

# UMA PROPOSTA DE BUSINESS INTELLIGENCE PARA ANÁLISE DO PERFIL DAS PESSOAS PRIVADAS DE LIBERDADE NO BRASIL

Victor Domingos Duque<sup>1</sup>, Emerson Augusto Priamo Moraes<sup>2</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais -  
Campus Juiz de Fora – Juiz de Fora, MG - Brasil

victor.duque.domingos@gmailcom, emerson.moraes@ifsudestemg.edu.br

***Abstract.** With the increase in the prison population over the years in Brazil, the information related to this has also increased. In this study, an analysis made with Microsoft Power BI will be presented using the data collected from a test aimed at people deprived of their liberty called the National Examination for the Certification of Youth and Adult Skills and the National Penitentiary Department. For this, a database was built and later integrated with a Business Intelligence tool. The available results, as the education level of people in this situation is still precarious, with a high rate of prisoners with incomplete basic education.*

***Resumo.** Com o aumento da população carcerária nos últimos anos no Brasil, aumentou-se também as informações que estão relacionadas a isso. Neste estudo, será apresentada uma análise feita com o Microsoft Power BI utilizando os dados coletados de uma avaliação voltada para pessoas privadas de liberdade, chamada de Exame Nacional para Certificação de Competências de Jovens e Adultos e do Departamento Penitenciário Nacional. Para isso, foi construído um banco de dados e posteriormente foi integrado com a ferramenta de Business Intelligence. Os resultados mostraram como ainda é precário o grau de escolaridade de pessoas nesta situação, possuindo um alto índice de prisioneiros com o ensino básico incompleto.*

## 1. Introdução

O Brasil apresenta números que o colocam em uma lista de países que mais possuem pessoas encarceradas. De acordo com a matéria publicada no site da emissora Globo [Velasco et al., 2020], o país possui 338 encarcerados para cada 100 mil habitantes, com estas estatísticas o país fica na 26<sup>a</sup> posição no mundo de um total de 222 analisados, em relação a número absoluto, o Brasil ocupa a 3<sup>a</sup> posição, ficando antes apenas de China e Estados Unidos. No que se refere a aprisionamento por estados, o Acre fica na 1<sup>a</sup> posição logo a frente do Mato Grosso do Sul.

De acordo com os dados e estatísticas levantadas por [Silva, 2017], em um estudo feito pelo Departamento Penitenciário Nacional (DEPEN), notamos que a escolaridade de pessoas privadas de liberdade no Brasil é muito baixa, possuindo apenas 9% da população carcerária, o Ensino Médio Completo. Este fato gera uma grande barreira para a reinserção social dos indivíduos que terminaram de cumprir suas penas,

uma vez que os mesmos não apresentam um grau de escolaridade mínima para exercerem grande parte das profissões, tendo que recorrer por empregos mais desvalorizados ou formas ilícitas de conseguirem ganhar dinheiro para conseguir sobreviver.

O ser humano privado da liberdade, como qualquer outro, possui o direito a educação. De acordo com [Graciano, 2005], o direito a educação possibilita e potencializa a garantir os demais direitos, o mesmo pode ser chamado de direito de síntese.

Diante disso, qual o nível de escolaridade dos cidadãos privados de liberdade no Brasil e qual impacto da prova do ENCCEJA na vida dos mesmos? Para responde-la, com base neste trabalho, encontramos no sistema prisional brasileiro dados que podem ser processados e analisados, como por exemplo a extração dos dados do Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA) [Agosto, 2017], que foi uma das fontes de conhecimento utilizadas neste trabalho. Este exame foi criado pelo Instituto Nacional de Ensino e Pesquisas Educacionais Anísio Teixeira, o mesmo possui como finalidade, a avaliação da população que voltou a estudar, por não ter conseguido completar o ensino básico, com o intuito de verificar e reconhecer a aprendizagem de todas as faixas etárias do Brasil.

A partir do ano de 2010, além do exame regular, o ENCCEJA começou a disponibilizar a prova para Pessoas Privados de Liberdade (PPL), incluídas neste grupo os adultos, que estão cumprindo pena nas penitenciárias e menores em situação de cumprimento de medidas socioeducativas no Brasil e no exterior. Este exame se mostrou extremamente importante para os mesmos, uma vez que eles possuem o benefício de conseguir remissão de pena e ajudando na reinserção social. O ENCCEJA PPL no Brasil é executado com o apoio dos órgãos estaduais que são responsáveis pela administração das unidades prisionais.

A motivação deste trabalho vem a partir destas constatações, pretendendo analisar o grau de ensino destas pessoas e as disciplinas que as mesmas possuem mais dificuldades, utilizando um software analítico, que consegue gerar informações com estes dados, através de análises em gráficos e estatísticas. Estas análises têm o intuito de ajudar os gestores de unidades prisionais a direcionar melhor políticas públicas, seus métodos de trabalho e materiais didáticos que os professores irão lecionar, auxiliando na formação dos indivíduos no período de cumprimento da pena. Atualmente, a reinserção social apresenta elevadas taxas de insucesso e é um tema de grande relevância na sociedade brasileira.

## **2. Referencial Teórico**

O objetivo deste capítulo é apresentar de maneira concreta os fundamentos teóricos e conceituais que são referentes ao tema deste trabalho. Os conteúdos que serão descritos posteriormente são ligados ao propósito da pesquisa, como por exemplo: Data warehouse, Modelagem Multidimensional e Power BI.

### **2.1. Modelagem Relacional**

Em 1970, um pesquisador da IBM, Edgar Frank Codd, difundiu os bancos de dados relacionais. De acordo com [Navathe, 2011], o que constitui uma modelagem relacional

é uma coleção de relações. Estas relações são similares a uma tabela de valores, em que cada linha desta tabela reflete uma coleção de valores associados. Os valores de cada linha de uma tabela são separados por colunas que possuem nomes, o que auxilia no entendimento.

Os bancos de dados deste modelo comportam processamento de transações on-line (OLTP), que contêm exclusões, inserções e atualizações, da mesma forma que possuem requisitos de informações sobre consultas. O mesmo permite a diminuição de vários problemas aos dados com o uso das regras de normalização.

Com o intuito de fazer com que o armazenamento e o acesso a dados em disco sejam mais simples, este modelo é gerenciado pelo Sistema Gerenciador de Banco de Dados (SGBD), trazendo diversos benefícios para a utilização de um grande banco de dados acessado por vários usuários.

Conforme [Ariolo, 2014] “SGBDRs são usados para armazenar a informação requerida por aplicações construídas usando tecnologias procedurais, tais como COBOL ou FORTRAN, tecnologias orientadas a objeto tais como Java e C# e tecnologias baseadas em componentes como Visual Basic”.

Em relação a organização dos dados, os usuários de um determinado sistema visualiza os mesmos em forma de tabelas, no qual essas tabelas possuem restrições de integridades, onde podem ser operados posteriormente a serem inseridos em nestas tabelas, conforme a figura 1, que representa um banco de dados relacional com as tabelas: pessoa, cliente e fornecedor.

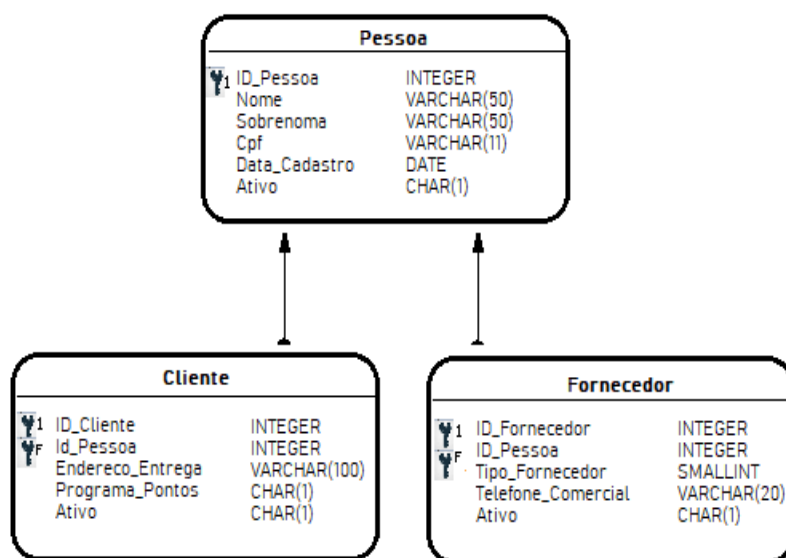


Figura 1. Modelagem Relacional. Fonte: O autor.

## 2.2. Modelagem Multidimensional

Conforme [Elias, 2014], uma modelagem multidimensional constitui-se como uma técnica de modelagem de banco de dados que auxilia consultas em um Data Warehouse.

Esta visão auxilia no processamento analítico dos dados pelas ferramentas OLAP (Online Analytical Processing).

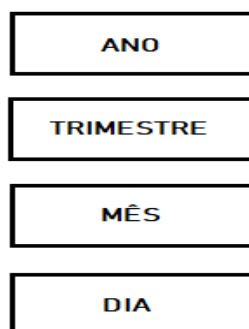
Esta modelagem é usada principalmente para resumir e reestruturar dados e mostrá-los de forma que seja possível a análise dos valores desses dados.

De acordo com [Machado, 2008], a modelagem relacional é uma forma de ilustrar o que desejamos concretizar, sendo uma maneira de abstrair o mundo real. Em relação ao DW esta modelagem é necessária atender dois quesitos, sendo eles performance para consultas analíticas e que o próprio usuário consiga realizar estas consultas.

O tratamento do dado através do OLAP pode ser realizado de diversas maneiras utilizando a visão multidimensional do Data Warehouse. As ferramentas OLAP podem ser de diferentes tipos: MOLAP, ROLAP ou HOLAP.

O modelo estrela, que é o modelo mais utilizado na modelagem multidimensional, é constituído por uma tabela principal no centro do esquema, a mesma é a única no esquema que se comunica com as demais. As outras tabelas são ligadas à tabela central por meio de uma única junção. A tabela central é nomeada de tabela de fato e as outras são as tabelas de dimensões.

A classificação de dados inclusa em uma dimensão se constitui como hierarquia de uma dimensão, desta forma, conforme representado na figura 2, é possível estruturar os membros de uma dimensão em uma ou mais hierarquias.



**Figura 2. Representação de uma hierarquia de dimensões. Fonte: O autor.**

### **2.2.1. Esquema Estrela**

É utilizado este nome porque o modelo dos dados se assemelha a uma estrela. As tabelas de fato e de dimensão são normalmente combinadas em um esquema em estrela.

O benefício de utilizar o mesmo é a efetividade na extração dos dados, o que se mostra um grande ganho se falando em Data Warehouse. Conforme [Machado 2008], neste modelo é possível equilibrar rapidez em consultas e na quantidade de dados em disco, ele também relata que este modelo é eleito pela grande parte dos autores como uma forma de modelagem lógica que expõe para os usuários os dados de uma forma mais óbvia.

Neste modelo, os dados não estão normalizados, o que evita joins entre tabelas, diminuindo o tempo de consultas. Contudo, devido a repetição de dados, necessita de

mais espaço em disco. O modelo está sendo representado na figura 3, ajudando a visualizar o que foi descrito anteriormente.

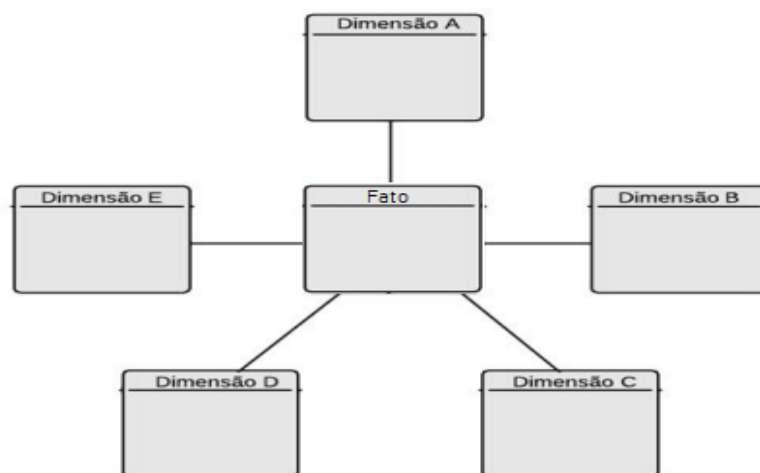


Figura 3. Exemplo de modelo de dados Esquema Estrela. Fonte: O autor.

### 2.2.2. Esquema Snowflake

Neste esquema são usadas tabelas normalizadas, as tabelas de Dimensão são criadas de uma forma que consigam atender a terceira forma normal. Este modelo possui algumas mudanças do esquema Estrela, sendo uma derivação do mesmo.

De acordo com [KIMBALL e ROSS 2002] o aumento da complexidade da estrutura de dados é uma desvantagem que o esquema floco de neves traz, proporcionando um grau de dificuldade a mais para alguns usuários a compreenderem a estrutura física da tabela, e isto não é justificado devido a representação das dimensões no Data Warehouse ocupar menos de dez por cento na área de armazenamento.

A figura 4 representa a estrutura do esquema Snowflake, mostrando a diferença entre ele e o esquema Estrela.

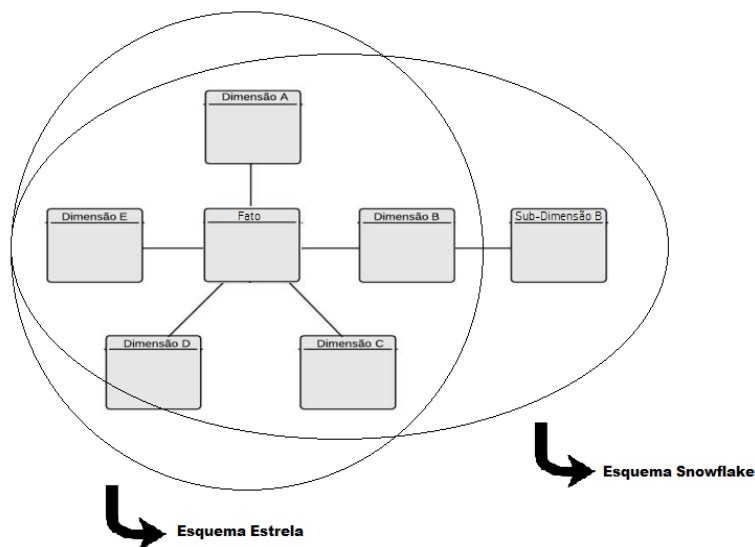


Figura 4. Diferença entre os Esquemas Estrela e Snowflake. Fonte: O autor.

## **2.3. Informação**

Uma informação bem lapidada nos permite tomar decisões com base em argumentos sólidos que nos ajudam a resolver problemas, sendo dados organizados e relacionados de uma forma significativa, a mesma possui como finalidade a redução das incertezas sobre algo. Um dado sozinho, por sua vez, é um fato isolado e que não apresenta valor, não conseguindo extrair uma informação relevante em cima do mesmo.

O processo de transformação dos dados em informação gera grande valor para os gestores de qualquer área, permitindo que os mesmos tenham muito mais convicção em uma futura decisão. Segundo [Barbieri, 2011] “A informação é considerada como um ingrediente básico do qual dependem os processos de decisão”.

Para [Capurro e Hjørland 2007] “Informação é qualquer coisa que é de importância na resposta a uma questão. Qualquer coisa pode ser informação. [...] Informação é o que pode responder questões importantes relacionadas às atividades do grupo-alvo”.

O estudo referente a informação tem se intensificado nos últimos anos, uma vez que quem possui um maior nível de informação consegue ter uma vantagem competitiva em cima dos demais. Isso proporcionou um rápido desenvolvimento científico e tecnológico, fazendo com que estudiosos e pesquisadores começassem a tentar desenvolver ferramentas e métodos que ajudassem a dar uma visão melhor para compreensão da informação. Diante disso, o estudo da Ciência da Informação se intensificou.

## **2.4. Business Intelligence**

A quantidade, variedade e velocidade de informação gerada na sociedade contemporânea é extremamente maior do que a de décadas anteriores, sendo necessário o auxílio de ferramentas (software) automatizadas para que seja possível manipular esse nível e quantidade de dados. A informação é algo essencial para apoiar a decisão, sendo matéria prima para obtenção de conhecimento. A informação está no nível operacional e gerencial, sendo que no nível operacional, a mesma é usada para que os sistemas funcionem.

Diante disso, encontramos diversas definições de BI, e em sua grande maioria refere-se a um conjunto de ferramentas e sistemas, que conectam dados de diferentes sistemas organizacionais, gerando informação e conhecimento a partir destes dados, que serão úteis para o processo de tomada de decisão.

Para [Popovic], “BI pode ser definido como informação de qualidade em armazéns de dados bem estruturados, [...] que fornecem aos trabalhadores do conhecimento acesso oportuno, análise efetiva e uma apresentação intuitiva da informação correta, habilitando-os a tomar ações ou decisões corretas”.

O termo BI também pode ser entendido com um software que é capaz de integrar várias fontes de informação para se definirem táticas adotadas pela respectiva empresa. Pode ser necessário também a utilização de ferramentas como o Data Warehouse e o data Mining (DM), devido ao uso de um grande nível de dados. É usufruído também recursos de análise conhecidos por On Line Analytical Processing

(OLAP), que viabiliza uma maior eficácia em consultas realizadas de maneira mais ágil e fácil.

## 2.5. Power BI

De acordo com o site oficial do Power BI, o mesmo é um pacote de BI e com ele é possível conectar e visualizar os dados usando uma plataforma única e escalonável para BI corporativo e por self-service no qual "é fácil de usar e ajuda você a obter insights mais aprofundados sobre os dados".

Este pacote possui vários serviços que o usuário pode utilizar, como a forma de conexão dos dados que pode ser feita de algumas formas, como por exemplo, via banco de dados ou arquivos de textos suportados pela ferramenta.

O Power BI possui tanto a versão paga quanto a gratuita, ajudando as empresas que não são de porte grande a poder também utilizar a ferramenta sem que tenham que arcar custos que não poderão arcar no futuro. Algumas diferenças entre as duas versões, são os fluxos de carga, o uso de aplicativos, como por exemplo, o Google Analytics, que são suportados só na versão paga.

Esta ferramenta auxilia na criação de dashboards, não precisando de conhecimentos aprofundados em tecnologia para conseguir manipular a mesma, estes dashboards criados pelo software são bastantes flexíveis e eficientes apresentando uma interface bem amigável e de fácil de criação deles.



Figura 5. Dashboard representando as funcionalidades do sistema. Fonte: <https://powerbi.microsoft.com/pt-br/power-bi-premium/>

## 2.6. Data Warehouse

Em um sistema de banco de dados, pode ser implementada uma tecnologia chamada de Data Warehouse (DW), que é hoje em dia um grande auxílio para as soluções de Business Intelligence (BI). Entretanto, para que ocorra um bom desempenho e estruturação, é necessário que a modelagem do banco tenha sido feita da forma multidimensional.

Para [Guimarães, 2018] “É nesse armazém de dados digitais que empresas e outras organizações acumulam informações detalhadas e sistematizadas, que dão origem a relatórios capazes de apoiar a tomada de decisões. O bom uso do Data Warehouse pode ajudar no controle de processos, por meio da disponibilização de informações estratégicas”.

Em um Data Warehouse os dados são originados de sistemas transacionais, bancos de dados relacionais entre outras fontes de dados. O mesmo é um repositório central de informações, que ajuda os analistas de dados em geral, através de ferramentas de inteligência de negócios (BI), a tomarem decisões com base em argumentos sólidos.

Segundo [Palestino, 2002], O Data Warehouse pode ser enxergado como um repositório de dados. O banco de dados é criado com a intenção de persuadir os clientes na área da informação, sendo essas informações periódicas e descritivas.

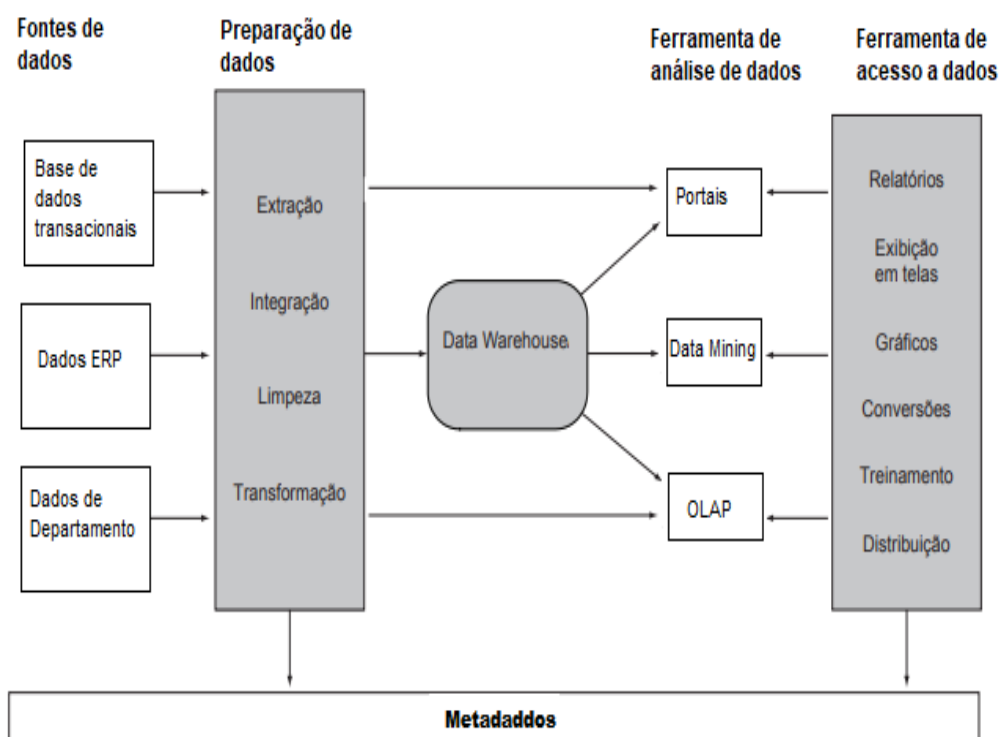


Figura 6. Arquitetura e processo de um Data Warehouse. Fonte: O autor.

## 3. Materiais e Métodos

Neste trabalho foi feita a construção de um sistema de Business Intelligence (BI), que através de um estudo de caso com dados disponibilizados pelo INEP, oriundos de uma



prova destinada a detentos, foram criados dashboards que apresentam métricas e indicadores importantes para alcançar objetivos e metas traçadas.

Para tanto, foi necessário realizar alguns passos que serão descritos no decorrer do artigo.

### **3.1. Base de dados**

Foram utilizadas como fonte de dados planilhas para construção deste trabalho, estando as duas no site do INEP, sendo a primeira o dicionário dos dados, que é um documento que centraliza informações sobre este conjunto de dados, e a segunda, as respostas que foram dadas por cada participante e as informações pessoais dos candidatos.

A planilha que contém as respostas e informações dos candidatos é constituída com as seguintes colunas: número de inscrição, ano que foi aplicado o exame, indicador do tipo de certificação pedida pelo participante, podendo ser ensino fundamental ou ensino médio, idade e sexo do participante, código da UF de residência, código da Unidade da Federação da aplicação da prova, solicitação de quais matérias deseja-se realizar a prova, presença ou não no dia da aplicação do exame, código da prova realizada, e as respectivas notas de cada disciplina.

#### **3.1.1. Criação do banco**

O banco de dados utilizado neste trabalho foi o Microsoft SQL Server que é um sistema gerenciador de Banco de dados relacional. A princípio foi feita uma identificação dos padrões existentes na planilha já citada, com o intuito de começar a esboçar uma modelagem para este banco, uma vez que as informações precisariam ser tratadas antes de serem inseridas no banco de dados.

Durante o processo de transformação de dados foram utilizados alguns processos tanto no Excel quanto no Editor do Power Query do próprio Excel para possibilitar que ao subir os dados para o banco de dados, os mesmos estivessem consistentes, como o fato de ter tido que criar um agrupamento de um participante para apenas uma área de conhecimento, uma vez que uma pessoa poderia fazer mais de uma prova e na planilha de origem que foi utilizada ter vindo todas essas possibilidades divididas em colunas na mesma linha do candidato que realizou ou não todas elas.

Para concluir este modelo de entidade relacionamento, foi preciso pensar nos três elementos principais de que ele é composto, que são os atributos, os relacionamentos e as entidades. Este processo é modelado a partir das entidades que serão ligadas entre si a partir dos relacionamentos, indicando as dependências entre elas, essas entidades podem conter vários atributos que é o que as caracterizam.

Na criação do banco foi gerado um Schema chamado de DASH\_PPL, que serve para agrupar objetos no nível de aplicação como também para simplesmente fazer divisões departamentais. Schemas são úteis para performance e segurança do banco.

Posteriormente a criação deste repositório, o mesmo foi populado com os dados que foram transformados em momento anterior, população essa que foi feita através da criação de uma fórmula no Excel com um INSERT sendo concatenado com as células que possuíam valores referentes às tabelas criadas no banco de dados, e replicada esta formula para todas as linhas do arquivo Excel.

Este modelo relacional criado é constituído de entidades, como já foi citado acima, sendo cada uma um tema ou tópicos diferentes. Foram criadas seis tabelas no mesmo que foram nomeadas como “t\_area\_conhecimento”, “t\_associacao\_area\_conhecimento\_prova”, “t\_associacao\_prova\_participante”, “t\_local\_prova”, “t\_participante”, “t\_prova” e estão sendo representadas conforme observamos na imagem 7, com seus respectivos atributos, que são os dados contidos nas tabelas e as ligações.

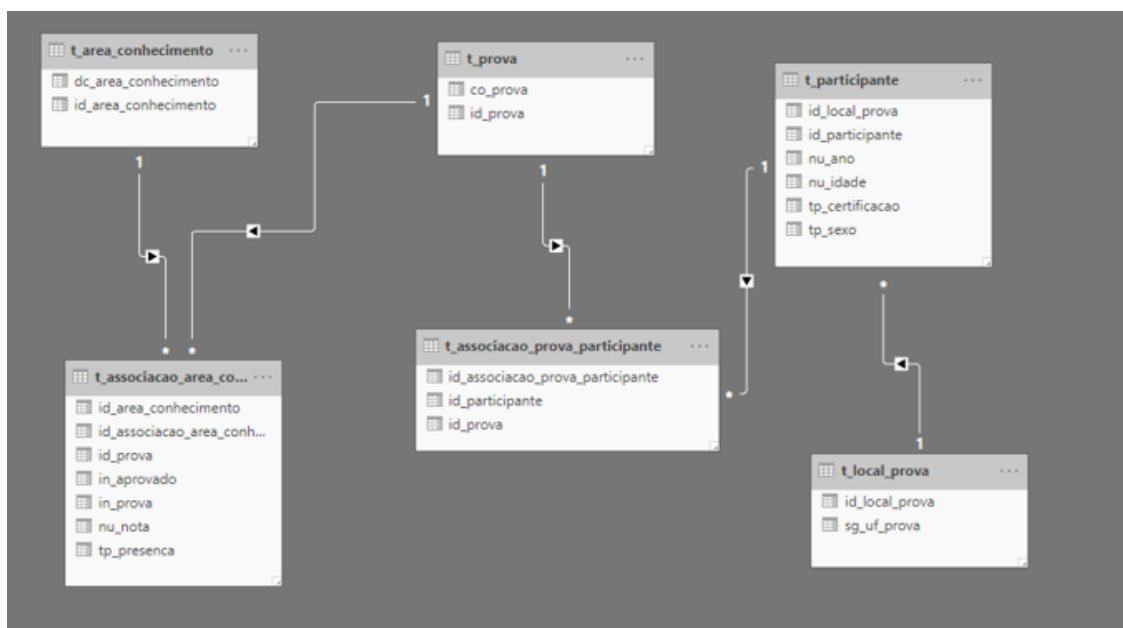


Figura 7. Modelo Relacional do Banco de dados Criado. Fonte: O autor.

### 3.2. Construção do Data Warehouse

Com o banco criado e devidamente populado, foi integrado com o serviço de análise de negócios da Microsoft chamado de Power BI, que após conectado com o banco foram feitas as adaptações necessárias para que a modelagem ficasse na forma multidimensional.

Foi utilizado o modelo dimensional tipo SnowFlake para realização do trabalho. Na modelagem Entidade Relacionamento, a tabela central chamada de fAvaliacao, contém os dados sumarizados do banco Relacional, e utilizará as chaves das tabelas “dLocalidade\_prova”, “dProva\_Enceja”, “dSituacao\_Aluno”, “dLocalidade\_Prova”, “dParticipante” e “dArea\_conhecimento” e os valores de medida “Media\_Nota” e “Quantidade\_Participantes”.

Os dados complementares são adicionados nas tabelas de dimensão, mas os mesmos são completamente relacionados com a tabela central.

No software foram criados gráficos com filtros e medidas, medidas estas que auxiliam na execução de cálculos em seus dados de acordo com a forma em que se deseja interagir com os relatórios, com o intuito de analisar o rendimento dos presidiários de diversas formas.

## 4. Resultados

### 4.1. Aumento da população prisional no Brasil

De acordo com os levantamentos feitos através do site do Departamento Penitenciário Nacional, foi possível observar o aumento significativo do número de presidiários no Brasil desde o ano de 2017, que é o ano que estamos analisando neste trabalho. Este aumento foi de 704.576 em dezembro de 2017 para 748.009 em dezembro de 2019, contabilizando um aumento de 6,16% em um intervalo de dois anos.

Deste total de pessoas privadas de liberdade, conseguimos destrinchar por regimes de cumprimento de pena, e a análise foi feita no mesmo período de tempo da anterior. No regime fechado, o aumento no período foi de 17,021% de pessoas, já o de regime semiaberto foi de 13,24%.

Este crescimento pode ser explicado pelas políticas punitivas no Brasil, e não só pelo aumento de crimes cometidos no país. Através dos dados coletados, também visualizamos que em 2017 a quantidade total de presos por tráfico de drogas chegava a mais de 176 mil pessoas, o que refletia na época cerca de 30% do número total de pessoas encarceradas, podendo concluir que esta política antidrogas se baseia em uma das políticas que mais causam prisões.

Diante do gráfico da população prisional por ano disponibilizado também no site do DEPEN, conforme a figura 8, conseguimos notar a constante crescente dos números de presos no Brasil, que multiplicou mais de três vezes do ano de 2000 para o ano de 2019, que foi o último ano que tivemos registros.

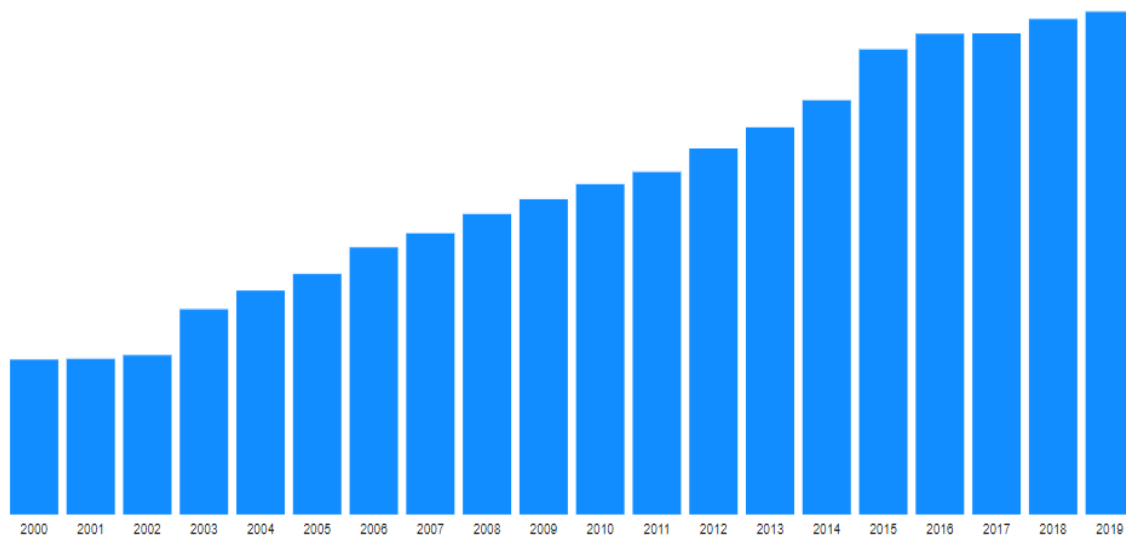


Figura 8. População Prisional por Ano. Fonte: DEPEN Dezembro 2019.

### 4.2. Análise no BI construído em relação a quantidade de participantes

De acordo com o documento internacional que foi aceito pelos órgãos da ONU (Organização das Nações Unidas) no ano de 1957, as pessoas que estão presas possuem o direito a terem acesso à educação. Desta forma [Carreira, 2009], relata que o documento reconhece que "[...] devem ser tomadas medidas no sentido de melhorar a educação de todos os reclusos, incluindo instrução religiosa. A educação de analfabetos

e jovens reclusos deve estar integrada no sistema educacional do país, para que depois da sua libertação possam continuar, sem dificuldades, a sua formação”.

Diante disso a prova do ENCEJA é uma ótima oportunidade para estas pessoas estudarem adquirirem cultura e informação e conseqüentemente conseguirem um certificado de conclusão de curso. Com base na planilha de dados utilizada, tivemos dados de 73.759 participantes oriundos de 27 estados brasileiros, e foram gerados gráficos que demonstrassem os estados que mais e menos possuíam participantes na realização da prova.

Foi possível visualizar através dos gráficos da figura 9 que o Estado de São Paulo foi o que de longe mais possuiu participantes no Exame ENCCEJA, com um total de 26.081 participantes, seguido pelo estado de Minas Gerais, que teve 11.007 pessoas participando do processo seletivo.

As duas primeiras posições, em relação a maior quantidade de participantes, nós conseguimos entender facilmente, devido ao fato de no ano de 2017, esses dois estados que ficaram nos primeiros lugares, terem sido os que mais possuíam pessoas encarceradas com o número de 227.701 pessoas em São Paulo e 74.805 em Minas Gerais.

Vale ressaltar que um participante pode escolher uma ou mais áreas de conhecimento para realizar a prova, dependendo da quantidade de disciplinas que estavam faltando para o mesmo concluir o ensino básico. Com as informações levantados através das análises feitas, o estado do Paraná, no ano do estudo feito neste trabalho, foi o estado que menos participantes teve, com um total de 146 participantes.

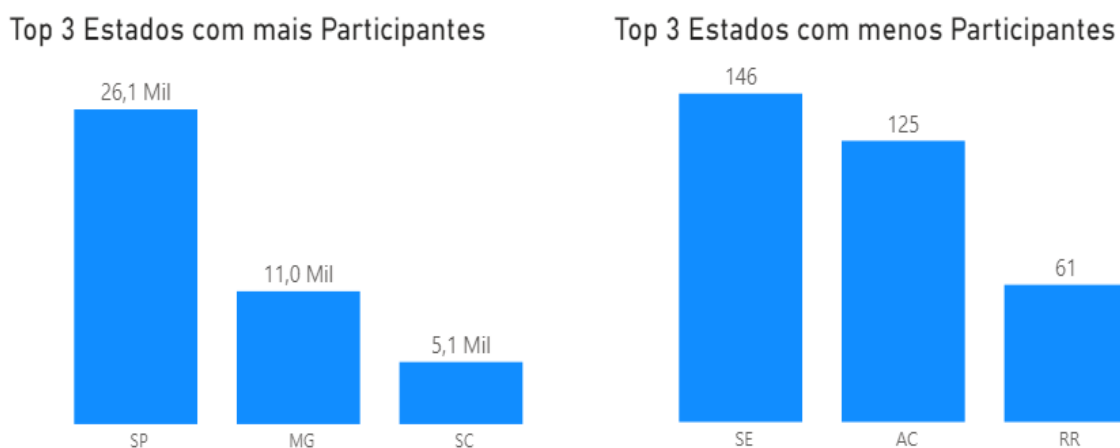


Figura 9. Estados com mais e menos participantes. Fonte: O autor.

### 4.3. Análise no BI construído em relação a taxa de reprovação por área de conhecimento

Esta análise foi feita englobando tanto os participantes que optaram por fazer as provas para conclusão do ensino fundamental quanto do ensino médio, unindo as áreas de conhecimento destas duas modalidades em apenas quatro, afim de ter uma visão mais geral dos dados.

Foi possível notar, com base no gráfico da figura 10, que as áreas relacionadas a matemática tiveram a maior taxa de reprovação, com o total de 23.082 participantes

reprovados. Já Ciências Humanas e suas tecnologias ficou em último lugar do posto, com o total de 14.361 participantes reprovados.

Estes altos números de reprovação em matemática nos comprovam novamente como o ensino é precário em grande parte das escolas no Brasil, para [Felicetti e Giraffa, 2011], “O sucesso ou insucesso na disciplina de Matemática está ligado não só naquilo que é ensinado, mas, principalmente, em como é ensinado de modo a consolidar os conteúdos matemáticos a cada nível de aprendizagem, uma vez que esta disciplina se torna mais complexa a cada nível de ensino”.

As altas taxas de reprovação escolar, como temos observado no decorrer do artigo, é um problema no Brasil, sendo o problema tanto no sistema prisional mas também quanto fora dele, estas reprovações na rede pública geram um alto desperdício de recursos financeiros para o país, uma vez que o aluno fará novamente o mesmo ano gastando mais uma vez com o material que é disponibilizado para ele, quanto o salário que terá que ser pago para os professores ministrarem suas aulas.

Taxa de Reprovação por Área de Conhecimento

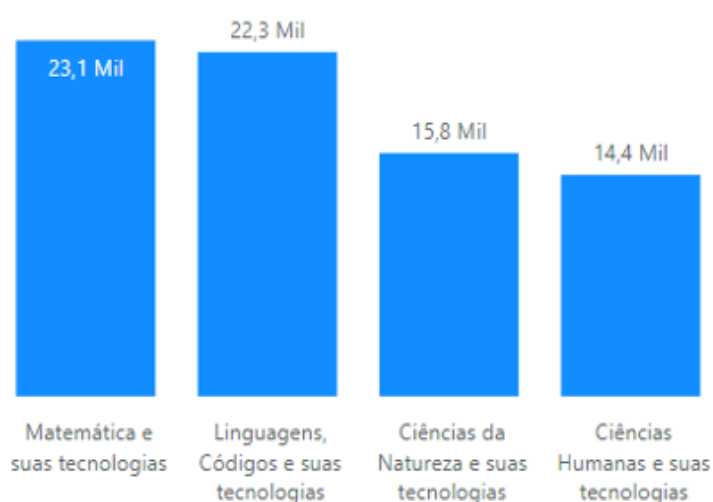


Figura 10. Gráfico representando a taxa de reprovação por área de conhecimento. Fonte: O autor.

#### 4.4. População prisional em atividade educacional

O estudo, além de ampliar os conhecimentos e gerar uma melhor perspectiva de vida de uma pessoa que está encarcerada, também diminui o tempo de cumprimento de pena, ajudando a reduzir a superlotação dos presídios. Conforme será evidenciado no presente trabalho, com os dados levantados e as análises feitas, a baixa escolaridade dos detentos é algo preocupante, e o estudo se mostra extremamente importante para a reinserção destas pessoas na sociedade.

Foi constatado, através do site do DEPEN, que 94.916 presidiários, ou seja, 13.47% do número total no ano de 2017 se encontrava em atividade educacional, sendo que dentro deste número, 12.654 encontrava-se em processo de alfabetização, 37.144 fazendo o ensino fundamental, 16.288 cursando o ensino médio, entre outras modalidades de ensino.

Este percentual poderia ser muito maior, pelo fato de 51%, ou seja, a maior parcela das pessoas privadas de liberdade no Brasil não possuírem ensino fundamental completo. Se for fazer o somatório das pessoas que não terminaram o ensino básico no sistema prisional, chegamos a assustadores 76%.

Através dos dados da prova do ENCCEJA que validamos com a ajuda do Power BI, a média de idade das pessoas que fizeram a prova para tentar completar pelo menos uma das etapas do ciclo escolar foi de 29 anos, sendo uma idade ainda muito produtiva e que esperamos que com a conclusão do ensino e o cumprimento da pena, estas pessoas consigam ter um novo começo.

Na figura 11 é representada a faixa etária de todas as pessoas privadas de liberdade no Brasil, o que nos revela um alto número de jovens nessa situação. Este número expressivo de pessoas com idade inferior aos 30 anos é muito preocupante, uma vez que o cenário educacional dos brasileiros dessa faixa etária, mesmo estando fora das grades, não é nem um pouco satisfatório, segundo os levantamentos do Instituto Brasileiro de Geografia e Estatística [Setembro, 2020]. Apesar da proporção da população de 25 anos ou mais com ensino médio completo ter crescido no país, aumentando de 45,05% em 2016 para 47,4% em 2018 e 48,8% em 2019, mais da metade (51,2 % ou 69,5 milhões) dos adultos não conseguiram concluir essa fase, possuindo o ensino médio incompleto.

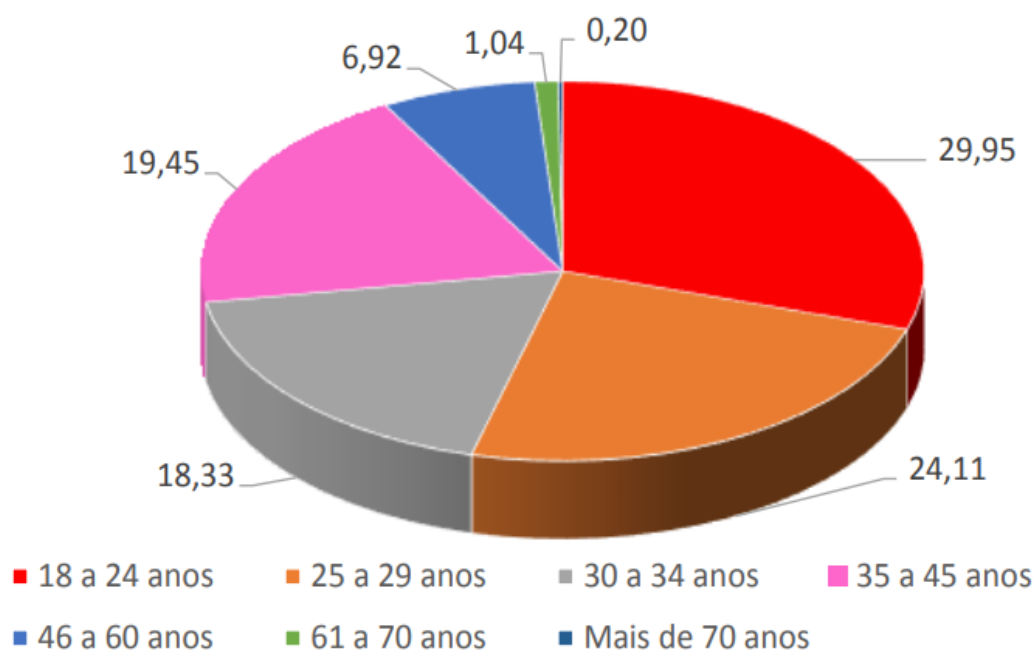


Figura 11. Faixa etária de pessoas encarceradas no Brasil. Fonte: [Silva, 2017].

#### 4.5. Análise no BI construído em relação ao sexo dos participantes

Como já mencionado acima, o total de presidiários no ano de 2017 foi de 704.576, e com esses dados foi possível também analisar o percentual de pessoas presas em relação ao seu respectivo sexo.

A quantidade de homens que estão nessa situação é muito maior do que a de mulheres, resultados estes que foram retirados do site do DEPEN, com um percentual de 94,67% de pessoas do sexo masculino e apenas 5,33% pessoas do sexo feminino.

As análises feitas em cima da prova do ENCCEJA com os dados representados pela figura 12 também demonstraram que a quantidade de participantes do sexo masculino foi muito maior do que os do sexo feminino, com um total de 90.09% dos participantes sendo homens, coincidentemente ficando com um percentual muito próximo do geral no Brasil, estes um pouco mais de 90% totalizaram 66.445 pessoas. Já as participantes do sexo feminino contabilizaram cerca de 7.312.

## Total de Participantes por Sexo

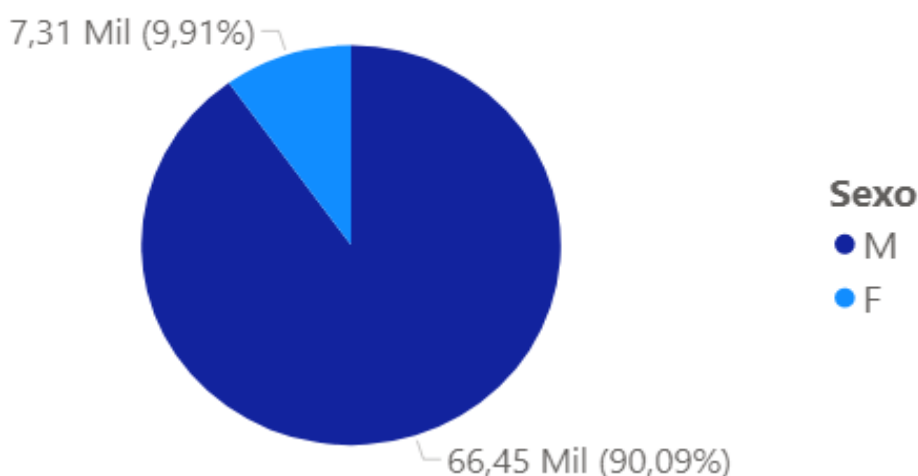


Figura 12. Gráfico representando o total de participantes por sexo que realizaram a prova do ENCEJA. Fonte: O autor.

## 5. Conclusão

Esse trabalho teve o objetivo de analisar o desempenho escolar das pessoas privadas de liberdade que realizaram o Exame Nacional para Certificação de Competências de Jovens e Adultos e tentar identificar também o perfil de todos os brasileiros que se encontram nesta situação. O estudo foi desenvolvido após ter sido encontrada uma base de dados governamentais e um painel interativo no site da DEPEN que apresentava diversas estatísticas em relação a situação dos presidiários no Brasil. Esta base de dados encontrada disponibilizava vários dados em relação ao participante, como as notas obtidas em cada disciplina feita e se a pessoa conseguiu ser aprovada neste processo seletivo.

Diante dos dados que foram exibidos através do Power BI, foi possível fazer algumas análises em relação ao contexto social que o Brasil se encontra e identificar algumas particularidades dos participantes.

Analisando o gráfico da população prisional por ano, conseguimos perceber que este crescimento que vem ocorrendo é um enorme problema que o nosso país possui, sendo que os números foram triplicados em um pequeno intervalo de tempo, o que gera uma despesa gigantesca para o país e uma dificuldade enorme em reinserir estas pessoas na sociedade. Um estado se destacou na quantidade tanto nos participantes que

realizaram a prova, que foi de 26.081, quanto em números de pessoas que estão privadas de liberdade, que no ano analisado, chegou a incríveis 227.701 pessoas.

Um dado que chamou bastante atenção também foi o percentual de pessoas reclusas da liberdade que não possuem o ensino básico completo, que seriam pessoas que não concluíram o ensino médio, percentual este que é de 76%. Em relação as pessoas que fizeram a prova para tentar concluir alguma etapa do ensino básico, notamos que este número chegou a 73.759 candidatos, número bem baixo comparado com o total de pessoas que poderiam ter feito a prova.

Com as análises realizadas, foi possível demonstrar também que o percentual de participantes do exame do sexo masculino foi muito maior que o do sexo feminino, com mais de 90% das pessoas sendo do sexo masculino, número explicado também pela enorme diferença da quantidade de pessoas destes dois sexos em todo o território brasileiro.

Com isso, conseguimos identificar que a maior parte do perfil de uma pessoa que está em cumprimento de pena no Brasil são jovens do sexo masculino sendo que a maioria deles não possui ensino médio completo, revelando que as pessoas que vivem à margem da sociedade necessitam de uma atenção especial para que elas consigam trilhar caminhos que sejam longe da criminalidade.

## Referências

- BARBIERI, Carlos. (2011) “BI2 -Business Intelligence: modelagem e qualidade”. p.392.
- BRASIL. DEPEN. Departamento Penitenciário Nacional. Disponível em: <<https://app.powerbi.com/view?r=eyJrIjojZTlkZGJjODQtNmJlMi00OTJhLWFlMDktNzRlNmFkNTM0MWI3IiwidCI6ImViMDkwNDIwLTQ0NGMtNDNmNy05MWYyLTRiOGRhNmJmZThlMSJ9>>. Acesso em: 10 set. 2020.
- CAPURRO, Rafael; HJORLAND, Birger. (2007) “O conceito de informação”. Revista Perspectivas em Ciência da Informação, v. 12, n. 1, p. 148-207.
- CARREIRA, Denise. (2009) “Relatório nacional para o direito humano à educação: educação nas prisões brasileiras”. São Paulo: Plataforma DhESCA.
- ELIAS, Diego. (2014) "Reportagem Entendendo a modelagem multidimensional". <https://canaltech.com.br/business-intelligence/entendendo-a-modelagem-multidimensional-19988/>.
- FELICETTI, V.L. e GIRAFFA, L. M. M. (2011) "Intervenientes na aprendizagem matemática". XII CIAEM-IACME, Recife, Brasil.
- GRACIANO, Mariângela. (2005) “Educação também é direito humano”. São Paulo: Ação Educativa, Plataforma Interamericana de Direitos Humanos, Democracia e Desenvolvimento.
- Guimarães. (2018) “O que é data warehouse: 7 vantagens que você precisa conhecer!”. <https://www.knowsolution.com.br/o-que-e-data-warehouse-3-vantagens-que-voce-precisa-conhecer/#close>.



- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Microdados do ENCCEJA 2017. Brasília: Inep, 2017. Disponível em: <<http://portal.inep.gov.br/microdados>>. Acesso em: 01 jan. 2020.
- KIMBALL, Ralph; ROSS, Margy. (2002) "DATA WAREHOUSE TOOLKIT - O Guia Completo para a Modelagem Multidimensional". Rio de Janeiro: CAMPUS.
- PNAD Educação 2019: Mais da metade das pessoas de 25 anos ou mais não completaram o ensino médio. Disponível em: <<https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/28285-pnad-educacao-2019-mais-da-metade-das-pessoas-de-25-anos-ou-mais-nao-completaram-o-ensino-medio>>. Acesso em: 10 set. 2020.
- Machado, F. N. (2008) "Tecnologia e Projeto do Armazém", Editora Érica São Paulo, SP.
- NETO, Arilo Cláudio Dias. (2014) "Bancos de Dados Relacionais". Revista SQL Magazine 86. <http://www.devmedia.com.br/bancos-de-dadosrelacionais-artigo-revista-sql-magazine-86/20401>.
- Palestino, Carlos. (2011) "Business Intelligence Modelagem e Qualidade", Editora Elsevier Rio de Janeiro, RJ.
- POPOVIC, A.; et al. (2012) "Towards business intelligence system success: effects of maturity and culture on analytical decision making". Decision Support Systems, v.54, n. 1, p.729-739.
- Silva, Marcos Vinícius Moura. (2017) "Levantamento Nacional de Informações Penitenciárias Atualização - Junho de 2017".
- Velasco, Clara. et al. (2020) "Brasil tem 338 encarcerados a cada 100 mil habitantes; taxa coloca país na 26ª posição do mundo". <https://g1.globo.com/monitor-da-violencia/noticia/2020/02/19/brasil-tem-338-encarcerados-a-cada-100-mil-habitantes-taxa-coloca-pais-na-26a-posicao-do-mundo.ghtml>.