



UNIVERSIDADE ESTADUAL DO NORTE DO PARANÁ
CAMPUS LUIZ MENEGHEL - CENTRO DE CIÊNCIAS TECNOLÓGICAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

JOÃO VITOR COPPI MARTINS

**ANÁLISE DE DADOS HISTÓRICOS PARA PREVISÃO E
VISUALIZAÇÃO DE INFORMAÇÕES SOBRE ACIDENTES
RODOVIÁRIOS.**

BANDEIRANTES-PR

2017

JOÃO VITOR COPPI MARTINS

**ANÁLISE DE DADOS HISTÓRICOS PARA PREVISÃO E
VISUALIZAÇÃO DE INFORMAÇÕES SOBRE ACIDENTES
RODOVIÁRIOS.**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Ciência da Computação da Universidade Estadual do Norte do Paraná para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Me. Thiago Adriano Coleti
Coorientador: Prof^a. Dra. Tamara Angélica Baldo

BANDEIRANTES-PR

2017

JOÃO VITOR COPPI MARTINS

**ANÁLISE DE DADOS HISTÓRICOS PARA PREVISÃO E
VISUALIZAÇÃO DE INFORMAÇÕES SOBRE ACIDENTES
RODOVIÁRIOS.**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Ciência da Computação da Universidade Estadual do Norte do Paraná para obtenção do título de Bacharel em Ciência da Computação.

BANCA EXAMINADORA

Prof. Me. Thiago Adriano Coleti
Universidade Estadual do Norte do Paraná
Orientador

Prof. Dr. Tamara Angélica Baldo
Universidade Estadual do Norte do Paraná

Prof. Dr. Bruno Squizato Faiçal
Universidade Estadual do Norte do Paraná

Bandeirantes-PR, ____ de novembro de 2017

AGRADECIMENTOS

Quero começar agradecendo meus pais, João e Edneia, que sempre me apoiaram, sempre me deram condições e principalmente nunca deixaram de acreditar em mim.

A minha namorada, Jade, que nunca falhava em me animar, sem ela eu não teria energias ou forças pra continuar sempre me esforçando.

Ao meu orientador, Prof. Me. Thiago Coleti, por seu apoio e dedicação, competência e não me deixar desistir de minhas ideias.

A minha coorientadora, Prof^ª. Dra. Tamara Baldo, que sempre me ajudou a me organizar, sem sua firmeza eu não teria nem mesmo feito a primeira entrega.

Ao Prof. Dr. Bruno Faiçal, por sempre me auxiliar nas reuniões de entregas e sempre dar aquela crítica construtiva, sem elas o TCC também não sairia.

Aos professores do Centro de Ciências Tecnológicas que, de alguma forma, contribuíram para minha formação.

As amizades que foram feitas durante esses anos, sem elas eu nunca teria chego até aqui.

Obrigados a todos os envolvidos na minha formação, seja de maneira direta ou indireta, vocês fazem parte dessa conquista.

MARTINS, J. V. C.. **Análise de dados históricos para previsão e visualização de informações sobre acidentes rodoviários..** 52 p. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Universidade Estadual do Norte do Paraná, Bandeirantes-PR, 2017.

RESUMO

O termo segurança nas estradas é algo muito discutido, principalmente no Brasil onde há um grande índice de acidentes. Tendo em vista tal constatação, este trabalho busca utilizar os dados históricos(acidentes que já ocorreram e tiveram suas informações armazenadas), disponibilizados pela Polícia Rodoviária Federal em sua base de dados, para prever os possíveis acidentes por toda extensão da BR-376 no Paraná por meio de métodos estatísticos de previsão, e então unir essas previsões com técnicas de *data visualization*, para transmitir essas informações a pessoas interessadas. Por meio dessa visualização, é possível que administradores de rodovias e órgãos do governo pudessem criar logísticas de distribuição de recursos rodoviários, de acordo com as previsões de acidentes, possivelmente diminuindo o tempo de atendimento à ocorrências e ociosidade de recursos. Ao final do trabalho, já com o protótipo da ferramenta criada, foi possível criar meios de avaliar a mesma, e assim buscar meios para melhora-lá no futuro.

Palavras-chave: análise de dados. dados históricos. data visualization. big data. data science. pesquisa operacional. estatística. matemática

MARTINS, J. V. C.. **Analysis of historical data for prediction and visualization of information of road accidents..** 52 p. Final Project (Bachelor of Science in Computer Science) – State University Northern of Parana , Bandeirantes–PR, 2017.

ABSTRACT

The term road safety is something much discussed, especially in Brazil where there is a high rate of accidents. According to these findings, this work seeks to use historical data (accidents that have already occurred and had their information stored in some database), made available by the Polícia Rodoviária Federal on its website, to predict possible accidents throughout the BR-376 in Paraná through statistical methods of forecasting, and then joining these predictions with data visualization techniques, and finally transmitting this information to people who are interested. Through this visualization, it would be possible for highway administrators and government agencies to create distribution logistics of road resources according to accident forecasts, possibly reducing time to attend to occurrences and idleness of resources. At the end of the work, already with the prototype of the tool created, it was possible to create means to evaluate it, and thus to seek ways to improve it in the future.

Keywords: data analysis. historical data. Data visualization. big data. Data science. operational research. statistics. mathematics.

LISTA DE ILUSTRAÇÕES

Figura 1 – Informações sobre acidentes no Brasil de julho de 2014 até junho de 2015.	10
Figura 2 – Classificação do estado geral.	11
Figura 3 – Exemplo 1 - Postos de atendimento.	13
Figura 4 – Exemplo 2 - Postos de atendimento.	13
Figura 5 – Exemplo 3 - Postos de atendimento.	14
Figura 6 – Aplicação de <i>Data Visualization</i> : transformação de informação em tabela para visualização em forma de gráfico.	20
Figura 7 – Exemplo de Infográfico.	21
Figura 8 – Exemplo de <i>Data Visualization</i>	22
Figura 9 – Processo de modelagem	23
Figura 10 – Métodos de previsão: Qualitativos e Quantitativos	24
Figura 11 – Média de acidentes por rodovia federal (2017, 2016, 2015 e 2014).	28
Figura 12 – Gráfico de acidentes por trecho da BR-376 no ano de 2014.	29
Figura 13 – Gráfico de acidentes por trecho da BR-376 no ano de 2015.	30
Figura 14 – Gráfico de acidentes por trecho da BR-376 no ano de 2016.	30
Figura 15 – Gráfico da média de causa dos acidentes da BR-376 nos últimos anos(2016, 2015, 2014).	31
Figura 16 – Camadas resultantes do processo de tratamento da tabela e aplicação do método de previsão.	33
Figura 17 – Ferramenta protótipo: visão geral.	34
Figura 18 – Ferramenta protótipo: barra de navegação lateral.	35
Figura 19 – Ferramenta protótipo: gráfico de linha.	36
Figura 20 – Ferramenta protótipo: gráficos de barras.	36
Figura 21 – Ferramenta protótipo: gráfico de pizza e de Gauge.	37
Figura 22 – Questionário parte 1.	47
Figura 23 – Questionário parte 2.	48
Figura 24 – Respostas parte 1.	49
Figura 25 – Respostas parte 2.	50
Figura 26 – Respostas parte 3.	51

LISTA DE TABELAS

Tabela 1 – Legenda da tabela de acidentes rodoviários.	32
Tabela 2 – Valor do erro de acordo com causas.	39

LISTA DE ABREVIATURAS E SIGLAS

CNT	Confederação Nacional de Transporte.
SAU	Serviço de Atendimento ao Usuário.
PRF	Polícia Rodoviária Federal.
EDA	<i>Exploratory Data Analysis.</i>
PO	Pesquisa Operacional.

SUMÁRIO

1	INTRODUÇÃO	10
1.1	Justificativa	12
1.2	Objetivos	14
1.3	Materiais e Métodos	15
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	Big Data	17
2.2	Data Science	18
2.3	Data Visualization	19
2.4	Métodos de previsão	22
3	DESENVOLVIMENTO	28
3.1	Análise dos dados	28
3.2	Análise da tabela	30
3.3	Método de previsão	31
3.4	Ferramenta de Visualização	34
4	VALIDAÇÃO	38
4.1	Validação do método de previsão	38
4.2	Validação da ferramenta de visualização	38
5	CONSIDERAÇÕES FINAIS	41
	REFERÊNCIAS	42
	APÊNDICES	44
	APÊNDICE A – TABELA RESULTADOS DA PREVISÃO 2017	45
	APÊNDICE B – QUESTIONÁRIO: FERRAMENTA DE VI- SUALIZAÇÃO	47
	APÊNDICE C – QUESTIONÁRIO: RESPOSTAS	49

1 INTRODUÇÃO

Muito se tem discutido, recentemente, acerca de segurança nas estradas, isso devido ao grande índice de acidentes recorrentes nas rodovias federais do Brasil. Dentre esses acidentes, 43% terminam com mortos ou feridos [1]. "Especialistas explicam que, quanto mais rápido o atendimento médico, há mais chances de sobrevivência e menos possibilidade de sequelas" [2]. Isso reafirma a importância de um atendimento rápido e eficaz, que presumivelmente, evitará maiores danos. Na Figura 1 são mostrados números sobre acidentes, referentes a um período de aproximadamente um ano, que aconteceram em toda malha rodoviária federal.

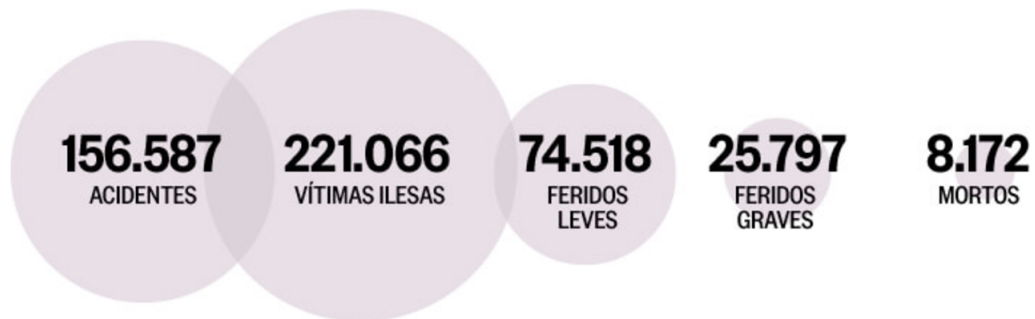


Figura 1 – Informações sobre acidentes no Brasil de julho de 2014 até junho de 2015.

Fonte: [1].

São mais de 150 mil acidentes em um período de aproximadamente um ano, dentre esses, foram mais de 8 mil mortos.

Todas as rodovias do Brasil possuem um padrão definido pelo Departamento Nacional de Infraestrutura de Transportes (DNIT). Este padrão infere que rodovias federais recebem a sigla BR seguida de 3 dígitos. A sigla representa sua característica nacional, o primeiro dígito representa a categoria da rodovia e, por fim, os últimos dois dígitos representam a posição geral da rodovia, relativa à Capital Federal e aos limites do país [3]. Sendo assim, o DNIT classifica as rodovias em 5 categorias, sendo elas:

1. **Rodovias Radiais** - Rodovias que tem início na Capital Federal e partem para os extremos do país, estas recebem como primeiro dígito o número 0. Exemplo: BR-040.
2. **Rodovias Longitudinais** - Rodovias que atravessam o país nas direções Norte e Sul, e recebem como primeiro dígito o número 1. Exemplo: BR-101, BR-153, BR-174.
3. **Rodovias Transversais** - São as rodovias que atravessam o país nas direções Leste e Oeste, e recebem como primeiro dígito o número 2. Exemplo: BR-230, BR-262,

BR-290.

4. **Rodovias Diagonais** - São as rodovias que atravessam o país nas direções Noroeste, Sudeste, Nordeste e Sudoeste e tem como primeiro dígito o número 3. Exemplos: BR-304, BR-369, BR-376.
5. **Rodovias de Ligação** - Estas rodovias não possuem direção definida, podendo seguir qualquer direção, são normalmente usadas para ligar rodovias com outras rodovias, cidades ou pontos importantes. Estas rodovias tem como o primeiro dígito o número 4. Exemplo: BR-401, BR-480.

Em uma pesquisa feita pela Confederação Nacional de Transporte (CNT), em 2016, foram analisados 103.259km de estradas, abrangendo toda a malha rodoviária federal e as principais rodovias estaduais, buscando apontar as características, deficiências e necessidades de melhoria da infraestrutura das rodovias do Brasil. De toda a extensão avaliada, como pode ser observado na Figura 2, 58,2% apresentaram algum tipo de deficiência, seja no pavimento, na sinalização ou na geometria da via [4], isso implica que, hipoteticamente, há uma chance maior de acidentes onde há tais deficiências. Portanto, mais de metade da malha rodoviária brasileira possui um risco maior de acidentes devido às deficiências apontadas pela pesquisa.

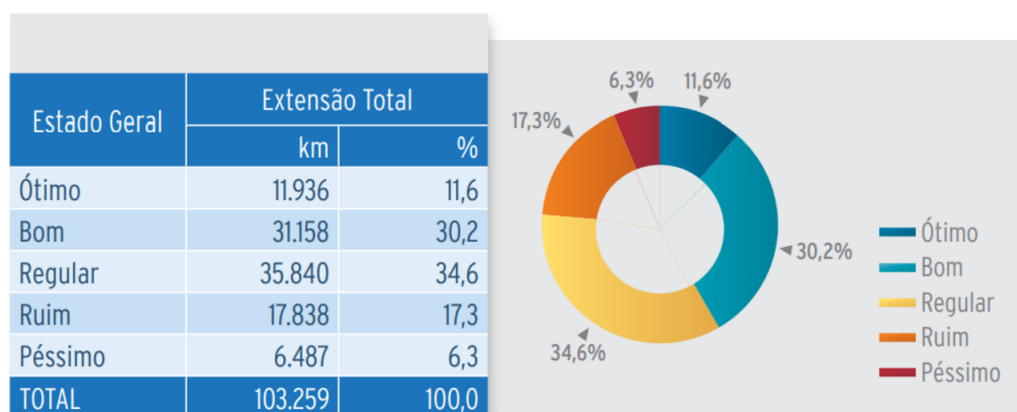


Figura 2 – Classificação do estado geral.

Fonte: [4].

A manutenção e conservação das rodovias federais é dividida entre órgãos públicos e privados. Ao buscar formas mais adequadas para financiar investimentos necessários, em virtude de restrições financeiras do Estado, o país iniciou o programa de privatização das rodovias. Portanto, as concessionárias tem o direito de administração da rodovia mediante a um prazo de tempo determinado pelo governo. "O modelo adotado foi o de concessão onerosa, na qual a concessionária adquire o direito de exploração da rodovia por um prazo fixo, estipulado pelo poder concedente"[5], e dentro deste prazo cabe à concessionária a

administração daquela rodovia. Porém, devido a inexistência da concorrência durante o período determinado na licitação, é necessário um contínuo monitoramento, cumprimento e renegociação dos contratos.

Por conseguinte, para a manutenção e conservação das rodovias, as concessionárias dispõem de pontos de apoio para equipes que realizam inspeção rodoviária e é onde ficam os guinchos pesados e superpesados, ambulâncias e os profissionais das equipes de resgate [6]. Estes postos tem diferentes alcunhas, variando de concessionária para concessionária. Por exemplo: na concessionária VIA040, responsável pelo trecho de Juiz de Fora-MG até Brasília-DF, os pontos são chamados de "postos de atendimento", já para a concessionária CART, responsável pelo Corredor Raposo Tavares, são chamados de "SAU"(Serviço de atendimento ao usuário). É a partir desses postos, localizados ao longo das rodovias, que ambulâncias e guinchos saem para o atendimento aos acidentes.

Os acidentes que ocorrem em rodovias federais são documentados e arquivados pela Polícia Rodoviária Federal (PRF). Essas informações são abstraídas para tabelas, que podem ser acessadas no site da PRF¹, e são disponibilizadas através da política de dados abertos [7]. Dentro das mesmas, existem dados referentes a acidentes que ocorreram em todas as rodovias federais. Tais devem ser tratados antes da utilização, para que se ajustem às necessidades do trabalho, assim há um aumento na precisão da informação gerada.

Neste sentido, assume-se que o uso de métodos matemáticos e estatísticos em união com a visualização de dados, pode tornar possível a avaliação e visualização do risco de acidentes em estradas, proporcionando ao administrador de rodovia, teoricamente, saber quais trechos poderão exigir uma demanda maior por recursos, de acordo com o relatório gerado pela análise.

Por fim, para gerar tal informação, serão utilizados métodos estatísticos que em conjunto aos dados históricos, subjetivamente, podem abstrair informações que serão necessárias para prever potenciais acidentes nos trechos de rodovia.

1.1 Justificativa

No Brasil, entre junho de 2014 e julho de 2015, ocorreram mais de 150 mil acidentes em rodovias federais, dentre as 221.066 pessoas envolvidas, 8.172 vieram a óbito [1]. O transporte rodoviário ainda é o principal responsável pela locomoção de bens e pessoas no Brasil [8], porém tem custos altos de manutenções de estradas, veículos e combustíveis. Com isso, a logística ao distribuir recursos nessas rodovias é muito importante, porque estes são limitados e o tempo de atendimento em acidentes é determinante na hora de salvar vidas, e também evitar novos acidentes.

¹ site: <https://www.prf.gov.br/portal>

As rodovias são administradas pelo governo (pública) ou por concessionárias (privada), e ao longo dessa rodovia o administrador conta com recursos que o auxilia nessa administração. Os recursos são: ambulâncias, guinchos, caminhões pipa, veículos de inspeção rodoviária, veículos de apreensão animal. Eles são distribuídos entre os postos de atendimento que estão situados ao longo das rodovias. Assim, quando um acidente acontece, o posto de atendimento mais próximo é informado, direcionando os recursos para o atendimento no local do acidente, os recursos são enviados de acordo o que foi informado na solicitação do serviço [9].

Por exemplo, existem dois postos de atendimento ao longo de uma rodovia, o "Posto de atendimento 1" e o "Posto de atendimento 2", o primeiro possui apenas uma ambulância e um guincho. Figura 3.

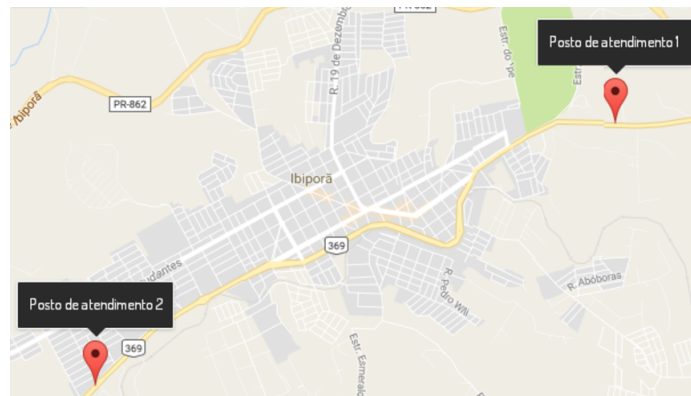


Figura 3 – Exemplo 1 - Postos de atendimento.

Um acidente acontece no km "x" de uma rodovia, o "Posto de atendimento 1" é o mais próximo então ele é acionado e envia sua única ambulância e único guincho para atender o acidente. Figura 4.

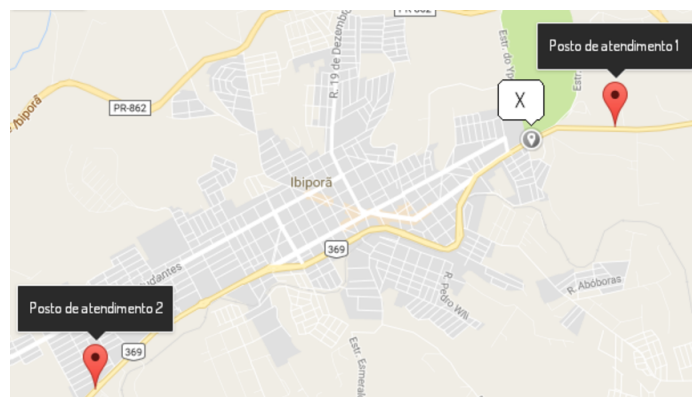


Figura 4 – Exemplo 2 - Postos de atendimento.

Infelizmente outro acidente acontece no km "y", e o posto de atendimento mais próximo é o mesmo que atendeu o acidente do km "x", o que fazer? Figura 5.

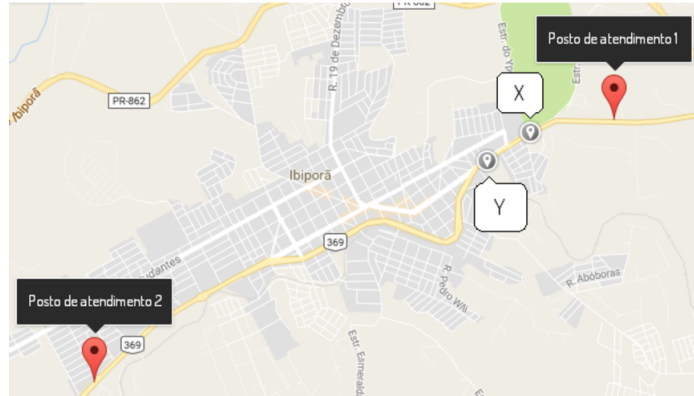


Figura 5 – Exemplo 3 - Postos de atendimento.

É necessário acionar outro posto de atendimento, mais longe, para que envie recursos para tratar do acidente “y”. O tempo é crucial para evitar mortes e novos acidentes, o atraso no atendimento pode ser fatal, a distribuição inteligente de recursos entre os postos de atendimentos pode aumentar a rapidez com que é tratado o problema, aumentando as chances de evitar vítimas fatais e novos acidentes causados pelo anterior.

Sendo assim, se houvesse uma ferramenta que pudesse oferecer um meio de prever os possíveis acidentes que iriam ocorrer nos trechos da rodovia, seria possível ter uma maior precisão ao distribuir os recursos disponíveis ao longo da rodovia para que assim, teoricamente, aumente a velocidade de atendimento às ocorrências.

1.2 Objetivos

O tempo entre a lesão, causada pelo acidente, e o primeiro atendimento é crucial para a vítima. Em grandes cidades como Toronto e Nova York, foi fixado como meta que 80% dos chamados sejam atendidos em até 10 minutos [2]. Em rodovias, teoricamente, o tempo de atendimento é determinado pela distância entre os postos de atendimentos, e os recursos disponíveis nestes postos.

Dado o problema, o objetivo deste trabalho é utilizar métodos matemáticos e estatísticos para análise de dados históricos e então prever possíveis acidentes ao longo da rodovia, e com isso em aplicação, criar uma ferramenta de visualização que ajudará possivelmente a diminuir o tempo de deslocamento até o local do acidente e evitar ociosidade de recursos.

A fim de atingir o objetivo principal, os seguintes objetivos específicos foram definidos:

- Estudar e analisar os dados históricos de acidentes em rodovias federais.

- Tratar os dados por meio de métodos matemáticos e/ou estatísticos, buscando condições que possam apontar similaridades nos cenários de ocorrências.
- Desenvolver uma ferramenta que auxilie na visualização das informações.
- Aplicar experimentos para validar a abordagem.

1.3 Materiais e Métodos

Com base em nestes objetivos, esta pesquisa possui natureza exploratória. "Estas pesquisas têm como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses. Pode-se dizer que estas pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições"[10]. Portanto, é através do levantamento de dados sobre acidentes e a análise desses dados que a hipótese é levantada.

Por outro lado, com base nos procedimentos técnicos utilizados, a pesquisa se encaixa com a descrição de pesquisas de levantamento. Isso porque segundo [10] a maior parte das pesquisas desse tipo envolve a seleção de uma amostra do universo estudado e esta amostra se torna o objeto de investigação do estudo. Ao contextualizar essa informação, é possível dizer que o universo seriam os acidentes, e o objeto de estudo escolhido são as tabelas que descrevem com dados tais acidentes. Também é válido denotar que as fases do desenvolvimento do levantamento se assemelham aos objetivos específicos deste trabalho, com fases para especificação do objetivo, levantamento de dados, análises, testes e apresentação de resultados [10].

Além disso, a pesquisa usa de métodos qualitativos e quantitativos para trabalhar com suas variáveis. "Os métodos qualitativos e quantitativos não se excluem. Embora difiram quanto à forma e a ênfase, os métodos qualitativos trazem como contribuição ao trabalho de pesquisa uma mistura de procedimentos de cunho racional e intuitivo capazes de contribuir para a melhor compreensão dos fenômenos"[11]. É possível distingui-los, mas é incorreto afirmar que são opostos.

A pesquisa quantitativa terá foco na análise de informações, visando compreender de forma objetiva os dados aplicando métodos matemáticos e estatísticos para executar tal tarefa. Assim, com os resultados da análise quantitativa, a análise qualitativa será capaz de determinar um "significado" à informação adquirida contextualizando-a. "Em certa medida, os métodos qualitativos se assemelham a procedimentos de interpretação dos fenômenos que empregamos no nosso dia-a-dia, que têm a mesma natureza dos dados que o pesquisador qualitativo emprega em sua pesquisa"[11].

Um exemplo seria: ao analisar uma tabela de acidentes de um ano qualquer, foi possível, através da pesquisa quantitativa, determinar um número de acidentes em um

determinado tempo em um trecho de rodovia, e por meio de métodos qualitativos foi possível determinar quantos acidentes podem vir a ocorrer e até mesmo suas causas em um período de tempo futuro.

2 FUNDAMENTAÇÃO TEÓRICA

Ao decorrer desta seção, serão apresentados conceitos relevantes à pesquisa, tais como: *Big data*, *data science*, visualização de dados, análises quantitativas de projeção (médias móveis, suavização exponencial, projeção de tendências) e correlação (regressão simples), além de análises qualitativas.

2.1 Big Data

Segundo [12] em apenas 2 anos (2008 e 2009) diariamente foi produzido por volta de 6,3 quintilhões de bytes de dados, e 90% dos dados do mundo foram criados nos últimos anos, isso devido a adesão à internet por parte de grandes empresas, a criação das redes sociais e a utilização de dispositivos móveis.

Na computação, a existência de grandes concentrações de dados despertou o estudo do que é chamado de *BigData*. "*Big data* é um termo que é usado para qualquer coleção de dados tão grande que se torna difícil processá-la por meio de técnicas tradicionais de gerenciamento de dados como, por exemplo, os banco de dados relacionais"[13].

"Com a globalização, e com o sistema de administração na produção que determina que nada deve ser produzido, transportado ou comprado antes da hora exata chamado *Just in time*, a expansão virtual se tornou necessária. A partir do início do século XXI, com a alta demanda e o grande anseio por tecnologia, a quantidade de dados e informações gerenciadas na rede mundial de internet se tornou volumosa e cresce rapidamente a cada dia, tornando-se assim o maior reservatório de informações existente no mundo, atualmente sendo necessário assim organizar e gerenciar todos esses dados para que eles não se percam ou se tornem lixo eletrônico [14]".

De acordo com [13], as características da *Big data* são comumente referenciadas como os três V's:

- Volume - Qual a quantidade de dados?
- Variedade - Quão diversos são os diferentes tipos de dados?
- Velocidade - Qual velocidade esses dados são gerados?

Segundo [13], para analisar essa grande quantidade de dados é feito o uso da *data science*, que envolve o uso de métodos que analisam e extraem o conhecimento que a

base de dados contém. Portanto, o processo de *data science* é vital para a interpretação e contextualização desse grande amontoe de dados.

2.2 Data Science

Esta seção é fortemente baseada nos conceitos apresentados no livro *Introducing Data Science* [13], o livro citado se encontra em inglês, logo, o conteúdo a seguir é uma tradução livre. É recomendada a leitura do livro citado para maior aprofundamento no assunto.

Tipicamente, o processo de *data science* consiste em seis passos: fixar o objetivo da pesquisa, obter os dados, preparar os dados, explorar os dados, modelar os dados e apresentar os dados.

Para fixar o objetivo da pesquisa é necessário preparar uma tabela para seu projeto, nesta tabela devem ser incluídas informações como: o que você quer pesquisar, quais os benefícios da pesquisa, quais dados e recursos serão necessários, prazo e derivados.

O segundo passo é obter os dados. Depois de definir quais dados são necessários e onde encontrá-los é preciso checar a existência dos mesmos, sua qualidade e o acesso. Esses dados podem ser obtidos de diversas fontes e possuírem diversos formatos, tudo depende do que você precisa e como vai usá-los.

Ao obter dados, os mesmos podem estar corrompidos, fora de contexto, ou conter informações erradas, tais dados possuem uma baixa qualidade. Como a obtenção de dados é uma tarefa propensa a erros, é possível aumentar a qualidade e preparar estes dados para o uso em passos subsequentes a obtenção. Neste passo, caso haja necessidade, os dados são preparados para uso, removendo inconsistências, enriquecendo as informações e garantindo que os dados estão no formato correto para uso da sua aplicação.

No terceiro passo, o foco é um entendimento profundo dos dados. É preciso entender como as variáveis interagem umas com as outras, a distribuição dos dados, e onde podem haver anexos. Esse passo frequentemente é chamado de *EDA*, *Exploratory Data Analysis*.

Este passo consiste no uso de modelos, domínio de conhecimento e informações sobre os dados encontrados nos passos anteriores para responder a pergunta da pesquisa. Sendo assim é preciso escolher uma técnica para, em seguida, construir um modelo. A construção de um modelo é considerado um processo iterativo que envolve a seleção das variáveis úteis, a execução do modelo e, por fim, o diagnóstico gerado.

O último passo é a apresentação dos dados. Esses são apresentados em variadas formas, como apresentações ou relatórios de pesquisa, procurando uma forma que seja mais adequada às informações que serão apresentadas.

É importante ressaltar que, embora dê a impressão, o processo de *data science* não é linear, na realidade, é necessário frequentemente retornar alguns passos em ordem de obter o resultado esperado.

2.3 Data Visualization

O *data visualization* faz parte do último passo do processo de *data science*, como o nome já diz, ele é responsável pela apresentação dos dados ao usuário. Mas por que usar o *data visualization*?

Como destacado no livro *Designing Data Visualization* [15], essas são algumas das razões do porque a visualização de dados tem uma importância tão grande:

- Potencializa a capacidade e quantidade de informações absorvidas pelo sistema visual para mover essa quantidade rapidamente ao cérebro.
- Toma vantagem da capacidade de nosso cérebro de identificar padrões e associar relacionamentos e significados.
- Pode inspirar novos questionamentos e uma exploração mais além.
- Ajuda a identificar subproblemas.
- É muito útil para identificar tendências e anexos, descobrindo ou procurando por pontos de informação interessantes ou específicos em um campo de informações muito grande.

Segundo [16], a visualização é uma das melhores maneiras de explorar e entender uma grande quantidade de dados. Disponibilizar um cenário visual e deixar que o cérebro do leitor encontre os padrões além de prático é uma maneira de aumentar o interesse do leitor. Visualizar as informações em seu estado "cru" é confuso, pois o usuário pode não saber o que está procurando em meio à tantas palavras e valores que não significam nada além do que se pode ver.

Após determinar qual informação será transmitida visualmente, é preciso decidir de qual maneira ela será exposta. Para isso, existem vários tipos diferentes de apresentações de dados, e cabe ao autor/pesquisador decidir de que forma ele irá apresentar essas informações. Segundo [15], a primeira regra do *Data Visualization* é: a função primeiro e depois aparência, isso quer dizer que, mesmo que sua aplicação fique linda, se ela não se comunicar com a informação que deve ser repassada, então ela não está sendo útil. Manter a visualização limpa e "bonita" é importante, mas é ainda mais importante que ela ajude o usuário a entender a informação exposta (e não o contrário).

É fácil observar a diferença que esse conceito traz, na Figura 6 estão as mesmas informações dispostas de maneiras diferentes, por exemplo: se o usuário estiver buscando saber qual mês tem o maior consumo, é muito mais fácil visualizar isso através do gráfico do que da tabela, e esse é apenas um exemplo de muitos.

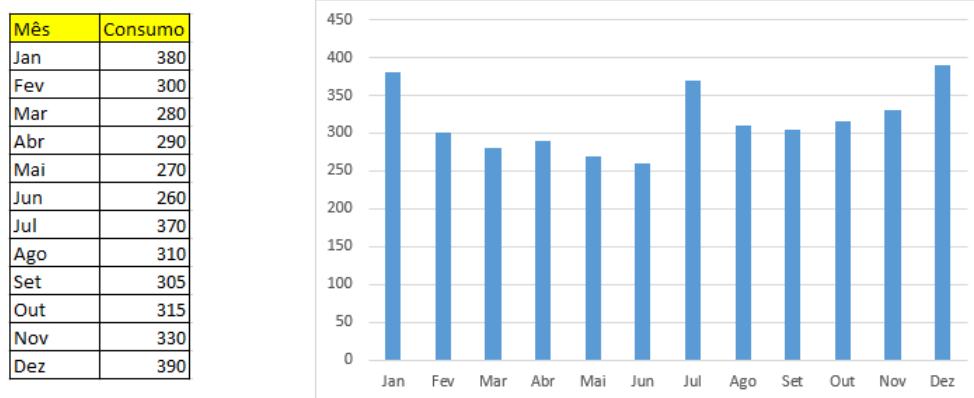


Figura 6 – Aplicação de *Data Visualization*: transformação de informação em tabela para visualização em forma de gráfico.

Ainda, segundo [13], a visualização depende de fatores como¹:

- Que tipo de decisão será apoiada? É uma decisão estratégica ou operacional? Decisões estratégicas, frequentemente, apenas requerem que você analise e reporte uma vez, por outro lado, decisões operacionais requerem que o relatório seja atualizado regularmente.
- Quão grande é sua organização? Em organizações menores, você estará a cargo de todo o ciclo: de coletar os dados até os relatórios. Em organizações maiores um time pode estar disponível para fazer o *dashboard* para você. Mas mesmo nessa última situação, entregar um protótipo do *dashborad* pode ser benéfico, porque serve de exemplo e diminui o tempo de entrega.

Fica inferido então que, dadas as informações apresentadas, o *data visualization* tem o propósito de criar um meio de visualização que será bem recebido e facilmente entendível ao usuário, e todas as escolhas de design e implementações devem servir esse propósito [15].

A representação de dados estatísticos de forma visual é comum há centenas de anos, porém o vocabulário deste campo ainda está evoluindo e se ajustando, por isso é comum muitas vezes associar, e até usar para o mesmo fim, os termos infográfico e

¹ Tradução livre feita a partir do livro *Introducing Data Science*. cap.9. pag.254

visualização de dados, quando na verdade são distintos em suas aplicações em diversos pontos[15].

O autor do livro *Designing Data Visualizations* [15] sugere que um infográfico é determinado por qualquer representação visual de dados que seja:

- Gerado manualmente;
- Específico para dados tratados (não é útil para dados diferentes);
- Esteticamente rico (visual forte, feito para chamar a atenção e manter interesse);
- Relativamente pobre de dados (porque toda informação é gerada manualmente);



Figura 7 – Exemplo de Infográfico.

Fonte: <http://www.naovoupagaropato.com.br/>

Os infográficos são ilustrações, ou desenhos, criados para chamar e manter a atenção do alvo, onde os dados são predispostos e por isso tem pouca manuseabilidade, ou seja, estão limitados aos dados informados na ilustração, tornando a manutenção ou atualização dos mesmos trabalhosa, visto que tem de ser feitas manualmente.

Para simplificar observe a Figura 7, a figura representa um infográfico sobre os impostos que o brasileiro paga em diversos produtos, nela é possível observar as características denotadas anteriormente, principalmente o ponto que destaca a relativa pobreza de dados e sua especificidade ao tratar dados, visto que se os impostos ou os preços mudarem, o infográfico se torna desatualizado.

Em contraste ao infográfico, o livro continua a sugerir que o termo *data visualization* deve ser usado para representações visuais onde os dados são:

- Gerados a partir de um algoritmo;

- Fácil de recriar usando diferentes dados;
- Esteticamente "pobre"(os dados não são adornados);
- Relativamente rico em dados (grandes volumes de dados são viáveis, ao contrário de infográficos);

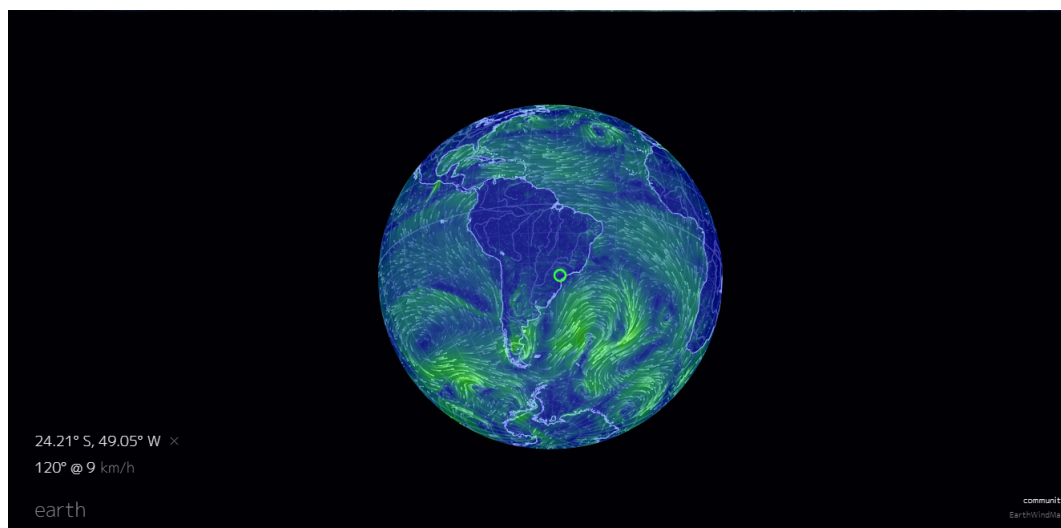


Figura 8 – Exemplo de *Data Visualization*.

Fonte: <https://earth.nullschool.net/current/wind/surface/level/orthographic>

Data Visualization, ou visualização de dados, é gerado por meio de um algoritmo que utiliza softwares de gráficos e diagramas. A grande vantagem é a facilidade de atualização e recuperação de visualização com mais, ou novos, dados. Embora a visualização contenha mais volume de dados, ela é, geralmente, menos atraente esteticamente do que os infográficos.

A Figura 8 é uma visualização de dados criada a partir de supercomputadores que predizem o clima a cada 3 horas, essa visualização mostra as condições do clima, ventos e oceanos por todo o globo. A característica mais clara da visualização de dados nessa aplicação é a facilidade de gerar novas informações e atualizar o gráfico, além do grande volume de dados que são gerados à cada 3 horas.

2.4 Métodos de previsão

Os métodos estatísticos são fundamentais para análises de dados, e através deles é possível, utilizando o comportamento dos dados, chegar a uma previsão do que, possivelmente, poderia se tornar realidade. Isso torna viável para o responsável ter uma previsão do que poderia vir a acontecer antes de tomar uma decisão. Dentre as técnicas de previsão dos dados tem-se métodos de previsão quantitativos e qualitativos. Os métodos qualitativos usam de argumentos subjetivos, especulações, experiência e opiniões de profissionais

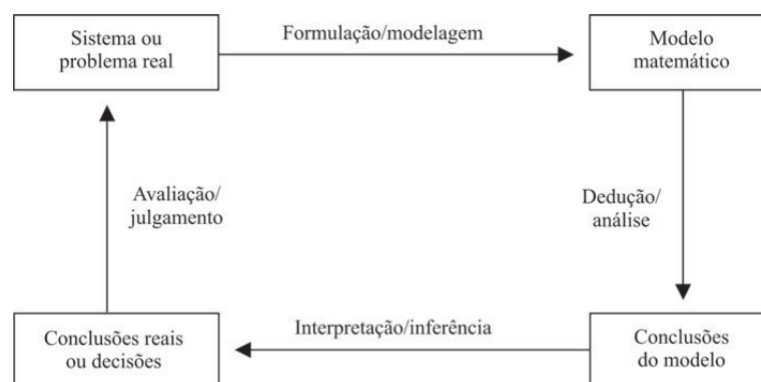


Figura 9 – Processo de modelagem

Fonte: [17].

especialistas. Já os métodos quantitativos buscam padrões de comportamento em séries de dados históricos, para então utilizar esses padrões para projetar dados futuros [18]. Sendo estes últimos explorados nesse trabalho.

"Os métodos causais envolvem a determinação de fatores que se relacionam com a variável a ser prevista enquanto que os métodos de previsão de séries temporais envolvem a projeção de valores futuros de uma variável, com base, inteiramente, em observações do presente e do passado dessa variável"[18].

Sendo assim, séries temporais são definidas como uma sequência de dados obtidos em intervalos regulares de tempo durante um determinado período [19]. Portanto existem duas subcategorias de métodos quantitativos a serem explorados para previsão de dados, os de projeção (que utilizam-se de séries temporais) e os de correlação (que usam relações entre as variáveis), como pode ser visualizado na Figura 10.

A Figura 10 aborda de maneira esquemática métodos de previsão que serão elucidados nas próximas seções a fim de apresentar ao leitor os métodos que podem ser usados no trabalho.

Segundo [21], existem três condições que indicam se um método quantitativo pode ser aplicado, são elas:

- A informação do passado estar disponível.
- Essa informação pode ser quantificada na forma de dados numéricos.
- Pode ser assumido que alguns aspectos padrões do passado irão continuar no futuro.

Um método quantitativo muito utilizado é o de médias móveis. Esse método, como citado por [21], pode ser aplicado apenas em séries estacionárias, isto é, séries em que os

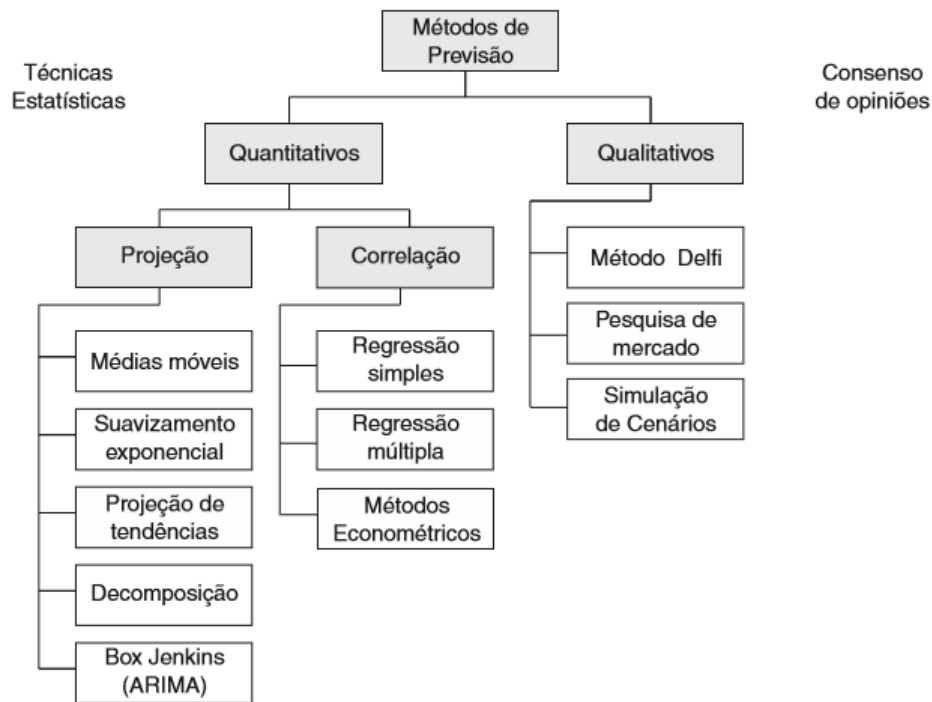


Figura 10 – Métodos de previsão: Qualitativos e Quantitativos

Fonte: [20].

dados flutuam em torno de um valor médio, neste método a previsão para o período futuro corresponde à uma média das observações passadas recentes. A fórmula que define a média móvel para um período de tempo t é:

$$P_t = \frac{x_{t-1} + x_{t-2} + \dots + x_{t-n}}{n}$$

Onde n é o número de observações incluídas na média M_t que origina a previsão P_t . Por exemplo: Imagine que meu n é 5, e a média móvel será aplicada sobre os valores 10, 13, 14, 11 e 12, então temos que a média é 12:

$$P_t = \frac{10 + 13 + 14 + 11 + 12}{5}$$

$$P_t = 12$$

Porém, em médias móveis, a ideia é ignorar valores antigos e acrescentar novos. Nesse caso, se um outro valor fosse adicionado a sequência anterior, o valor mais antigo seria esquecido, e o novo entraria na equação, supondo que a sequência recebesse um novo valor 15, o valor mais antigo (10) seria retirado da equação e portanto o valor da média móvel seria alterado:

$$P_t = \frac{13 + 14 + 11 + 12 + 15}{5}$$

$$P_t = 13$$

Caso novos valores fossem adicionados o mesmo processo iria se repetir, devido a mudança dos valores levados em consideração, esse método recebe o nome de médias móveis.

Como é descrito no livro [20], outro método quantitativo de projeção é a suavização exponencial. Na média móvel as parcelas do cálculo tem o mesmo peso, porém, na suavização há uma ponderação na média, onde valores mais recentes possuem um peso maior. Existem três tipos de suavização exponencial: simples, com tendência e com tendência e sazonalidade.

Na suavização exponencial simples, assume-se que há uma constante de uma demanda. Partindo de um valor inicial o valor base é corrigido a cada período, de acordo com a entrada de novos dados de demanda na série histórica. O cálculo da base pode ser feita através da fórmula:

$$B_t = \alpha * D_t + (1 - \alpha) * B_{t-1}$$

Onde B_t representa a base ao final do instante t e α é a constante de suavização para base. Quando α é fixado em um ou zero significa que a base será exatamente igual ao último valor da demanda ou a base permaneceria sempre com seu valor inicial sem sofrer correções.

Já na suavização exponencial com tendência uma segunda variável é adicionada, essa variável é responsável por refletir o crescimento da demanda de um período para outro. Essa variável será atualizada exponencialmente e aplicada no cálculo seguindo a seguinte formula:

$$B_t = \alpha * D_t + (1 - \alpha) * (B_{t-1} + T_{t-1})$$

Onde T_t representa a tendência ao final do instante t , e β é a constante de suavização para tendência. Essa equação calcula uma média ponderada entre a demanda real e a nova base, que aumenta ou reduz o valor da demanda expressa na tendência. Na próxima equação acontece a suavização da tendência:

$$T_t = \beta(B_t - B_{t-1}) + (1 - \beta) * T_{t-1}$$

E, por fim, a última equação fornece uma previsão para k períodos futuros:

$$F_t(t+k) = B_t + kT_t \quad k = 1, 2, \dots$$

O último dos três é a suavização exponencial com tendência e sazonalidade. Além da tendência, esse modelo incorpora a sazonalidade em seus cálculos, isto é dado através de um índice de sazonalidade para cada período, esse índice representa a proporção entre a demanda média do período e a demanda média anual. Por exemplo, se a demanda anual é igual a 1, e a demanda do período é 1,2 significa que a demanda média desse período é 20% maior que a demanda média anual.

O princípio desse modelo consiste em prever a demanda base, removendo o efeito da sazonalidade. A equação que representa esse modelo é:

$$B_t = \alpha * \left(\frac{D_t}{I_{t-L}} \right) + (1 - \alpha) * (B_{t-1} + T_{t-1})$$

$$T_t = \beta(B_t - B_{t-1}) + (1 - \beta) * T_{t-1}$$

$$I_t = \gamma * \frac{D_t}{B_t} + (1 - \gamma) * I_{t-1}$$

$$F_t(t+k) = (B_t + kT_t) * I_{t-L+k} \quad k = 1, 2, \dots$$

Onde I é o índice de sazonalidade e γ é a constante de sazonalidade, enquanto F , como em todos os modelos de sazonalidade, representa a previsão final. Para que se dê início a esse processo é preciso definir os índices de sazonalidade iniciais de cada período. Ao comparar este modelo com os anteriores, este modelo precisa de uma série de dados mais longa para que hajam, ao menos, três ciclos sazonais completos.

Nos métodos que utilizam relações causais (correlação), assume-se que há uma relação de causa-efeito entre algumas variáveis e, portanto, essa relação gera a previsão, um exemplo seria o número de vendas de um produto e seu preço. Os métodos de regressão e os métodos econométricos utilizam essas relações para fazerem suas previsões. A análise de regressão tem como objetivo prever o comportamento futuro de um fator por meio do relacionamento dele com outros fatores de naturezas diferentes e que possam ter influência sobre o valor numérico desse fator [21].

A regressão linear simples pode ser resumida em "modelo de previsão que permite a determinação de previsões futuras em séries com tendência, baseada no Método dos Mínimos Quadrados"[21]. Este modelo busca, através da equação da reta: $Y = a + bx$, ajustar os parâmetros. Os valores de a e b podem ser encontrados a partir do Método de Mínimos Quadrados:

$$a = \frac{\Sigma x^2 * \Sigma y - \Sigma x * \Sigma x * y}{n * \Sigma x^2 - (\Sigma x)^2}$$

$$b = \frac{n * \Sigma x * y - \Sigma x * \Sigma y}{n * \Sigma x^2 - (\Sigma x)^2}$$

$$Y = a + bx$$

Onde x e y representam valores de variáveis independentes (tempo, vendas), n é o número de observações, a o intercepto do eixo vertical, b inclinação da linha de regressão e Y valores de y que se situam na linha de tendência $Y = a + bx$.

Para avaliar um relacionamento, e se ele é positivo (se um valor sobre o outro sobe também) ou negativo (se um valor diminui o outro diminui também), utiliza-se um coeficiente (r), que permite examinar a extensão do relacionamento entre duas variáveis, onde r pode variar de -1 até 1 .

$$r = \frac{n * \Sigma x * y - \Sigma x * \Sigma y}{[n * \Sigma x^2 - (\Sigma x)^2] * [n * \Sigma y^2 - (\Sigma y)^2]}$$

"O coeficiente de determinação (r^2) pode representar a proporção de variância em Y (independente), que pode ser explicada pelas variáveis que compõem a regressão"[21].

3 DESENVOLVIMENTO

O desenvolvimento do presente trabalho, se deu por meio de duas abordagens principais, a estatística e a visualização. A parte estatística foi responsável por tratar os dados, aplicar um método de previsão sobre os mesmos e por fim gerar uma previsão sobre os possíveis acidentes na BR-376. A visualização então foi responsável por pegar essas previsões e transformá-las em informação visual, para o maior entendimento de qualquer individuo interessado em utilizar essas informações para um determinado fim.

3.1 Análise dos dados

Normalmente a análise de dados tem início com o dado em sua forma bruta, precisando ser trabalhado passo a passo até se tornar "utilizável", mas felizmente o site da PRF¹ disponibiliza esses dados já na forma de tabelas que serão utilizadas neste trabalho. A área de atuação da PRF cobre toda a malha rodoviária federal, sendo assim, o número de entradas na base de dados é enorme, por isso é necessário realizar a seleção de uma pequena parte dos elementos que compõe a base de dados, essa pequena parte é chamada de amostra.

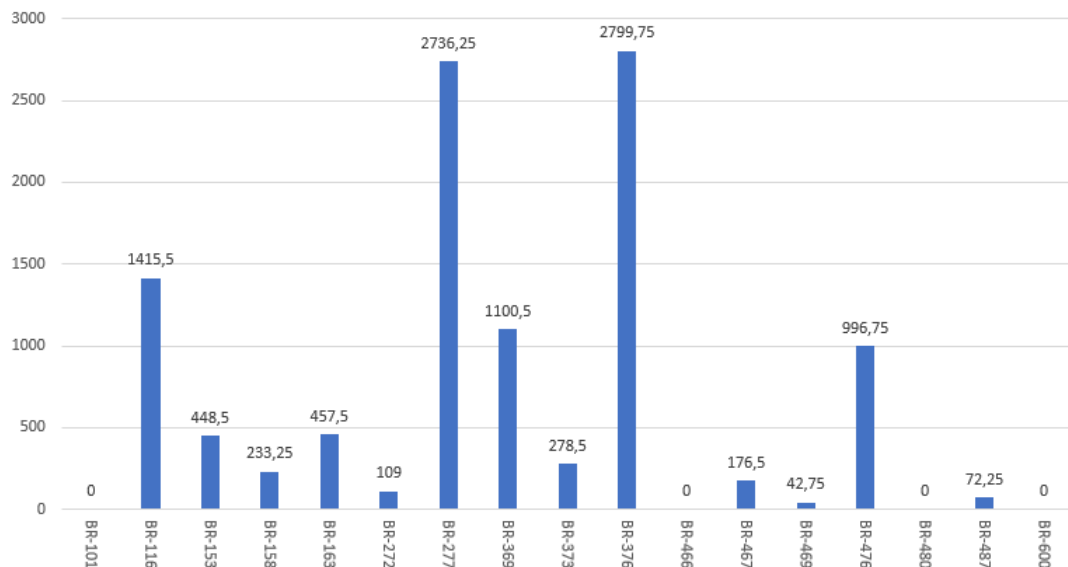


Figura 11 – Média de acidentes por rodovia federal (2017, 2016, 2015 e 2014).

De acordo com [10], em certas pesquisas, o levantamento de dados abrange um universo muito grande que torna impossível considerar sua totalidade e, por essa razão, é

¹ site: <https://www.prf.gov.br/portal>

frequente o uso de amostras, ou seja, utilizar apenas uma parte dos elementos que aquele universo contém. A amostra, no contexto deste trabalho, será uma rodovia em específico.

Para selecionar a amostra a ser utilizada, foram definidos dois critérios principais:

- Localização da rodovia;
- Número de acidentes;

Sendo assim, a rodovia escolhida como amostra foi a BR-376, ou melhor, a extensão da BR-376 que se encontra no estado do Paraná. A escolha é justificada em seus critérios, a BR-376 tem aproximadamente 682 quilômetros de rodovia no Paraná, dada a natureza estadual da universidade, uma rodovia que se encontra no estado seria a melhor opção de escolha. De acordo com o segundo critério, a BR-376 é a rodovia com mais acidentes no Paraná, ao observar a Figura 11 pode-se perceber que a média nessa rodovia é de aproximadamente 2800 acidentes por ano, desta maneira, ela é considerada a rodovia no Paraná com maior média de acidentes anuais.

Após decidir que a BR-376 seria a amostra, sua extensão foi dividida em 34 trechos, cada trecho possui 20 quilômetros com exceção do último que tem 22 km. Essa divisão se dá para possibilitar uma análise mais específica dos trechos, evitando que a rodovia seja tratada como apenas um grande trecho. Com isso uma análise sobre o número de acidentes, suas causas e a ocorrência de vítimas ou não foi gerada de acordo com cada trecho. Nas Figuras 12, 13 e 14 são exibidos os números de acidentes por trechos de rodovia nos anos 2014, 2015 e 2016.

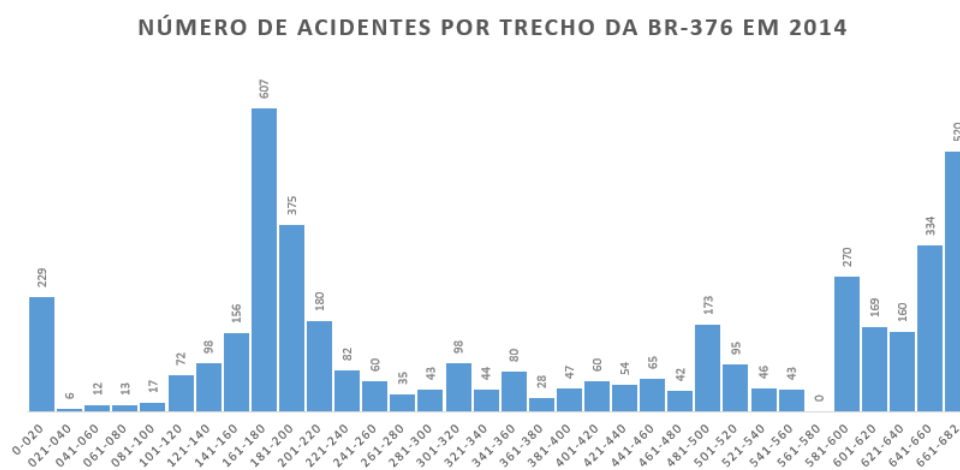


Figura 12 – Gráfico de acidentes por trecho da BR-376 no ano de 2014.

O eixo x apresenta os trechos da BR-376 e o eixo y traz o número de acidentes, com isso é possível observar como o número de acidentes, em geral, vem diminuindo e que o trecho mais perigoso fica entre o quilometro 161-180, seguido do trecho do quilometro

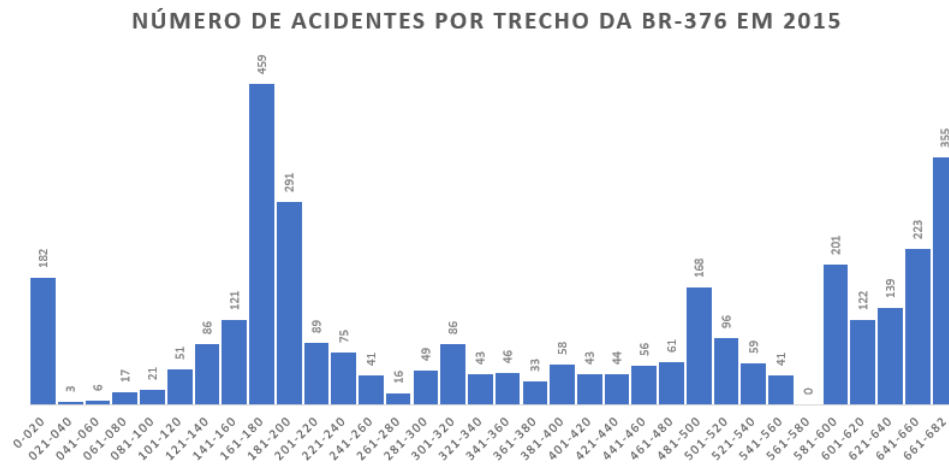


Figura 13 – Gráfico de acidentes por trecho da BR-376 no ano de 2015.

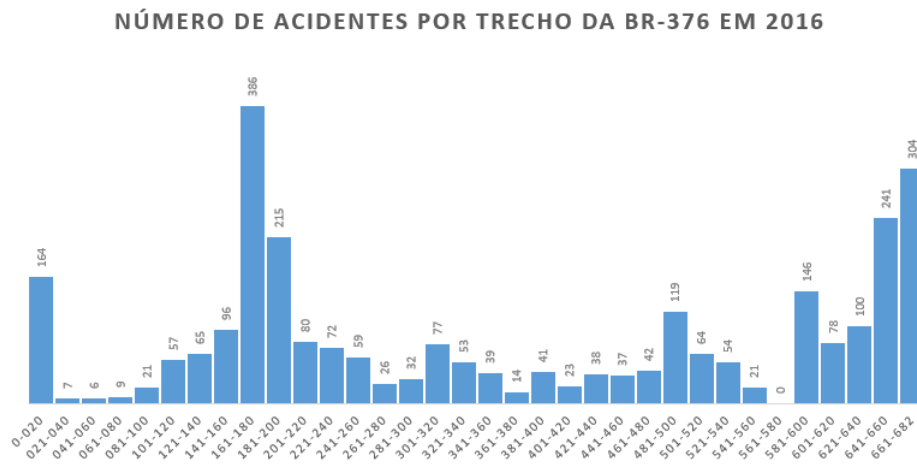


Figura 14 – Gráfico de acidentes por trecho da BR-376 no ano de 2016.

661-682 e o quilometro 181-200, sendo que os trechos 161-180 e 181-200 ficam um em seguida do outro, e a cidade mais próxima é Maringá-PR, já o trecho 661-682 corresponde ao último trecho da BR-376 e a cidade mais próxima é Guaratuba, que chama atenção por ser uma cidade litorânea.

Além de tais informações, também foi possível determinar quais as maiores causas de acidentes nessa rodovia, e ao fazer uma média dos últimos 3 anos foi possível determinar quais causas de acidentes são mais comuns, isso pode ser observado na Figura 15.

3.2 Análise da tabela

Toda informação coletada pela PRF se encontra na forma de tabelas. Nessas tabelas, cada coluna representa uma variável, e cada linha um acidente reportado, preenche o valor das variáveis das colunas. A tabela 1 que faz referência a tabela que contém os dados sobre os acidentes, possui uma legenda para cada variável.

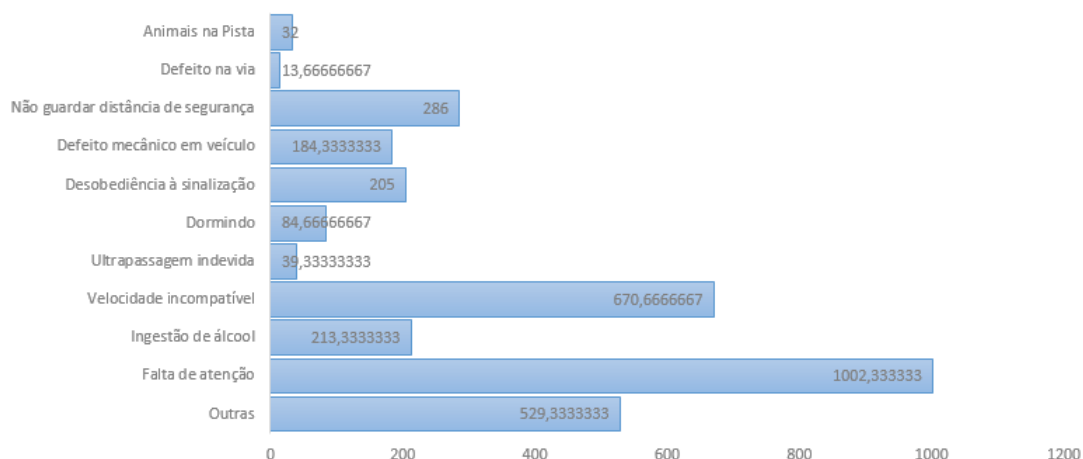


Figura 15 – Gráfico da média de causa dos acidentes da BR-376 nos últimos anos(2016, 2015, 2014).

Na tabela existem 24 variáveis, partindo do identificador do acidente(*id*) até o número de veículos envolvidos no acidente(*veiculos*). Agora que os dados já haviam sido tratados e as informações relevantes selecionadas, partiu-se para a aplicação do método de previsão.

3.3 Método de previsão

Separar as informações desejadas de acordo com o trecho era o primeiro passo para aplicar o método, isso porque com as informações escolhidas para serem utilizadas, em comunhão com o método, não era possível a predição do quilômetro que o acidente poderia vir a ocorrer, apenas se ele iria acontecer ou não. Limitar essa chance de ocorrência em trechos de 20 quilômetros possibilita que ao menos haja uma expectativa de onde aquele acidente ocorrerá, mesmo que essa margem não seja um dado exato, ao menos há um foco dessas ocorrências em um trecho conhecido.

Como citado na seção ??, as médias móveis estabelecem previsões correspondentes à uma média das informações de um "passado recente", isto é, ela utiliza as últimas entradas de uma base de dados para gerar uma média, ignorando dados antigos (obsoletos). Isto posto, às entradas correspondem à acidentes, e os mesmos foram considerados obsoletos, ou antigos, a partir de 2010.

Em sequência a seleção dos anos que seriam utilizados, as informações importantes foram destacadas, sendo elas:

- Acidentes com vítimas ou sem vítimas;
- Acidentes de acordo com suas causas;

Tabela 1 – Legenda da tabela de acidentes rodoviários.

Nome da Variável	Descrição
id	Representa o identificador do acidente.
data_inversa	Data da ocorrência no formato dd/mm/aaaa.
dia_semana	Dia da semana da ocorrência.
horario	Horário da ocorrência no formato hh:mm:ss.
uf	Unidade da Federação.
br	Representa o identificador da BR do acidente.
municipio	Nome do município de ocorrência do acidente.
causa_acidente	Causa presumível do acidente.
tipo_acidente	Identificação do tipo de acidente.
classificação_acidente	Classificação quanto à gravidade do acidente.
fase_dia	Fase do dia no momento do acidente.
sentido_via	Sentido da via considerando o ponto de colisão
condição_meteorologica	Condição meteorológica no momento do acidente.
tipo_pista	Tipo da pista considerando a quantidade de faixas.
tracado_via	Descrição do traçado da via.
uso_solo	Descrição sobre as características do local do acidente.
pessoas	Total de pessoas envolvidas na ocorrência.
mortos	Total de pessoas mortas envolvidas na ocorrência.
feridos_leves	Total de pessoas com ferimentos leves envolvidas na ocorrência.
feridos_graves	Total de pessoas com ferimentos graves envolvidas na ocorrência.
Ilesos	Total de pessoas ilesas envolvidas na ocorrência.
Ignorados	Total de pessoas envolvidas que não se soube o estado físico.
feridos	Total de pessoas feridas envolvidas na ocorrência.
veiculos	Total de veículos envolvidos na ocorrência.

Por exemplo, neste ponto era possível saber quantos acidentes causados por falta de atenção ocorreram no ano de 2015 entre 1 de abril e 23 de julho. Diante disso, foi possível gerar uma previsão dos acidentes que podem vir a ocorrer em 2017, além de saber suas possíveis causas e o número de acidentes que poderiam ter vítimas, por meio das entradas dos anos de 2010 até 2016, ou seja, sete anos de dados sobre acidentes. Essa previsão se dá através da aplicação do método de previsão sobre os dados de acidentes destacados.

Para fazer a previsão, primeiro era selecionado uma data de início e uma data final, estas representavam o período a ser previsto e por meio delas que os dados específicos daquelas datas eram buscados na tabela.

Depois de buscar os dados sobre acidentes, os mesmos eram divididos entre os quilômetros que ocorreram, gerando assim uma nova classificação, acidentes por trecho. Os dados da tabela continham colunas que os dividiam entre causas, daí surgiu uma nova camada, além de um acidente ser classificado pelo trecho que ocorreu, também era separado pela causa. Por fim, o passo anterior era repetido, gerando uma polarização sobre os acidentes, os que continham vítimas e os sem vítimas. A Figura 16 exemplifica o resultado desse processo.

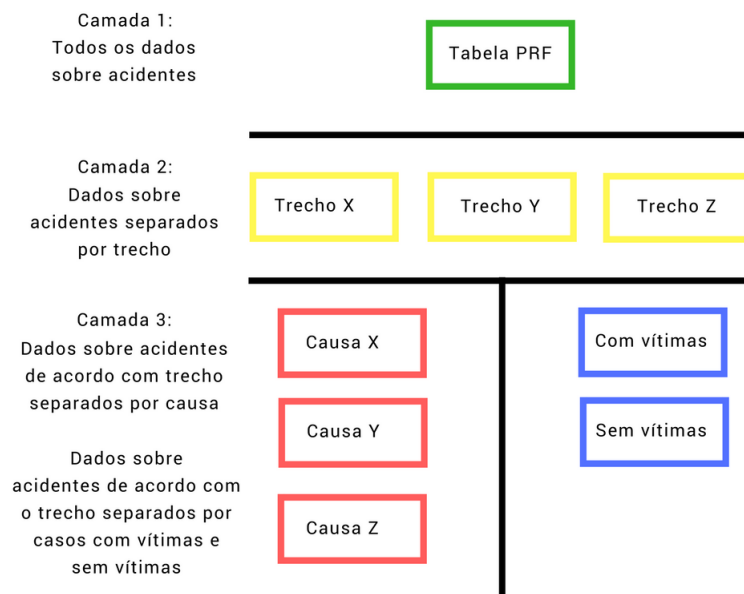


Figura 16 – Camadas resultantes do processo de tratamento da tabela e aplicação do método de previsão.

O próximo passo é aplicação da média móvel, como foi decidido que serão utilizados dados de 2010 até 2016, o valor da variável n será igual a 7, a fórmula a seguir é um exemplo de como foi gerada a previsão de acidentes que podem ocorrer em algum trecho da BR-376 causados por "falta de atenção":

$$FA_t = \frac{FA_{2010} + FA_{2011} + FA_{2012} + FA_{2013} + FA_{2014} + FA_{2015} + FA_{2016}}{7}$$

Nessa fórmula temos FA_{ano} que representa o número de ocorrências de acidente por "falta de atenção" em um determinado trecho da rodovia, a soma dessas variáveis é dividida por 7, que representa o número de anos levados em conta nesta equação (2010-2016), e por fim FA_t que representa a previsão de acidentes por "falta de atenção" que podem vir a ocorrer no trecho contestado.

A mesma lógica pode ser implementada para gerar previsões referentes ao número de acidentes que potencialmente irão conter, ou não, vítimas:

$$CV_t = \frac{CV_{2010} + CV_{2011} + CV_{2012} + CV_{2013} + CV_{2014} + CV_{2015} + CV_{2016}}{7}$$

Onde CV_t é a previsão que representa o número de acidentes que potencialmente irão conter vítimas.

A tabela disponível no Apêndice A é um exemplo do resultado de todo esse processo, nela estão os resultados de previsões do número de acidentes de acordo com a causa e trecho no período de 1 de janeiro até 31 de março de 2017. Com o resultado do método de previsão em mãos, da-se início à visualização de dados.

3.4 Ferramenta de Visualização

A ferramenta de visualização vem suprir a necessidade de transmitir ao usuário uma percepção geral da rodovia analisada. Transformar as tabelas de previsões em algo compreensível e visual era parte fundamental desse processo, sendo assim deu-se início ao processo de visualização.

É necessário que o usuário possa selecionar o período de tempo que ele quer que haja uma previsão, e após isso, selecionar os trechos de acordo com a sua vontade. Partindo desse princípio a logística do site deve permitir que ele altere esses valores de maneira fácil, e vislumbrar os mesmos com facilidade. A Figura 17 é o resultado dessa logística.

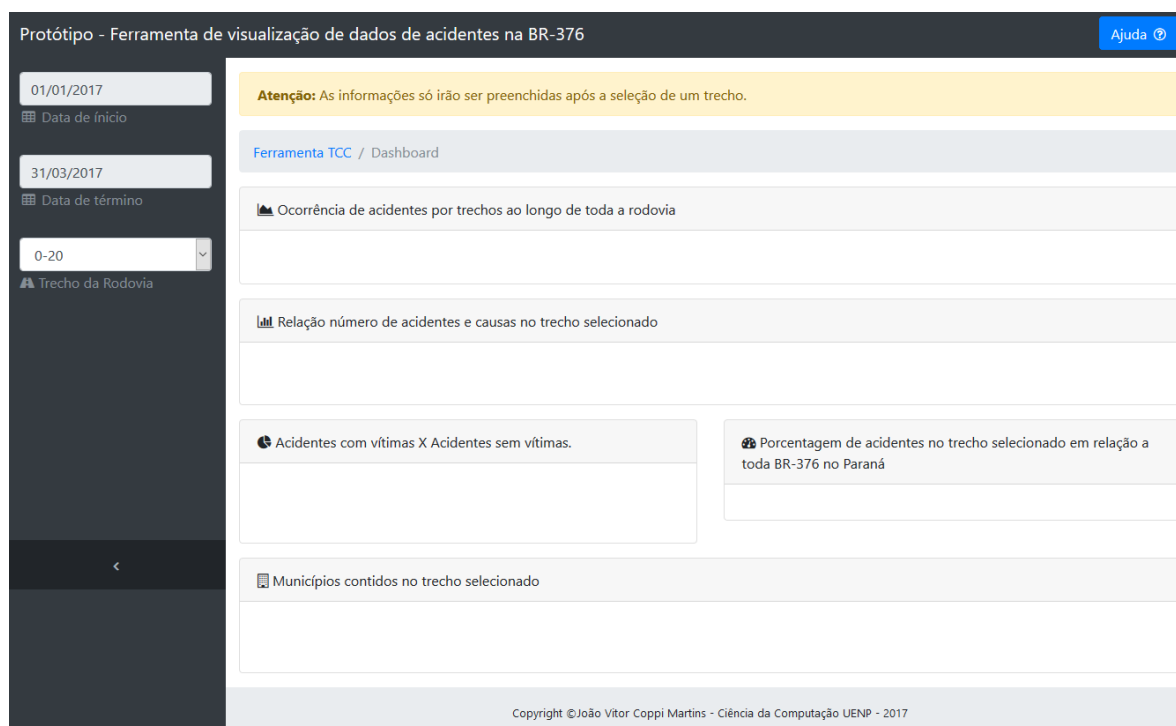
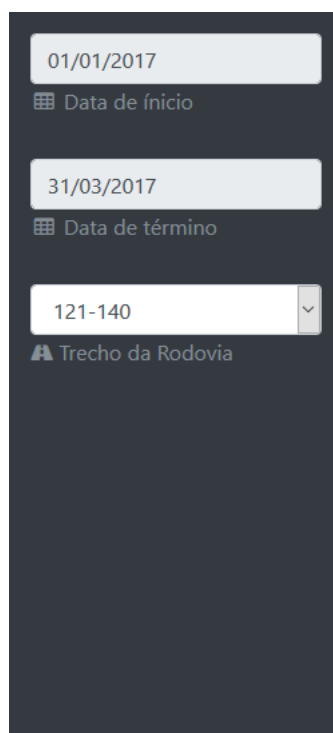


Figura 17 – Ferramenta protótipo: visão geral.

A barra de navegação superior contém um botão de ajuda que auxilia o usuário a entender o que é aquela ferramenta. Nesse botão é explicado o propósito do projeto, suas características e alguns termos utilizados na página. Assim, o usuário que não tem um conhecimento técnico dos termos também pode utilizar a ferramenta.

Na barra de navegação lateral estão os campos que devem ser selecionados para dar início à visualização, através desses campos é que o período de tempo é identificado e o trecho selecionado, só depois disso que os dados em si irão aparecer na tela.



01/01/2017
Data de início

31/03/2017
Data de término

121-140
Trecho da Rodovia

Figura 18 – Ferramenta protótipo: barra de navegação lateral.

O primeiro gráfico é o de linhas, nele estão dispostos os números de acidentes previstos em todos os trechos de rodovia da BR-376. Com isso, é possível ter uma visão geral e já identificar os trechos com maiores números de acidentes, permitindo que o usuário selecione os trechos com uma predisposição, isto é, selecionar um trecho com um objetivo e não apenas escolher ao acaso.

O próximo elemento da visualização é o gráfico de barras, este traz as informações relacionadas às causas de acidentes. O gráfico de barras é um ótimo gráfico para comparações, colocar uma causa ao lado da outra permite que o usuário, logo ao olhar para o gráfico, já identifique o maior valor e sua respectiva causa. Essa forte característica é potencializada com o botão "Comparar trechos" logo abaixo do gráfico, permitindo ao usuário uma comparação entre as causas mais recorrentes de dois diferentes trechos ao mesmo tempo.

Mais abaixo estão localizados mais dois elementos, o gráfico de pizza e o gráfico de

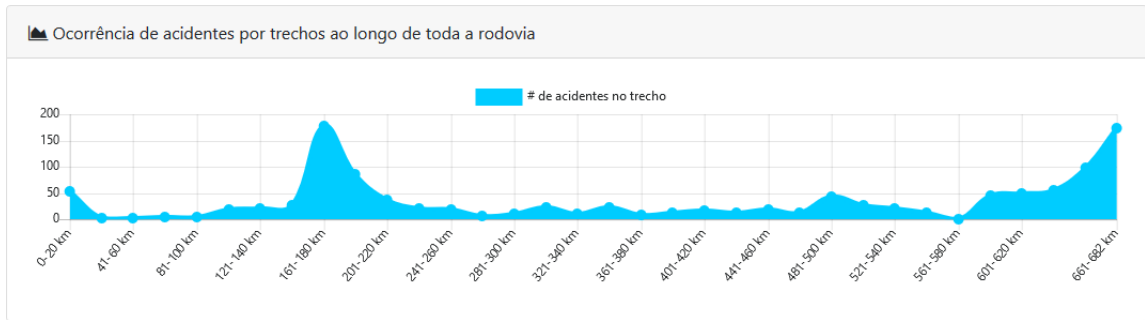


Figura 19 – Ferramenta protótipo: gráfico de linha.

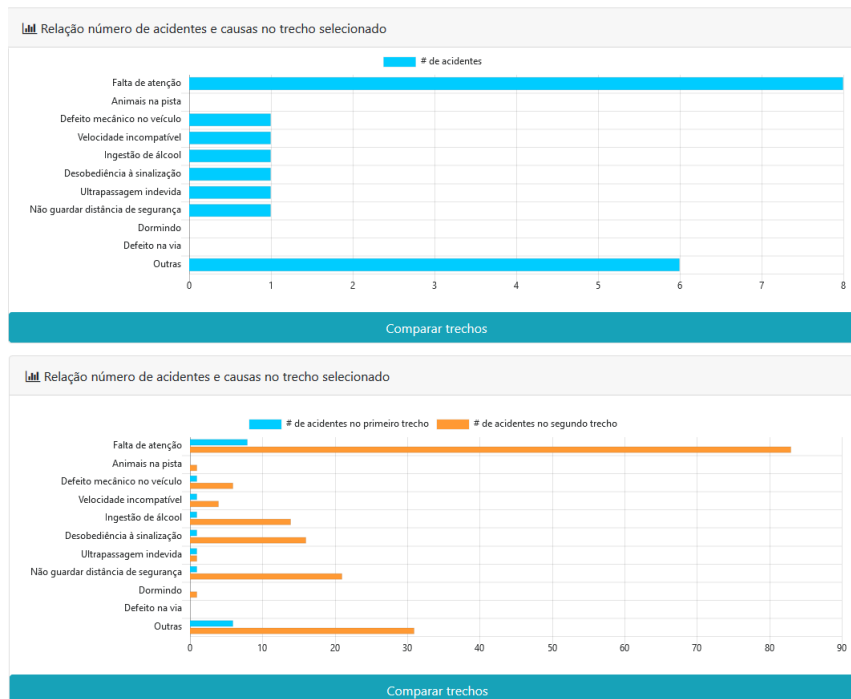


Figura 20 – Ferramenta protótipo: gráficos de barras.

Gauge. O gráfico de pizza é utilizado para representar o número de acidentes com vítimas e sem vítimas, o que é ideal devido ao fato da informação ser tão binária, fornecendo uma compreensão das proporções desses setores com grande eficiência visual.

O gráfico de Gauge tem características parecidas, nele é refletido a porcentagem que o número de acidentes naquele trecho representa em relação a toda sua extensão, nesse caso ele serve como medidor, mostrando quais trechos foram apontados, pela previsão, com o maior número de acidentes. Devido aos trechos possuírem apenas 20 quilômetros de extensão, foi usado como comparação o maior e o menor caso. O menor caso é o trecho que vai do quilômetro 561 até o 580, com 0 acidentes previstos, o maior é o trecho que vai do quilômetro 161 até o 180, que representa 15,4% dos acidentes da rodovia.

Usando um esquema de cores foi então definido que porcentagens menores que cinco teriam a cor verde, pois estes possuem relativamente poucos acidentes. A cor amarela

representa acidentes com a porcentagem menor que dez e maior que cinco. A cor laranja para porcentagens maiores que dez e menores que quinze. E por fim a cor vermelha para porcentagens maiores que 15. Estabelecendo assim uma distinção qualitativa dos dados.

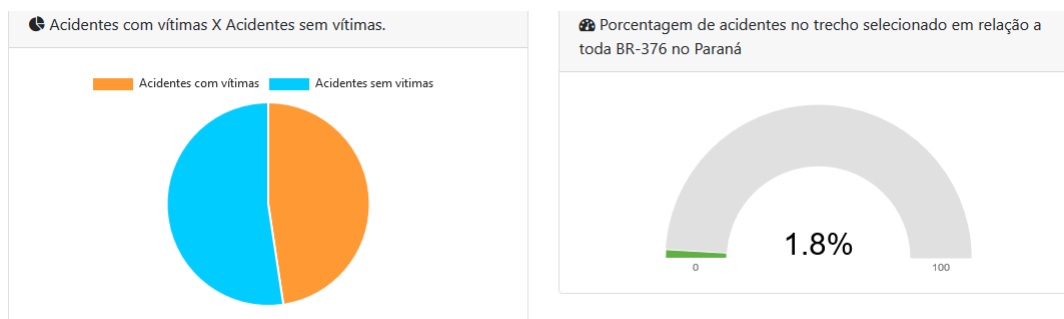


Figura 21 – Ferramenta protótipo: gráfico de pizza e de Gauge.

O último componente da visualização é uma tabela que exibe as cidades contidas em cada trecho selecionado, permitindo o usuário usar essas cidades como ponto de referência ao tentar se situar da localização dos trechos.

Concluindo a visualização, a ferramenta usou diversos elementos que possibilitam uma visão geral do número de acidentes por todos os trechos da rodovia, quais causas tem maior recorrência em cada trecho, o número de acidentes que geram vítimas, o quanto o número de acidentes naquele trecho representa na rodovia toda e por fim uma noção de geolocalização por meio das cidades como ponto de referência.

4 VALIDAÇÃO

Assim como o desenvolvimento do trabalho, a validação também é dividida entre a parte estatística e a visualização. Cada abordagem teve sua validação feita de maneira diferente, este capítulo é dedicado a explicá-las.

4.1 Validação do método de previsão

A validação com o método de previsão se dará por meio da comparação entre a previsão de um período de tempo que já aconteceu em 2017, com o que foi previsto pelo método para o mesmo período, uma comparação entre o real e a previsão.

A tabela disponibilizada pela PRF de 2017 contém dados de 1 de janeiro até 31 de março, ou seja, 3 meses de dados sobre acidentes. Estes dados foram tratados e moldados para uma tabela idêntica a de previsões, assim célula a célula, é feita a comparação gerando uma tabela de erros. Esses "erros" são as diferenças entre o número de acidentes real e o previsto, e pode ser observada na Tabela 2.

A tabela reflete os valores de diferença entre real e previsão. Valores iguais a 0 representam previsões corretas, onde real e previsão tiveram o mesmo resultado. Todo valor acima e abaixo de 0 são discrepâncias entre os resultados.

Essa diferença entre previsão e real é um reflexo da escolha do método de previsão. Neste caso, médias móveis não é um método eficiente para previsão de tais informações, pois é um método que deve ser aplicado a séries estacionárias, o que não é o caso dos acidentes utilizados, o que ocasionou a falta de precisão nas previsões. Isso pode ser corrigido através da aplicação de outros métodos, que levem em conta outras variáveis, as quais foram não utilizadas.

4.2 Validação da ferramenta de visualização

Afim de validar se a ferramenta de visualização cumpria com seus objetivos, um questionário foi aplicado. O objetivo do questionário era determinar se quem usou a ferramenta absorveu suas informações como era pretendido, e se ela apresentou algum defeito, seja estético ou funcional. O questionário está localizado no Apêndice B. Ele é dividido em duas partes, a primeira com relação a interface e a segunda com relação as informações, foram 7 perguntas simples e objetivas e, ao final, uma pergunta sobre a opinião dos usuários.

Visualmente a ferramenta se mostrou agradável na maioria dos casos, tendo 72,2% de "ótimo" nessa questão, a navegação ficou com sua maioria entre "bom" e "ótimo", com

Tabela 2 – Valor do erro de acordo com causas.

Trechos	FA	AP	DM	VI	IA	DS	UI	GD	DD	DV	OT
trecho1	7	-2	0	1	1	2	-1	3	-2	-2	0
trecho2	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	0
trecho3	0	-1	-1	-1	-2	-1	-1	-1	-2	-1	-1
trecho4	-7	-1	-1	-1	-2	-2	0	-1	-1	-1	0
trecho5	-3	-1	-2	-2	-2	-1	-1	-1	-2	-1	1
trecho6	1	-1	0	-2	0	-1	-1	0	-1	-1	-2
trecho7	4	-1	0	-2	-1	-2	0	0	-2	-1	1
trecho8	6	-1	0	-3	-1	-1	-1	2	-1	-1	3
trecho9	52	-1	2	-3	7	3	0	15	-1	-1	22
trecho10	14	-1	0	-2	0	1	-2	1	0	-2	15
trecho11	5	-1	-1	-2	0	1	0	2	-2	-1	4
trecho12	1	0	1	-1	0	0	-1	-1	-1	-1	3
trecho13	-2	0	0	0	0	-2	-2	-2	0	-1	3
trecho14	-1	-1	0	-2	-2	-1	-1	0	-1	-1	0
trecho15	1	-2	0	-1	-1	0	-1	0	-1	-1	1
trecho16	-1	-1	-5	2	-1	-1	-1	-3	-1	-1	-4
trecho17	0	-1	-1	-2	-1	-2	0	-1	-2	-2	0
trecho18	3	-1	0	5	0	-1	-2	0	-2	-1	0
trecho19	3	-1	-2	-1	-1	-1	-1	-1	-2	-1	0
trecho20	-4	-1	0	0	-3	-1	-2	-1	-2	-1	-1
trecho21	-4	-1	-1	4	-1	-1	-1	-1	-2	-1	2
trecho22	0	-2	-1	3	-1	-2	0	-2	0	-1	0
trecho23	1	-1	-3	1	-2	-1	0	-1	-1	-1	2
trecho24	-1	-1	-1	-5	-1	-1	0	0	-1	-2	-1
trecho25	3	0	1	-1	-3	0	-1	-2	-1	-2	3
trecho26	0	-1	-1	0	0	-1	-1	0	-3	-1	5
trecho27	-2	-1	-1	-1	-2	-1	-1	-1	0	-2	4
trecho28	0	-1	0	1	-1	-1	-1	0	-1	-1	3
trecho29	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
trecho30	8	-1	-2	0	-1	0	-1	-2	0	-1	3
trecho31	4	-1	-1	3	-1	-2	-1	9	-1	0	4
trecho32	7	-1	0	0	2	-1	-2	5	0	-1	5
trecho33	8	-1	-1	13	-2	-3	0	7	2	-1	4
trecho34	22	-1	-2	54	0	-1	0	13	1	0	-2

iguais 38,9% em cada, por fim o botão de ajuda teve um resultado completamente positivo, não havendo votos em "ruim" ou "péssimo", e sua maioria, 61,1%, foram em "ótimo".

A análise da informação transmitida continha perguntas binárias, no qual apenas respostas "sim" ou "não" estavam disponíveis. Em 100% dos casos os usuários puderam identificar qual trecho possuía mais acidentes e concordaram que as informações permitiam uma análise geral da BR-376. Porém, as duas últimas perguntas tiveram respostas negativas apesar da maioria ainda ser positiva.

Dando seguimento, a última pergunta do questionário, esta relacionada a opinião dos usuários com o que poderia ser melhorado na ferramenta. A maioria das respostas se atentava a questões referentes a rotulação dos dados, e aperfeiçoamento da interface, alegando uma certa "pobreza" nesse quesito. Outra opinião frequente era a possibilidade de mais comparações.

No geral, o questionário apresentou um *feedback* positivo, porém denotou diversos pontos em que a ferramenta pode melhorar. Os resultados do questionário estão disponíveis no Apêndice C.

5 CONSIDERAÇÕES FINAIS

O foco deste trabalho foi unir a previsão e a visualização de dados sobre acidentes rodoviários. O método usado, médias móveis, não proporcionou a precisão necessária, apresentando grande discrepância com relação aos dados reais. Novos métodos podem ser aplicados, e outras variáveis levadas em consideração, a tabela original da PRF tem muitas informações que podem ser trabalhadas, como o clima na hora do acidente, horário do acidente, tipo de via, época do ano e outros. Quanto mais variáveis, mais precisos serão os cálculos e mais confiável será a ferramenta.

Ao final do desenvolvimento, a ferramenta não era completamente funcional, não existia um banco de dados, os métodos de previsão eram aplicados diretamente na tabela, e por questões como essa foi levantada a hipótese que talvez a visualização na verdade era um infográfico, todavia, essa hipótese não reflete a realidade. Em parte os dados visualizados são "estáticos", mas apenas com relação ao período do tempo no qual a previsão é feita, nas demais funcionalidades eles refletem as características de uma visualização de dados, e portanto é de fato uma.

Contudo, não foi possível que a ferramenta fosse testada por profissionais do ramo rodoviário, como administradores de rodovia. Portanto é impossível dizer com convicção se ela realmente pode ser usada com o propósito de distribuir recursos orientados pela ferramenta. Todavia, baseado nos resultados do questionário, há sim uma passagem de informação, e há lógica em dizer que as chances da ferramenta atingir tal propósito não são nulas. A análise geral da rodovia pode ser feita através dos gráficos e informações da ferramenta, a maneira como essa análise é utilizada depende do usuário e sua própria capacidade.

O trabalho durante todo seu desenvolvimento se manteve pontual aos seus objetivos sempre em busca de alcançá-los, oferecendo aos interessados um primeiro passo do que pode se tornar uma ferramenta da sociedade para um melhoramento da logística de recursos rodoviários.

REFERÊNCIAS

- [1] VASCONCELLOS FÁBIO E LIMA, D. *Mapa dos acidentes nas rodovias Federais*. 2016. [Http://infograficos.oglobo.globo.com/brasil/mapa-dos-acidentes-das-rodovias-federais.html](http://infograficos.oglobo.globo.com/brasil/mapa-dos-acidentes-das-rodovias-federais.html). Disponível em: [<http://infograficos.oglobo.globo.com/brasil/mapa-dos-acidentes-das-rodovias-federais.html>](http://infograficos.oglobo.globo.com/brasil/mapa-dos-acidentes-das-rodovias-federais.html). Acesso em: 14.6.2017. 10, 12
- [2] BERNARDES, A. *Vítimas de acidentes têm mais chance de sobreviver quando socorro é rápido*. 2011. Correio Brasileiro. Disponível em: [<http://www.correiobraziliense.com.br/app/noticia/cidades/2011/02/28/interna_cidadesdf,240145/vitimas-de-acidentes-tem-mais-chance-de-sobreviver-quando-socorro-e-rapido.shtml>](http://www.correiobraziliense.com.br/app/noticia/cidades/2011/02/28/interna_cidadesdf,240145/vitimas-de-acidentes-tem-mais-chance-de-sobreviver-quando-socorro-e-rapido.shtml). Acesso em: 14.6.2017. 10, 14
- [3] DNIT. *Nomenclatura das rodovias federais 2017*. 2017. Disponível em: [<http://www.dnit.gov.br/rodovias/rodovias-federais/nomeclatura-das-rodovias-federais>](http://www.dnit.gov.br/rodovias/rodovias-federais/nomeclatura-das-rodovias-federais). 10
- [4] TRANSPORTE, C. N. D. *Pesquisa CNT de Rodovias 2016: Relatório gerencial*. 2016. 20.ed. – Brasília : CNT : SEST : SENAT. Disponível em: [<http://pesquisarodovias.cnt.org.br/>](http://pesquisarodovias.cnt.org.br/). Acesso em: 14.6.2017. 11
- [5] PIRES, J. C. L.; GIAMBIAGI, F. Retorno dos novos investimentos privados em contextos de incerteza: uma proposta de mudança do mecanismo de concessão de rodovias no brasil. *Revista de Administração Pública*, v. 34, n. 4, p. 27–64, 2000. 11
- [6] CART. *CART, SAU24h, serviço ao usuário*. 2017. CART website. Disponível em: [<http://www.cart.invepar.com.br/nossos-servicos/sau/>](http://www.cart.invepar.com.br/nossos-servicos/sau/). 12
- [7] CATIA. *Dados Abertos Governamentais*. 2016. Disponível em: [<https://www.governoeletronico.gov.br/eixos-de-atuacao/cidadao/dados-abertos/dados-abertos-governamentais>](https://www.governoeletronico.gov.br/eixos-de-atuacao/cidadao/dados-abertos/dados-abertos-governamentais). 12
- [8] SCHROEDER, E.; CASTRO, J. C. d. Transporte rodoviário de carga: situação atual e perspectivas. *Revista do BNDES*, 1996. 12
- [9] VIA040. *Via40, Nossos Serviços*. 2017. Site via40. Disponível em: [<http://via040.com.br/pages/nossos-servicos>](http://via040.com.br/pages/nossos-servicos). 13
- [10] GIL, A. C. Como elaborar projetos de pesquisa. *São Paulo*, v. 5, p. 61, 2002. 15, 28
- [11] NEVES, J. L. Pesquisa qualitativa: características, usos e possibilidades. *Caderno de pesquisas em administração, São Paulo*, v. 1, n. 3, p. 2, 1996. 15
- [12] HILBERT, M.; LÓPEZ, P. The world’s technological capacity to store, communicate, and compute information. *science*, American Association for the Advancement of Science, v. 332, n. 6025, p. 60–65, 2011. 17
- [13] CIELEN, D.; MEYSMAN, A.; ALI, M. *Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools*. Manning Publications Company, 2016. ISBN 9781633430037. Disponível em: [<https://books.google.com.br/books?id=zYbisgEACAAJ>](https://books.google.com.br/books?id=zYbisgEACAAJ). 17, 18, 20

- [14] TESSAROLO, W. B. M. P. H. A era do big data no conteúdo digital: Os dados estruturados e não estruturados. 2015. 17
- [15] ILIINSKY, N.; STEELE, J. *Designing Data Visualizations: Representing Informational Relationships*. O'Reilly Media, 2011. ISBN 9781449317065. Disponível em: <<https://books.google.com.br/books?id=ygLH4qrGm-wC>>. 19, 20, 21
- [16] YAU, N. *Visualize This!* Wiley-VCH, 2012. ISBN 9783527760220. Disponível em: <<https://books.google.com.br/books?id=H7ctBAAAQBAJ>>. 19
- [17] ARENALES, M. et al. *Pesquisa Operacional: Para cursos de engenharia*. Elsevier Brasil, 2015. ISBN 9788535281835. Disponível em: <<https://books.google.com.br/books?id=aZbpCgAAQBAJ>>. 23
- [18] VERÍSSIMO, A. J. et al. Métodos estatísticos de suavização exponencial holt-winters para previsão de demanda em uma empresa do setor metal mecânico. *Revista Gestão Industrial*, v. 8, n. 4, 2013. 23
- [19] LATORRE, M. d. R. A. D. d. O.; CARDOSO, M. R. A. Análise de sÃtemporais em epidemiologia: uma introduÃ§Ãsobre os aspectos metodolÃ. *Revista Brasileira de Epidemiologia*, sciELO, v. 4, p. 145 – 152, 11 2001. ISSN 1415-790X. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-790X2001000300002&nrm=iso>. 23
- [20] LUSTOSA, L. J.; MESQUITA, M. A. de; OLIVEIRA, R. J. *Planejamento e controle da produção*. [S.l.]: Elsevier Brasil, 2008. 24, 25
- [21] GUERRINI, F. M.; JUNIOR, W. A.; RENATO, V. B. *Planejamento e Controle da Produção: Projecto e Operação de Sistemas*. [S.l.]: Elsevier Brasil, 2013. v. 1. 23, 26, 27

Apêndices

APÊNDICE A – TABELA RESULTADOS DA PREVISÃO 2017

Legenda:

Falta de atenção = FA;

Animais na pista = AP;

Defeito mecânico em veículo = DM;

Velocidade incompatível = VI;

Ingestão de álcool = IA;

Desobediência à sinalização = DS;

Ultrapassagem indevida = UI;

Não guardar distância de segurança = GD;

Dormindo = DD;

Defeito na via = DV;

Outras = OT;

Trechos	FA	AP	DM	VI	IA	DS	UI	GD	DD	DV	OT
Trecho1	21	0	1	4	5	6	0	8	0	0	8
Trecho2	0	0	0	0	0	0	0	0	0	0	1
Trecho3	1	0	0	0	0	0	0	0	0	0	1
Trecho4	1	0	0	0	0	0	1	0	0	0	1
Trecho5	1	0	0	0	0	0	0	0	0	0	2
Trecho6	8	0	1	1	2	0	0	1	0	0	4
Trecho7	8	0	1	1	1	1	1	1	0	0	6
Trecho8	11	1	1	0	2	2	0	3	0	0	6
Trecho9	83	1	6	4	14	16	1	21	1	0	31
Trecho10	37	1	2	4	5	7	0	10	1	0	20
Trecho11	16	0	1	2	2	4	1	3	0	0	8
Trecho12	6	1	2	1	1	2	0	1	0	0	5
Trecho13	6	1	1	1	2	1	0	2	1	0	5
Trecho14	2	0	1	0	0	0	0	1	0	0	2
Trecho15	3	0	1	1	1	1	0	1	0	0	3
Trecho16	3	0	2	10	0	0	0	1	1	0	5
Trecho17	2	0	0	3	0	0	1	0	0	0	4
Trecho18	8	0	1	6	1	0	1	2	0	0	5
Trecho19	4	0	0	1	0	0	0	0	0	0	2
Trecho20	5	0	1	2	1	0	0	0	0	0	2
Trecho21	6	0	1	5	0	0	0	0	1	0	3
Trecho22	3	0	0	4	0	0	1	0	1	0	2
Trecho23	6	0	0	6	0	1	1	1	0	0	4
Trecho24	4	0	0	3	0	0	1	1	1	0	3
Trecho25	19	1	2	4	2	1	0	6	1	0	9
Trecho26	9	0	1	4	1	0	0	3	2	0	7
Trecho27	6	0	1	4	0	0	0	3	1	0	5
Trecho28	2	0	1	3	0	0	0	1	1	0	4
Trecho29	0	0	0	0	0	0	0	0	0	0	0
Trecho30	17	0	1	6	2	1	0	8	1	1	10
Trecho31	19	0	2	6	3	2	0	10	0	1	6
Trecho32	21	0	2	8	3	1	0	9	2	1	8
Trecho33	23	0	3	43	2	1	1	9	3	0	14
Trecho34	33	0	15	82	2	1	1	17	5	1	16

APÊNDICE B – QUESTIONÁRIO: FERRAMENTA DE VISUALIZAÇÃO

Questionário: Ferramenta de visualização de dados sobre acidentes da BR-376.

Questionário relacionado a aplicação web: http://cct.uenp.edu.br/thiago/tcc_joao/

*Obrigatório

Análise da interface.

Responda de acordo com a seguinte escala:

- 1- Péssimo
- 2- Ruim
- 3- Razoável
- 4- Bom
- 5- Ótimo

O site é agradável visualmente? *

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Como você avalia a navegação pelo site? *

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Em relação ao conteúdo apresentado pelo botão de ajuda, o mesmo foi? *

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 22 – Questionário parte 1.

Análise das Informações

Você conseguiu identificar qual trecho possui mais acidentes? *

- Sim
 Não

As informações disponibilizadas permitem uma análise geral da BR-376? *

- Sim
 Não

As informações apresentadas sobre o trecho que você selecionou foram claras e permitiram você analisar a incidência de acidentes??

- Sim
 Não

Você utilizaria o site para auxiliar sua decisão de viajar ou não? *

- Sim
 Não

Qual sua opinião sobre o que poderia ser melhorado? *

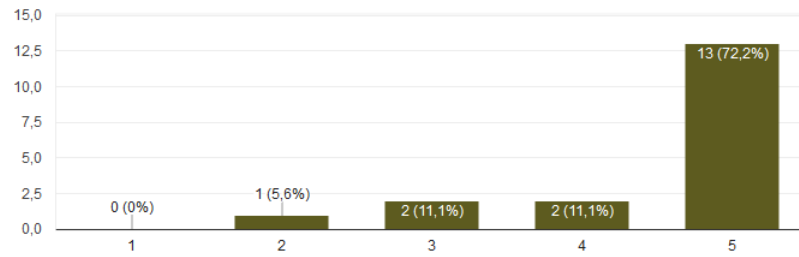
Sua resposta

Figura 23 – Questionário parte 2.

APÊNDICE C – QUESTIONÁRIO: RESPOSTAS

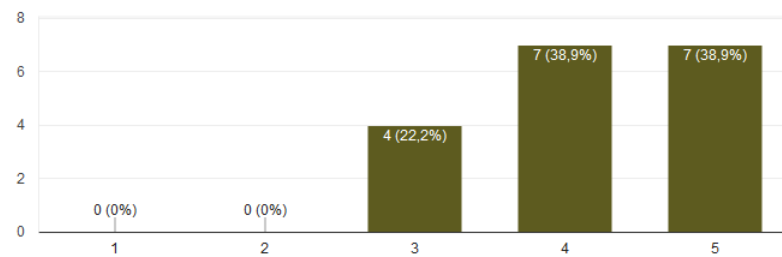
O site é agradável visualmente?

18 respostas



Como você avalia a navegação pelo site?

18 respostas



Em relação ao conteúdo apresentado pelo botão de ajuda, o mesmo foi?

18 respostas

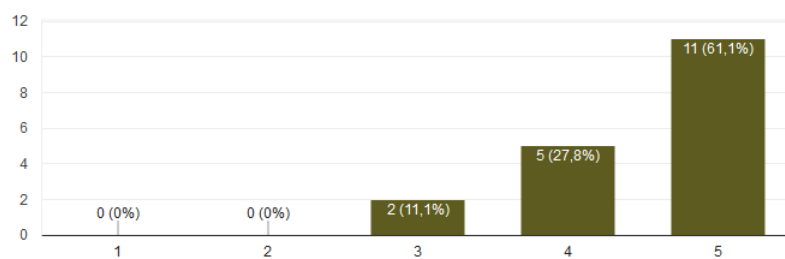
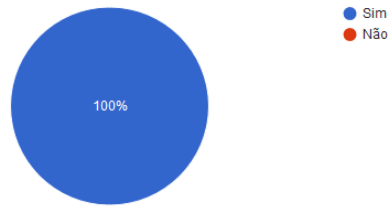


Figura 24 – Respostas parte 1.

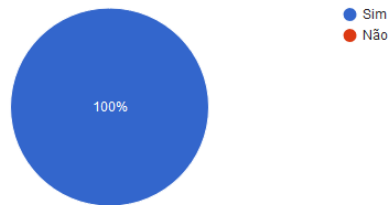
Você conseguiu identificar qual trecho possui mais acidentes?

18 respostas



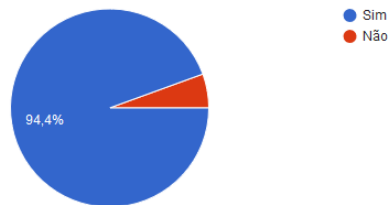
As informações disponibilizadas permitem uma análise geral da BR-376?

18 respostas



As informações apresentadas sobre o trecho que você selecionou foram claras e permitiram você analisar a incidência de acidentes??

18 respostas



Você utilizaria o site para auxiliar sua decisão de viajar ou não?

18 respostas

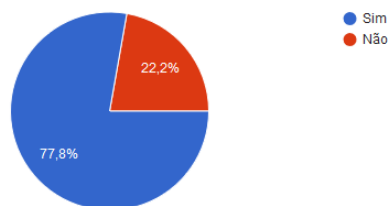


Figura 25 – Respostas parte 2.

Qual sua opinião sobre o que poderia ser melhorado?

18 respostas

Faltou um pouco de design (2)

Minha vida

A visibilidade dos dados e mais tipos de comparações

A ferramenta é boa, porém acho que ao me deparar com uma rota no aplicativo e ver que ela possui vários acidentes com vítimas, por exemplo, ter a opção de saber se existiria uma rota alternativa que não possuísse tantos acidentes.

Poderia ter um menu lateral com funções específicas, não apenas uma interface todas as informações. Também a parte das datas poderia mostrar um mini calendário para o usuário selecionar a data sem a necessidade de digitar.

Os gráficos poderiam apresentar rótulos de dados, como por exemplo, o gráfico de pizza apresentando a porcentagem (60% pra uma ocorrência, 40% pra outra). Os gráfico que tem a informação sobre toda a rodovia também, apresentar um pequeno campo de texto acima do pico demonstrando visualmente o número exato de acidentes sem que o cliente tenha que passar o mouse. Grato!

Poderia colocar "nº" ou "número" ao invés de "#" nas descrições dos gráficos. #ForaTemer

Como é para viagens, fazer um site para celular seria bom

Figura 26 – Respostas parte 3.

