

The background of the slide is a grayscale photograph of a library aisle, with rows of bookshelves filled with books on either side. A solid blue horizontal banner is positioned across the middle of the image, containing the main title in white text.

Painel em Python do Centro de Inteligência

Kallil de Araújo Bezerra
Universidade Federal do Rio Grande do Norte
Outubro 2020

- ① Histórico do BI e aplicações na JFRN
- ② Construção do painel
- ③ Detecção de anomalias
- ④ Conclusões

① Histórico do BI e aplicações na JFRN

② Construção do painel

③ Detecção de anomalias

④ Conclusões

Conceitos básicos de *Business Intelligence*

- ▶ Termo apresentado em 1865 por Richard Millar Devens
- ▶ Apareceu novamente em 1958 num artigo de Hans Peter Luhn
- ▶ Difundido em várias instituições - computadores menos caros
- ▶ Uso da informática para suporte às **tomadas de decisões**
 - > Visualização de dados
 - > Análise de dados
 - > Armazenamento de dados

1 Aplicação na JFRN

| 4

- ▶ Usado por servidores e magistrados
- ▶ Acessado através dos navegadores
- ▶ Desenvolvido com Qlikview
- ▶ Problemas de entendimento entre painéis
- ▶ TRF5 recebe as demandas

Portal de Business Intelligence da 5ª Região

Versão 1.7.8 (21/09/2020 18:00)

[Início](#) / [Gestão Processual](#) / [Relatórios Judiciais](#)

Relatórios Judiciais

1º Grau

2º Grau

JEF

Relatórios


 Atos Magistrados Estatísticas de Distribuição Prazos Vencidos/Sem Movimentação Processos Migrados Relatório Dinâmico de Processos

Figura: Portal BI do TRF5

1 Ferramentas disponíveis

| 6

As ferramentas mais usadas:

- ▶ Qlikview
- ▶ PowerBI
- ▶ Pentaho
- ▶ Metabase



Figura: Ferramentas BI

- ① Histórico do BI e aplicações na JFRN
- ② Construção do painel
- ③ Detecção de anomalias
- ④ Conclusões

2 Escolhendo as ferramentas

| 8

Pontos que devem ser levados em consideração na escolha da ferramenta:

- ▶ Pago vs gratuito
- ▶ Pronto vs próprio
- ▶ Facilidade de se desenvolver



Figura: Comprado vs construído

2 Python escolhido!

| 9

- ▶ Linguagem de programação
- ▶ Lançada em 1991 por Guido van Rossum
- ▶ Sintaxe fácil de ler
- ▶ Grande quantidade de bibliotecas *open-source*
- ▶ Usado em grandes empresas:
 - > Netflix
 - > Spotify
 - > Google
 - > Instagram
 - > Facebook

► Prós:

- > Gratuito
- > Desenvolvido *in-house*
- > Fácil de se encontrar desenvolvedores
- > Extremamente customizável

► Contras:

- > Gratuito
- > Desenvolvido *in-house*

- ▶ Python
 - > Dash - visualização dos dados
 - > Numpy - análise dos dados
- ▶ Qlikview - conversão dos dados para formato que possa ser lido pelo painel

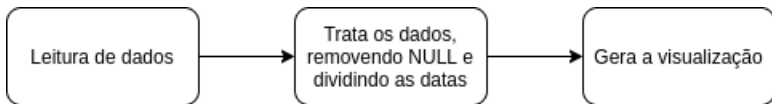


Figura: Fluxo do painel

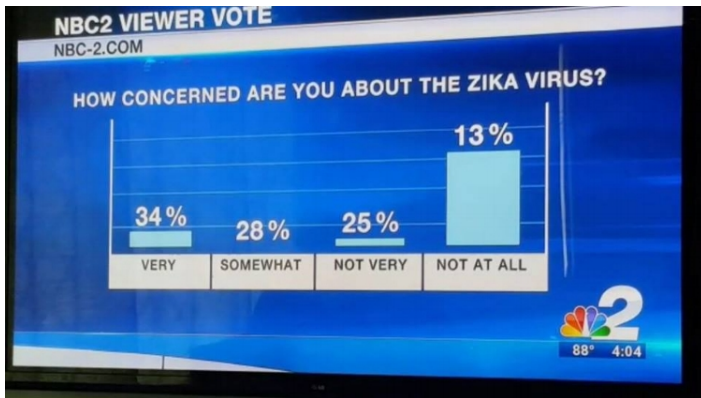


Figura: Gráfico tendencioso

- ▶ Mistura de arte com ciência
- ▶ Deve transmitir o comportamento dos dados
- ▶ Nunca distorcer ou enganar
- ▶ Deve apresentar elementos balanceados
- ▶ Boa escolha de cores

Uma visualização de dados mal feita pode levar a conclusões erradas, resultando em uma má tomada de decisões.

2 Visualização de dados

| 14

Julho: Total-Jul	
Dívida Ativa	482
Competência da Justiça Federal	301
Cartão de Crédito	186
Profissional	97
Lei de Imprensa	87
Efeito Suspensivo / Impugnação / Embargos à Execução	46

Figura: 6ª Vara em 2015



- ▶ Entender o comportamento
- ▶ Relacionar com os negócios
- ▶ A análise NÃO é isolada, precisa de integração entre diferentes áreas e disciplinas
- ▶ Está ligada à visualização de dados
- ▶ É importante entender as causas de determinados comportamentos

2 Análise de dados - exemplo

| 17



É importante entender a causa do comportamento observado nos dados para poder agir melhor.

Em linhas gerais a estrutura do painel fica da seguinte forma:

- ▶ Carregamento dos dados em formato .csv, convertido pelo Qlikview
 - > O QlikView possui um formato próprio de calcular o tempo, e ele é necessário porque os dados do TRF5 são tratados nele
- ▶ Filtragem dos dados, pegando a partir de 2014 e removendo *Null*
 - > Alguns dados são carregados com problemas, seja no número do processo, seja na data, eles precisam ser removidos para não atrapalharem a análise e visualização dos dados
- ▶ Associação dos dados com a filtragem feita pelo usuário do painel
 - > Nesse ponto a visualização é gerada e apresentada de acordo com o que o usuário escolher (combinação de Vara e Ano)

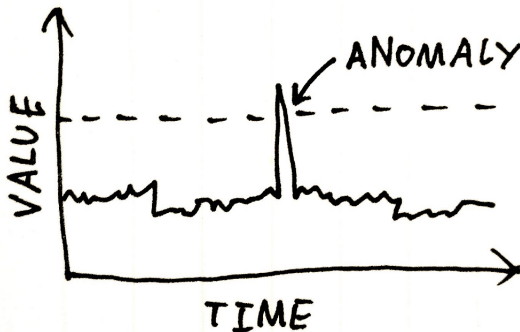
É importante lembrar que as análises são dinâmicas, ou seja, as médias e o que é considerado anômalo muda de acordo com o período selecionado.

- 1 Histórico do BI e aplicações na JFRN
- 2 Construção do painel
- 3 Detecção de anomalias
- 4 Conclusões

3 Objetivo da detecção de anomalias

| 20

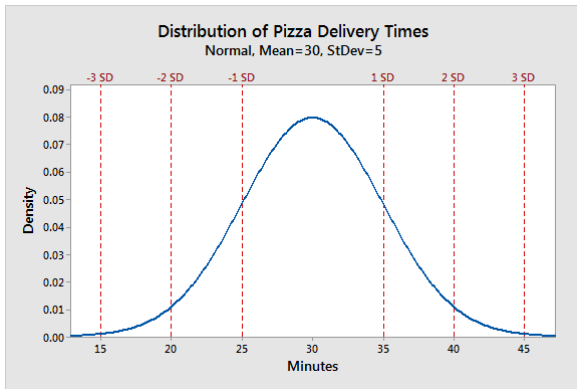
- ▶ Detecção de valores destoantes
- ▶ Análise pelo Centro de Inteligência
 - > O Centro de Inteligência trabalha para diminuir a multiplicação das demandas repetitivas, evitando uma sobrecarga na JFRN



3 Distribuição dos dados

| 21

Existem diversos métodos de detecção de anomalias, mas eles quase sempre são focados em distribuição normal.



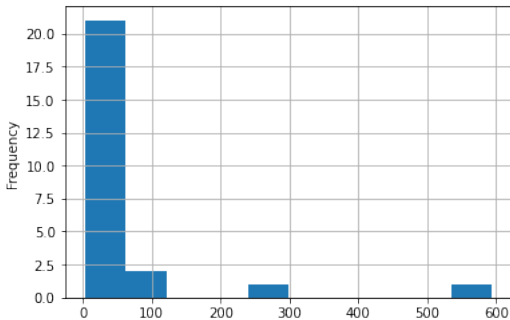
As Varas apresentam um comportamento um pouco diferente por causa das competências.

3 Distribuição dos dados

| 23

- 1 Dívida ativa - 594
- 2 Profissional - 266
- 3 Competência da Justiça Federal - 111
- 4 Efeito Suspensivo/Impugnação/Embargos à Execução - 102

6ª Vara - fevereiro de 2015

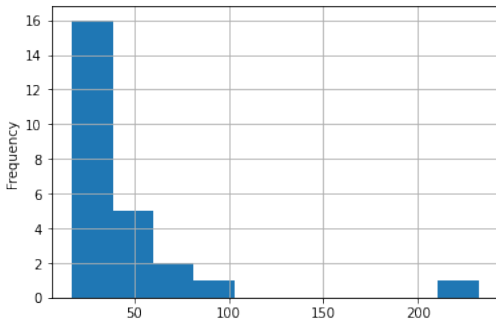


3 Distribuição dos dados

| 24

- 1 Vícios de Construção - 232
- 2 Contratos Bancários - 94

4ª Vara - junho de 2016

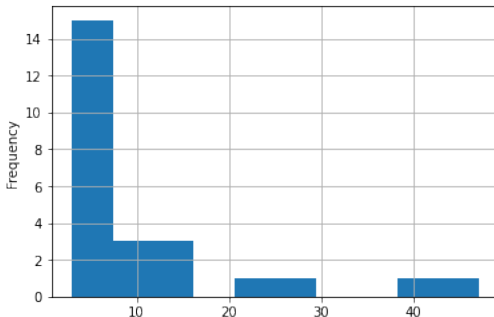


3 Distribuição dos dados

| 25

- 1 Competência da Justiça Federal - 47
- 2 Pena de Multa - 41
- 3 Prestação de Serviços à Comunidade - 29

14ª Vara - julho de 2018



3 Distribuição dos dados

| 26

- 1 Competência da Justiça Federal - 47
- 2 Pena de Multa - 41
- 3 Prestação de Serviços à Comunidade - 29



3 Encontrando os Assuntos mais frequentes

| 27

A definição do que seria uma anomalia passou a levar em consideração a frequência dos Assuntos e a distância dele em relação à média. Então, a partir disso foram elaborados três níveis diferentes:

- ▶ $media_{assuntos}$ que é a média simples dos assuntos, a cada dois anos:

$$media_{assuntos} = \sum_{ano}^{ano-1} \frac{assuntos}{total_{meses}}$$

- ▶ $anomalia_2$ definida como: $anom_2 = media_{assuntos} + (2 * \sigma)$

- ▶ $anomalia_1$ definida como: $anom_1 = media_{assuntos} + \sigma$

- > Vermelho se $total \geq anom_2$
- > Amarelo se $total \geq anom_1$ e $total < anom_2$
- > Verde se $total \geq media_{assuntos}$ e $total < anom_1$

3 Distribuição dos dados

| 28

Resultado:

Dezembro:	Total-Dez
Dano ao Erário	121
Violação aos Princípios Administrativos	106
Enriquecimento ilícito	49
Dívida Ativa	30
Profissional	19
Piso Salarial	17

- ① Histórico do BI e aplicações na JFRN
- ② Construção do painel
- ③ Detecção de anomalias
- ④ Conclusões

As ferramentas gratuitas e o time de TI podem contribuir bastante para a gestão da JFRN, gerando análises que auxiliem os magistrados e servidores da Justiça a ter uma melhor visão do funcionamento das Varas e sem pagar por licenças de *software*. Além disso, os painéis poderiam ser desenvolvidos e distribuídos usando o pessoal e a infraestrutura da JFRN. A manutenção dessas ferramentas também é simples porque usam Python, uma das linguagens de programação mais usadas no mundo todo e relativamente fácil de se aprender.

