



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA  
UNIDADE ACADÊMICA DE SISTEMAS E COMPUTAÇÃO  
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**HELDER MACHADO DE LIMA**

**OBTENÇÃO DE INFORMAÇÕES PARA GUIAR  
O APRIMORAMENTO DO ENSINO BASEADO NO  
QUESTIONÁRIO DO ENADE**

**CAMPINA GRANDE - PB**

**2019**

**HELDER MACHADO DE LIMA**

**OBTENÇÃO DE INFORMAÇÕES PARA GUIAR  
O APRIMORAMENTO DO ENSINO BASEADO NO  
QUESTIONÁRIO DO ENADE**

**Trabalho de Conclusão Curso  
apresentado ao Curso Bacharelado em  
Ciência da Computação do Centro de  
Engenharia Elétrica e Informática da  
Universidade Federal de Campina  
Grande, como requisito parcial para  
obtenção do título de Bacharel em Ciência  
da Computação.**

**Orientador: Professor Dr. Nazareno Andrade.**

**CAMPINA GRANDE - PB**

**2019**



L732o Lima, Helder Machado de.  
Obtenção de informações para guiar o aprimoramento do ensino baseado no questionário do Enade. / Helder Machado de Lima. - 2019.

10 f.

Orientador: Prof. Dr. Nazareno Andrade.

Trabalho de Conclusão de Curso - Artigo (Curso de Bacharelado em Ciência da Computação) - Universidade Federal de Campina Grande; Centro de Engenharia Elétrica e Informática.

1. Exame Nacional de Desempenho de Estudantes - Enade. 2. Ciência de Dados - técnicas. 3. Dados públicos. 4. Limpeza de dados. 5. Dados - filtragem e armazenamento. 6. Desenvolvimento de ferramenta. 7. Opinião de estudantes universitários. I. Andrade, Nazareno. II. Título.

CDU:004(045)

**Elaboração da Ficha Catalográfica:**

Johnny Rodrigues Barbosa  
Bibliotecário-Documentalista  
CRB-15/626

**HELDER MACHADO DE LIMA**

**OBTENÇÃO DE INFORMAÇÕES PARA GUIAR  
O APRIMORAMENTO DO ENSINO BASEADO NO  
QUESTIONÁRIO DO ENADE**

**Trabalho de Conclusão Curso  
apresentado ao Curso Bacharelado em  
Ciência da Computação do Centro de  
Engenharia Elétrica e Informática da  
Universidade Federal de Campina  
Grande, como requisito parcial para  
obtenção do título de Bacharel em Ciência  
da Computação.**

**BANCA EXAMINADORA:**

**Professor Dr. Nazareno Andrade  
Orientador – UASC/CEEI/UFCG**

**Professor Dr. Dalton Dario Serey Guerrero  
Examinador – UASC/CEEI/UFCG**

**Professor Dr. Tiago Lima Massoni  
Examinador – UASC/CEEI/UFCG**

**Trabalho aprovado em: 25 de novembro 2019.**

**CAMPINA GRANDE - PB**

# Obtenção de informações para guiar o aprimoramento do ensino baseado no questionário do enade

## Trabalho de Conclusão de Curso

Helder Machado de Lima

Universidade Federal de Campina Grande

Campina Grande, Paraíba, Brasil

helder.machado.lima@ccc.ufcg.edu.br

Nazareno Andrade

Laboratório de Sistemas Distribuídos

Universidade Federal de Campina Grande  
Campina Grande, Paraíba, Brasil

nazareno@computacao.ufcg.edu.br

### Resumo

O contato aluno-professor muitas vezes não é da forma que ambos gostariam, e ao mesmo tempo, fazer comentários positivos ou negativos sobre uma disciplina, curso ou a universidade em geral tende a ser uma grande barreira para os alunos. A falta de conhecimento sobre as necessidades/opiniões dos estudantes, por sua vez, dificulta a implementação de melhorias por parte dos professores que atendam necessidades dos alunos.

Nesse contexto, este projeto propõe o desenvolvimento de uma ferramenta que auxiliará professores a analisar a opinião dos próprios alunos sobre seus cursos a fim de encontrar respostas que auxiliem na melhoria do ensino. Para isso, será usado o resultado do Exame Nacional de Desempenho dos Estudantes (Enade), que é não só uma ferramenta avaliadora dos estudantes mas também um mecanismo de coleta de informações e opiniões dos estudantes e seu ambiente universitário.

Utilizamos técnicas de visualização e ciência de dados para gerar gráficos auto-descritivos e quadros comparativos das respostas dos alunos em categorias como: ensino, aprendizado além do técnico, atuação dos professores, curso em geral e infraestrutura. Espera-se que seja possível identificar o que mudou nas avaliações dos estudantes durante as últimas edições do exame, apontando de forma clara melhoras, piores e pontos não observados, informando esforços de aprimoramento de ensino e em outras categorias para quem utilizar a ferramenta.

### 1. Introdução

O Enade é um exame obrigatório aplicado em todas as áreas de conhecimento do ensino superior Brasileiro, implementado em 2004 como ferramenta de avaliação dos estudantes em relação a conteúdos programáticos, habilidades e competências adquiridas durante a formação dos estudantes (INEP, 2019).

Um ponto importante do exame é o Questionário do Estudante, um instrumento de coleta de informações para construção de um perfil socioeconômico e uma forma de adquirir dados sobre o processo de formação dos estudantes de ensino superior (INEP, 2019). As 68 questões encontradas neste questionário são fonte de um potencial aparentemente inexplorado quando o assunto é entender as necessidades dos alunos ou até mesmo saber o que é possível melhorar em cursos e na própria universidade na visão dos seus estudantes.

*Os autores retêm os direitos, ao abrigo de uma licença Creative Commons Atribuição CC BY, sobre todo o conteúdo deste artigo (incluindo todos os elementos que possam conter, tais como figuras, desenhos, tabelas), bem como sobre todos os materiais produzidos pelos autores que estejam relacionados ao trabalho relatado e que estejam referenciados no artigo (tais como códigos fonte e bases de dados). Essa licença permite que outros distribuam, adaptem e evoluam seu trabalho, mesmo comercialmente, desde que os autores sejam creditados pela criação original.*

Com isto, as hipóteses deste trabalho surgiram diante de uma experiência pessoal e empírica, sendo constituída através do contato com os colegas. Porquanto, pude perceber que os métodos tradicionais de ensino tais como: apresentações de slides, aulas expositivas tradicionais etc.; não são mais efetivas como foram outrora. Muitas vezes subutilizando os recursos que a universidade dispõe.

Nesse contexto, é notório que o ambiente universitário sofreu grandes alterações na última década, mas será que as necessidades dos alunos foram atendidas ou ajustadas nesse tempo? Houve adequação no modo de ensino para novas tecnologias? As opiniões dos alunos foram levadas em consideração para ajustes no ensino? O método de ensino se adequou às necessidades? A estrutura do curso melhorou? Nas atividades fora de sala, houve melhoria?

Na ausência de estudos científicos que ajudem a responder estas perguntas, ou que, proporcionem uma melhor visualização e compreensão dos dados, fez-se necessário uma ferramenta que possa auxiliar o aprimoramento do ensino, fornecendo dados para quem precisa destas informações: os professores e gestores. Também é importante salientar que esse processo foi baseado nas opiniões dos mais interessados em melhorias e adequações, os alunos.

Pensando nisso, este trabalho combina a obtenção e extração de informações do questionário com um exemplo prático de utilização dos dados, uma ferramenta de visualização foi implementada para permitir demonstrar possíveis técnicas para auxiliar os professores, gestores e alunos a entender o ponto de vista dos estudantes sobre seus cursos através do ENADE. Essa ferramenta baseando-se na ampla visão do ambiente universitário pelos olhos dos estudantes, proporcionando adequações, sugestões ou correções nos métodos de ensino baseado nas respostas dos concluintes/ingressantes.

### 2. Trabalhos Relacionados

Ao que tange a parte de Libertação de dados e Qualidade dos dados o [brasil.io](http://brasil.io) é um portal que obtém, limpa, transforma e disponibiliza dados abertos do governo em formato estruturado e não proprietário. O site conta com algumas bases de dados mas ainda não conta nada sobre o ENADE.

Uma análise que inspirou a realização deste trabalho foi feita por Andrade (2018), também disponível na plataforma

observablehq.com, onde o mesmo destaca pontos positivos e negativos em cada um dos cinco quesitos selecionados por ele a partir das respostas dos alunos ao questionário do ENADE, totalizando 40 questões de 67 concluintes do ano de 2017. Uma análise importante, já que foi feita por um alvo que esta ferramenta pretende auxiliar.

### 3. Solução

#### 3.1 Verificação e Limpeza dos dados

Para dados públicos não existe nenhum tipo de padrão, podendo eles serem disponibilizados em qualquer formato, *CSV - Comma Separated Values* (Valores Separados por Vírgulas), *TXT* - Arquivo de Texto, *PDF - Portable Document Format*, *XLS - formato do Microsoft Excel* ou até mesmo em formato proprietário. Muitos vêm com um manual de como manipulá-los, ou seja, para utilização dos dados quase sempre será necessário algum tipo de adequação para obtenção de qualquer tipo de informação dos mesmos.

O processo de verificação e limpeza dos dados é uma etapa longa, o que dificulta a observação imediata. Durante o desenvolvimento desta solução, o procedimento de verificação e limpeza durou cerca de 2 meses até que fosse possível fazer as primeiras investigações. Segundo Álvaro Justen (2019), este processo demanda entre 75% a 80% do tempo de uma análise.

Este método foi desenvolvido utilizando *Python* e *Pandas*, a combinação se mostrou bastante eficaz nesta etapa, pois o grande volume de informação dificulta qualquer tentativa de manipulação de dados com algum tipo de problema. No andamento deste trabalho me deparei com uma vírgula fora de contexto nos dados por volta da linha 210.000, que só foi possível de contornar utilizando este conjunto recursos. Um problema como este, se não for contornado corretamente pode comprometer as informações a partir do ponto ocorrido.

#### 3.2 Liberação dos dados

Ainda, segundo Justen, existem 4 problemas de acessibilidade dos dados públicos: Formato, Dispersão, Quantidade de dados e Domínio. Uma das soluções para este problema chama-se Liberação de dados, define-se por tornar os dados menos dispersos e em um formato aberto mais acessível. Este não era o objetivo inicial deste trabalho, mas foi de longe o mais importante alcançado. Os dados públicos do ENADE são encontrados em um formato de difícil interpretação em um arquivo *TXT*, necessitando de múltiplos arquivos auxiliares para que depois de algumas transformações ele esteja passível de exploração. Foi possível libertar até o momento os dados dos anos de 2014, 2015, 2016 e 2017. Uma decisão tomada durante este processo foi limitar o escopo para os anos de 2014 e 2017, e analisar somente os cursos da própria universidade que fizeram o exame nestes anos.

Para o processo de limpeza foi utilizadas a biblioteca *Python Data Analysis Library (Pandas)* do *python*, que se mostrou robusta para leitura, conversão e exportação dos dados.

#### 3.3 Armazenamento e Filtragem dos dados

Depois de libertar os dados foi a vez armazená-los, buscando mais opções de importação resolvi salvá-los em *CSVs* e tabelas *MySQL*,

o que facilitou bastante o processo de migração de sistemas operacionais e de verificação da qualidade dos dados. Visto que com a ferramenta *MySQL Workbench* torna-se possível visualizar as colunas, tipos de cada coluna, identificar erros e fazer consultas nos dados.

Para a filtragem foram utilizadas a biblioteca *Pandas*. Neste tipo de biblioteca os arquivos de textos possuem mais de 250 mil linhas e mais de 120 colunas, ou seja, editores de texto mais simples e bibliotecas menos poderosas não conseguem lidar com este tamanho de dados tão facilmente. Para tanto, este processo se deu de duas formas: 1) inicialmente para os dados de 2015, 2016 e 2017 utilizou-se o próprio *R* para importar os dados das tabelas *MySQL*, filtrá-los e depois salvá-las em *CSVs* para reutilizar nas visualizações; 2) já nos dados de 2014 o filtro foi feito com consultas nas tabelas e depois salvando da mesma forma, pois ao utilizar o mesmo processo dos anos anteriores algum lixo gerado na exportação fez com que fosse necessário uma abordagem diferente, além do tempo de limpeza e verificação que foi mais curto que o dos demais. Vale salientar que para os microdados do ENADE nem sempre é bom remover as linhas, pois mesmo que alguma coluna tenha ficado sem resposta ainda é possível reaproveitá-la. Caso contrário isso diminuiria bastante o volume de dados, inviabilizando a análise.

#### 3.4 Ferramenta interativa e Relatório estático

Usando a plataforma *Observablehq.com* foi criada uma visualização interativa utilizando a biblioteca *Vega Lite* onde é possível visualizar as categorias: Ensino apoio e avaliação; Aprendizagem além do técnico; Curso em geral; Infraestrutura etc. Dois tipos de visualização estão disponíveis no momento: Média das respostas e *Rank* do curso. Em cada linha do gráfico é possível ver informações sobre as perguntas de cada categoria e suas respostas para os cursos que fizeram o ENADE nos anos em questão (2014 e 2017), além de ano de realização, média do curso, rank, diferença em relação a média e a média para todos os cursos no ano, além de ser possível destacar o curso desejado.

Um relatório estático comparando computação com os outros cursos em categorias distintas também foi gerado, utilizando *R* e hospedando no *RPubs*. Para este caso, o confronto foi feito explorando as proporções de cada valor que foi respondido para computação em relação a proporção da média dos demais. Rank e média não estão visivelmente disponíveis neste modelo.

#### 3.5 Análises

Este trabalho gerou duas análises, uma interativa encontrada no *observable.com*<sup>1</sup> e uma outra visão disponibilizada no *RPubs*<sup>2</sup>, ambas contendo quadros comparativos do curso de Ciência da Computação em relação aos demais cursos para os mesmos anos de 2014 e 2017 na UFCG.

Várias divisões do formulário do ENADE foram contempladas pelas duas análises, separamos em categorias distintas para cada estudo, mas para questões de comparação, uma categoria foi escolhida para demonstração.

<sup>1</sup> Acessar: <https://observablehq.com/d/a8b6afb13c694549>

<sup>2</sup> Acessar: <http://rpubs.com/helderm/547406>

### 3.6 Complementos

Para complementar este trabalho, disponibilizei uma pasta<sup>3</sup> contendo um *dump MySQL*, com as tabelas referentes aos dados de 2014, 2015, 2016, 2017, posteriormente irei adicionar no mesmo diretório dados de outros anos. Também criei um site<sup>4</sup> contendo informações adicionais e as outras categorias do questionário passíveis de análise.

## 4. Exemplos de análises

As imagens encontradas nesta seção comparam o Curso de Ciência da Computação da UFCG com os demais cursos da universidade. Para exemplificação apenas as respostas para as perguntas do questionário que abrangem o tópico “Ensino” estão destrinchadas abaixo. As respostas em amarelo representam os demais cursos, as respostas em outras cores são os valores de computação. Um quadro comparativo para todas as categorias pode ser encontrado acessando o *RPUBS*. As figuras mostram no eixo X os valores das respostas e no eixo Y a proporção das respostas, ex: 0.3 = 30%.

### 4.1 Imagens 2014 - “Sobre o Ensino”

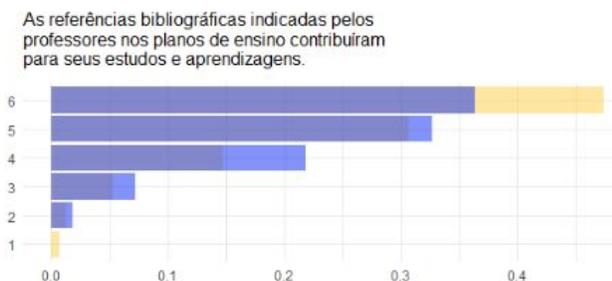


Figura 1: Pergunta 1, categoria “Sobre o Ensino em 2014”.

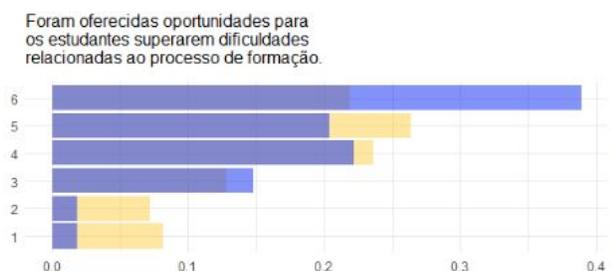


Figura 2: Pergunta 2, categoria “Sobre o Ensino em 2014”.



Figura 3: Pergunta 3, categoria “Sobre o Ensino em 2014”.

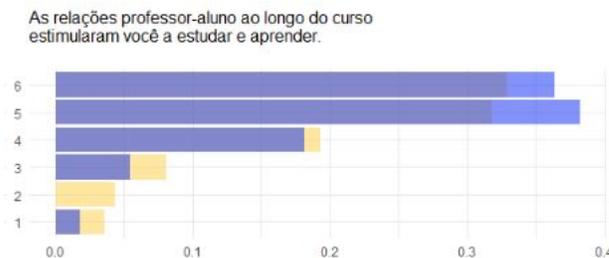


Figura 4: Pergunta 4, categoria “Sobre o Ensino em 2014”.

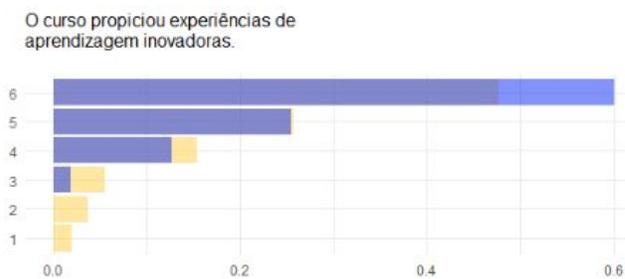


Figura 5: Pergunta 5, categoria “Sobre o Ensino em 2014”.

#### 4.1.1 Observações 2014

Verificando as respostas dos concluintes para 2014 podemos inferir que para a pergunta 1 “As referências bibliográficas indicadas pelos professores nos planos de ensino contribuíram para seus estudos e aprendizagens” o curso de computação apresentou a maior distância entre os demais cursos para a nota máxima de avaliação, algo que posteriormente poderia ser melhor observado.

Nas demais perguntas computação se destaca bem, quase sempre liderando o gráfico nas melhores respostas. Como na pergunta 2 “Foram oferecidas oportunidades para os estudantes superarem dificuldades relacionadas ao processo de formação”, em que computação obteve quase o dobro de respostas com avaliação máxima em relação aos demais.

### 4.2 Imagens 2017 - “Sobre o Ensino”

<sup>3</sup> Acessar: [Dados Enade](#)

<sup>4</sup> Acessar: [Site Complementar TCC](#)

As referências bibliográficas indicadas pelos professores nos planos de ensino contribuíram para seus estudos e aprendizagens.

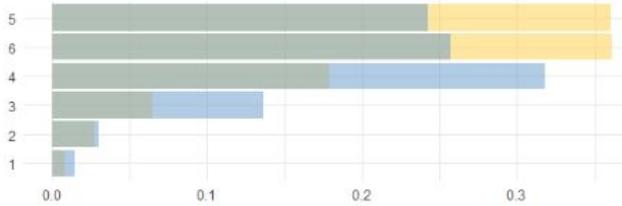


Figura 6: Pergunta 1, categoria "Sobre o Ensino em 2017".

Foram oferecidas oportunidades para os estudantes superarem dificuldades relacionadas ao processo de formação.

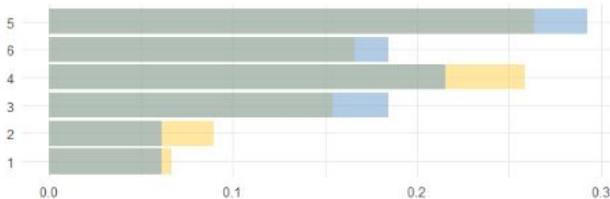


Figura 7: Pergunta 2, categoria "Sobre o Ensino em 2017".

Os planos de ensino apresentados pelos professores contribuíram para o desenvolvimento das atividades acadêmicas e para seus estudos.

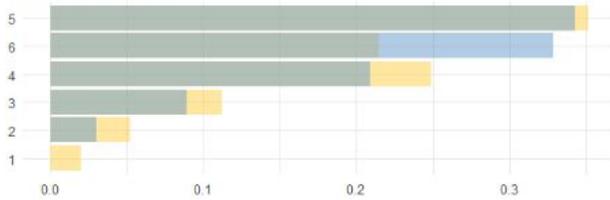


Figura 8: Pergunta 3, categoria "Sobre o Ensino em 2017".

As relações professor-aluno ao longo do curso estimularam você a estudar e aprender.

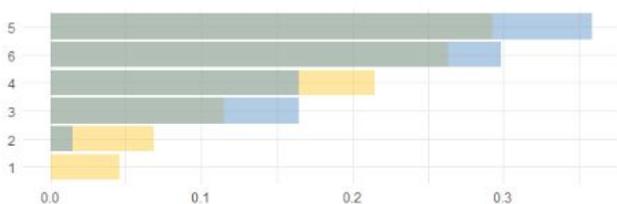


Figura 9: Pergunta 4, categoria "Sobre o Ensino em 2017".

O curso propiciou experiências de aprendizagem inovadoras.

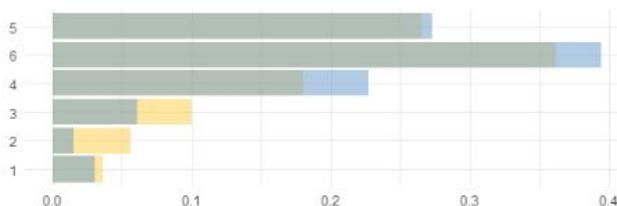


Figura 10: Pergunta 5, categoria "Sobre o Ensino em 2017".

#### 4.2.1 Observações 2017

Ao se observar as respostas para o ano em questão é possível inferir que computação se destacou negativamente na pergunta 1, "As referências bibliográficas indicadas pelos professores nos planos de ensino contribuíram para seus estudos e aprendizagens", onde é possível perceber que os cursos em geral foram bem melhores nesse quesito.

Nas demais perguntas computação tem um número menor de notas altas (5 e 6), mas se mantém em um bom patamar de respostas.

#### 4.3 Comparativo 2014 - 2017

##### 4.3.1 As referências bibliográficas indicadas pelos professores nos planos de ensino contribuíram para seus estudos e aprendizagens:

- Computação mostrou uma regressão considerável, o restante dos cursos teve uma regressão leve, mas perceptível.

##### 4.3.2 Foram oferecidas oportunidades para os estudantes superarem dificuldades relacionadas ao processo de formação:

- Computação teve uma regressão em torno de 10% para a nota máxima, o restante também regrediu de forma moderada.

##### 4.3.3 Os planos de ensino apresentados pelos professores contribuíram para o desenvolvimento das atividades acadêmicas e para seus estudos:

- Todos mostraram um pequeno retrocesso aqui.

##### 4.3.4 As relações professor-aluno ao longo do curso estimularam você a estudar e aprender:

- Todos os cursos se mantiveram sem grandes mudanças neste quesito.

##### 4.3.5 O curso propiciou experiências de aprendizagem inovadoras :

- Neste ponto computação mostrou um retrocesso significativo na nota máxima, -20%, já os outros cursos mostraram uma melhora significativa.

#### 4.4 Visualização Observablehq.com

Filtros:

Visualizar a média das respostas ▼

Destacar: Arquitetura e Urbanismo (1129147) ▼

Figura 11: Opções para filtragem, média ou rank e destacar um curso.

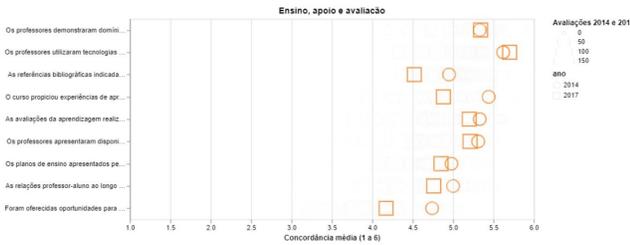


Figura 12: Informações sobre o curso quando selecionado no filtro e ao colocar o mouse sobre o valor no gráfico. Quanto maior a forma geométrica, mais respostas aquele curso teve. Círculos representam o ano de 2014 e Quadrados o de 2017.



Figura 13: Informações sobre o curso de Ciência da Computação quando selecionado no filtro de média e ao colocar o mouse sobre o valor no gráfico.

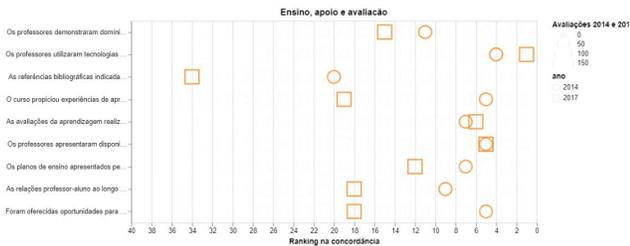


Figura 14: Informações sobre o curso de Ciência da Computação quando selecionado no filtro de rank.

#### 4.4.1 Análise Observable

Neste tipo de visualização, com a auxílio da legenda, do destaque de um curso e das informações adicionais, exibidas ao posicionarmos o mouse sobre as formas geométricas, é possível extrair informações e gerar conclusões para cada questão de forma rápida e precisa. Assim, é possível inferir que para as perguntas, 1, 3, 4, 7, 8 e 9 computação desceu no ranking dos cursos destas questões. Em relação à média, as respostas para 3, 4, 5, 6, 7, 8 e 9 houve um decréscimo na pontuação. Sendo possível também comparar os quesitos de forma separada, possibilitando outras deduções.

#### 4.5 Diferenças Visuais

Ao compararmos as imagens do instrumento de auxílio com as figuras de 1 à 10, é possível notar uma grande diferença de imediato. O processo de inferência demanda mais tempo, paciência e cuidado. Estas diferenças visuais da ferramenta em relação aos gráficos de barra tradicionais, facilitam o entendimento, a extração de informações e características, permitindo uma rápida comparação em cada pergunta. Desta forma, qualquer pessoa pode inferir sobre os dados de forma prática, tirando suas próprias conclusões sem a necessidade de comparar múltiplos gráficos, tornando este tipo de análise um procedimento mais prazeroso.

## 5. Experiência

A experiência de desenvolvimento deste trabalho foi um processo de aprendizado e superação constante, cada passo foi executado utilizando diversas técnicas adquiridas e lapidadas ao longo da minha graduação em Ciência da Computação. Começando com o aprendizado assimilado em Programação 1, que foi extremamente útil para trabalhar com as diversas bibliotecas que manipulam *Big Data*.

Pensando em como manter as informações/dados obtidos no processo, os conceitos vistos em *Banco de Dados* foram utilizados para armazenar, consultar, verificar e exportar todo o material gerado neste trabalho.

Também foram úteis técnicas diversas e capacidades desenvolvidas nos meus estágios no Embedded e na Signove, que me proporcionaram durante toda minha graduação a oportunidade de melhorar minhas *skills* em diversas áreas da computação.

Culminando na grata experiência em Ciência de Dados Descritiva, onde surgiu a ideia deste trabalho, formas de visualização, extração de informação e utilização de estatística foram bastante exploradas, contemplando a parte de análise aqui executada.

### 5.1 Processo de desenvolvimento

O processo de desenvolvimento foi definido por mim com o aval do Prof. Dr. Nazareno Andrade, orientador deste trabalho. Inicialmente, definimos quais dados utilizaríamos e os problemas que gostaríamos de resolver com esta análise. Após esta definição começou a parte de análise de projeções feitas a partir dos dados disponibilizados pelo professor Nazareno. Esta solução foi necessária para um melhor entendimento dos elementos desse projeto, para então e avaliar o que precisaria ser feito para dar início ao desenvolvimento.

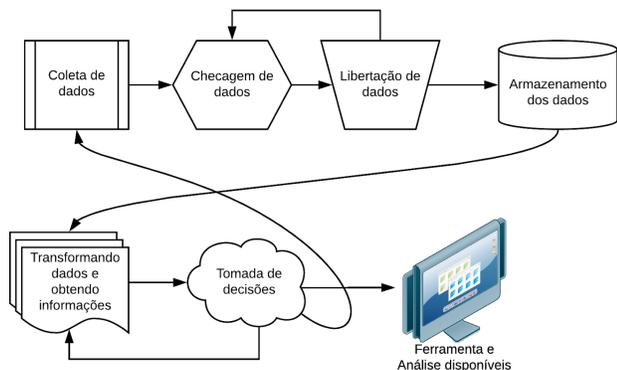


Figura 15: Etapas realizadas no processo de concepção da ferramenta.

## 6. Desafios e suas soluções

### 6.1 Limitações

Uma limitação encontrada durante o desenvolvimento foi a obtenção de informações sobre os cursos, cidades, regiões e universidades para que fosse possível transformar o dados em informação. O MEC (Ministério da Educação) mantém um portal online chamado e-mec onde é possível fazer consultas sobre instituições. Contudo, é praticamente inviável obter dados sobre todos os cursos e instituições pois o trabalho manual demandaria tempo demasiado, o que não era uma opção. Esta limitação foi crucial para a redução do escopo da ferramenta.

### 6.2 Processo

O primeiro desafio começou no dia seguinte às primeiras definições do projeto, visto que, já existia um código que fazia o tratamento dos microdados do Enade. Tentei utilizar o mesmo para extrair dados dos anos de 2015, 2016 e os mais atuais de 2017, sem sucesso por um longo período de tempo. A ferramenta *Tidyverse* se mostrou ineficiente na leitura de dados “sujos”, dados sem nenhum tratamento, da forma que foram disponibilizados pelo governo. Por não haver padrão definido para leitura dos dados, nem a utilização do código sugerido no *leia-me* dos dados foi eficiente para tal.

Diante disso, demorei alguns meses para conseguir terminar o tratamento dos dados, já que depois de algum tempo os dados foram atualizadas por parte do MEC e o *parser* não funcionava corretamente. Não consegui encontrar ajuda em nenhum fórum para a solução dos problemas encontrados. Para contornar estas mudanças fiz alguns *scripts* em *Python* na tentativa de ler os *TXTs* que continham os dados brutos e transformá-los em *CSVs* para começar as transformações. Depois de algumas tentativas foi possível observar alguns problemas nos dados como: falta de aspas nos campos de texto, quebra incorreta de linha no meio dos dados, ponto flutuante e valores nulos possuíam o mesmo caracter na definição, o ponto.

A partir deste movimento, descobri que os dados estavam muito mal formatados e nem ferramentas conhecidas para transformar *TXTs* em *CSVs* conseguiam ler e converter os dados de forma simples. Por fim, consegui fazer estas conversões utilizando a

biblioteca *PANDAS* do *Python* e depois de análises minuciosas, a mudança de sistema operacional se fez necessária, pois auxiliou no *backup* de meus dados em banco de dados *MySQL* e resolver possíveis problemas de incompatibilidade que foram apresentados no *Xubuntu* (*Ubuntu 16.04 based*), foi uma solução satisfatória.

Assim, foi apenas depois de tais mudanças que consegui contornar os problemas com o *R* e com o meu sistema operacional anterior. Por fim, resolvi salvar os dados em dois formatos: *CSV* e tabelas *MySQL*, para facilitar o meu reuso e tornar viável a disponibilidade por minha parte destes dados e a rápida análise e utilização deste dados a posteriori.

## 7. Conclusões e Trabalhos Futuros

Conforme a seção 4 deste trabalho, fica perceptível que o produto concebido pode, sem dúvida, auxiliar em comparações nas diversas áreas que formam as universidades, possibilitando um amplo campo de pesquisas na área. Os dados fornecidos ao final desta jornada estarão disponíveis para reutilização em uma qualidade superior ao que nos foi proporcionado. Além disso, as técnicas utilizadas no desenvolvimento também estarão acessíveis para auxiliar qualquer um que queira trabalhar com o ENADE no futuro.

Com a disponibilização dos dados em uma qualidade superior, um leque de possibilidades para trabalhos futuros se abre, cada uma das áreas: Ensino apoio e avaliação, Aprendizagem além do técnico, Curso em geral e Infraestrutura pode ser explorada tanto para melhorias dentro da ferramenta como para análises/pesquisas mais detalhadas sobre as causas das variações em cada quesito.

- Melhorias nos filtros
- Número maior de opções de seleção
- Novos parâmetros
- Mais visualizações
- Visualizações dinâmicas
- Mais cursos
- Mais anos

Estas são algumas opções possíveis para incorporar na ferramenta do *observable*, ou até mesmo em outra plataforma. Como exemplo, poderíamos gerar visualizações para todos os cursos e ao selecionar um curso mostrar todos os anos em que ele foi selecionado para participar do ENADE, contendo um gráfico para cada pergunta e mostrando a evolução/regressão para cada questionamento.

Ou até pensar em algo mais ambicioso, no qual várias universidades e seus cursos, comparando cursos de diversas universidades entre vários anos disponíveis, campus, regiões, etc.

## 8. Referência

- [1] ANDRADE, Nazareno. Computação@UFCG no questionário do ENADE 2017. Campina Grande, 2018. Disponível em: <https://observablehq.com/@nazareno/a-computacao-ufcg-no-questionario-do-enade-2017>. Acesso em: 11 nov. 2019.

- [2] BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Enade. Brasília, DF, 2019
- [3] JUSTEN, Álvaro. BRASIL.IO: Libertação de Dados Públicos, com Python. [S. l.], 2019. Disponível em: <https://www.youtube.com/watch?v=2NLJfDycbGg>. Acesso em: 12 jun. 2019.
- [4] Lima, Priscila. Análise das Provas do Enade Agrupadas por Tema: Um Estudo de Caso para Estudantes de Ciência da Computação . Goiânia, 2018. 94p. Dissertação de Mestrado. Instituto de Informática, Universidade Federal de Goiás. Disponível em: [http://www.ie.uftj.br/images/grupo\\_cede/publica%C3%A7%C3%B5es/site\\_antigo/tds/td73\\_9438c.pdf](http://www.ie.uftj.br/images/grupo_cede/publica%C3%A7%C3%B5es/site_antigo/tds/td73_9438c.pdf). Acesso em: 12 jun. 2019.