remoção

Merge

Append

Split

Compactação via script

AGENDA

MÓDULO 4

Carga / Importação

Estruturação de Rotinas

Boas práticas

Dados

Tipos de Dados

Metadados e rastreabilidade

Limpeza de Dados

Filtro de Dados

Operações com Variáveis

Coerções

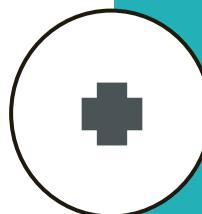
Pitfalls

Outliers

Padrões de Informação

Introdução

- Proposta do Curso e objetivos
- Material do Curso e dinâmica
- Recursos



Fundamentos sobre Auditoria de Dados

Conceitos gerais e definições

- Auditoria de dados é definida como o processo de inspeção, limpeza, transformação, e modelagem de dados com o objetivo de destacar informações úteis, sugerindo conclusões, e apoiar a tomada de decisões. (Diversas fontes)
- Auditoria de dados é um processo analítico pelo qual ideias são extraídas de dados eletrônicos internos e externos a partir de fontes operacionais, financeiras entre outras. Esses insights podem ser históricos, em tempo real, ou preditivos e também pode ser focadas em risco (por exemplo, controle de eficácia, fraude, desperdício, abuso, política de não conformidade regulamentar) ou o desempenho, centrado (por exemplo, aumento de vendas, redução de custos, melhorou rentabilidade) e frequentemente fornecer o "Como?" e "Por quê?" respostas para a inicial "o quê?" perguntas frequentemente encontradas na informação inicialmente extraídas dos dados. (KPMG)

Histórico

As aplicações de auditoria são um tipo aplicação voltadas especificamente para processos de auditoria tanto para empresas médias e grandes, especialmente quando algumas das aplicações são desenvolvidas in-house.

Há alguns princípios básicos de aplicações de auditoria que os auditores de TI precisam saber e entender.

- No fim dos anos 1980 – criados por empresas voltadas para o desenvolvimento de softwares genéricos para a área de auditoria
 - ACL, 1987
 - Caseware, 1988
- Charles Carslaw, *Applying Benford's Law to Accounting*, 1988
- Continuous Process Auditing System, AT&T Bell Laboratories, 1989

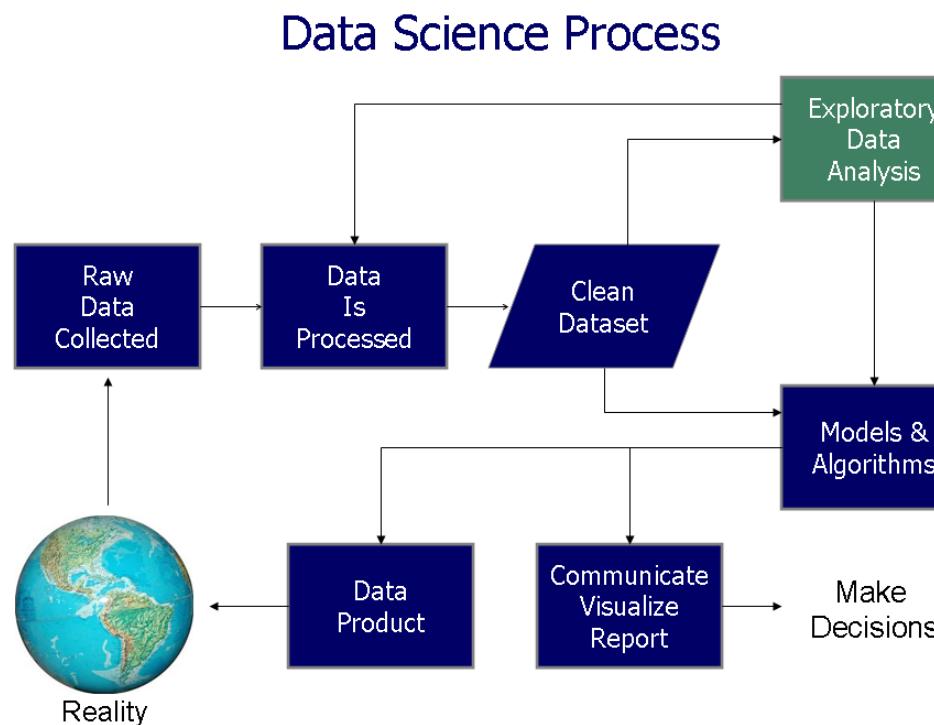
Aplicações Corporativas

- TIPO
 - Sistemas Transacionais
 - Monitoramento
 - Análise em Tempo Real
 - Predição
- ÁREAS
 - TI
 - Financeira
 - RH
 - Governo
 - Demais Indústrias



Mercados de Trabalho

- Data Auditor
- Data Analyst
- Data Scientist



Utilizações Funcionalidades Desejadas nas Ferramentas

- Capazes de manipular com grandes bases de dados
- Vasto número de funcionalidades estatísticas e analíticas
- Programável
- Registro de procedimentos – Rastreabilidade
- Facilidade de execução na mudança de dados

```
ngScope, element, attr, ngSwitchController) {  
  var = attr.ngSwitch || attr.on,  
  transcludes = [],  
  elements = [],  
  scopes = [];  
  
  switchExpr, function ngSwitchWatchAction(value) {  
    i = previousElements.length; i < ii; ++i) {  
      elements[i].remove();  
  
      ts.length = 0;  
  
      selectedScopes.length; i < ii; ++i) {  
        selectedElements[i];  
        [i].$destroy();  
        ts[i] = selected;  
        selected, function() {  
          ts.splice(i, 1);  
  
          length = 0;  
          th = 0;  
  
          udes = ngSwitchController.cases['!' + value] || ngSwitchController.  
          includes, function(selectedTransclude) {  
            = scope.$new();  
            h(selectedScope);  
          }  
        }  
      }  
    }  
  }  
}
```

Tipos de Ferramenta

- Excel
 - Bancos de Dados Corporativos
 - Desktop software (Microsoft Access)
 - Software genérico para auditoria
 - ACL (1987)
 - IDEA DATA ANALYTICS



```
ngSwitch, element, attr, ngSwitchController) {
  var = attr.ngSwitch || attr.on,
  transcludes = [],
  elements = [],
  elements = [],
  scopes = [];
  switchExpr, function ngSwitchWatchAction(value) {
    i = previousElements.length; i < ii; ++i) {
      elements[i].remove();
    }
    elements.length = 0;
    selectedScopes.length; i < ii; ++i) {
      selectedElements[i];
      destroy();
      = selected;
      selected, function() {
        elements.splice(i, 1);
      }
    }
    length = 0;
    length = 0;
    includes = ngSwitchController.cases['!' + value] || ngSwitchController
    transcludes, function(selectedTransclude) {
      = scope.$new();
      (selectedScope).
```

Alguns Desafios para a Auditoria de Dados



- Escolha do Software
- Fonte de Dados
- Segurança e Privacidade das Informações
- Fonte de Dado e Saída de Informação (Visualização)
- Qualificação e Experiência da Equipe
- Poder de Processamento

Software

EXCEL

Especificações e limites de planilha e de pasta de trabalho

Recurso	Limite máximo
Pastas de trabalho abertas	Limitado pela memória disponível e pelos recursos do sistema
Tamanho da planilha	1.048.576 linhas por 16.384 colunas



- Capacidade de Processamento

Especificações do banco de dados

Lista de tabelas a seguir se aplica aos bancos de dados do Access 2016:

ACCESS

Geral

Atributo	Máximo
Tamanho total para um banco de dados do Access 2016 (.accdb), incluindo todos os objetos de banco de dados e dados	2 gigabytes menos o espaço necessário para objetos do sistema.

OBSERVAÇÃO : Você pode contornar esta limitação de tamanho vinculando a tabelas em outros bancos de dados do Access. Você pode vincular a tabelas em vários arquivos de banco de dados, cada um dos quais pode ser tão grande quanto 2GB.

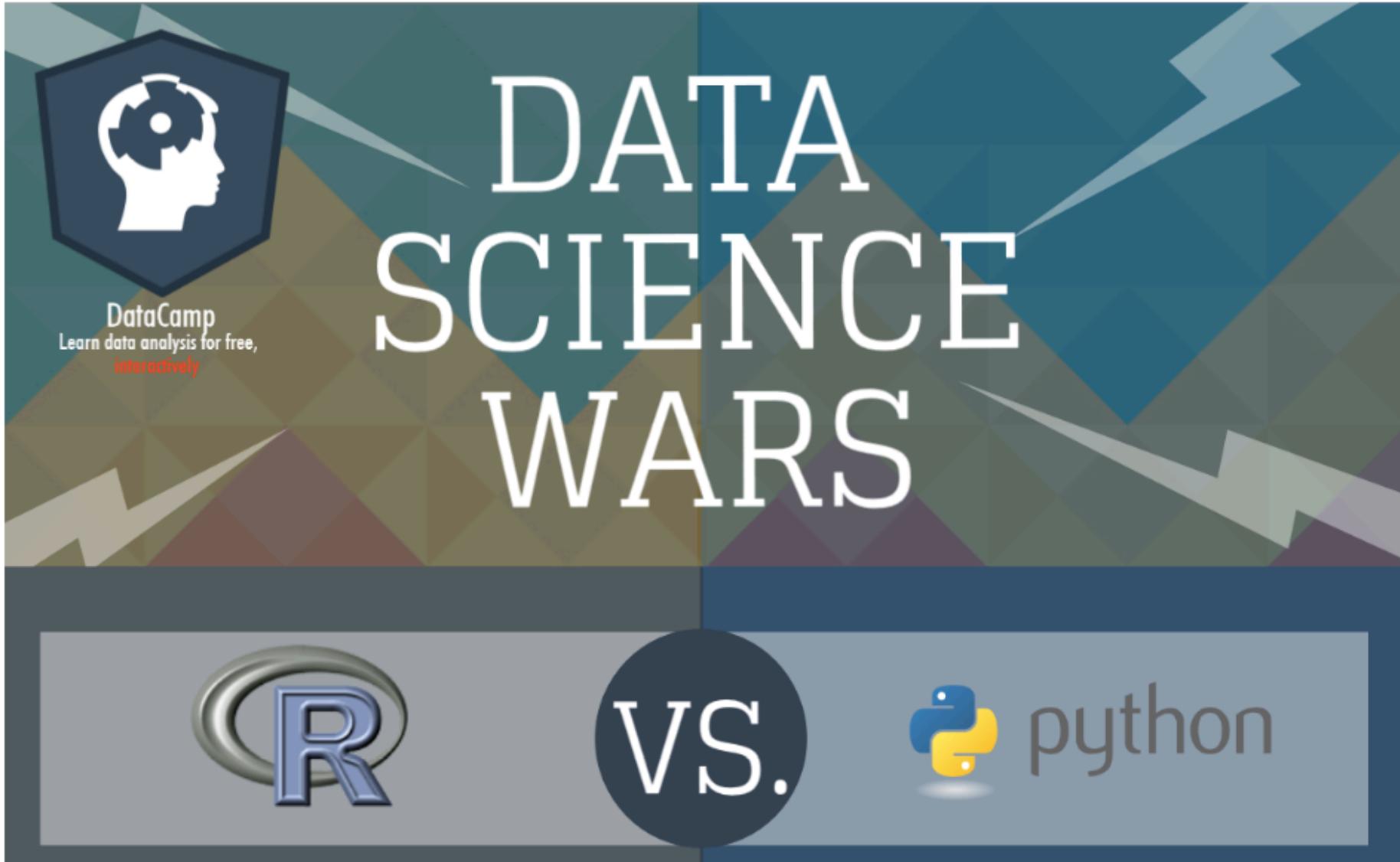
DICA : Para obter mais informações sobre como reduzir o tamanho do seu banco de dados, consulte [impedir e corrigir problemas de arquivo de banco de dados usando compactar e reparar](#).

Tipos de Ferramenta

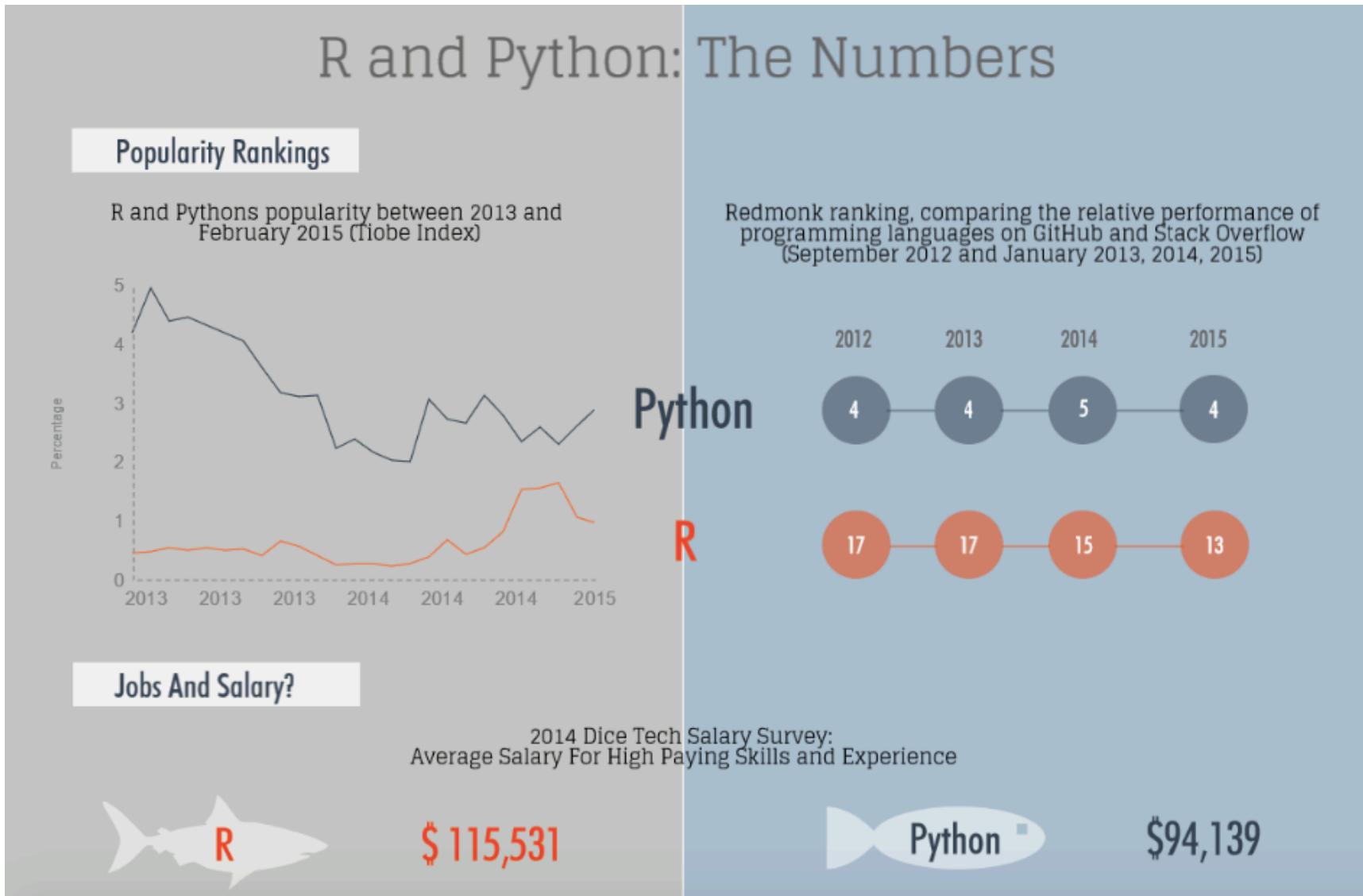


- Ferramentas Estatísticas e de Análise de Dados
 - R (1993)
 - SAS (1976)
 - Python (1991) – Pandas e Scikit-learnn (2011)

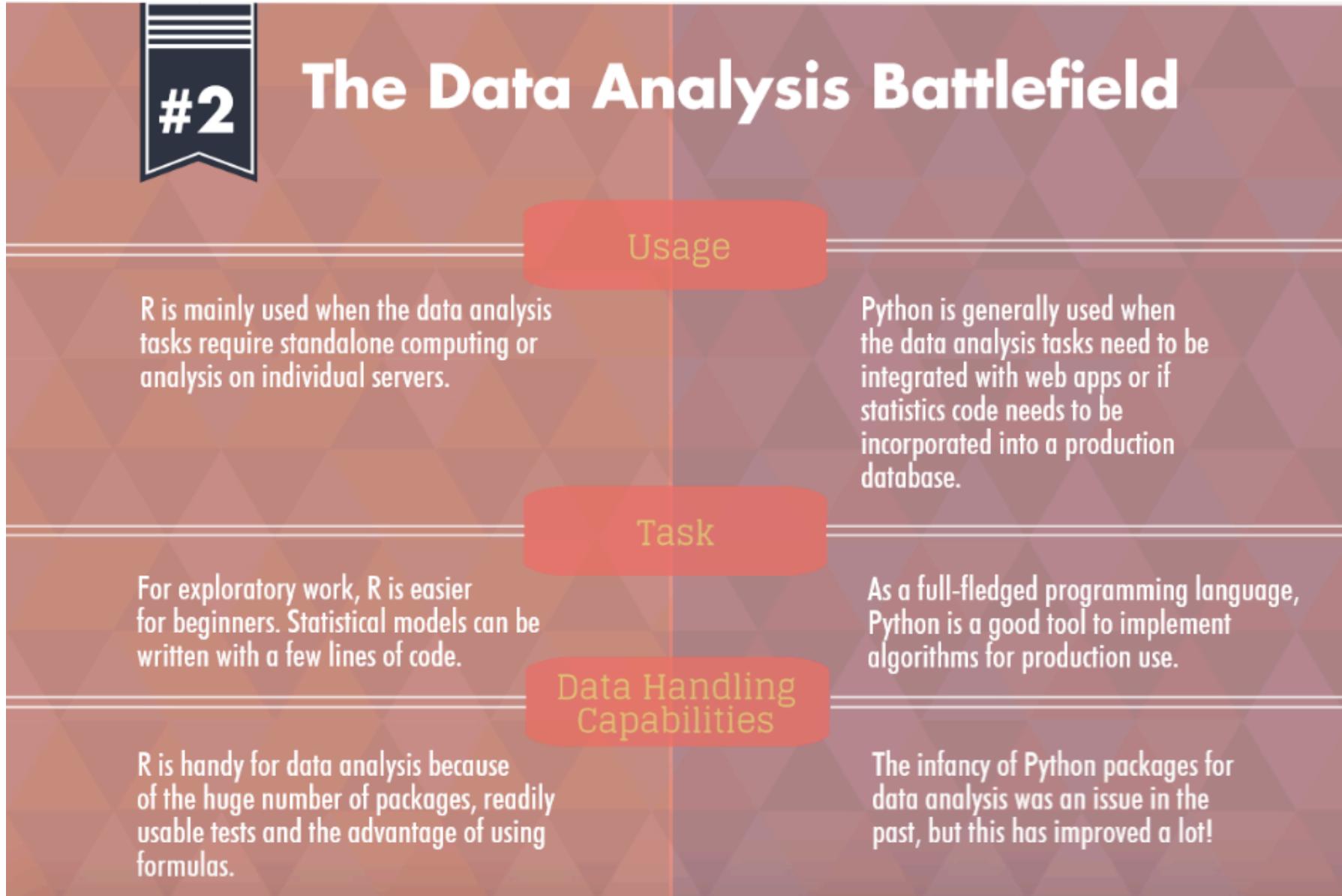
Software



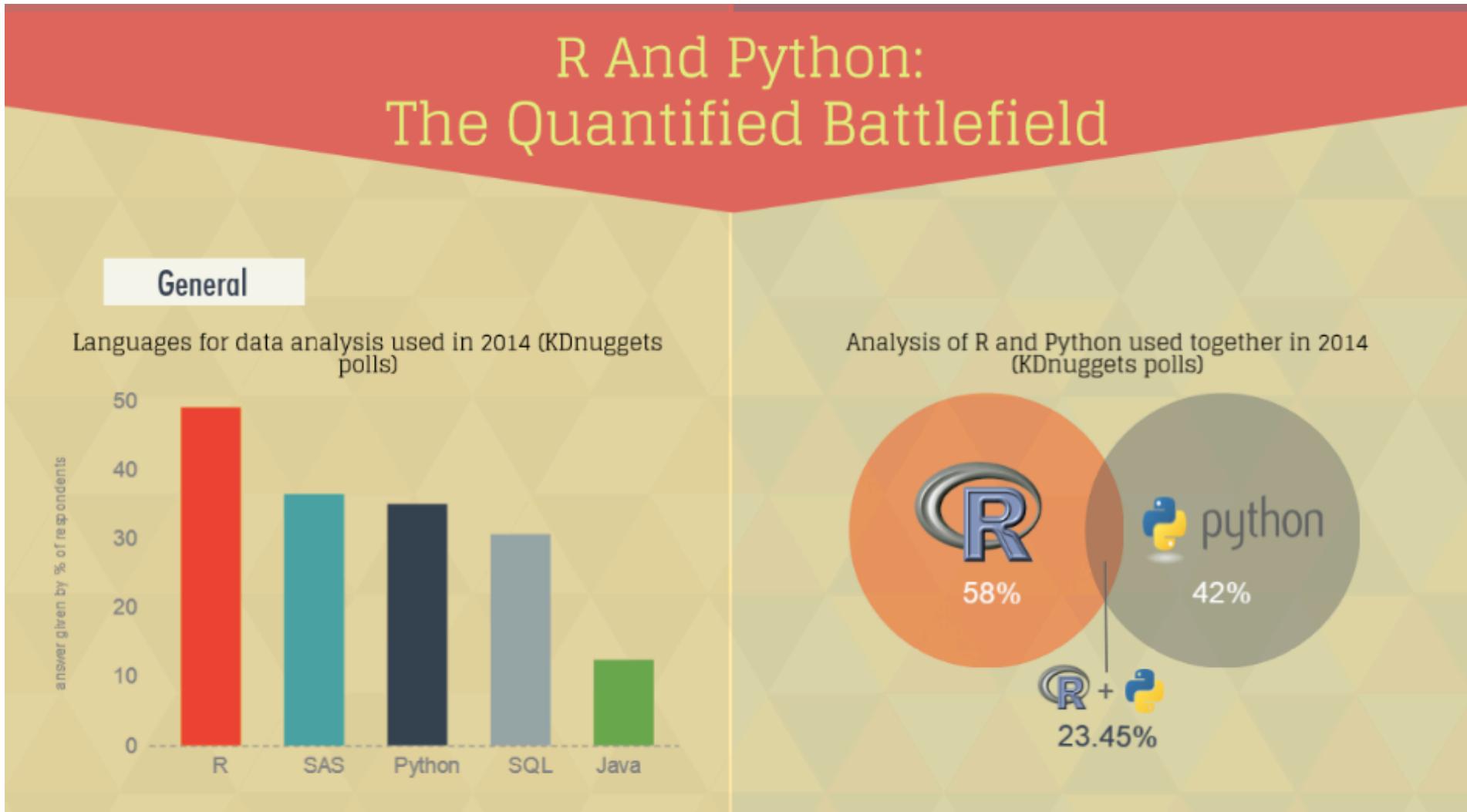
Software



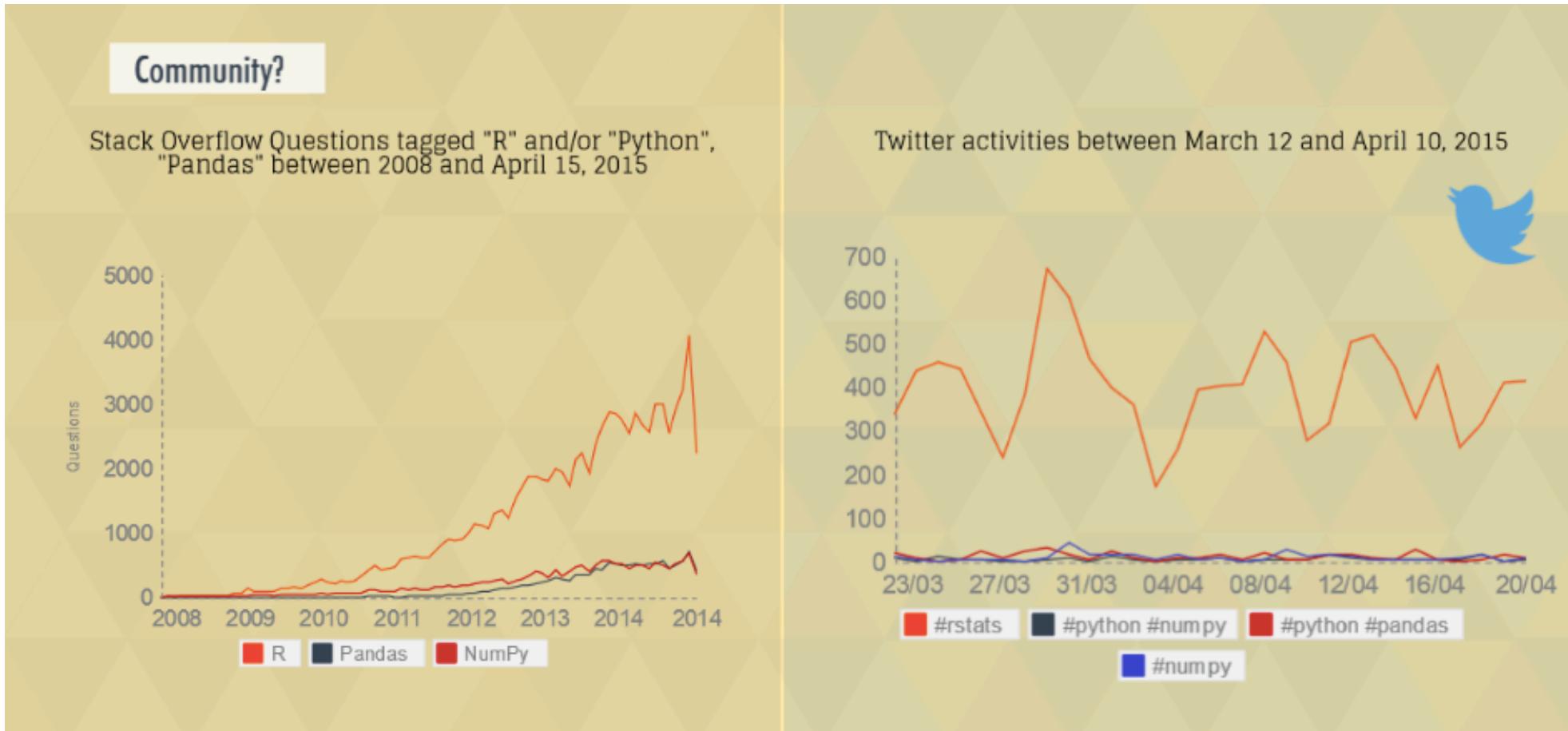
Software



Software

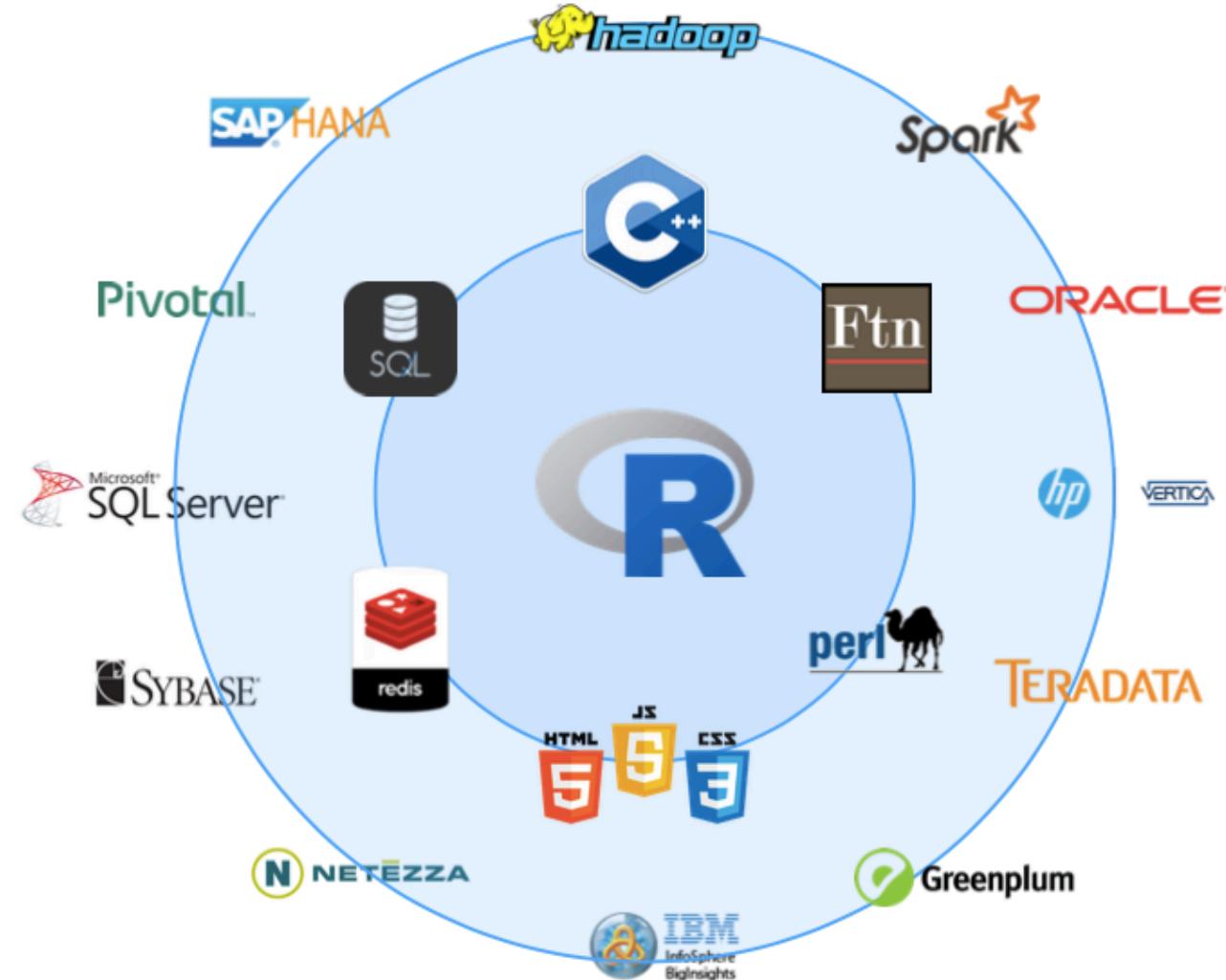


Software



Software

R Studio



Planejamento

Audiência	Aquisição	Valor
Demografia Interesse Geo Localização Linguagem Dispositivo Rede Comportamento / Engajamento Tempo para análise	Busca Orgânica Acesso Direto ou por Referência (WS) Palavras-chave / Indexação	Custo da Busca e Análise
<h2>Comportamento</h2> <p>Rollback mais fácil Entender os pontos fracos Enteder os porntos Fortes Acelerar os resultados</p>		Conversão (Sucesso) Caminho de Conevrção Perfomenace Fluxo de Objetivos

Direto ao Ponto

PROBLEMA	SOLUÇÃO: Dividir para Conquistar
<p>Dados</p> <ul style="list-style-type: none">- Acesso<ul style="list-style-type: none">- Segurança- Periodicidade- Tipo de ...- Metadados- Workflow- Carga- Disponibilidade- Crescimento dos Dados	<p>Planejamento</p> <ul style="list-style-type: none">• Data sources• Databases metadata• Métricas <p>COMPORTAMENTO</p> <ul style="list-style-type: none">• KPI• Analysis & Modeling to Predict• Armazenamento <p>Rollback mais fácil Entender os pontos fracos Enteder os porntos Fortes Acelerar os resultados</p>

Data Audit pode lhe dizer

Datasources	Databases & Arquivos de metadados	KPI
Acesso Disponibilidade Qualidade dos Metadados Qualidade dos Dados Qualidade do Armazenamento Tempo de Carga Tempo de Disponibilidade de Conexão Espaço de Armazenagem Crescimento dos Dados Tamanho da Base	O propósito da criação da base Propósito do Dado Hora e Data de Criação Criador e Responsáveis pela manutenção do Dado Localização, Arquivos, Tamanho, Granularidade Métricas Total de Metadados / Campos / Bancos / Arquivos Inconsistentes Total de Registros processados % de Falhas / Objeto Taxa de Auditoria Total de Dados Missing / Retenção de Qualidade Custo por Processamento de Extração / Auditoria ROI	Crescimento da Base Engajamento para Acerto de Dados sem Gravidade Custo/Benefício por Auditoria Efetividade da Auditoria

O Auditor de
Dados também
pode lhe falar ...

- **Predição e Estimativa**
- **Modelagem**



```
function(element, attr, ngSwitchController) {
    var previousElements = [];
    var selectedElements = [];
    var selectedScopes = [];

    if (attr.ngSwitch || attr.on,
        attr.ngTransclude || attr.ngSwitchController) {
        var previousElements = [];
        var selectedElements = [];
        var selectedScopes = [];

        if (attr.ngSwitchExpr, function ngSwitchWatchAction(value) {
            var i, ii;
            for (i = 0, ii = previousElements.length; i < ii; ++i) {
                previousElements[i].remove();
            }
            previousElements.length = 0;

            for (i = 0, ii = selectedScopes.length; i < ii; ++i) {
                var selected = selectedElements[i];
                selectedScopes[i].$destroy();
                previousElements[i] = selected;
                $animate.leave(selected, function() {
                    previousElements.splice(i, 1);
                });
            }

            selectedElements.length = 0;
            selectedScopes.length = 0;

            if ((selectedTranscludes = ngSwitchController.cases['!' + value] || ngSwitchController.cases[value])) {
                var selectedScope = scope.$new();
                selectedScopes.push(selectedScope);
                selectedScope.$eval(attr.change);
                forEach(selectedTranscludes, function(selectedTransclude) {
                    var selectedScope = scope.$new();
                    selectedScopes.push(selectedScope);
                    selectedScope.$eval(selectedTransclude);
                });
            }
        });
    }
}
```