

**FACULDADE SENAC GOIÁS**

Bruno Camargo Manso

Guilherme Mello de Santana

Luys Fernnando Ribeiro Caetano Brasil

Matheus Vieira Tavares

Rodolfo Franco de Paula Silveira

**PROJETO INTEGRADOR II**

Ana Flávia Marinho de Lima Garrote

GOIÂNIA,

2019

Bruno Camargo Manso

Guilherme Mello de Santana

Luys Fernnando Ribeiro Caetano Brasil

Matheus Vieira Tavares

Rodolfo Franco de Paula Silveira

**PROJETO ANALISADOR DE ARQUIVOS EM FORMATO CSV (COMMA SEPARATED VALUES)**

Relatório apresentado como requisito parcial para obtenção de aprovação na disciplina Projeto Integrador II, no Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas, na Faculdade SENAC Goiás.

Ana Flávia Marinho de Lima Garrote

GOIÂNIA,

2019

**JUSTIFICATIVA**

Em tempos de excesso de dados e informações, saber filtrar, ou seja, isolar variáveis significativas, com o objetivo de extrair de maneira correta e rápida conteúdos sensíveis pertencentes a uma instituição, é condição essencial para sobrevivência e saúde dessas.

Estamos, de alguma forma, imersos em uma quantidade absurda de dados os quais têm amplo significado, não somente às instituições mas principalmente às pessoas que, de alguma maneira, as gerenciam.

Pessoas são recursos, e tais recursos precisam de ferramentas ágeis, mesmo que estas estejam em grande número, muitas vezes os dados apresentam características pouco inteligíveis aos olhos humanos, no entanto são facilmente compreendidas pelas máquinas. Todavia, assim como os humanos, máquinas precisam de ‘treinamentos’, treinamentos estes elaborados por humanos que aos poucos vêm elaborando linguagens e lógicas para propiciar a devida comunicação com aquelas que impreterivelmente ainda precisam de nós.

Evasões são um problema em qualquer instituição educacional, e saber o por quê é fundamental. Não somente os porquês, mas também quem foi, quando foi, em qual a área de atuação, qual curso e quais matérias estavam sendo cursadas quando houve a evasão, bem como demais dados pessoais que somam a esses volumes, como por exemplo o sexo da pessoa, seu estado civil, o turno e o mais importante: qual o motivo da evasão.

Nota-se então a necessidade do cruzamento desses dados. Ora, uma pessoa casada certamente têm seus motivos pessoais para evadir, assim como o sexo pode de repente elucidar propensão ou não à adaptabilidade de um curso, tendo em vista, por exemplo, que cursos de exatas costumam conter mais homens, isso é indiscutível. Assim como observar qual turno têm mais evasão pode reafirmar suposições que fazemos intuitivamente, tais dados serveriam então para confirmar tais suspeitas.

As máquinas têm o poder de reunir informações e tratá-las, de forma que, ajudando a elucidar tais fatos, fica bem mais fácil, para a instituição e seus recursos humanos, criarem estratégias contra as nefastas evasões que corroem todo um sistema, isso inclui não somente questões financeiras mas a própria reputação da instituição.

Por fim, vale ressaltar que o resultado desse trabalho pode ter consequências nos planos de marketing, nos treinamentos dos funcionários, na tratativa com os alunos e até mesmo o valor pago por eles, entre outros aspectos, a saber.

**OBJETIVO**

Fora proposto pela Faculdade de Tecnologia Senac e seus professores, que tratássemos um amplo espectro de informações advindas de um arquivo que é extremamente significativo à instituição. Tal arquivo está em formato csv, sigla em inglês que significa: Comma Separated Values, ou seja valores separados por vírgulas. O arquivo se assemelha às planilhas, na verdade é uma planilha, porém seus dados são separados por vírgulas de modo que, as máquinas reconheçam como tal. É então, um banco de dados, contendo informações terrivelmente importantes ao Senac, pois se referem primordialmente às evasões dos alunos.

O artefato deve ser criado com a finalidade de gerar um relatório, relatório esse que contêm dados estatísticos sobre os cruzamentos entre dados, juntamente com gráficos que ajudariam na visualização e a inteligibilidade do dilema.

**ESCOPO**

Neste momento será exposto os limites do trabalho. Essa é uma parte importante no processo de criação de um software pois devemos nos atentar aos prazos que entram em embate com as implementações que hora fazem parte dos quesitos que foram impostos pelo corpo docente da faculdade, hora sugeridos pelos participantes do grupo, tais aspectos precisam de equalizações, optamos então por processos ágeis com simplicidade e coesão.

Aqui também temos uma boa oportunidade de observar a integração entre matérias do curso de Análise e Desenvolvimento de Sistemas. A partir da recomendação de cada professor, é notório a qualidade multidisciplinar da matéria Projeto Integrador II, em especial, quando observamos o link estabelecido por cada módulo estudado dentro de cada matéria e sua aplicabilidade em um case real.

A seguir detalharemos cada aspecto referente ao que foi proposto juntamente com o que decidimos, enquanto um grupo, implementar além do esperado.

**Sobre os conteúdos de cada matéria e o que foi proposto por cada professor**

Começamos falando sobre este documento que faz parte da matéria de Engenharia de Requisitos ministrada pela Professora Ana Flávia Marinho de Lima Garrote. Foi proposto pela ilustre Professora que comentássemos os requisitos referentes ao Projeto Integrador. Atendendo então as recomendações, elaboramos contendo (logo acima) o objetivo do projeto, seu escopo, descobertas de requisitos (página XX), modelos de caso de uso (apêndice XX ) e por fim a tabela de rastreabilidade (apêndice XX ). Também adicionamos uma justificativa no início para fazer um breve resumo da ideia sobre o que se trata o projeto.

Ainda sobre a Professora Ana Flávia, porém em agora na matéria de Controle de Versão, nosso trabalho era o de versionar tanto documentos relativos ao projeto, como os códigos-fonte envolvidos produzidos em Java.

Professora Isabel Moreno, que ministra a matéria de Gestão de Pessoas sugeriu que elaborássemos um funcionograma (página XX, apêndice XX) contendo cargos e tarefas do setor de T.I. (apêndice XX ) .Também propôs que elaborássemos um programa de capacitação e treinamento de cada cargo estabelecido previamente pelo grupo que contém o levantamento de necessidades, programação e planejamento, juntamente com a execução e avaliação (apêndice XX ), por fim, tudo deve ser demonstrado em funcionograma (apêndice XX ).

Em relação aos artefatos produzidos deve ser produzido um organograma (apêndice XX ) e em seguida um novo funcionograma (apêndice XX ) com base nesse organograma, para que no fim seja produzido o desenho de cargos de tarefas do setor de T.I. (apêndice XX ).

Sobre a matéria de Sistemas em Computação devidamente ministrada pela Professora Kelly Alves Martins de Lima foi sugerido que levantássemos requisitos de Software e de Hardware capazes de produzir e processar os artefatos produzidos, isso envolve também a identificação dos Sistemas Operacionais e aplicativos necessários ao programa.

Um relatório (apêndice XX) sobre Análise de Informações de Sistemas também foi exigido com a finalidade de detalhamento do Sistema Operacional e os recursos utilizados pelo programa, quanto ao processamento e memórias utilizadas.

Partindo agora para a matéria Estatística Aplicada orientada pela Professora Luciene Nunes Ribeiro, ficamos comprometidos a tratar os dados da tabela contendo o Banco de Dados em formato csv. Um relatório contendo a Estatística Descritiva do Banco de Dados será então implementado no software, tais dados incluem: cálculos de média, moda, mediana, e frequência das colunas contidas na tabela. Tal relatório com o resultado estatístico também deve ser devidamente persistido em um arquivo .PDF e também persistido em um Banco de Dados em MySql, cumprindo o requisito que veremos a seguir.

Nosso Professor Gildenor de Sousa Amorim Cavalcante, como forma de garantir o bom cruzamento da matéria ministrada “Projeto de Banco de Dados” e o projeto desenvolvido, definiu que o relatório estatístico gerado pelo artefato desenvolvido em Java, precisaria de ser persistido em um Banco de Dados de um Servidor MySql hospedado na própria máquina.

Por fim e não menos importante, devemos seguir também todas as recomendações levantadas pelo Professor Marcelo Rodrigues Faustino, que ministra brilhantemente a matéria de Programação Orientada a Objetos, todas recomendações, por sua vez, estão diretamente ligadas aos requisitos estritamente funcionais ligadas a programação.

Os requisitos exigidos foram: implementação de Analisador de dados de extensão .CSV contendo inclusive a parte estatística; criar uma classe que contém as colunas da tabela e seus campos; uma classe chamada ‘Arquivo’ com dois métodos estáticos que irão retornar um ArrayList da leitura do arquivo .CSV e cuja finalidade será de receber tratamento estatístico e também de preenchimento das tabelas e do relatório em .PDF e em um banco de dados MySql; deve conter uma classe chamada ‘Executável’ que utilize de Interface Gráfica, esta também deve ser responsável pela manipulação estatística do arquivo juntamente com a classe principal ‘Main’ para devida inicialização do programa.

**DESCOBERTA DE REQUISITOS**

**1 - Requisitos não Funcionais**

Como já dito na anteriormente, é um software com a função de analisar dados institucionais, que funcione em arquiteturas diferentes. Deverá tratar tais dados a partir de um arquivo amplamente utilizado nessas instituições, auxiliando tanto a parte administrativa quanto a parte executiva/estratégica da mesma.

Deverá classificar, enumerar e quantificar tanto dados qualitativos como dados quantitativos dentro do arquivo proposto. Isto irá produzir, por fim, relatórios que poderão ser apresentados em reuniões de variadas ordens.

Inicialmente o software deverá suprir a necessidade de análise de dados massivos sobre evasão, porém o grupo de desenvolvedores também sentiu que este pode, de repente, ter aplicabilidade à outras instituições e também, com outros fins e não somente sobre evasão.

Em relação à linguagem de programação, o uso do JAVA é uma condição devido a sua adaptabilidade em computadores pessoais e até mesmo, quem sabe, ser utilizado em plataformas móveis como tablets e celulares, facilitando assim a apresentação dos resultados para um corpo qualquer de *stakeholders*.

Outra condição é sobre sua fluidez e facilidade, sendo bastante importante nos dias de hoje, tendo em vista que programas não devem ser lentos, devem se portar, de certa forma, como autômatos com inteligência própria, exigindo sempre o mínimo dos seus usuários. Considera-se que tenha uma interface intuitiva, fácil e *plug-and-play* com o intuito de ajudar a vida, não só do corpo de funcionários mas também dos executivos.

Então, para desenvolvimento do software fizemos observações e nos debruçamos também sobre códigos fontes no site da Oracle que continham modelos e exemplos principalmente referente a parte visual, sobre os fundamentos do Jframe aplicado a tabelas e gráficos.

**2 - Requisitos Funcionais**

O artefato criado deve ser elaborado em linguagem JAVA, para que seja aplicado na maioria das arquiteturas vigentes, incluindo sistemas Apple, Microsoft e Linux. Por isto a escolha desta linguagem, pois ela é perfeitamente utilizável nos sistemas citados. Todavia está restrito à linguagem JAVA, mas que costuma não decepcionar os desenvolvedores por conter uma gama enorme de funções e APIs que podem tratar com eficácia tipos diferentes de dados.

Então, o programa deverá reconhecer o sistema operacional vigente, com o objetivo de se auto-configurar frente aos diferentes sistemas de arquivos. Também deverá ler formatos diferentes de CSV, pois às vezes costumam, não somente serem separados por vírgulas, mas por ponto e vírgulas (;) também.

Enquanto desenvolvedores, não queremos restringir tal programa à apenas este arquivo de banco de dados em específico, evitamos então predefinir seu cabeçalho e a quantidade de colunas envolvidas, de forma que tal software consiga ler também outras planilhas. Imagine que, diferentes planilhas contêm quantidades diferentes de colunas e certamente nomes diferentes, assim tal artefato poderá ser aplicado a outro tipo de conteúdo além das evasões. Logicamente estamos restritos a esse formato .CSV, sendo assim necessária a conversão de um banco de dados qualquer para tal formato, tarefa essa relativamente fácil para pessoas que entendam o mínimo de informática.

Nosso programa deverá então: ler o arquivo .CSV; mostrar seu cabeçalho; solicitar ao usuário a escolha das colunas que deseja cruzar; depois de escolhidas, gerar em sua interface gráfica valores estatísticos de média, desvio padrão, moda, mediana e frequência sobre as colunas selecionadas; produzir histograma e gráfico de pizza; persistir um relatório em arquivo com extensão .PDF; também persistir o resultado dos dados estatísticos para formato MySQL devidamente armazenado em um servidor na própria máquina, isso para que, o relatório seja usado em outras análises futuras.

Devemos destacar, a respeito aos cruzamentos das informações em colunas, que apenas duas delas poderão ser cruzadas entre si, pois estamos consideramos os eixos X e Y dos gráficos, que obedecem apenas à duas dimensões.

**CLASSIFICAÇÃO DOS REQUISITOS**

**1- Requisitos não Funcionais**

**2 - Requisitos Funcionais**

O requisito principal, dentro de uma cadeia de processos seria a leitura correta das informações de um arquivo de formato CSV. Em seguida, podemos classificar como importante requisito a leitura dos cabeçalhos, tendo em vista que tal informação deverá ser mostrada ao usuário, e que os cabeçalhos definem o nome de cada coluna da planilha. Tal característica está diretamente ligada à uma interface gráfica amistosa que propiciará a escolha das colunas que serão cruzadas.

Com as colunas devidamente escolhidas, automaticamente dados estatísticos devem surgir e consequentemente gráficos, em seguida, um botão que terá a função de gerar um relatório, relatório esse que não deverá sobrescrever o arquivo de texto previamente criado pois, no arquivo, deverá conter múltiplos resultados de cruzamentos diferentes conforme escolhido pelo usuário.

Por último, a interface gráfica deverá conter um menu irá mostrar uma opção de gerar um relatório que deverá ser persistido em um Banco de Dados para um arquivo de banco de dados mais apropriado como o MySQL dentro do próprio computador que irá gerar processar o Software.

**MODELOS DE CASO DE USO**

<< Definir quem vai ficar com essa parte >>

**TABELA DE RASTREABILIDADE**

<< Definir quem vai ficar com essa parte >>

*Para que serve: Rastreabilidade é um conceito que pode ser considerado chave em projetos de desenvolvimento de software. Este artigo trata de seus principais conceitos*.

*Em que situação o tema é útil: Para aqueles que pretendem ter um maior controle sobre os artefatos gerados ao longo do desenvolvimento facilitando sua manutenção e análise de impacto a partir de solicitações de alteração*.