SENAC

AMEAÇAS GLOBAIS À SEGURANÇA CIBERNÉTICA: APOIO DECISÓRIO EM SEGURANÇA CIBERNÉTICA

AYRA WANDERLEY HIDASI
GABRIELA GUEDES DA CRUZ
RAFAEL DE LIMA SANTOS
VITOR BORDIN GOMES

2025

AYRA WANDERLEY HIDASI
GABRIELA GUEDES DA CRUZ
RAFAEL DE LIMA SANTOS
VITOR BORDIN GOMES

Ameaças globais à segurança cibernética: apoio decisório em segurança cibernética

Projeto integrador: apoio decisório aos negócios apresentado ao Senac, como exigência para obtenção de nota no curso de Tecnologia em Banco de Dados.

2025

AGRADECIMENTOS

À equipe de professores e colaboradores do SENAC.

RESUMO

Este projeto tem como objetivo analisar dados sobre ameaças cibernéticas globais, com foco no apoio à tomada de decisões estratégicas em segurança da informação. A partir da base de dados Global Cybersecurity Threats (2015–2024), foram desenvolvidas atividades analíticas que envolvem a identificação de padrões de ataques, setores mais vulneráveis, perdas financeiras, tipos de vulnerabilidades exploradas e eficácia dos mecanismos de defesa utilizados. Para viabilizar essas análises, foi estruturado um processo completo de Extração, Transformação e Carga (ETL), utilizando ferramentas como MySQL Workbench, Microsoft Excel e Power BI. Os dados foram organizados em um modelo dimensional estrela, com tabelas de dimensão e uma tabela de fatos central, permitindo a realização de operações OLAP e a construção de dashboards interativos. Os resultados obtidos demonstram o potencial da análise de dados como instrumento estratégico para a formulação de políticas de segurança cibernética mais eficazes e direcionadas, contribuindo para a prevenção de riscos e o fortalecimento da resiliência digital das organizações.

Palavras-chave: 1. Segurança cibernética. 2. Análise de dados. 3. ETL. 4. Tomada de decisão. 5. Ameaças digitais.

ABSTRACT

This project aims to analyze data on global cybersecurity threats, focusing on supporting strategic decision-making in information security. Based on the Global Cybersecurity Threats (2015–2024) dataset, analytical activities were developed to identify attack patterns, the most vulnerable sectors, financial losses, types of exploited vulnerabilities, and the effectiveness of defense mechanisms. To enable these analyses, a complete Extract, Transform, and Load (ETL) process was structured using tools such as MySQL Workbench, Microsoft Excel, and Power BI. The data was organized into a star schema dimensional model, with dimension tables and a central fact table, allowing OLAP operations and the creation of interactive dashboards. The results demonstrate the potential of data analysis as a strategic tool for developing more effective and targeted cybersecurity policies, contributing to risk prevention and strengthening organizational digital resilience.

Keywords: 1. Cybersecurity. 2. Data analysis. 3. ETL. 4. Decision-making. 5. Digital threats.

SUMÁRIO

1. INTRODUÇÃO

A crescente digitalização dos processos organizacionais tem promovido avanços significativos em produtividade, conectividade e inovação. No entanto, esse processo também tem aumentado a exposição a riscos cibernéticos, tornando a segurança da informação um dos principais desafios para as organizações. Nesse contexto, a análise de dados sobre ameaças cibernéticas torna-se uma ferramenta estratégica para a formulação de políticas de prevenção e resposta a incidentes.

Este projeto tem como objetivo principal explorar a base de dados "*Global Cybersecurity Threats* (2015–2024)" disponibilizada na plataforma Kaggle. Essa base reúne informações detalhadas sobre incidentes de segurança cibernética registrados em diversos países ao longo de uma década, incluindo variáveis como país afetado, ano do incidente, tipo de ataque, setor alvo, perdas financeiras, número de usuários impactados, origem do ataque, tipo de vulnerabilidade explorada, mecanismos de defesa utilizados e tempo de resolução.

A proposta consiste em extrair insights estratégicos que possam subsidiar a tomada de decisões no campo da segurança cibernética, destacando-se a identificação de padrões de ataques, a análise de setores mais vulneráveis, a determinação das origens mais recorrentes das ameaças e a avaliação da eficácia dos mecanismos de defesa empregados. Tais informações são fundamentais para orientar investimentos em infraestrutura de segurança, capacitação de equipes e definição de políticas de proteção digital.

Para viabilizar a análise estratégica desses dados, foi implementado um processo ETL, utilizando ferramentas como MySQL Workbench, Microsoft Excel e Power BI. A estruturação dos dados seguiu um modelo dimensional em estrela, com a criação de tabelas de dimensão e uma tabela de fatos central (fact_incident), permitindo a realização de operações OLAP e a construção de dashboards interativos.

Visando a colaboração acadêmica, o projeto foi publicado publicamente no GitHub <https://github.com/ayrahidasi/PI_SENAC_GRUPO2_AMEACAS_GLOBAIS_A_SEGURANCA_CIBERNETICA>. O repositório contém todos os arquivos relacionados ao desenvolvimento do projeto, incluindo os scripts SQL de criação e inserção de dados, a estrutura do modelo dimensional, a fonte de dados original em CSV, e instruções para replicação do processo ETL e visualizações no Power BI.

2. FONTE DE DADOS

A seleção da base de dados é um crucial para a qualidade e a relevância das análises desenvolvidas em projetos de ciência de dados e apoio à decisão. Para este trabalho, optou-se pela utilização da base intitulada "*Global Cybersecurity Threats* (2015–2024)"[1], disponibilizada na plataforma Kaggle, um repositório amplamente reconhecido na comunidade de ciência de dados por hospedar conjuntos de dados públicos de alta qualidade e diversidade temática.

Trata-se de um conjunto de dados estruturado e bem documentado, contendo informações detalhadas sobre incidentes de segurança cibernética ocorridos em vários países ao longo de uma década. As variáveis disponíveis abrangem aspectos essenciais para a análise estratégica da segurança da informação, como o tipo de ataque, setor afetado, perdas financeiras, número de usuários impactados, origem da ameaça, tipo de vulnerabilidade explorada, mecanismos de defesa utilizados e tempo de resposta. A base possui abrangência temporal e geográfica, permitindo a identificação de padrões e tendências globais, bem como a comparação entre diferentes contextos.

Dessa forma, a base "*Global Cybersecurity Threats* (2015–2024)" apresenta-se como uma fonte de dados robusta, confiável e alinhada aos objetivos do presente trabalho, possibilitando a análise de dados para a geração de *insights* relevantes no contexto da segurança cibernética.

Tabela 1 - Variáveis contidas na base "*Global Cybersecurity Threats* (2015–2024)"

Nome da Variável	Descrição	Tipo de Dado
country	País afetado pelo incidente	String
year	Ano do incidente	Integer
attack_type	Tipo de ataque cibernético	String
target_sector	Setor alvo do ataque	String
financial_loss	Perda financeira resultante do ataque	Float
users_affected	Número de usuários afetados	Integer
attack_origin	Origem do ataque	String
vulnerability_type	Tipo de vulnerabilidade explorada	String
defense_mechanism	Mecanismo de defesa utilizado	String
resolution_time	Tempo de resolução do incidente	Integer

3. DEFINIÇÃO DAS TECNOLOGIAS

A escolha das tecnologias utilizadas neste projeto foi orientada pela necessidade de garantir eficiência, acessibilidade e integração entre as etapas do processo de análise de dados. As ferramentas selecionadas permitiram a execução do processo ETL, a modelagem dimensional, a análise exploratória e a visualização interativa dos dados.

Para o armazenamento e manipulação dos dados, foi utilizado o MySQL Workbench, uma ferramenta robusta de gerenciamento de banco de dados relacional. Através dela, foi possível importar a base de dados original em formato CSV, estruturar as tabelas de dimensão e fato, e executar os scripts SQL necessários para a carga e organização dos dados conforme o modelo estrela.

A ferramenta Power BI, da Microsoft, foi empregada para a construção dos dashboards analíticos. Sua capacidade de integração com bancos de dados relacionais e sua interface intuitiva permitiram a criação de relatórios interativos, com indicadores-chave de desempenho (KPIs) e recursos OLAP, como drill-down, slicing e pivotagem. Esses recursos foram fundamentais para viabilizar análises estratégicas e operacionais no contexto da segurança cibernética.

Além disso, o Microsoft Excel foi utilizado como apoio nas etapas iniciais de exploração da base de dados, permitindo uma visualização rápida dos registros e facilitando a identificação de padrões e categorias relevantes para a modelagem dimensional.

A combinação dessas tecnologias proporcionou uma solução completa e integrada, desde a preparação dos dados até a geração de insights visuais, contribuindo para a eficácia do apoio decisório em segurança da informação.

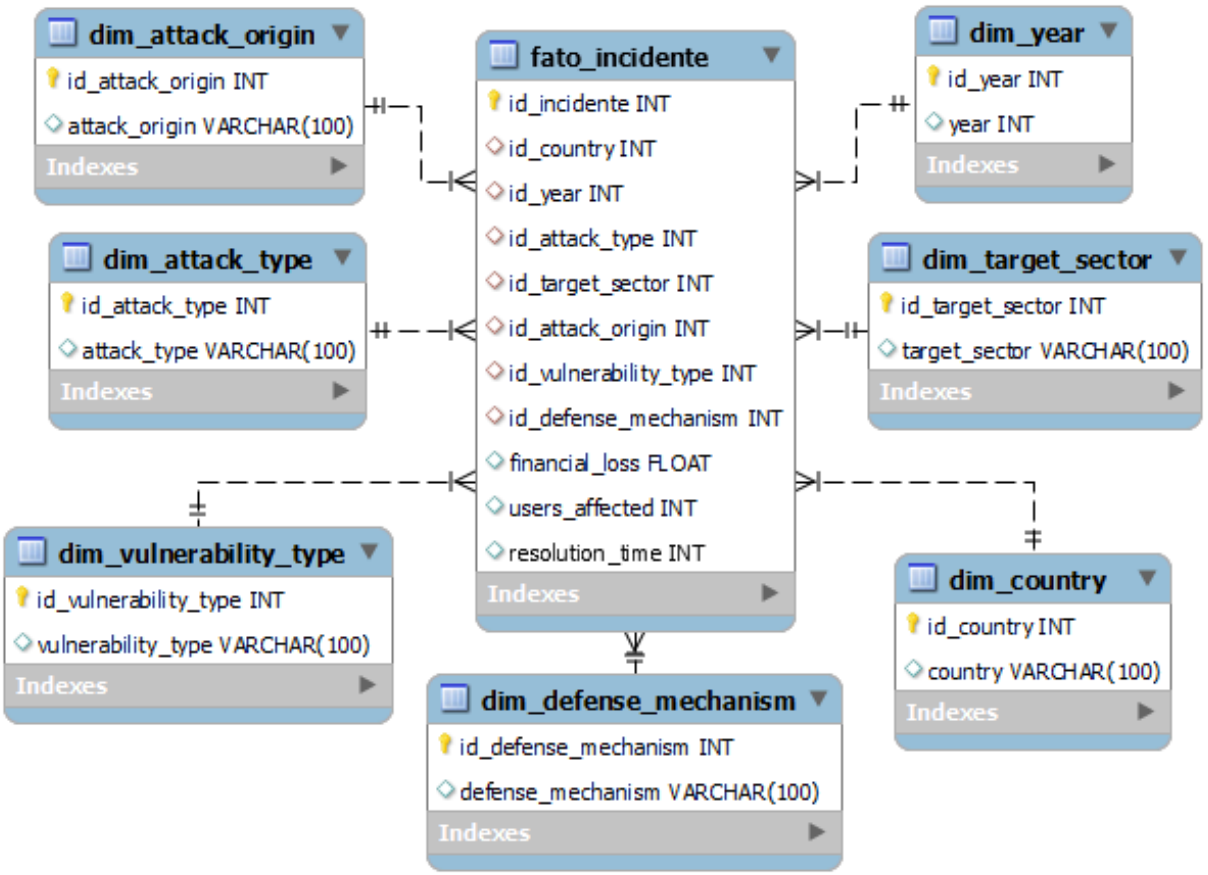
4. MODELAGEM DE DADOS: MODELO ESTRELA

O modelo estrela é a arquitetura mais comum para Data Warehouses, sendo composto por uma tabela central, chamada Tabela de Fatos, que se conecta diretamente a várias Tabelas de Dimensão.

A modelagem dimensional no formato estrela foi adotada neste projeto como base

para a estruturação dos dados, visando facilitar análises multidimensionais e operações OLAP (Online Analytical Processing), fundamentais para o apoio à tomada de decisão em segurança cibernética.

No centro do esquema estrela está a tabela de fatos, que armazena os dados quantitativos dos incidentes cibernéticos, como perdas financeiras, número de usuários afetados e tempo de resolução. Circundam a tabela de fatos as tabelas de dimensões, que descrevem os atributos qualitativos dos ataques, como país, tipo de ataque, setor alvo, origem da ameaça, tipo de vulnerabilidade e mecanismo de defesa. Essa separação clara entre fatos e dimensões permite consultas eficientes e flexíveis, como agregações por país, comparações entre tipos de ataques ou análises temporais, separando as informações quantitativas (fatos) das descritivas (dimensões), tornando a base de dados mais intuitiva e fácil de gerenciar.



4.1. SCRIPTS DDL: CRIAÇÃO DAS TABELA

```
CREATE DATABASE AmeacasGlobaisSegurancaCibernetica;
USE AmeacasGlobaisSegurancaCibernetica;
-- Tabelas de Dimensão
CREATE TABLE dim_country (
  id_country INT AUTO_INCREMENT PRIMARY KEY,
   country VARCHAR(100));
CREATE TABLE dim_year (
   id_year INT AUTO_INCREMENT PRIMARY KEY,
   year INT);
CREATE TABLE dim_attack_type (
   id_attack_type INT AUTO_INCREMENT PRIMARY KEY,
   attack_type VARCHAR(100));
CREATE TABLE dim_target_sector (
   id_target_sector INT AUTO_INCREMENT PRIMARY KEY,
   target_sector VARCHAR(100));
CREATE TABLE dim_attack_origin (
   id_attack_origin INT AUTO_INCREMENT PRIMARY KEY,
   attack_origin VARCHAR(100));
CREATE TABLE dim_vulnerability_type (
   id_vulnerability_type INT AUTO_INCREMENT PRIMARY KEY,
   vulnerability_type VARCHAR(100));
CREATE TABLE dim_defense_mechanism (
   id_defense_mechanism INT AUTO_INCREMENT PRIMARY KEY,
   defense_mechanism VARCHAR(100));
```

-- Tabela de Fatos

```
CREATE TABLE fact_incident(
    id_incidente INT AUTO_INCREMENT PRIMARY KEY,
    id_country INT,
    id_year INT,
    id_attack_type INT,
    id_target_sector INT,
    id_attack_origin INT,
    id_vulnerability_type INT,
    id_defense_mechanism INT,
    financial_loss FLOAT,
    users_affected INT,
    resolution_time INT,
    FOREIGN KEY (id_country) REFERENCES dim_country(id_country),
    FOREIGN KEY (id_year) REFERENCES dim_year(id_year),
    FOREIGN KEY (id_attack_type) REFERENCES dim_attack_type(id_attack_type),
    FOREIGN KEY (id_target_sector) REFERENCES dim_target_sector(id_target_sector),
    FOREIGN KEY (id_attack_origin) REFERENCES dim_attack_origin(id_attack_origin),
    FOREIGN KEY (id_vulnerability_type) REFERENCES
dim_vulnerability_type(id_vulnerability_type),
    FOREIGN KEY (id_defense_mechanism) REFERENCES
dim_defense_mechanism(id_defense_mechanism));
```

4.2. SCRIPTS DML: INSERÇÃO DOS DADOS NAS TABELAS

-- Inserir dados nas Tabelas Dimensão

```
USE ameacasglobaissegurancacibernetica;
INSERT INTO dim_country (country) VALUES ('Australia');
INSERT INTO dim_country (country) VALUES ('Brazil');
INSERT INTO dim_country (country) VALUES ('China');
INSERT INTO dim_country (country) VALUES ('France');
INSERT INTO dim_country (country) VALUES ('Germany');
INSERT INTO dim_country (country) VALUES ('India');
INSERT INTO dim_country (country) VALUES ('Japan');
INSERT INTO dim_country (country) VALUES ('Russia');
INSERT INTO dim_country (country) VALUES ('UK');
INSERT INTO dim_country (country) VALUES ('USA');
INSERT INTO dim_year (year) VALUES ('2015');
INSERT INTO dim_year (year) VALUES ('2016');
INSERT INTO dim_year (year) VALUES ('2017');
INSERT INTO dim_year (year) VALUES ('2018');
INSERT INTO dim_year (year) VALUES ('2019');
INSERT INTO dim_year (year) VALUES ('2020');
INSERT INTO dim_year (year) VALUES ('2021');
INSERT INTO dim_year (year) VALUES ('2022');
INSERT INTO dim_year (year) VALUES ('2023');
INSERT INTO dim_year (year) VALUES ('2024');
INSERT INTO dim_attack_type (attack_type) VALUES ('DDoS');
INSERT INTO dim_attack_type (attack_type) VALUES ('Malware');
INSERT INTO dim_attack_type (attack_type) VALUES ('Man-in-the-Middle');
```

```
INSERT INTO dim_attack_type (attack_type) VALUES ('Phishing');

INSERT INTO dim_attack_type (attack_type) VALUES ('Ransomware');

INSERT INTO dim_attack_type (attack_type) VALUES ('SQL Injection');

INSERT INTO dim_target_sector (target_sector) VALUES ('Banking');

INSERT INTO dim_target_sector (target_sector) VALUES ('Education');

INSERT INTO dim_target_sector (target_sector) VALUES ('Government');

INSERT INTO dim_target_sector (target_sector) VALUES ('Healthcare');

INSERT INTO dim_target_sector (target_sector) VALUES ('IT');

INSERT INTO dim_target_sector (target_sector) VALUES ('Retail');

INSERT INTO dim_target_sector (target_sector) VALUES ('Telecommunications');

INSERT INTO dim_attack_origin (attack_origin) VALUES ('Hacker Group');

INSERT INTO dim_attack_origin (attack_origin) VALUES ('Insider');

INSERT INTO dim_attack_origin (attack_origin) VALUES ('Nation-state');

INSERT INTO dim_attack_origin (attack_origin) VALUES ('Unknown');

INSERT INTO dim_vulnerability_type (vulnerability_type) VALUES ('Social Engineering');

INSERT INTO dim_vulnerability_type (vulnerability_type) VALUES ('Unpatched Software');

INSERT INTO dim_vulnerability_type (vulnerability_type) VALUES ('Weak Passwords');

INSERT INTO dim_vulnerability_type (vulnerability_type) VALUES ('Zero-day');

INSERT INTO dim_defense_mechanism (defense_mechanism) VALUES ('AI-based Detection');

INSERT INTO dim_defense_mechanism (defense_mechanism) VALUES ('Antivirus');

INSERT INTO dim_defense_mechanism (defense_mechanism) VALUES ('Encryption');

INSERT INTO dim_defense_mechanism (defense_mechanism) VALUES ('Firewall');

INSERT INTO dim_defense_mechanism (defense_mechanism) VALUES ('VPN');
```

```
-- Inserir dados na Tabela Fatos
USE ameacasglobaissegurancacibernetica;
INSERT INTO fact_incident(
    id_country, id_year, id_attack_type, id_target_sector,
    id_attack_origin, id_vulnerability_type, id_defense_mechanism,
    financial_loss, users_affected, resolution_time
)
SELECT
    c.id_country,
    y.id_year,
    a.id_attack_type,
    s.id_target_sector,
    o.id_attack_origin,
    v.id_vulnerability_type,
    d.id_defense_mechanism,
    g.`Financial Loss (in Million $)` AS financial_loss,
    g.`Number of Affected Users` AS users_affected,
    g.`Incident Resolution Time (in Hours)` AS resolution_time
FROM pi_grupo2_seguranca_cibernetica.`global_cybersecurity_threats_2015-2024` g
JOIN dim_country c ON g.Country = c.country
JOIN dim_year y ON g.Year = y.year
JOIN dim_attack_type a ON g.`Attack Type` = a.attack_type
JOIN dim_target_sector s ON g.`Target Industry` = s.target_sector
JOIN dim_attack_origin o ON g.`Attack Source` = o.attack_origin
JOIN dim_vulnerability_type v ON g.`Security Vulnerability Type` = v.vulnerability_type
JOIN dim_defense_mechanism d ON g.`Defense Mechanism Used` = d.defense_mechanism;
```

5. PROCESSO ETL - EXTRACT, TRANSFORM, LOAD

O processo de extração, transformação e carga **(**ETL) é uma etapa fundamental na preparação de dados para análise, especialmente em projetos que envolvem grandes volumes de informações. No contexto deste trabalho, o ETL tem como objetivo organizar, limpar e estruturar os dados da base CSV "*Global Cybersecurity Threat*s (2015–2024)" para utilização em um modelo dimensional e visualizações analíticas no Power BI como foco em atividades de a apoio decisório aos negócios.

5.1. Extração

A etapa de extração consistiu na obtenção dos dados a partir de um arquivo CSV, disponibilizado na plataforma Kaggle. Esse arquivo contém registros de incidentes de segurança cibernética ocorridos entre 2015 e 2024, com informações como país afetado, tipo de ataque, setor alvo, perdas financeiras, número de usuários impactados, entre outros. A extração foi realizada diretamente no MySQL Workbench, por meio da importação do arquivo CSV para estruturação em tabela por meio do “Table Data Import Wizard”.

5.2. Transformação

Como a base de dados original estava bem estruturada e padronizada, não foi necessário realizar etapas de limpeza, transformação ou enriquecimento. A fonte de dados estava consistente, sem valores nulos ou duplicados, e com categorias bem definidas. Esta consistência permitiu a transformação direta dos dados em tabelas de dimensão e fato, conforme o modelo estrela.

5.3. Carga

A carga dos dados foi realizada em um novo banco de dados chamado AmeacasGlobaisSegurancaCibernetica, com a seguinte estrutura:

Tabela 2 - Tabelas de Dimensões

dim_country	País afetado pelo incidente
dim_year	Ano do incidente
dim_attack_type	Tipo de ataque cibernético
dim_target_sector	Setor alvo do ataque
dim_financial_loss	Perda financeira resultante do ataque
dim_users_affected	Número de usuários afetados
dim_attack_origin	Origem do ataque
dim_vulnerability_type	Tipo de vulnerabilidade explorada
dim_defense_mechanism	Mecanismo de defesa utilizado
dim_resolution_time	Tempo de resolução do incidente

Tabela 3 - Tabela de Fatos

fact_incident	Consolida os dados quantitativos dos incidentes

A carga dos dados foi realizada por meio de comandos INSERT que inseriram os registros nas tabelas de dimensão e, posteriormente, através de comandos INSERT com JOINs para abastecer a tabela fato com base nos dados da tabela importada (global_cybersecurity_threats_2015-2024).

Os dados transformados foram usados em ferramenta de análise, Power BI, para consulta e visualização. Ao importar os dados limpos e estruturados para o Power BI, foram criados dashboards interativos com indicadores, facilitando a interpretação dos dados.

6. ATIVIDADES DE APOIO DECISÓRIOS AOS NEGÓCIOS E OPERAÇÕES OLAP

As atividades de apoio decisório consistem na aplicação de técnicas de análise de dados com o objetivo de gerar informações relevantes para a formulação de estratégias organizacionais. No contexto da segurança cibernética, essas atividades são fundamentais para antecipar riscos, alocar recursos de forma eficiente e fortalecer a resiliência digital das instituições.

A partir da base de dados "*Global Cybersecurity Threats* (2015–2024)", foi possível definir um conjunto de análises que contribuem diretamente para a tomada de decisões em níveis estratégico e operacional.

O modelo de dados é estruturado em um esquema estrela, , tendo como núcleo a tabela de fatos fact_incident, que consolida as métricas quantitativas dos incidentes.Essa tabela se conecta às tabelas de dimensão, que fornecem o contexto categórico para as análises.

A camada de apresentação do relatório foi estruturada em duas perspectivas analíticas, conforme visuais disponibilizados:

6.1. Visão Estratégica: Página "Overview"

Esta página concentra-se na agregação de alto nível e no monitoramento de indicadores-chave de desempenho (KPIs). Através de cartões informativos, são apresentadas métricas consolidadas como Total Financial Loss, Total Users Affected, Number of Incidents e Average Resolution Time. A aplicação da medida dinâmica permite uma análise flexível de tendências ao longo do tempo, capacitando o usuário a alternar entre diferentes métricas em um mesmo visual para uma avaliação estratégica e comparativa.

6.2. Análise Multidimensional: Página "Detailed Analysis"

Esta seção é o núcleo da aplicação OLAP. A matriz central e os parâmetros de campo na lateral permitem a exploração multidimensional dos dados, viabilizando operações analíticas essenciais:

Slicing e Dicing: a utilização de slicers permite que o usuário filtre e refine o conjunto de dados em qualquer dimensão (por exemplo, analisar incidentes apenas na Alemanha ou no setor de Educação).

Pivotagem e Drill-down: a funcionalidade de parâmetros de campo eleva a análise a um novo patamar. O usuário pode dinamicamente definir a granularidade das linhas da matriz, transformando a visualização de, por exemplo, Total Users Affected por País para Setor Alvo com um simples clique. Essa capacidade de pivotar e aprofundar (drill-down) a análise em múltiplas dimensões é fundamental para identificar padrões, vulnerabilidades e direcionar ações de forma precisa.

6.3. Detalhamento do Modelo Analítico e de Negócios

Relacionamentos do Modelo de Dados:

O modelo de dados é estruturado em um esquema estrela, arquitetura de modelagem mais eficiente e recomendada para análise de dados e aplicações OLAP (Online Analytical Processing).

O núcleo do modelo é a tabela de fatos, fact_incident, que registra as métricas quantitativas de cada incidente. Ela se conecta a todas as tabelas de dimensão, que fornecem o contexto descritivo. Todos os relacionamentos são do tipo muitos para um (* para 1), garantindo que os dados possam ser agregados e filtrados de forma eficiente a partir das dimensões.

dim_country: A tabela fact_incident se relaciona com dim_country através da chave id_country. Isso permite segmentar e analisar incidentes por país.

dim_year: A tabela fact_incident se relaciona com dim_year através da chave id_year. Essencial para analisar as tendências temporais e a evolução dos ataques ano a ano.

dim_attack_type: A tabela fact_incident se relaciona com dim_attack_type através da chave id_attack_type. Isso possibilita analisar as métricas por tipo de ataque (ex: Ransomware, Phishing).

dim_target_sector: A tabela fact_incident se relaciona com dim_target_sector através da chave id_target_sector. Permite a análise do impacto e da frequência dos ataques por setor da indústria.

dim_attack_origin: A tabela fact_incident se relaciona com dim_attack_origin através da chave id_attack_origin. Útil para entender a fonte ou o autor dos ataques.

dim_vulnerability_type: A tabela fact_incident se relaciona com dim_vulnerability_type através da chave id_vulnerability_type. Permite a análise dos ataques com base na vulnerabilidade explorada.

dim_defense_mechanism: A tabela fact_incident se relaciona com dim_defense_mechanism através da chave id_defense_mechanism. Conecta os incidentes ao tipo de defesa que foi utilizado.

Medidas Analíticas (DAX Measures)

As medidas a seguir convertem dados brutos em indicadores-chave de desempenho (KPIs) com relevância estratégica e operacional.

- **Total Financial Loss (Millions USD)**

Propósito: Quantificar o impacto econômico cumulativo dos incidentes de segurança cibernética. Esta medida é crucial para justificar orçamentos de segurança, avaliar riscos e priorizar investimentos em mitigação.

Fórmula: SUM('fact_incident'[financial_loss])

- **Total Users Affected**

Propósito: Medir o alcance e a escala humana dos ataques. O resultado é a soma total de todos os usuários afetados, fornecendo um indicativo do impacto direto sobre a base de clientes ou colaboradores.

Fórmula: SUM('fact_incident'[users_affected])

- **Average Resolution Time**

Propósito: Avaliar a eficiência operacional da resposta a incidentes. Esta medida calcula o tempo médio, em horas, necessário para conter e resolver os ataques. Um tempo de resolução menor indica maior resiliência e preparo da equipe.

Fórmula: AVERAGE('fact_incident'[resolution_time])

- **Number of Incidents**

Propósito: Determinar a frequência dos ataques. Conta o número total de incidentes, servindo como uma métrica fundamental para a análise de tendências ao longo do tempo.

Fórmula: COUNTROWS('fact_incident')

- **Dynamic Measure**

Propósito: Centralizar a análise. Esta medida age como um hub, permitindo que um único visual (como gráficos de barras ou de linha) exiba qualquer uma das métricas principais de

acordo com a seleção do usuário. Isso melhora a interatividade do relatório e otimiza o espaço do dashboard.

Fórmula: Utiliza a função SWITCH para mapear a seleção de um Parâmetro de Campo para a medida correspondente.

- **Chart Title**

Propósito: Aprimorar a experiência do usuário. Esta medida altera o título dos visuais dinamicamente para refletir a medida atualmente selecionada pelo usuário, tornando o relatório intuitivo e autoexplicativo.

Parâmetros de Campo (Field Parameters)

O Parâmetro de Campo é um recurso essencial do seu modelo de dados, responsável por viabilizar a análise interativa e multidimensional.

Nome: Parameter

Função: Permite que o usuário **dynamically** altere a dimensão de análise em visuais como a matriz na página Detailed Analysis. Em vez de fixar a análise por País, o usuário pode "pivotar" a visualização para analisar os dados por Setor, Tipo de Ataque, Origem do Ataque, etc., com um simples clique em um slicer.

Campos Incluídos no Parâmetro:

- País (dim_country[country])
- Setor (dim_target_sector[target_sector])
- Origem do Ataque (dim_attack_origin[attack_origin])
- Tipo de Ataque (dim_attack_type[attack_type])
- Tipo de Vulnerabilidade (dim_vulnerability_type[vulnerability_type])
- Mecanismo de Defesa (dim_defense_mechanism[defense_mechanism])
- Ano (dim_year[year])

7. CONCLUSÃO

A crescente frequência dos incidentes de segurança cibernética tem exigido das organizações uma postura cada vez mais estratégica na proteção digital. Nesse cenário, a análise de dados se consolida como uma ferramenta essencial para compreender o panorama global das ameaças e embasar decisões que visem reduzir riscos, otimizar recursos e fortalecer a resiliência cibernética.

Este projeto teve como foco a exploração da base de dados Global Cybersecurity Threats (2015–2024), com o objetivo de identificar padrões, tendências e fatores críticos relacionados a ataques cibernéticos registrados em diferentes países e setores ao longo de uma década. A partir dessa base, foram definidas atividades de apoio decisório que abrangem desde a identificação dos tipos de ameaças mais recorrentes até a avaliação da eficácia dos mecanismos de defesa utilizados.

Para viabilizar essas análises, foi estruturado um processo completo de Extração, Transformação e Carga (ETL), utilizando ferramentas acessíveis como Microsoft Excel, MySQL Workbench e Power BI. Os dados foram organizados em um modelo dimensional estrela, que permitiu a construção de dashboards interativos e a aplicação de operações OLAP, como slicing, dicing, drill-down e pivotagem.

A definição clara das etapas do ETL, aliada à modelagem eficiente e à visualização analítica, reforça a importância da governança de dados e da qualidade da informação em projetos de ciência de dados aplicados à segurança da informação. Conclui-se, portanto, que a integração entre dados estruturados, ferramentas analíticas e metodologias de apoio à decisão representa um caminho promissor para enfrentar os desafios impostos pelas ameaças cibernéticas globais.

Visando a colaboração acadêmica, o projeto foi publicado publicamente no GitHub <https://github.com/ayrahidasi/PI_SENAC_GRUPO2_AMEACAS_GLOBAIS_A_SEGURANCA_CIBERNETICA>. O repositório contém todos os arquivos relacionados ao desenvolvimento do projeto, incluindo os scripts SQL de criação e inserção de dados, a estrutura do modelo dimensional, a fonte de dados original em CSV, e instruções para replicação do processo ETL e visualizações no Power BI.

8. REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **NBR ISO/IEC 17799:2005**: Tecnologia da informação e comunicação: práticas para a gestão de segurança da informação. Rio de Janeiro, 2005.

BUENO, P. H. M. **Avaliação das Estruturas de Segurança Cibernética**. Universidade de Brasília (UnB). Brasília, 2021. Disponível em: <https://bdm.unb.br/bitstream/10483/28937/1/2021_PauloHenriqueMendoncaBueno_tcc.pdf> . Acesso em 21/05/2025.

KIMBALL, Ralph; ROSS, Margy. **The Data Warehouse Toolkit**: The Definitive Guide to Dimensional Modeling. 3. ed. New York: Wiley, 2013. ISBN 9781118530801.

MACHADO, Á. M. **Construção de um Processo ETL Automatizado em Dados de Campanhas de uma Empresa no Setor Bancário**. Universidade Federal de Uberlândia, 2023.

NOGUEIRA, M., BORGES, L. F., NEIRA, A. B., ALBANO, L., & COELHO, K. K. **Ciência de Dados Aplicada à Cibersegurança: Teoria e Prática.** Disponível em: < https://www.researchgate.net/publication/384533168_Ciencia_de_Dados_Aplicada_a_Cibersegurança_Teoria_e_Pratica>. Acesso em 21/05/2025.

ZORZO, A. L. **ETL 2.0**: Uma proposta de extensão ao processo de extração, transformação e carga voltada à integração de dados estruturados e não estruturados. Universidade Federal de Santa Catarina, 2009. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/184522/Projeto_Andre_Zorzo.pdf?sequence=-1>. Acesso em 21/05/2025.