

UNIVERSIDADE NOVE DE JULHO
DIRETORIA DOS CURSOS DE INFORMÁTICA

NAIRA ALVES FERREIRA SILVA - RA 2223200384
WESLEY MESQUITA LOPES - RA 2223203032
PAULO HICARO SILVA NUNES AMORIN - RA 2223202942
IELMA MARTINS DOS SANTOS - RA 2223204028
GABRIELLE UCHOA DE CASTRO MOREIRA - RA 2223203550
CLEYTON DOS SANTOS COSTA - RA 2222103392
MARCELA SIANA SILVA - RA 2222103070
GABRIEL FERREIRA - RA 2222106567

PLANEJAMENTO EM INFORMÁTICA:
PINGTECH

PLANEJAMENTO EM INFORMÁTICA:
PINGTECH

Trabalho apresentado à Universidade Nove de Julho, UNINOVE, em cumprimento parcial às exigências da disciplina de Planejamento em Informática, sob orientação do Prof. **Felipe Santos de Jesus**

SÃO PAULO
2024

RESUMO

Esse é o projeto dado para os alunos, no 1 semestre de 2024, com o intuito de fazer com que os alunos entendessem mais sobre como funciona uma empresa. O projeto tem como prioridade sua maior parte, relacionada na área de TI, com base no curso que estamos a estudar. O intuito deste trabalho é criar uma empresa, a critério dos alunos, e implantar técnicas de tecnologia análise e desenvolvimento de software que faça com que ela funcione sem grandes riscos.

Foi criada uma empresa de vendas de produtos eletroeletrônicos, chamada "PINGTECH", que tem prioridade na venda de produtos online. Dentro do contexto foi feita métricas para um bom funcionamento da empresa, tais com medidas de segurança, técnicas de machine learning, a escolha dos periféricos de redes (computadores, classe de redes etc.), uma modelagem de dados da empresa, e como lidar com grande número de dados.

Por fim o projeto entrega bastante conhecimento, não só para quem fez, mas acredito que para quem está lendo também.

SUMÁRIO

OBJETIVOS.....	5
ESTRUTURA INTERNA DA EMPRESA.....	6
APRENDIZADO DE MÁQUINA.....	8
CIÊNCIA DE DADOS.....	12
ESTRUTURA FÍSICA DA EMPRESA.....	23
MODELAGEM DE DADOS.....	25
SEGURANÇA DA INFORMAÇÃO.....	34
REFERÊNCIAS.....	41
CONCLUSÃO.....	42

OBJETIVOS

PINGTECH/eletrônicos em geral

A PINGTECH foi criada em 2024, apesar de ser uma loja nova não é insuficiente, trazendo o melhor da tecnologia atual com um preço que cabe no seu bolso. Nossos profissionais muito bem capacitados, para melhor atender nossos clientes, farão com que você tenha uma experiência única em nossa loja. Entrega rápida e com rastreamento, produtos bem embalados para que chegue sem danos em sua casa. Então seja BEM-VINDO e BOAS COMPRAS.

Os serviços que a empresa oferecerá e as tecnologias que serão utilizadas:

Entrega local feita por empresas terceirizadas;

Entrega para outros estados feita pelos Correios;

Site de fácil acesso;

Segurança com dados de clientes;

informações seguras;

SAC. (e-mail, telefone, site, WhatsApp.);

Nossa empresa tem como objetivo principal a venda online, por conta disso a segurança dos dados de nossos clientes esta como prioridade. por conta disso será implementado medidas de segurança para o acesso do mesmo seja limitado a contribuintes específicos que saibam lidar com a segurança de dados.

Entender como funciona uma empresa, é crucial para que tais medidas possam ser incrementadas dentro do negócio, portanto estaremos abertos a sugestões para melhor funcionamento de nossa empresa.

ESTRUTURA INTERNA DA EMPRESA

Este projeto apresenta as atividades realizadas em um projeto abrangente de ADS, que cobriu vários tópicos fundamentais na área. O projeto envolveu o estudo e a aplicação prática de:

- Aprendizado de Máquina
- Ciência de Dados
- Modelagem de Dados
- Redes de Computadores
- Segurança da Informação

Atividades Realizadas

Aprendizado de Máquina

- Estudo de diferentes algoritmos de aprendizado de máquina, incluindo regressão e agrupamento.
- Implementação de modelos de aprendizado de máquina para resolver problemas possíveis problemas da empresa.

Ciência de Dados

- Coleta e preparação de dados para análise.
- Exploração e visualização de dados para identificar padrões e tendências.
- Utilização de técnicas estatísticas e de aprendizado de máquina para extrair insights valiosos dos dados.

Modelagem de Dados

- Criação de modelos de dados relacionais e não relacionais.
- Normalização e otimização de modelos de dados para garantir integridade e desempenho.

Redes de Computadores

- Estudo da arquitetura e protocolos de redes de computadores.
- Configuração e gerenciamento de redes locais (LANs) e redes de longa distância (WANs).
- Análise de desempenho de rede e solução de problemas.

Segurança da Informação

- Identificação e avaliação de ameaças à segurança da informação.
- Implementação de medidas de segurança, como criptografia, firewalls e sistemas de detecção de intrusão.
- Auditoria de sistemas de informação para garantir conformidade e segurança.

Resultados e Impacto

O projeto forneceu aos participantes uma compreensão abrangente dos fundamentos da área de Análise e desenvolvimento de Sistemas e suas aplicações práticas em vários domínios. As atividades realizadas contribuíram para:

Desenvolvimento de habilidades em técnicas de aprendizado de máquina e ciência de dados.

Melhoria da capacidade de projetar e implementar modelos de dados eficientes.

Fortalecimento das competências em redes de computadores e segurança da informação.

Preparação dos participantes para carreiras em Ciência da Computação e áreas relacionadas.

1. APRENDIZADO DE MÁQUINA

1.1 Exploração de Dados e Pré-processamento:

Coleta de Dados:

O Dataset utilizado nesse projeto foi retirado do site:

<https://www.kaggle.com/datasets/datacertlaboratoria/projeto-3-segmentao-de-clientes-no-ecommerce>

Lista de variáveis/features incluídas:

- N° da fatura: Um número atribuído exclusivamente a cada transação. Se começar com "c", indica uma substituição.
- Data da fatura: Data e hora da fatura, dia e hora em que a transação foi gerada.
- ID Cliente: Número de clientes. Um número integral de 5 dígitos atribuído exclusivamente a cada cliente.
- País: O nome do país onde cada cliente reside.
- Quantidade: As quantidades de cada produto (item) por transação.
- Valor: Valor da fatura em libras esterlinas.

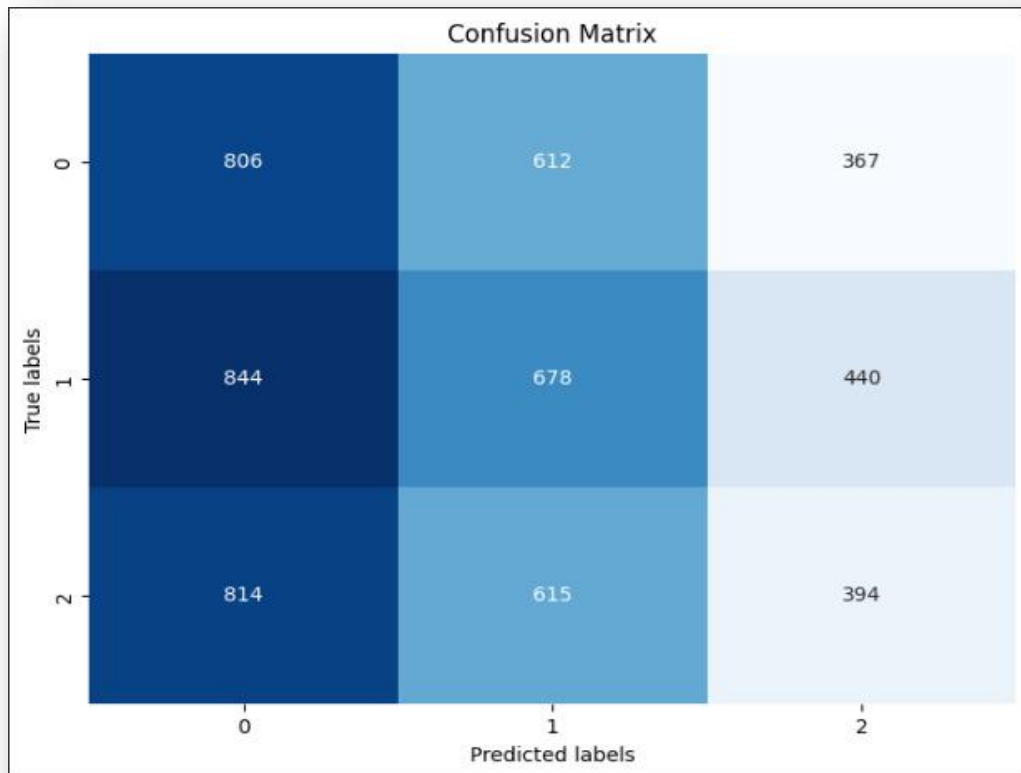
Limpeza e Pré-processamento:

Primeiramente foi retirado registros nulos da coluna "ID Cliente", valores em que não era possível a identificação do cliente. Logo após a remoção de dados duplicados do dataset caso houvesse.

Após foi feita a padronização da coluna "N° da fatura", para remover dados começados com "C". Depois foi alterado os tipos de algumas variáveis. Por exemplo: object -> datetime64[ns]. Também a alteração ',' para '.'. E com isso foi adicionada mais uma coluna denominada "Ano mês".

Matriz Confusão:

Com a conclusão da limpeza dos dados, foi realizado a matriz confusão, para verificar o desempenho de compras de clientes em um determinado período de 2024. Gerando esse gráfico:



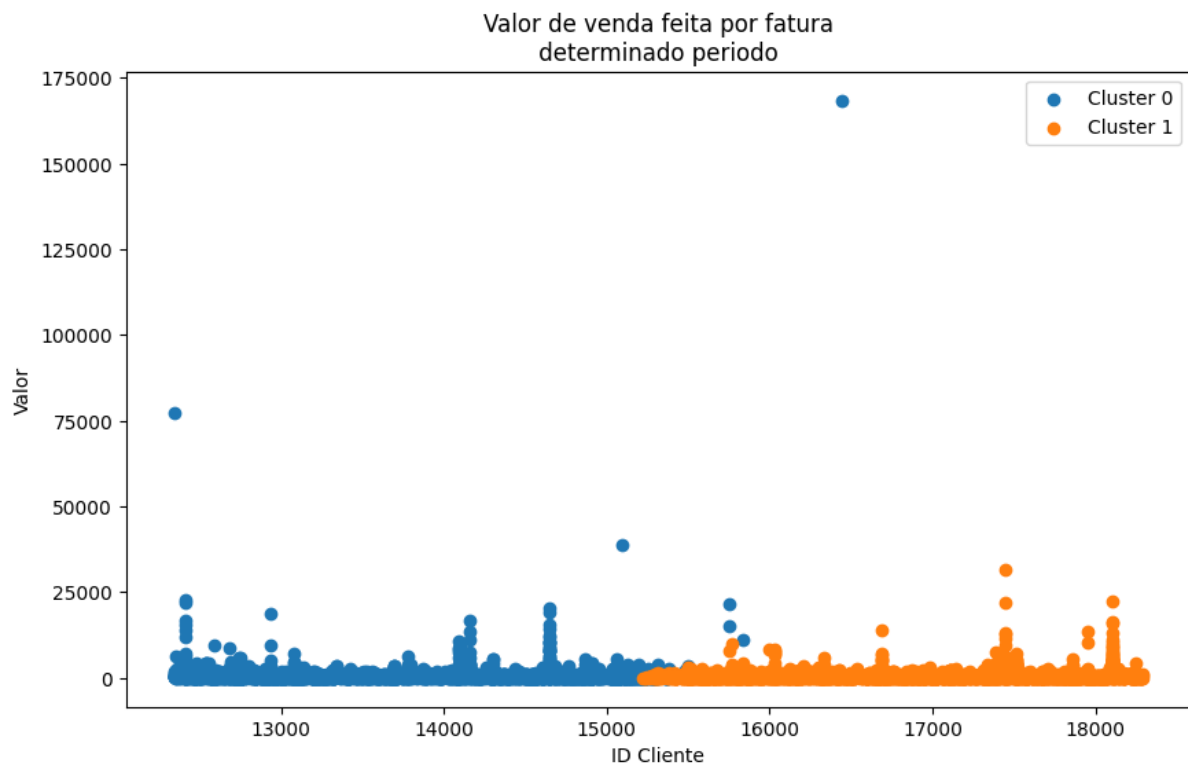
1.2 Implementação de Modelos de Aprendizado de Máquina:

Escolha de Algoritmos:

Nesta etapa foi feita a utilização do algoritmo do KMeans, para a geração de um gráfico de fácil entendimento. Relacionado a vendas por fatura.

Implementação:

Primeiro passo foi fazer a importação das bibliotecas apropriadas para o algoritmo "KMeans", depois foi adicionado o dataset na pasta, para ser utilizado dentro do algoritmo em questão. Após essa etapa temos que dar início ao treinamento do dataset, foi utilizado dois "2" clusters para ser visualizado no gráfico, gerando resultados dos centros dos clusters. E por último foi gerado o gráfico com o resultado do centro do cluster visivelmente espalhado e colorido dentro do gráfico.



1.3 Otimização e Validação do Modelo:

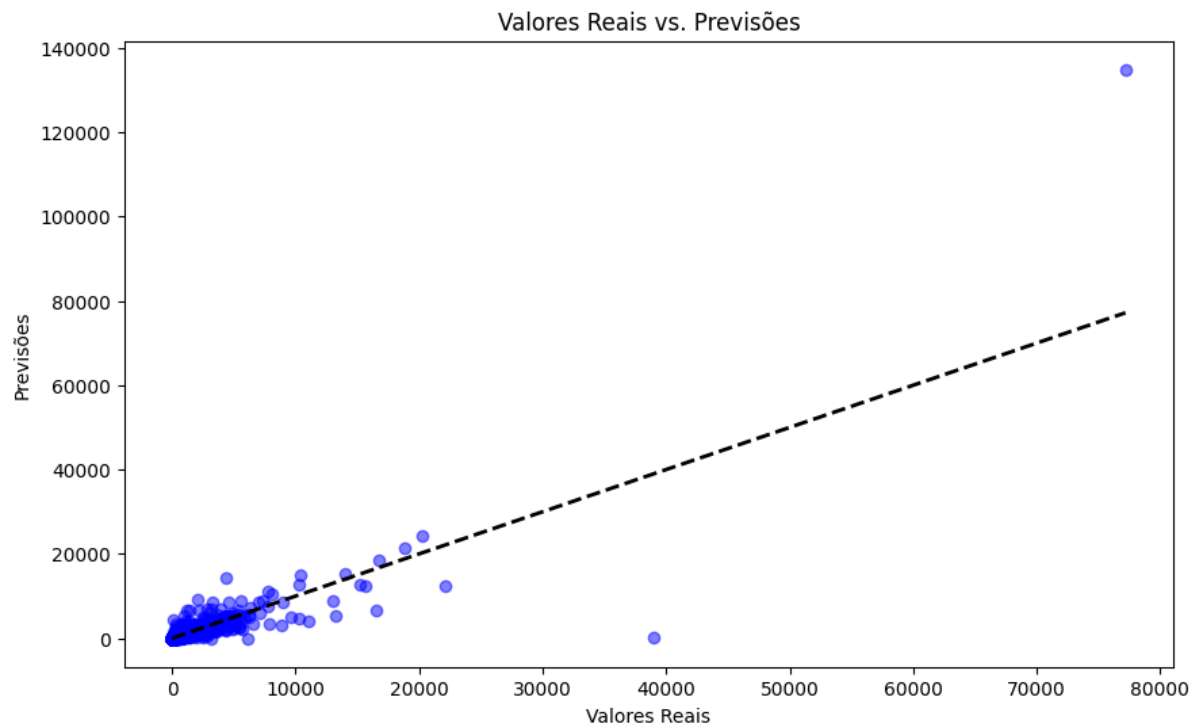
Otimização de Hiper parâmetros:

Nesta etapa i utilizado o algoritmo de "LinearRegression", fazendo a prevenção de futuras vendas, de acordo com as vendas já efetuadas.

Lista dos hiper parâmetros ajustados.

- Valor
- ID Cliente
- Data da fatura
- País
- Ano mês

gerando um gráfico que mostra as previsões das vendas:



Validação Cruzada:

Os parâmetros utilizados para fazer a validação cruzada foi o algoritmo de "LinearRegression", e utilizando os mesmos parâmetros.

Resultados obtidos:

Esses foram os resultados obtidos:

Validação cruzada:

[0.53619067, 0.680733, 0.57093758, 0.6932718, 0.81977309]

2.CIÊNCIA DE DADOS

Análise Descritiva dos Dados

A análise descritiva dos dados é uma etapa fundamental em qualquer projeto de análise de dados. Vamos explorar alguns pontos importantes para análise de dados:

2.1 Técnicas Estatísticas Básicas para Descrever os Dados:

- A análise descritiva envolve calcular medidas estatísticas que resumem as características dos dados. Alguns exemplos incluem:
 - Média (ou valor médio): Calcula a média dos valores em uma variável.
 - Mediana: Encontra o valor central dos dados.
 - Desvio padrão: Mede a dispersão dos dados em torno da média.
 - Quartis: Divide os dados em quartis (25%, 50% e 75%).
- Essas estatísticas nos ajudam a entender a distribuição dos dados e identificar possíveis outliers.

2.2 Visualização de Dados com Matplotlib e Seaborn:

- O Matplotlib e o Seaborn são bibliotecas populares em Python para criar gráficos e visualizações.
- Com o Matplotlib, você pode criar gráficos de linhas, barras, dispersão, histogramas e muito mais.
- O Seaborn é uma extensão do Matplotlib que oferece estilos predefinidos e funções mais simples para criar gráficos estatísticos, como gráficos de caixa e violino.
- Essas bibliotecas nos permitem visualizar padrões, tendências e relações nos dados.

2.3 Identificação de Padrões e Tendências nos Dados:

- Ao analisar os dados, procuramos por padrões e tendências que possam nos fornecer insights. Isso pode incluir:

- Identificar sazonalidades em séries temporais.
 - Encontrar correlações entre variáveis.
 - Detectar clusters ou agrupamentos.
 - Observar mudanças ao longo do tempo.
- Gráficos e estatísticas descritivas nos ajudam a identificar esses padrões.

Com tudo, podem acontecer alguns erros. Como por exemplo:

overfitting

Overfitting é quando, após o treino, o teste e a validação, seja feita ele demonstra padrões muito corretos, fazendo que as próximas validações tenham seus resultados sempre semelhantes ou iguais. Isso ocorre por alguns motivos tais como, poucos dados para teste, outra causa possa ser que haja muitos parâmetros trazendo uma complexidade para tais modelos.

Para saber se o modelo está em overfitting, pode ser quando há muitos parâmetros, fazendo com que o modelo memorize padrões gerando previsões erradas, alto desempenho e baixo desempenho também é uma questão de overfitting pois também podem gerar padrões errados.

Métricas para evitar o overfitting, é ter uma boa base de dados grande o suficiente, e manter o modelo com poucos parâmetros para ser treinado, outras técnicas que podem ser utilizadas são: regularização, validação cruzada, seleção de recursos ou parar o treinamento antecipadamente.

Modelagem Estatística

Nesta etapa usaremos, algumas técnicas, em "Phyton", para fazer validações preditivas para modelos de banco de dados. Como por exemplo regressão linear, classificação de dados, estatísticas e previsões.

2.4 Aplicação de técnicas estatísticas avançadas para modelagem dos dados:

A regressão Linear é um dos modelos preditivos mais simples. Seu objetivo é prever um determinado valor Y baseado em um valor X.

```
pip install pandas
```

```
pip install -U scikit-learn
```

```
import pandas as pd
```

```
from sklearn.linear_model import LinearRegression
```

```
from google.colab import files
```

```
uploaded = files.upload()
```

```
filename = next(iter(uploaded))
```

```
df = pd.read_csv(filename)
```

```
X = df[['area']]
```

```
y = df['price']
```

```
modelo = LinearRegression()
```

```
modelo.fit(X, y)
```

```
area = 500 # Exemplo de área em sqft
```

```
preço_previsto = modelo.predict([[area]])
```

```
print(f"O valor aproximado da casa com uma área de {area} sqft é de  
${preço_previsto[0]:.2f}")
```

2.5 Uso de ferramentas como regressão linear, classificação etc.

Regressão Linear Simples

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
from sklearn.linear_model import LinearRegression
```

```
x = np.array([1, 2, 3, 4, 5])
```

```
y = np.array([2, 4, 5, 4, 5])
```

```
model = LinearRegression()
```

```
model.fit(x.reshape(-1, 1), y)
```

```
beta_0 = model.intercept_
```

```
beta_1 = model.coef_[0]
```

```
plt.scatter(x, y, label="Dados")
```

```
plt.plot(x, model.predict(x.reshape(-1, 1)), color='red', label="Regressão Linear")
```

```
plt.xlabel("Variável Independente (x)")
```

```
plt.ylabel("Variável Dependente (y)")
```

```
plt.title("Regressão Linear Simples")
```

```
plt.legend()
```

```
plt.show()
```

```
print(f"Equação da reta:  $y = \{beta\_0:.2f\} + \{beta\_1:.2f\}x$ ")
```

Classificação:

```
from sklearn.datasets import load_iris
```

```
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
```

```
iris = load_iris()
```

```
X, y = iris.data, iris.target
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
```

```
k = 3
```

```
knn_model = KNeighborsClassifier(n_neighbors=k)
```

```
knn_model.fit(X_train, y_train)
```

```
accuracy = knn_model.score(X_test, y_test)
```

```
print(f"Acurácia do modelo k-NN: {accuracy:.2f}")
```

2.6 Avaliação da adequação dos modelos estatísticos aos dados.

Calculando o valor de R2 utilizando a biblioteca SkLearn

```
import pandas as pd
```

```
import numpy as np
```

```
dados = pd.read_csv('Calc_R2.csv')
```

```
Y = dados['Y'].values
```

```
Y1 = dados['Y1'].values
```

```
Y2 = dados['Y2'].values
```

```
Y3 = dados['Y3'].values
```



```
Y4 = dados['Y4'].values
```

```
Y5 = dados['Y5'].values
```

```
from sklearn.metrics import r2_score
```

```
r2_1 = r2_score(Y,Y1)
```

```
r2_2 = r2_score(Y,Y2)
```

```
r2_3 = r2_score(Y,Y3)
```

```
r2_4 = r2_score(Y,Y4)
```

```
r2_5 = r2_score(Y,Y5)
```

```
print("R2 (Grau 1) = ",r2_1)
```

```
print("R2 (Grau 2) = ",r2_2)
```

```
print("R2 (Grau 3) = ",r2_3)
```

```
print("R2 (Grau 4) = ",r2_4)
```

```
print("R2 (Grau 5) = ",r2_5)
```

```
R2 (Grau 1) = 0.7712730521486528
```

```
R2 (Grau 2) = 0.8003177439393002
```

```
R2 (Grau 3) = 1.0
```

```
R2 (Grau 4) = 1.0
```

```
R2 (Grau 5) = 1.0
```

2.7 Implementação de modelos preditivos utilizando Python.

Um modelo para prever o preço de casas com base em suas áreas.

Usamos uma lista fictícia de áreas e preços de casas.

Dividimos os dados em conjuntos de treinamento e teste.

Treinamos um modelo de regressão linear usando a área como variável independente e o preço como variável dependente.

Fizemos previsões para os dados de teste.

Avaliamos o modelo usando o erro médio quadrático.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error

area = [100, 150, 200, 120, 180, 160, 140, 210]
preco = [250000, 350000, 450000, 280000, 400000, 380000, 320000, 480000]

df = pd.DataFrame({'Área': area, 'Preço': preco})

X = df[['Área']]
y = df['Preço']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

modelo = LinearRegression()
modelo.fit(X_train, y_train)

y_pred = modelo.predict(X_test)

erro_medio_quadratico = mean_squared_error(y_test, y_pred)
```

```
print(f"Erro médio quadrático: {erro_medio_quadratico:.2f}")
```

```
plt.scatter(X_test, y_test, color='blue', label='Dados reais')  
plt.plot(X_test, y_pred, color='red', label='Predições')  
plt.xlabel('Área (m²)')  
plt.ylabel('Preço')  
plt.title('Modelo de Regressão Linear para Preço de Casas')  
plt.legend()  
plt.show()
```

2.8 Avaliação da performance dos modelos preditivos:

O que é o Erro Médio Quadrático (MSE)?

O MSE é calculado como a média dos quadrados das diferenças entre os valores previstos e os valores reais.

Quanto menor o valor do MSE, melhor o desempenho do modelo, pois indica que as previsões estão mais próximas dos valores reais.

Exemplo:

Usamos uma variável independente X com valores de 1 a 5.

Os valores verdadeiros da variável dependente y_verdadeiro são [2, 4, 5, 7, 8].

O modelo de regressão linear é treinado com base nesses dados.

As previsões são feitas para os mesmos valores de X.

O MSE é calculado e exibido.

```
import numpy as np  
import matplotlib.pyplot as plt  
from sklearn.linear_model import LinearRegression  
from sklearn.metrics import mean_squared_error
```

```
X = np.array([1, 2, 3, 4, 5]).reshape(-1, 1) # Variável independente (amostras)
y_verdadeiro = np.array([2, 4, 5, 7, 8])    # Variável dependente (valores
verdadeiros)
```

```
modelo = LinearRegression()
modelo.fit(X, y_verdadeiro)
previsoes = modelo.predict(X)
```

```
erro_medio_quadratico = mean_squared_error(y_verdadeiro, previsoes)
print(f'Erro Médio Quadrático: {erro_medio_quadratico:.2f}')
```

2.9 Comparação entre diferentes abordagens de análise preditiva.

Análise Descritiva:

Suponha que temos um conjunto de dados com as vendas mensais de um produto. Queremos calcular a média, a variância e criar um gráfico de barras para visualizar as vendas ao longo do tempo.

Exemplo:

```
import numpy as np
import matplotlib.pyplot as plt
```

```
vendas_mensais = [8000, 9500, 10500, 9000, 11000, 8500]
```

```
media_vendas = np.mean(vendas_mensais)
print(f"Média de vendas: ${media_vendas:.2f}")
```

```
variancia_vendas = np.var(vendas_mensais)
print(f"Variância das vendas: ${variancia_vendas:.2f}")
```

```
meses = ["Jan", "Fev", "Mar", "Abr", "Mai", "Jun"]
plt.bar(meses, vendas_mensais)
plt.xlabel("Mês")
plt.ylabel("Vendas ($)")
plt.title("Vendas Mensais")
plt.show()
```

Análise Preditiva:

Agora, vamos prever as vendas para o próximo trimestre usando regressão linear. Suponha que temos dados históricos e queremos prever as vendas para os próximos 3 meses:

```
from sklearn.linear_model import LinearRegression
```

```
X = np.arange(1, 7).reshape(-1, 1) # Mês (1 a 6)
y = np.array(vendas_mensais)
```

```
modelo = LinearRegression()
modelo.fit(X, y)
```

```
# Previsão para os próximos 3 meses
```

```
proximos_meses = np.arange(7, 10).reshape(-1, 1)
previsao_vendas = modelo.predict(proximos_meses)
print(f"Previsão de vendas para o próximo trimestre: ${previsao_vendas.sum():.2f}")
```

Análise Prescritiva:

Digamos que a previsão para o próximo trimestre seja inferior à meta desejada. Podemos sugerir uma ação para melhorar as vendas:

```
meta_desejada = 35000
```

```
diferenca = meta_desejada - previsao_vendas.sum()
```

```
if diferenca > 0:
```

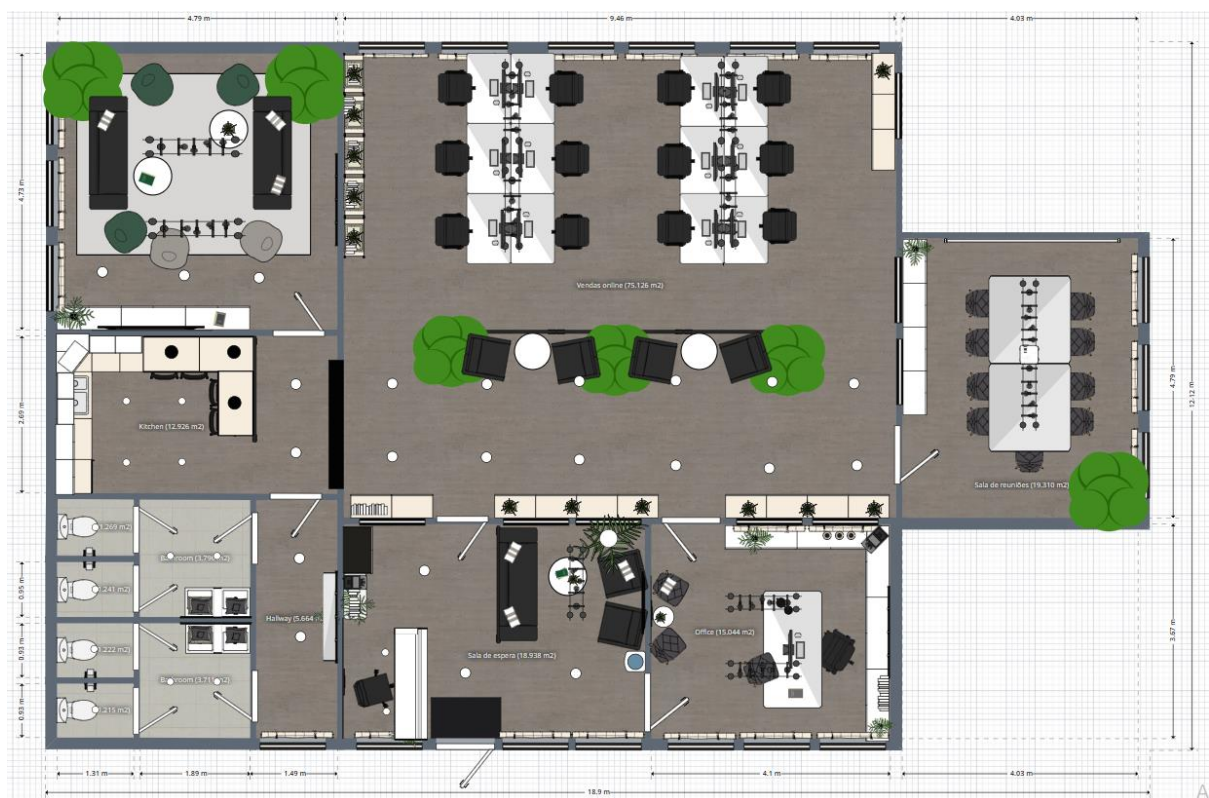
```
    print(f"Sugestão: Lançar uma campanha de marketing direcionada a clientes  
    existentes.")
```

```
    print(f"Ação recomendada: Oferecer descontos especiais para compras acima de  
    ${meta_desejada:.2f}.")
```

```
else:
```

```
    print("As previsões já atendem à meta desejada. Nenhuma ação necessária.")
```

3.REDES DE COMPUTADORES



Planta baixa da "PINGTECH", e seus departamentos separados.

3.1 Departamentos:

- Vendas

- Ti
- Finanças
- Recursos Humanos
- Sala de reunião

3.2 Equipamentos:

- Notebook Inspiron 15 com Intel® Core™ i5-1235U (12ª geração)
- placa de vídeo Intel® Iris® Xe Graphics
- 16 GB DDR4
- 512 GB SSD

3.3 Classe de rede

Classe de rede C:

192.168.1.0, sub rede: 255.255.255.0

O roteador é 192.168.1.1, que é o gateway padrão.

3.4 Padrão de rede de cada departamento:

- Vendas: 192.168.1.2 e 192.168.1.12 (DHCP) - 192.168.1.13, estático.
- Ti: 192.168.1.14 - DHCP e 192.168.1.15 - Estático.
- Finanças: 192.168.1.16, estático.
- RH: 192.168.1.17, estático.
- Sala de reunião: 192.168.1.18, estático.

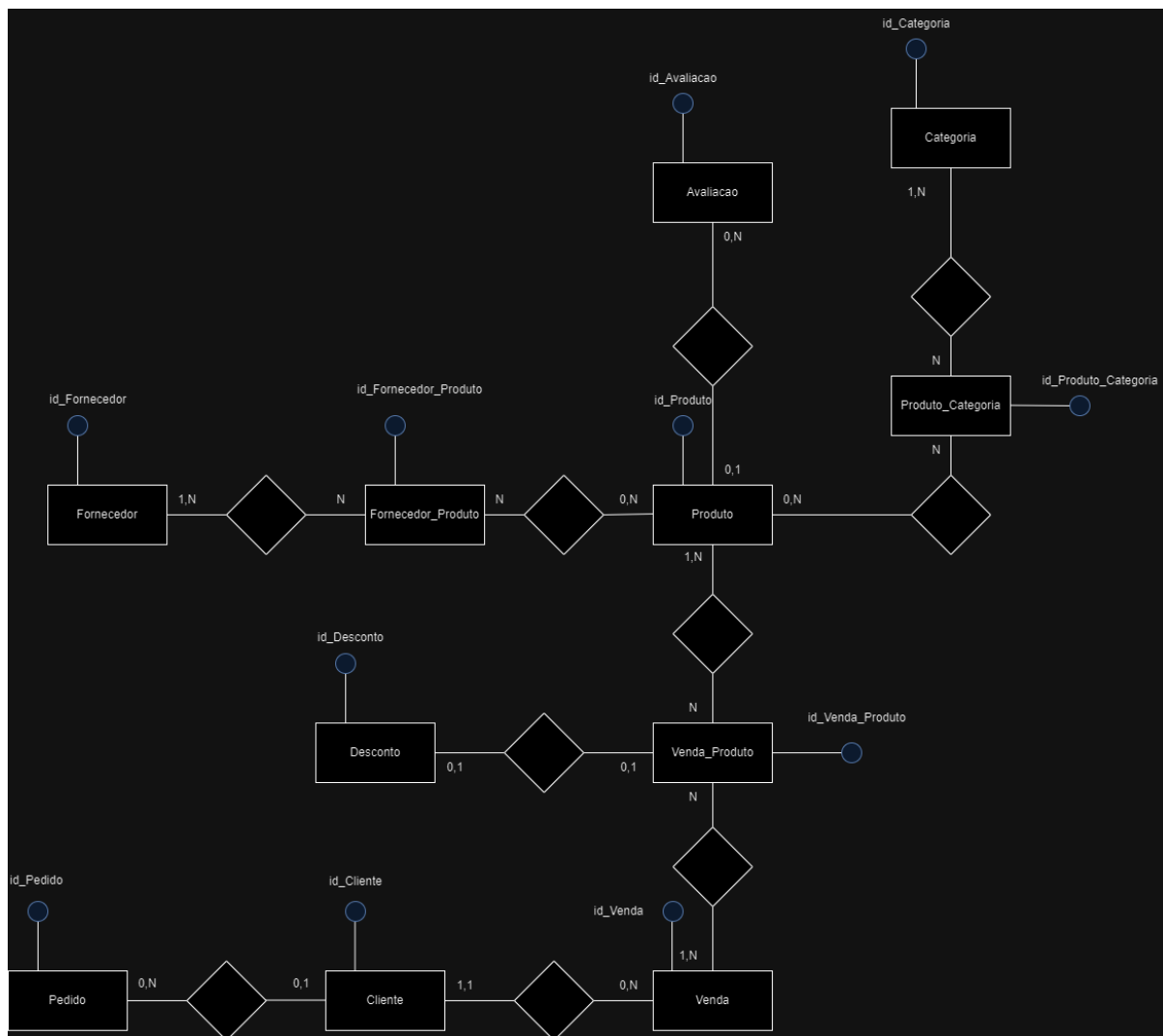
Obs.: O gateway padrão de cada departamento deve ser o IP do roteador.

4.MODELAGEM DE DADOS

4.1 Entidades e atributos:

- Produto - (id_Produto, Produto_Nome, quantidade, custo, preco, descricao, id_Fornecedor, id_Avaliacao)
- Categoria - (id_Categoria, descricao)
- Avaliacao - (id_Avaliacao, tipo)
- Produto_Categoria - (id_Produto_Categoria, id_Produto, id_Categoria)
- Fornecedor - (id_Fornecedor, nome, telefone, email)
- Fornecedor_Produto - (id_Fornecedor_Produto, id_Produto, id_Fornecedor)
- Venda_Produto - (id_Venda_Produto, id_Produto, id_Venda)
- Desconto - (id_desconto, nome, valor, data, status)
- Venda - (id_Venda, data, Valor_Total, Forma_Pagamento, id_Desconto, parcelas, id_Cliente, id_Pedido)
- Pedido - (id_Pedido, Numero_Pedido, status, valor, data, id_Produto)
- Cliente - (id_Cliente, Cliente_Nome, email, senha, CPF, telefone, endereco, id_Desconto, id_Pedido)

Modelo Entidade Relacionamento-MER do sistema da empresa:



Temos a tabela Produto no centro do diagrama, onde tem uma ligação direta com mais 4 entidades, sendo elas Avaliacao, Produto_Categoria, Fornecedor_Produto e Venda_Produto. Sendo essas 3 últimas uma relação N para N.

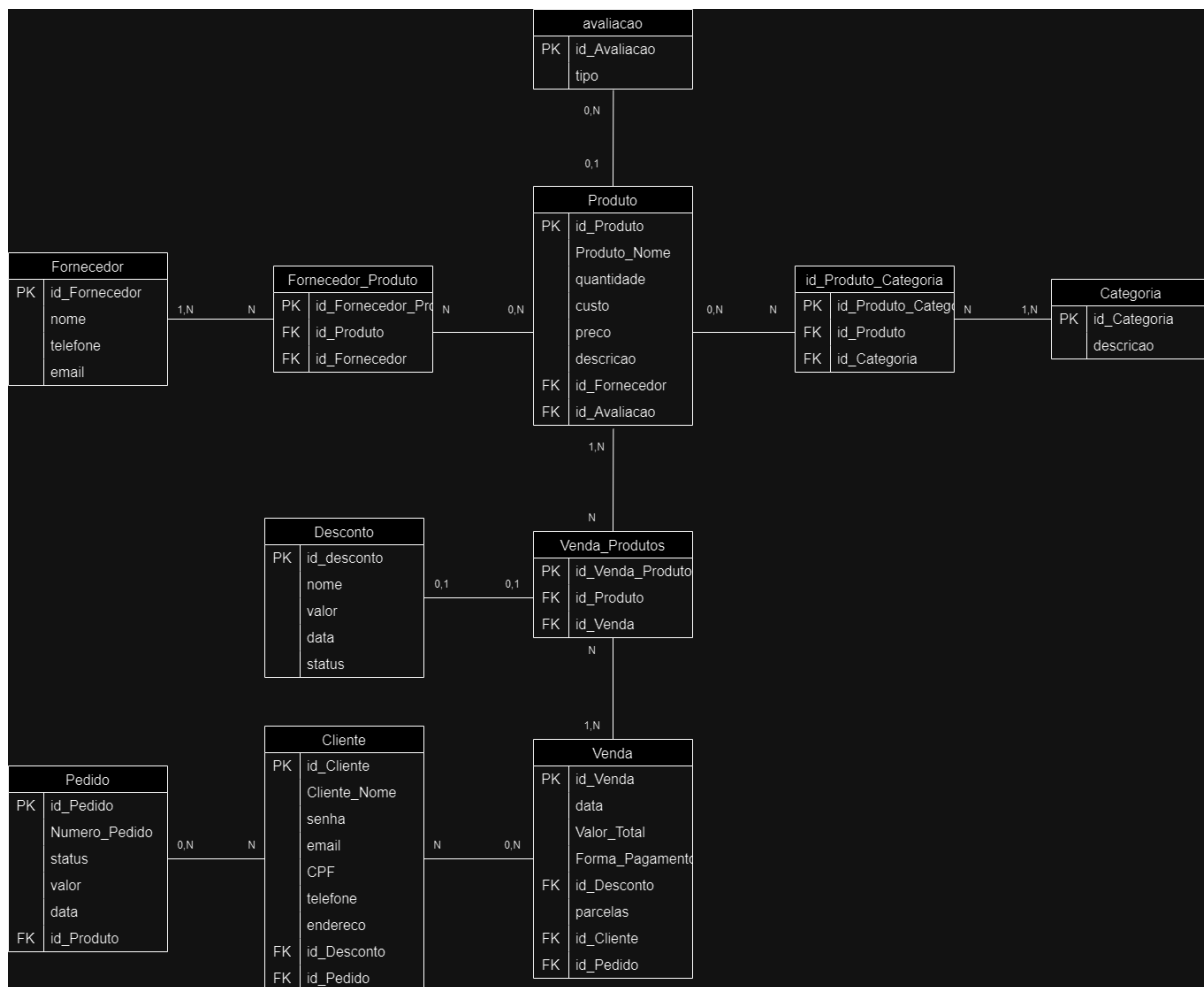
Produto_Categoria foi gerada da relação das entidades Categoria e Produto. Já a entidade Fornecedor_Produto, foi da ligação das tabelas Fornecedor e Produto. E a relação das tabelas Venda e Produto, gerou a nova tabela Venda_Produto. Essa última ainda faz ligação com a tabela Desconto.

Por fim, a tabela Venda, se relaciona com a tabela Cliente. E a tabela Cliente está ligada a entidade Pedido.

As cardinalidades estão em seus respectivos relacionamentos

4.2 Diagrama Entidade Relacionamento-DER:

Com base nos dados do modelo conceitual, abaixo temos a representação do modelo lógico do banco de dados do site.



No diagrama de Entidade Relacionamento Lógico, está organizado da mesma maneira do diagrama de entidade Conceitual.

Em cada tabela foi estabelecido as Chaves Primarias (PK) e Chaves Estrangeiras (FK) quando necessário.

As cardinalidades também estão em seus respectivos relacionamentos.

4.3 Dicionário de dados das tabelas do site:

O "tipo" listado abaixo foi usado como referência o SQL

Produto

Atributo	Tipo	Descrição	Chave
id_Produto	serial	Identificado único do produto	Primaria
Produto_Nome	varchar(200)	Nome único do produto	-
quantidade	int	Quantidade do produto	-
custo	money	Valor pago ao fornecedor pelo produto em reais (R\$)	-
preco	money	Preço de revenda do produto em reais (R\$)	-
descricao	varchar(200)	breve descrição do produto	-
id_Fornecedor	int	Chave estrangeira da tabela Fornecedor	Estrangeira
id_Avaliacao	int	Chave estrangeira da tabela Avaliação	Estrangeira

Categoria

Atributo	Tipo	Descrição	Chave
id_Categoria	serial	Identificado único da categoria	Primaria
descricao	varchar(100)	decrição da categoria (periféricos, acessórios, celulares)	-

Avaliação

Atributo	Tipo	Descrição	Chave
id_Avaliacao	serial	Identificado único da avaliação	Primaria
tipo	varchar(200)	campo para receber avaliações (positivas, negativas ou neutras)	-

Produto_Categoria

Atributo	Tipo	Descrição	Chave
id_Produto_Categoria	serial	Identificado único da tabela Produto_Categoria	Primaria
id_Produtos	int	Chave estrangeira da tabela Produtos	Estrangeira
id_Categoria	int	Chave estrangeira da tabela Categoria	Estrangeira

Fornecedor

Atributo	Tipo	Descrição	Chave
id_Fornecedor	serial	Identificado único do fornecedor	Primaria
nome	varchar(100)	Nome do fornecedor	-
telefone	bigint	Telefone do fornecedor	-
email	varchar(200)	Email do fornecedor	-

Fornecedor_Produto

Atributo	Tipo	Descrição	Chave
id_Fornecedor_Produtos	serial	Identificado único da tabela Fornecedor_Produtos	Primaria
id_Produtos	int	Chave estrangeira da tabela Produtos	Estrangeira
id_Fornecedor	int	Chave estrangeira da tabela Fornecedor	Estrangeira

Venda_Produtos

Atributo	Tipo	Descrição	Chave
id_Venda_Produtos	serial	Identificado único da tabela Venda_Produtos	Primaria
id_Produtos	int	Chave estrangeira da tabela Produtos	Estrangeira
id_Venda	int	Chave estrangeira da tabela Venda	Estrangeira

Desconto

Atributo	Tipo	Descrição	Chave
id_Desconto	serial	Identificado único do desconto	Primaria
nome	varchar(100)	Nome atribuido ao desconto	-
valor	money	Valor do desconto	-
data	date	Data em que o desconto foi aplicado	-
status	int	status do desconto (vencido, utilizado)	-

Venda

Atributo	Tipo	Descrição	Chave
id_Venda	serial	Identificado único da venda	Primaria
data	date	Data em que a venda foi realizada	-
Valor_Total	money	Valor total da venda	-
Forma_Pagamento	varchar(100)	Forma de pagamento escolhida na venda (Crédito, PIX, boleto)	-
id_Desconto	int	Chave estrangeira da tabela Desconto	Estrangeira
parcelas	int	Quantidade de parcelas da venda (1x, 2x... 10x)	-
id_Cliente	int	Chave estrangeira da tabela Cliente	Estrangeira
id_Pedido	int	Chave estrangeira da tabela Venda	Estrangeira

Pedido

Atributo	Tipo	Descrição	Chave
id_Pedido	serial	Identificado único do pedido	Primaria
Numero_Pedido	int	Numero único do pedido para ser encaminhado ao cliente	-
status	int	Status do pedido (em preparação, em trânsito, entregue)	-
valor	money	Valor do pedido	-
data	date	Data em que o pedido foi realizado	-
id_Produtos	int	Chave estrangeira da tabela Produtos	Estrangeira

Cliente

Atributo	Tipo	Descrição	Chave
id_Cliente	serial	Identificado único do cliente	Primaria
Cliente_Nome	varchar(200)	Nome do cliente	-
email	varchar(200)	E-mail do cliente	-
senha	int	Senha do cadastro do cliente	-
CPF	int	CPF do cliente	-
telefone	bigint	Telefone do cliente	-
endereco	varchar(200)	Endereço do cliente	-
id_Desconto	int	Chave estrangeira da tabela Desconto	Estrangeira
id_Pedido	int	Chave estrangeira da tabela Pedido	Estrangeira

No dicionário de dados, foi preenchido o tipo de cada atributo, sua descrição/função e sua chave, seja ela primária ou estrangeira.

4.4 Simulação de cadastro de cada uma das tabelas criadas:

Produto

id_Produto	Produto_Nome	quantidade	custo	preco	descricao	id_Fornecedor	id_Avaliacao
00001	Notebook e14 ryzen-3 5300u 8gb 256gb ssc	36	R\$ 3.399,99	R\$ 3.899,99	Com um design leve e ultrafino de apenas 19,8 mm, este not	3001	5001
00002	Teclado Alloy MKW100, HKBM1-R-EUA/G -	67	R\$ 199,99	R\$ 249,99	Com uma estrutura de alumínio sólida, o Alloy MKW100 foi p	3002	5002
00003	Mouse G502 LIGHTSPEED RGB LIGHTSY	85	R\$ 138,90	R\$ 199,90	Agora você pode jogar mais rápido e com mais precisão com	3003	5003
00004	Fone de Ouvido Tah2005bk/00, Over-ear, D	56	R\$ 219,99	R\$ 289,99	O fone de ouvido tah2005bk/00 é um modelo totalmente mod	3004	5004
00005	Cadeira Gamer GAMING Tempest CG600	41	R\$ 589,90	R\$ 749,90	Conheça a cadeira GAMING Tempest CG500! Perfeita para d	3005	5005

Categoria

id_Categoria	descricao
2001	Periféricos
2002	Computadores
2003	Escritório
2004	Smartphone

Avaliação

id_Avaliacao	tipo
5001	Positivas
5002	Positivas
5003	Positivas
5004	Neutras
5005	Positivas

Produto_Categoria

id_Produto_Categoria	id_Produto	id_Categoria
001	0001	2002
002	0002	2001
003	0003	2001
004	0004	2001
005	0005	2003

Fornecedor

id_Fornecedor	nome	telefone	email
001	Fornecedor Um	41 2223-6669	fornecedorum@email.com
002	Fornecedor Dois	51 6633-6855	fornecedordois@email.com
003	Fornecedor Tres	91 7888-4411	fornecedortres@email.com

Fornecedor_Produto

id_Fornecedor_Produtos	id_Produto	id_Fornecedor
3001	0001	003
3002	0002	003
3003	0003	002
3004	0004	001
3005	0005	002

Venda_Produtos

id_Venda_Produtos	id_Produto	id_Venda
800001	00001	700001
800002	00003	700002
800003	00004	700003
800004	00005	700004

Desconto

id_Desconto	nome	valor	data	status
001	Primeira Compra	R\$ 50,00	12/04/2024	Ativo
002	Desconto acima de R\$ 299	R\$ 50,00	12/04/2024	Ativo
003	Desconto acima de R\$ 599	R\$ 100,00	12/04/2024	vencido

Venda

id_Venda	data	Valor_Total	Forma_Pagamento	id_Desconto	parcelas	id_Cliente	id_Pedido
700001	12/04/2024	R\$ 700,00	Boleto Bancário	001	-	6005	1700004
700002	14/04/2024	R\$ 239,99	Crédito	001	1x	6003	1700003
700003	17/04/2024	R\$ 150,00	Crédito	001	6x	6001	1700002
700004	24/04/2024	R\$ 3.299,99	PIX	002	-	6002	1700001

Pedido

id_Pedido	Numero_Pedido	status	valor	data	id_Produtos
1700001	100001	em preparação	R\$ 700,00	12/04/2024	00001
1700002	100002	em preparação	R\$ 239,99	14/04/2024	00003
1700003	100003	entregue	R\$ 150,00	17/04/2024	00004
1700004	100004	em trânsito	R\$ 3.299,99	24/04/2024	00005

Cliente

id_Cliente	Cliente_Nome	email	senha	CPF	telefone	endereço	id_Desconto	id_Pedido
6001	Paulo	paulo1@email.com	paulo1@	879.365.245-32	33 9 2456-3659	rua v n1 36598-460	001	1700002
6002	Wesley	Wesley2@email.com	Wesley2@	325.241.365-00	41 9 6325-3654	rua w n2 39587-365	002	1700001
6003	Ielma	Ielma3@email.com	Ielma3@	154.797.732-20	58 9 9875-3654	rua x n3 36254-751	001	1700003
6004	Gabrielle	Gabrielle4@email.com	Gabrielle4@	689.656.368-52	54 9 3654-6547	rua y n4 66598-345	003	-
6005	Naira	Naira5@email.com	Naira5@	469.365.254-32	25 9 5478-3658	rua z n5 98756-400	001	1700004

Na simulação de cadastro apresentada acima, foi realizada quatro vendas onde demonstramos todo o funcionamento das tabelas, enquanto estão relacionadas. Utilizamos cinco clientes, onde quatro deles realizaram os respectivos pedidos.

Cada cliente comprou um produto, dos cinco que foram preenchidos na tabela.

Cada produto, tem sua devida categoria, fornecedor, e um campo para receber avaliações.

A tabela pedido irá receber os dados principais da compra e todo pedido será relacionado a tabela cliente.

5. SEGURANÇA DA INFORMAÇÃO

5.1 Identificação e avaliação dos riscos de segurança para a empresa

- Roubo de dados:

Visto que os dados são os principais insumos do meio digital, seja quando partem do usuário ou da própria empresa, são também um dos recorrentes alvos de ataques cibernéticos. Quando uma empresa estabelece operações conectadas aos seus serviços de TI e, aliado a isso, implementa tecnologia aos processos internos, o grau de informações retidas virtualmente é muito expressivo.

- Espionagem industrial:

Essa é uma das principais ameaças para a cibersegurança, que pode envolver: pessoas insatisfeitas com o local de trabalho – passando informações da empresa para terceiros; ameaça interna na figura de um funcionário recém-contratado que já tinha o intuito investigativo; ou até mesmo pode ocorrer por meio de táticas de engenharia social capazes de enganar um funcionário e fazê-lo divulgar dados internos sigilosos.

- Hackers de senhas:

A quebra da senha pode provocar sérios prejuízos para as organizações, visto que, uma vez dentro do sistema das mesmas, os cibercriminosos podem roubar os dados armazenados e mexer nas configurações dos servidores, por exemplo. Dentre as formas de corromper uma senha está o ataque de força bruta, que visa a violação de um nome de usuário ou senha, ou ainda a descoberta de uma página de web oculta ou de uma chave de criptografia. Para isso, os hackers utilizam uma metodologia de tentativa e erro, até que seja possível encontrar alguma lacuna no sistema de segurança da empresa.

- Funcionários não especializados / erros humanos:

Essa é uma ameaça de cibersegurança que deve ser observada bem de perto, pois é o que gera, na maior parte das vezes, danos quase irreversíveis para as empresas.

- Softwares vulneráveis:

Contar com uma infraestrutura digital atualizada é um dos pontos essenciais para se evitar ameaças à segurança da informação. Uma vez que softwares estejam defasados, os mesmos irão possibilitar erros de código e brechas no sistema, prejudicando a produtividade do usuário e potencializando os ataques cibernéticos.

- Ataques de Ransomware:

Por meio de um malware, que é qualquer software intencionalmente feito para causar danos a um computador, servidor, cliente, ou a uma rede de computadores, os cibercriminosos podem capturar informações e infectar diversos documentos acessíveis impedindo que estes sejam acessados.

- Phishing:

Em linhas gerais, o phishing é uma técnica de engenharia social que se vale de mensagens/plataformas que parecem ser verdadeiras e vindas de “lugares comuns”, como redes sociais, e-mail, sites de leilões, bancos, processadores de pagamento on-line ou administradores de TI, para atrair a atenção do usuário e o direcionar a links maliciosos capazes de executar diversas funções indevidas nos servidores.

- Ataques direcionados:

Demonstrando a evolução dos ataques cibernéticos e a necessidade de se pensar com mais cautela sobre o ambiente digital e a cibersegurança, o ataque direcionado é a prática de se estudar previamente uma empresa ou organização, conseguir seus dados, e utilizar essa informação de forma planejada, em direção a um alvo específico, com objetivos específicos.

- Adware:

Muito ligado à ameaça à segurança da informação que envolve o mau conhecimento dos perigos das redes pelos funcionários da empresa, está o Adware. Isso, pois esse tipo de risco requer que exista o clique humano em um anúncio “infectado” por um malware. A partir daí, os anúncios irão levar os usuários para outros sites maliciosos, abrir outras abas ou alterar a original de navegação. Essa é uma das ameaças à segurança da informação mais populares, presentes em quase toda a internet.

- Ataques “man-in-the-middle” (homem-no-meio):

Este tipo de ataque é bastante grave, e cada vez mais comum. Consiste em interceptar as comunicações entre duas pessoas e atuar como “homem-no-meio” da conversa. Nestes ataques, o criminoso pode tanto ficar passivamente bisbilhotando as mensagens trocadas quanto atuar de forma ativa alterando, excluindo ou inserindo frases enquanto os interlocutores acreditam estar travando uma conversa privada legítima.

- Fileless:

Como o nome já propõe, os ataques fileless são uma variante de malware que não utilizam arquivos para a invasão dos sistemas. Diferentemente das outras modalidades, este malware existe unicamente como um “artefato baseado em memória”, e geralmente desaparece após uma reinicialização de sistema.

Entretanto, por não contar com arquivos, é um dos tipos de ataque mais difíceis de detectar, mesmo utilizando as mais modernas ferramentas de anti-malware forense.

- Botnets:

Botnets são redes de dispositivos de computador sequestradas e usadas para realizar vários golpes e ciberataques. O termo "botnet" é formado pela junção das palavras "robot" (robô) e "network" (rede). A montagem de um botnet geralmente é a etapa de infiltração de um esquema de várias camadas. • Remote access trojan: são trojans que geralmente se disfarçam como aplicativos legítimos e podem ser compostos de funcionalidades maliciosas adicionadas a um aplicativo real. Monitore o aplicativo em busca de comportamento anormal, como notepad.exe gerando tráfego de rede.

- Cryptojacking:

São utilizados códigos ocultos em sites para explorar a capacidade de processamento dos equipamentos, podendo usar qualquer dispositivo com acesso à internet. Não há necessidade de instalação, pois o código é executado no acesso a uma página web infectada. • Spyware: As informações financeiras podem drenar contas bancárias ou podem servir de base para formulários de fraude usando contas bancárias legítimas. As informações obtidas através de documentos roubados, imagens, vídeos ou outros itens digitais podem até mesmo ser usados para fins de extorsão.

- Ataque força bruta:

Um ataque de força bruta consiste em todo e qualquer método usado por um invasor para descobrir uma senha, uma chave criptográfica ou outro tipo de informação sigilosa, por meio de tentativa e erro. • Hardware e Softwares desatualizados: Sistemas desatualizados representam um risco significativo de segurança cibernética, pois se tornam alvos fáceis para ataques maliciosos. Essa vulnerabilidade ocorre devido à ausência de atualizações que corrijam falhas conhecidas.

- Engenharia social:

É uma técnica de manipulação que explora erros humanos para obter informações privadas, acessos ou coisas de valor. No crime cibernético, esses golpes de "hacking humano" tendem a atrair usuários desavisados para expor dados, espalhar infecções por malware ou dar acesso a sistemas restritos.

- DDoS:

É um tipo de ataque cibernético, um invasor sobrecarrega um website, servidor ou recurso de rede com tráfego malicioso. Como resultado, o alvo trava ou não consegue operar, negando serviço a usuários legítimos e impedindo que o tráfego legítimo chegue ao seu destino.

- Ataques a cadeia de suprimentos:

Os ataques à cadeia de suprimentos são indiretos: eles visam as dependências de terceiros das quais seus alvos finais dependem (muitas vezes sem saber). Uma dependência é um programa ou parte de código (geralmente escrito em JavaScript) de provedores terceiros que aprimora a funcionalidade do aplicativo. Uma dependência usada por um varejista de comércio eletrônico, por exemplo, pode ajudar a executar chatbots de atendimento ao cliente ou capturar informações sobre a atividade do visitante no site.

5.1 TABELA GUT

Nesta etapa foi feita a "Tabela Gut", para identificar as vulnerabilidades que uma empresa pode sofrer, o que pode ser perdido, danificado ou danificado, e definir qual o grau de tais ameaças.

A tabela foi baseada na, identificação e avaliação dos riscos de segurança para a empresa. A tabela em si é de grande ajuda para fazer uma análise das ameaças, e definir o impacto que elas podem causar dentro de locais de trabalho dentro de uma empresa.

Tabela Gut	impacto	probabilidade
Ataques a cadeia de suprimentos	Perda de faturamento	Alta
Engenharia social	Implicações legais, financeiras e reputacionais para organizações	Alta
DDoS	Trava de sistemas	Alta
Hardware e Software desatualizados	Queda no sistema, levando a perda de trabalhos que não foram salvos	Alta
Ataques força bruta	Pode causar interrupções operacionais e financeiras	Alta
Spyware	Afeta a confiança dos clientes que pode levar à perda de negócios	Alta
Cryptojanking	Perda de produtividade	Alta
Botnets	Roubar dados ou dinheiro, interromper negócios, danificar sistemas.	Alta
Roubo de dados	Multas regulamentares e uma profunda erosão da confiança dos consumidores	Alta
Fileless	Baixa reputação da empresa pela violação de dados	Alta
Ataques "man-in-the-middle"(homem-no-meio)	Prejuízos financeiros ou roubo de dados	Alta
Adware	Lentidão nos sistemas	Alta
Ataques direcionados	Inatividade, perda de dados e perda de dinheiro	Alta
Phishing	Perda de confiabilidade perante os consumidores	Alta
Ataques de Ransomware	Danos à reputação da empresa	Alta
Softwares vulneráveis	Sequestros de informações e roubo de dados	Alta
Funcionários não especializados/ erros humanos	Danos quase irreversíveis para as empresas	Alta
Hackers de senhas	Lentidão anormal	Alta
Espionagem industrial	Desvantagem competitivas para as organizações	Alta
Remote access Trojan	Roubo de informações e dados financeiros	Alta

5.2 Implementação de políticas de controle de acesso aos sistemas e dados

Controle de Acesso

- Tag ou cartão de proximidade
- Leitura de impressão digital
- Leitura facial e de Íris • Sistema de acesso por senha
- Controle de acesso semiautomático
- Controle de acesso manual
- Sensor de movimento
- Autorização restrita a terceiros
- Controle de acesso físico
- Controle de acesso logico
- Registro online e offline

O controle de acesso evita que informações confidenciais, como dados de clientes e propriedade intelectual, sejam roubadas por pessoas mal-intencionadas ou outros usuários não autorizados. Também reduz o risco de exfiltração de dados pelos funcionários e mantém as ameaças baseadas na web afastadas. Em vez de gerenciar permissões manualmente, a maioria das organizações voltadas para a segurança recorrem a soluções de gerenciamento de identidade e acesso para implementar políticas de controle de acesso.

A ISO 27000 é um conjunto de certificações de segurança da informação e proteção de dados para empresas e órgãos públicos. Elas servem como base para a criação de um Sistema de Gestão de Segurança da Informação (SGSI) em organizações de pequeno, médio e grande porte.

O SGSI reúne políticas, procedimentos, diretrizes e recursos de proteção de informação de uma organização. O sistema deve estar alinhado aos objetivos de negócio e ser gerenciado de forma conjunta pela empresa.

As certificações da família ISO 27000 foram desenvolvidas em parceria entre a ISO e a International Electrotechnical Commission (IEC), outra organização internacional dedicada à padronização, porém voltada a tecnologias relacionadas à energia, multimídia, telecomunicações e eletroeletrônica. A primeira versão foi publicada em

2005, como uma atualização do BS 7799, padrão britânico de segurança da informação. Hoje está em vigor a ISO/IEC 27000:2018.

Ela traz alguns princípios norteadores da segurança da informação. Os principais são:

1. Confidencialidade
2. Integridade
3. Disponibilidade
4. Autenticidade

O documento ainda traz termos e definições usados no processo de padronização de um SGSI. Ele serve como uma introdução para as outras normas do grupo.

5.3 Configuração de sistemas de detecção de intrusão e prevenção de ataques

Implementação de Medidas de Segurança:

Firewalls de Próxima Geração:

- implemente firewalls avançados que possam identificar e bloquear ameaças em tempo real.

Sistema de Detecção de Intrusão (IDS):

- Utilize IDS para monitorar o tráfego de rede em busca de atividades suspeitas ou padrões maliciosos.

Sistema de Prevenção de Intrusão (IPS):

- Complemente o IDS com IPS para bloquear automaticamente tráfego malicioso identificado.

Análise de Logs:

- Implemente sistemas de análise de logs para identificar atividades anômalas ou potencialmente maliciosas.

Atualizações e Patches:

- Mantenha todos os sistemas e software atualizados com os patches de segurança mais recentes para corrigir vulnerabilidades conhecidas.

Segmentação de Rede:

- Divida a rede em segmentos menores e restrinja o tráfego entre eles para limitar a propagação de ataques.

Autenticação Multifator (MFA):

- Implemente MFA para adicionar uma camada extra de segurança ao verificar a identidade dos usuários.

Controle de Acesso:

- Adote políticas de controle de acesso para garantir que apenas usuários autorizados tenham permissão para acessar recursos críticos.

Monitoramento de Endpoint:

- Utilize soluções de segurança de endpoint para detectar e responder a ameaças nos dispositivos finais da rede

Treinamento de Conscientização em Segurança:

- Educar os funcionários sobre práticas seguras de computação e conscientização sobre phishing e outras técnicas de ataque.

Elaborar uma política de segurança da informação de uma parte da empresa.

REFERÊNCIAS

[Kaggle: Your Machine Learning and Data Science Community](#)

[APRENDIZADO DE MAQUINA TURMA 40 - SA \(google.com\)](#)

[\(14\) Validação Cruzada: Aprenda de forma Simples Como Usar essa Técnica. - YouTube](#)

[BR O que é Markdown? Tutorial Fácil para Iniciantes \(youtube.com\)](#)

CONCLUSÃO

Este projeto de Análise e desenvolvimento de Sistemas foi um sucesso, proporcionando aos participantes uma experiência prática valiosa em uma ampla gama de tópicos fundamentais na área.

Os resultados do projeto irão beneficiar os participantes em seus estudos e carreiras futuras, capacitando-os a resolver problemas complexos e contribuir para o avanço de todos os tópicos citados.