CENTRO PAULA SOUZA FACULDADE DE TECNOLOGIA DE FRANCA "Dr. THOMAZ NOVELINO"

TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

CAIRO DE PAULA VIEIRA

DOCUMENTAÇÃO DO SOFTWARE DE MINI BI E ESTATÍSTICA GRANGER

Trabalho apresentado à Faculdade de Tecnologia de Franca - "Dr. Thomaz Novelino", como parte dos requisitos obrigatórios para as disciplinas de Estatística, Engenharia de Software II, Interface Humano-Computador e Estrutura e Dados no curso de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Orientadores: Prof. Me. Ely Fernando do Prado

Prof. Me. Fausto Gonçalves Cintra Prof. Me. Maria Luísa Cervi Uzun

FRANCA/SP 2020

DOCUMENTAÇÃO DO SOFTWARE DE MINI BI E ESTATÍSTICA

GRANGER

Cairo de Paula Vieira¹

Resumo

Empresas produzem uma infinidade de dados e com o intuito de otimizar os processos de tomada de decisão, buscam cada dias mais ferramentas da categoria de Inteligiencia de Negócios Business Intelligence (B.I.). O B.I. é um conceito utilizando tanto na academia quanto nas indústrias voltado para diversas áreas do negócio, gerando valor e auxiliando no processamento de big data. Este trabalho teve como intuito desenvolver um Software de Mini B.I. empregando conceitos básicos da Estatística Aplicada. Utilizou-se o levantamento de requisitos como direcionamento, Diagrama de Processes BPMN e Diagrama de Casos de Uso como meio de melhor visualizar os fluxos do sistema. Utilizou-se também JavaScript como principal linguagem e bibliotecas como Bootstrap, Node e React. Como resultado, obteve-se um sistema simples e intuitivo com funções simples, mas capazes de gerar valor e auxiliar no processo de qual caminho seguir rumo ao crescimento de determinado empreendimento ou atividade. Para um software mais completo, fazse mister aprofundar os conceitos estudados e realizar um novo desenvolvimento.

Palavras-chave: Inteligência de Negócios. Sistemas de Informação Gerencial. Estratégia Organizacional.

Abstract

Companies generate a lot of data and in order to optimize decision-making process, they look for further tools in the field of Business Intelligence (B.I.). B.I. is a concept used both in academia and in industries, pointed for different areas of business, generating value and assisting in the processing of big data. This work aimed to develop a Mini B.I. employing basic concepts of Applied Statistics. The requirements survey was used as direction, BPMN Processes Diagram and Use Case Diagram as a way to get a better view of the system flow. JavaScript was used as the main language and libraries such as Bootstrap, Node and React was also used. As a result, a simple and intuitive system with simple functions was obtained, but capable of generating value and helping in the process of which path to follow towards the growth of a specific enterprise or activity. For a more complete software, it is required to deepen the studied concepts and carry out a new development.

Keywords: Business Intelligence. Management Information Systems. Organizational Strategy.

¹ Graduando em Análise e Desenvolvimento de Sistemas pela Fatec Dr Thomaz Novelino – Franca/SP. Endereço eletrônico: cairo.vieira@fatec.sp.gov.br.

1 Introdução

Empresas produzem diariamente uma infinidade de dados com padrões, relações, tendências, dentre outras informações. No intuito de otimizar os processos de tomada de decisão, buscam-se cada vez mais ferramentas que auxiliem nesse processo. A principal categoria destas ferramentas é a chamada Inteligência de Negócios - *Business Intelligence* (B.I.) (BACH; ZOROJA; ČELJO, 2017).

Business Intelligence é hoje um importante conceito tanto na academia, quanto nas indústrias, haja visto que são amplamente utilizadas nas mais diversas áreas de negócios, na geração de valor e também para o processamento de *big-data*. Por definição, B.I. é o conjunto de sistemas e métodos usados para permitir a análise de informações de um determinado negócio, objetivando o melhor direcionamento de estratégias a fim de gerar lucro, atuando efetivamente no mercado (GRUBLJEŁIČ; JAKLIČ, 2015; TRIEU, 2017).

Quando unimos o B.I. ao constante desenvolvimento tecnológico, contando com computadores de última geração e linguagens avançadas, alcançamos um processo mais acessível, utilizado por profissionais dos mais diferentes campos de atuação.

Assim, o intuito deste trabalho foi desenvolver um *software* de Mini B.I. chamado Granger que contemplasse processos simples, mas eficazes para o aprendizado dos conceitos básicos das disciplinas que envolvem este projeto. Apesar de possuir a possibilidade de auxiliar em pequenos empreendimentos e atividades, visou-se apenas fins didáticos.

2 Levantamento de Requisitos

2.1 Elicitação e especificação dos Reguisitos

A elicitação e especificação dos requisitos foi feita por meio de discussões e assessorias durante as aulas e por meio da implementação parcial do software, junto aos *stakeholders* do projeto, que neste caso são os docentes das disciplinas Estatística, Engenharia de Software II, Estrutura de Dados e Interação Humano Computador, do curso Tecnólogo em Análise e Desenvolvimento de Sistemas da Faculdade de Tecnologia de Franca – "Dr. Thomaz Novelino.

2.2 BPMN

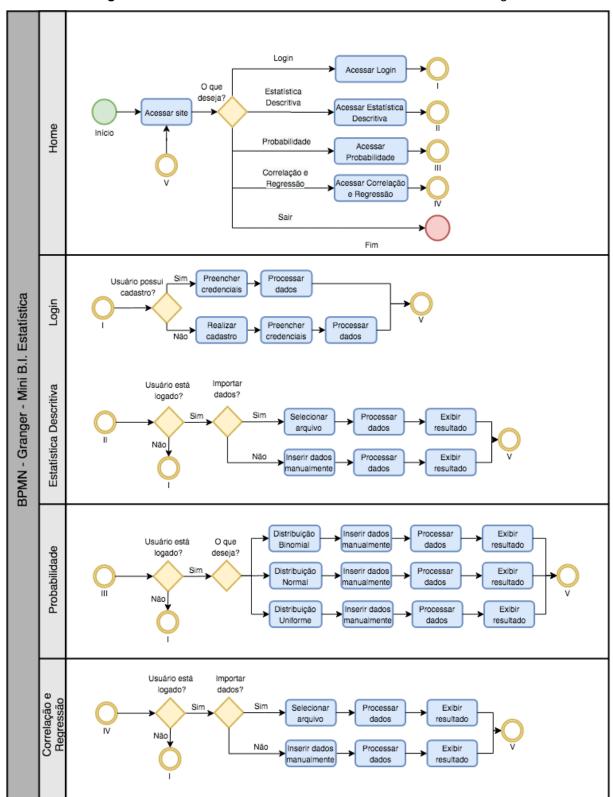


Figura 1 - Fluxo de Processos BPMN do Software de Mini B.I. Granger

Fonte: Cairo de Paula Vieira (construído em Draw.io)

2.3 Requisitos Funcionais

Quadro 1 - Requisitos Funcionais do sistema

4				
REQUISITOS FUNCIONAIS	CATEGORIA	PRIORIDADE		
RF001 - Escolher opções	Evidente	Altíssima		
	nitir que o usuário escolha a opção idade" ou "Correlação e Regressã			
RF002 - Inserir dados	Evidente	Altíssima		
Descrição: O sistema deve pern	nitir a inserção manual de dados p	elo usuário.		
RF003 - Importar dados	Evidente	Altíssima		
Descrição : O sistema deve permitir a importação/upload de dados pelo usuário em Estatística Descritiva e Correlação e Regressão.				
RF004 - Organizar dados	Evidente	Altíssima		
Descrição : O sistema deve organizar os dados para a realização dos cálculos de acordo com o tipo da variável (Qualitativa ou Quantitativa).				
RF005 - Tabular os dados de acordo com a Distribuição de Frequências	Evidente	Altíssima		
Descrição: O sistema deve tabular os dados de acordo com o subtipo da variável (Qualitativa				

nominal, Qualitativa Ordinal, Quantitativa Discreta, Quantitativa Contínua).

RF006 - Realizar a distribuição de frequências Evidente Altíssima

Descrição: O sistema deve fazer a distribuição das frequências, exibindo uma coluna para a variável pesquisada (ou intervalo de classe no caso de Quantitativa Contínua) pelo usuário, uma coluna para a frequência simples (número de vezes que cada dado apareceu), uma coluna para a frequência relativa percentual (valor percentual da frequência simples), uma coluna para a frequência acumulada (soma de todas as frequências até o ponto presente) e uma coluna para a frequência acumulada percentual (valor percentual da frequência acumulada).

RF007 - Gerar os gráficos Evidente Alta

Descrição: O sistema deve gerar os gráficos de acordo com subtipo da variável sendo do tipo "Pizza" para as Qualitativas, do tipo "Colunas" para a Quantitativa Discreta, do tipo "Histograma" para a Quantitativa Contínua e do tipo "Dispersão" para Correlação e Regressão.

RF008 - Calcular as Medidas de Tendência Central Evidente Altíssima

Descrição: O sistema deve realizar os cálculos das Medidas de Tendência Central (Média, Mediana e Moda) utilizando as seguintes fórmulas.

- **Média:** Para variável quantitativa discreta multiplicar o valor de cada elemento pesquisado pela frequência simples, somar os resultados e dividir pelo somatório da frequência simples.
 - Para variável quantitativa contínua multiplicar a frequência simples pelo ponto médio do intervalo de classe, somar os resultados de todas as linhas e dividir pelo somatório da frequência simples.
- **Mediana:** Para variáveis qualitativas e quantitativa discreta exibir o elemento da posição central dos dados.
 - Para quantitativa contínua utilizar a seguinte fórmula: $Md = I + (((\Sigma fi/2)-Fant)/fimd),$ Sendo Md - Mediana; fi - frequência simples; Fant - frequência acumulada anterior e;

fimd – frequência simples do intervalo de classe pesquisado.

Moda: - Para variáveis qualitativas e quantitativa discreta – exibir o elemento que mais se repete.
 - Para quantitativa contínua – exibir o ponto médio do intervalo de classe com maior de frequência simples.

RF009 - Calcular as Medidas Separatrizes Evidente Alta

Descrição: O sistema deve realizar os cálculos das Medidas Separatrizes (Quartil, Quintil, Decil e Percentil).

Quartil: Divide-se a sequência em quatro partes iguais, onde Q1=25%; Q2=50%; Q3=75%; Q4=100%.

Quintil: Divide-se a sequência em cinco partes iguais, onde K1=20%; K2=40%; K3=60%; K4=80%; K5=100%.

Decil: Divide-se a sequência em dez partes iguais onde, D1=10%; D2=20%; D3=30%; D4=40%; D5...; D10=100%.

Percentil: Divide-se a sequência em cem partes iguais onde, P1=1%; P2=2%;; P100=100%.

Para variáveis qualitativas e quantitativa discreta: Primeiro deve-se encontrar a posição em relação ao Σ fi. Na linha referente à posição encontrada, mostrar a variável pesquisada.

Para variável quantitativa contínua: Primeiro deve-se encontrar a posição em relação ao Σ fi. Depois, aplicar a fórmula I+ (Posição – Fac anterior) / fi*h, onde I é o limite inferior da linha da posição encontrada, Fac é a frequência acumulada anterior, fi é a frequência simples e h é o intervalo de classe.

RF010 - Calcular as Medidas de Dispersão Evidente Altíssima

Descrição: O sistema deve realizar os cálculos das Medidas de Dispersão (Desvio Padrão e Coeficiente de Variação).

Para o desvio padrão: Encontra-se inicialmente, a média. Depois, aplicar a seguinte fórmula: $\sigma=v((\Sigma(xi-x)^2/\Sigma fi)^*fi)$, onde "xi" é o valor da variável pesquisa na quantitativa discreta e o ponto médio na quantitativa contínua; e "x" é a média. Caso a pesquisa seja uma amostra: $\sigma=v((\Sigma(xi-x)^2/\Sigma fi-1)^*fi)$.

Para o coeficiente de variação: Divide-se o desvio padrão pela média e depois multiplica-se por 100.

RF011 - Escolher entre amostra e população Evidente Altíssima

Descrição: O sistema deve permitir ao usuário escolher se os dados inseridos ou importados se referem a uma amostra ou a uma população.

RF012 - Calcular a Distribuição
Binomial em Probabilidade

Evidente

Altíssima

Descrição: O sistema deverá calcular a distribuição binomial por meio da seguinte fórmula: $P(x=k)=(n!/k!)*p^k*q^{n-k}$, onde p=sucesso, q=fracasso, k=evento e n=tamanho da amostra.

RF013 - Calcular a Distribuição Normal em Probabilidade. Evidente Altíssima

Descrição: O sistema deve calcular a distribuição normal obtendo o valor Z por meio da seguinte fórmula: $z=(x-\mu)/\sigma$. Sendo "x" o valor a ser transformado, " μ " o valor da média e " σ " o valor do desvio padrão. Para localizar o valor na Tabela Z, deve-se pegar o inteiro e a primeira casa decimal e cruzar com a segunda casa decimal na tabela. Feito isso, deve-se realizar o cálculo propriamente dito de acordo com a área da Curva de Gauss.

RF014 - Calcular a Distribuição Uniforme em Probabilidade Evidente

Altíssima

Descrição: O sistema deve calcular a distribuição uniforme por meio da seguinte fórmula: f(x)=1/(b-a). Sendo "b" o ponto máximo e "a" o ponto mínimo a ser pesquisado. Caso a pesquisa seja em cima de um intervalo contido entre "a" e "b" deve-se seguir a seguinte fórmula: : f(x)=1/(b-a)*intervalo.

RF015 - Calcular a Correlação e Regressão

Evidente

Altíssima

Descrição: O sistema deve calcular a Correlação entre duas variáveis por meio da seguinte fórmula: $r=(n\Sigma xy - (\Sigma x)(\Sigma y))/v(n\Sigma x^2 - (\Sigma x)^2)*v(n\Sigma y^2 - (\Sigma y)^2)$. Sendo "r" o coeficiente de relação, "n" o tamanho da amostra, "x" a variável independente e "y" a variável dependente. A regressão deve ser calculada por meio da seguinte fórmula: y = a*x + b. Sendo $a = (n\Sigma xy - \Sigma x\Sigma y)/(n\Sigma xi^2 - (\Sigma x)^2)$ e b = y - ax, onde $x = \Sigma x/n$ e $y = \Sigma y/n$. O sistema deve exibir

2.4 Requisitos Não Funcionais

Quadro 2 - Requisitos Não Funcionais do sistema

Requisito Não Funcional	Categoria	Obrigatoriedade	Permanência	
RNF001 - Executar o sistema	Tecnologia	Obrigatório	Permanente	
Descrição: O sistema deverá ser executado r	no ambiente da i	nternet		
RNF002 - Ser responsivo	Interface	Obrigatório	Permanente	
Descrição: O sistema deverá ser responsivo, smartphones.	adaptando sua	interface para compu	ıtadores, tablets e	
RNF003 - Controlar acesso	Segurança	Obrigatório	Permanente	
Descrição: Usuários não cadastrados e não lo do sistema.	ogados não pod	erão executar nenhu	ma funcionalidade	
RNF004 - Gerar Curva de Gauss na Distribuição Normal da Probabilidade	Interface	Desejável	Permanente	
Descrição: O sistema deverá exibir a curva de Gauss na Distribuição Normal da Probabilidade.				
RNF005 - Ter landing page	Interface	Desejável	Permanente	
Descrição: O sistema deverá ter uma landing page.				
RNF006 - Usar JavaScript como linguagem de programação	Tecnologia	Obrigatório	Permanente	
Descrição: O sistema deverá usar JavaScript como linguagem de programação.				
RNF007 - Gerar relatório das análises	Usabilidade	Desejável	Permanente	

Descrição: O sistema deverá gerar um relatório da análise realizada.				
RNF008 - Ter paleta de cores azul, rosa, amarelo e cinza Interface Desejável Permanente				
Descrição: A paleta e cores do sistema deve ser azul, rosa, amarelo e cinza.				
RNF009 - Ter mais de um idioma	Usabilidade	Desejável	Permanente	
Descrição: O sistema deverá ter a possibilidade de mudança de idioma.				

2.5 Regras de Negócio

Como citado no item 1 deste trabalho, o sistema não tem o intuído de atender, tampouco envolve um cliente/empresa. Portanto, não existem regras de negócio.

2.6 Matriz de Rastreabilidade

Tabela 1 – Matriz de Rastreabilidade entre Requisitos Funcionais e Não Funcionais

	RNF001	RNF002	RNF003	RNF004	RNF005	RNF006	RNF007	RNF008	RNF009
RF001									
RF002									
RF003									
RF004									
RF005									
RF006									
RF007									
RF008									
RF009									
RF010									
RF011									
RF012									
RF013									
RF014									
RF015									

2.7 Casos de Uso

2.7.1 Índice de Casos de Uso e Diagrama de Casos de Uso

- UC 001: Escolher Estatística Descritiva;
- UC 002: Escolher Probabilidade;
- UC 003: Escolher Correlação e Regressão.

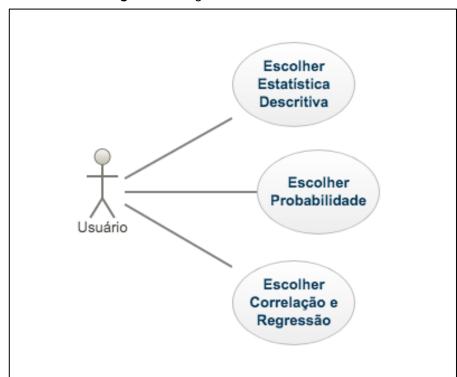


Figura 2 - Diagrama de Casos de Uso

Fonte: Cairo de Paula Vieira (construído em GenMyModel)

2.7.2 Especificação dos Casos de Uso

Quadro 4 – Use Case Escolher Estatística Descritiva

Caso de Uso – Escolher Estatística Descritiva		
ID	UC 001	
Descrição	Este caso de uso tem por objetivo permitir o acesso à Estatística Descritiva	
Ator Primário	Usuário do sistema	
Pré-condição	Estar logado	
Cenário Principal	O use case inicia quando o usuário seleciona a opção Estatística Descritiva no Menu superior principal ou no botão "Comece agora" imediatamente abaixo da descrição de Estatística Descritiva.	

	 O sistema carrega o formulário para inserção manual dos dados. O sistema carrega o formulário para inserção do nome da variável pesquisada, dados propriamente ditos, seleção entre amostra e população e botão calcular. O sistema valida os campos obrigatórios (nome da variável e dados). O sistema exibe o gráfico correspondente à variável pesquisada, a tabela com
	distribuição de frequências, as medidas centrais, as medidas de dispersão e as medidas separatrizes.
Pós-condição	Nenhuma
Cenário Alternativo	*a – O usuário pode sair do sistema a qualquer momento.
	2a – O sistema carrega o formulário para importação dos dados.
	3a – O sistema exibe o gráfico correspondente à variável pesquisada, a tabela com distribuição de frequências, as medidas centrais, as medidas de dispersão e as medidas separatrizes diretamente se o usuário importar os dados.
	 4a – Dados obrigatórios não preenchidos: 4a.1 O sistema informa ao usuário sobre campos obrigatórios que não foram preenchidos (nome da variável e dados). 4a.2 Retorna ao passo 2 do cenário principal.

Quadro 5 – Use Case Escolher Probabilidade

Caso de Uso – E	scolher Probabilidade
ID	UC 002
Descrição	Este caso de uso tem por objetivo permitir o acesso à Estatística Descritiva
Ator Primário	Usuário do sistema
Pré-condição	Estar logado
Cenário Principal	 O use case inicia quando o usuário seleciona a opção Probabilidade no Menu superior principal ou no botão "Comece agora" imediatamente abaixo da descrição de Probabilidade. O sistema carrega o formulário para escolha da Distribuição Binomial. O sistema carrega o formulário para inserção do valor da amostra, do valor do sucesso, do valor do fracasso, valor do evento e botão calcular. O sistema valida os campos obrigatórios (todos). O sistema exibe o resultado correspondente à Probabilidade e também a média, desvio padrão e coeficiente de variação.
Pós-condição	Nenhuma
Cenário Alternativo	 *a – O usuário pode sair do sistema a qualquer momento. 2a – O sistema carrega o formulário para escolha da Distribuição Normal. 2b – O sistema carrega o formulário para escolha da Distribuição Uniforme. 3a – O sistema carrega o formulário para inserção da média, do desvio padrão, escolha entre menor, entre e maior, valor a ser calculado e botão calcular. 3b – O sistema carrega o formulário para inserção do ponto mínimo, ponto máximo, escolha entre menor, entre e maior e botão calcular. 4a – Dados obrigatórios não preenchidos: 4a.1 O sistema informa ao usuário sobre campos obrigatórios que não foram preenchidos (nome da variável e dados). 4a.2 Retorna ao passo 2 do cenário principal. 5a – O sistema exibe o resultado correspondente à Probabilidade.

Quadro 6 – Use Case Escolher Correlação e Regressão

Caso de Uso – Escolher Correlação e Regressão			
ID	UC 003		
Descrição	Este caso de uso tem por objetivo permitir o acesso à Correlação e Regressão		
Ator Primário	Usuário do sistema		
Pré-condição	Estar logado		
Cenário Principal	 O use case inicia quando o usuário seleciona a opção Correlação e Regressão no Menu superior principal ou no botão "Comece agora" imediatamente abaixo da descrição de Correlação e Regressão. O sistema carrega o formulário para inserção manual dos dados. O sistema carrega o formulário para inserção do nome das variáveis pesquisadas, dados propriamente ditos e botão calcular. O sistema valida os campos obrigatórios (todos). O sistema exibe o gráfico de dispersão, o coeficiente de correlação linear, a equação de regressão e as projeções das variáveis pesquisadas. 		
Pós-condição	Nenhuma		
Cenário Alternativo	 *a – O usuário pode sair do sistema a qualquer momento. 2a – O sistema carrega o formulário para importação dos dados. 4a – Dados obrigatórios não preenchidos: 4a.1 O sistema informa ao usuário sobre campos obrigatórios que não 		
	foram preenchidos (nome da variável e dados). 4a.2 Retorna ao passo 2 do cenário principal.		

3 Ferramentas e Métodos ou Desenvolvimento

3.1 Ferramentas

Para a execução do trabalho, as ferramentas utilizadas foram o Visual Studio Code versão 1.46.1 para escrita dos códigos propriamente ditos sob o contrato de licença do MIT disponível em https://github.com/Microsoft/vscode/blob/master/LICENSE.txt.

O desenvolvimento foi feito utilizando as linguagens HTML, CSS e JS, bem como as bibliotecas React versão 16.13.0 sob licença do MIT - Copyright (c) Facebook, Inc. e seus afiliados, Bootstrap versão 4.1.3 sob licença do MIT disponível em https://github.com/twbs/bootstrap/blob/master/LICENSE e Node versão 12.16.1 para a API.

Foi utilizado também o navegador Google Chrome versão 83.0.4103.97 para visualização parcial e testes do software e o repositório GitHub versão 2.21.1 para versionamento do software.

Foi utilizado o software Adobe Illustrator CC versão 22.0.1 e Adobe Photoshop CC versão 19.1.5 para construção das imagens exibidas na home.

Por fim, para hospedagem do software, foi utilizado a plataforma Umbler sob licença pessoal, mundial, isenta de royalties, não atribuível e não exclusiva para utilizar o software fornecido.

3.2 Métodos ou Desenvolvimento

O processo de desenvolvimento do projeto baseou-se nos materiais de estudo fornecidos pelas disciplinas de Estatística Aplicada, Engenharia de Software II, Interação Humano Computador e Linguagem de Programação.

Também tomou como referência, trabalhos de alunos anteriores, vídeos disponíveis na plataforma YouTube, materiais compartilhados em fóruns e comunidades como StackOverflow e vídeo-aulas de plataformas de ensino à distância como Allura.

O projeto foi dividido ao longo de todo o semestre e feito em etapas de acordo com os prazos estipulados no Diário de Bordo disponível na plataforma AVA.

4 Resultados e Discussão

Depois de muitos testes, o *software* mostrou uma ótima performance, atingindo os objetivos propostos e podendo servir como plataforma para docentes e discentes que se relacionem à disciplina de Estatística e afins.

O desenvolvimento do Granger promoveu a aquisição de diversos conhecimentos, tanto no âmbito teórico, quanto no prático. Além disso, também foi uma grande oportunidade para fixar e aprimorar o que já foi estudado até o presente semestre.

As principais dificuldades enfrentadas consistiram no processo de visualizar o sistema como um todo, já que os conteúdos das disciplinas foram graduais e utilização do React, que apesar de ser uma ótima biblioteca, foi algo novo a ser explorado. Porém, essas adversidades foram contornadas com organização, construção de histórias e planejamento semanal das atividades a serem resolvidas e bugs a serem corrigidos.

Considerações finais

Unir o conhecimento estatístico à tecnologia é abraçar os melhores parâmetros de investigação e proporcionar a melhor tomada de decisão nas mais diversas esferas do conhecimento.

O intuito do trabalho era desenvolver um *software* de Mini B.I. que atendesse aos conteúdos das disciplinas de Estatística, Engenharia de Software II, Interação Humano Computador e Estrutura de Dados e servisse como uma ferramenta educacional.

Com os resultados obtidos, vemos que os objetivos foram alcançados e muitas funcionalidades implementadas, contudo, é visível os diversos pontos que podem ser melhorados e aprofundados, como:

- Implementar mais tipos de análises estatísticas;
- Permitir a tradução do sistema para outros idiomas;
- Implementar um banco de dados;
- Gerar relatórios das análises e permitir seu compartilhamento;
- Aprimorar a importação de dados;
- Permitir a filtragem dos resultados obtidos;
- Implementar funções de acessibilidade.

Referências

BACH, Mirjana Pejić; ZOROJA, Jovana; ČELJO, Amer. An extension of the technology acceptance model for business intelligence systems: project management maturity perspective. **Ijispm - International Journal Of Information Systems And Project Management**, [s.l.], n. 52, p. 5-21, 2017.

GRUBLJEŁIč, Tanja; JAKLIč, Jurij. Business Intelligence Acceptance: the prominence of organizational factors. **Information Systems Management**, [s.l.], v. 32, n. 4, p. 299-315, 14 ago. 2015.

TRIEU, Van-hau. Getting value from Business Intelligence systems: a review and research agenda. **Decision Support Systems**, [s.l.], v. 93, p. 111-124, jan. 2017.