

Apresentação de briefing de projeto



Unidade 3

- * **Visão geral**
- * **Objetivo**
- * **Base de Dados**
- * **Entregáveis**
- * **Especificações**
- * **Entrega e Apresentação**

AGENDA



Olá, aluno.. da turma mais top!

Você chegou ao momento de aplicar, na prática, tudo o que aprendeu em Python para Análise de Dados.

Nesta unidade, cada aluno irá trabalhar **individualmente** com um projeto em uma área de negócio específica. Essa escolha individual tem um motivo: dar **início ao seu portfólio de dados**, com um projeto completo, autoral e bem estruturado.

💡 Na próxima unidade (Unidade 4), essa mesma base será utilizada novamente, mas com foco em visualização e storytelling no Power BI — assim, seu projeto evolui e ganha ainda mais força como case de portfólio.

Objetivo



A proposta é simular um desafio real do mercado de dados, onde você será responsável por explorar uma base de dados do seu segmento escolhido, identificar padrões, responder perguntas relevantes e transformar dados brutos em informações estratégicas.

Neste processo, você irá:

- ✓ Aplicar os principais conceitos de Python para análise de dados, usando bibliotecas como pandas, numpy, matplotlib, seaborn e/ou plotly;
- ✓ Demonstrar organização e clareza no seu notebook, com códigos bem comentados e estruturados;
- ✓ Documentar o projeto com um raciocínio lógico, incluindo introdução, análise, visualizações e conclusão;
- ✓ Praticar a comunicação de resultados com impacto, por meio de uma apresentação final em formato de pitch, onde os principais insights extraídos da base serão destacados e contextualizados.

O foco não está apenas no código, mas na forma como você pensa, estrutura e apresenta a análise — exatamente como acontece no mercado profissional.



 **Nome:** base_financeiro_cliente



Clique aqui para acessar a pasta com a base

Segmento: Financeiro

Tipo de análise: Perfil de clientes adimplentes + comportamento financeiro

Objetivo: Compreender padrões e gerar recomendações para retenção, engajamento e crédito

Entendimento do Negócio

Este projeto tem como objetivo realizar uma análise detalhada dos dados de clientes de uma instituição financeira fictícia, com foco na compreensão do perfil demográfico e financeiro dos clientes adimplentes, bem como na identificação de padrões de uso, engajamento e risco futuro.

A proposta é que você atue como um analista de dados da área de CRM ou crédito, respondendo perguntas como:

- Quais perfis movimentam mais valor?
- Qual a relação entre escolaridade, renda e engajamento?
- Quem são os clientes mais inativos? Por quê?
- Há potencial para campanhas específicas ou upgrades de cartão?

Essa análise servirá de base para decisões estratégicas e será estruturada com foco em visualização, narrativa e impacto, simulando uma entrega real da área de dados.



 **Nome:** base_comercial_ecommerce



Clique aqui para acessar a pasta com a base

Segmento: Comercial / E-commerce

Tipo de análise: Comportamento de compra e perfil de consumidores

Objetivo: Analisar hábitos de consumo e padrões de comportamento em compras online

Entendimento do Negócio

A base representa dados de compras realizadas em um e-commerce fictício, com informações sobre produtos, categorias, perfil do cliente e avaliações.

Você atuará como um analista de dados da área comercial, responsável por extrair padrões de comportamento de compra, segmentar o público e identificar oportunidades de otimização de vendas, produtos e experiência do cliente.

Algumas perguntas que podem guiar sua análise:

- Quais categorias de produtos têm maior aceitação por faixa etária ou gênero?
- Existe correlação entre faixa de preço e avaliação?
- Quais subcategorias têm maior recorrência de compras?
- Há padrões nos sentimentos das avaliações por tipo de produto?
- Como o histórico de compras influencia a nota das avaliações?

Essa análise servirá como base para recomendações ao time de marketing e produto, simulando uma entrega de dados para melhorar a experiência e conversão de vendas.



 **Nome:** base_rh_trabalhista



Clique aqui para acessar a pasta com a base

Segmento: Recursos Humanos / People Analytics

Tipo de análise: Rotatividade de colaboradores (Attrition)

Objetivo: Analisar os fatores associados à saída de colaboradores e o perfil de permanência

Entendimento do Negócio

Essa base simula dados de colaboradores de uma empresa de médio-grande porte, trazendo características demográficas, profissionais e comportamentais, com o objetivo principal de entender por que os colaboradores pedem demissão ou permanecem.

Você atuará como analista de dados do time de People Analytics e deverá identificar fatores associados à rotatividade (attrition), explorar perfis mais propensos à saída e propor recomendações baseadas em dados para retenção de talentos.

Algumas perguntas que podem nortear sua análise:

- Existe um padrão de idade, cargo ou setor entre quem mais pede demissão?
- A distância de casa, o nível de satisfação ou o tempo de empresa influenciam na decisão de sair?
- Há diferença entre os gêneros ou níveis de educação em relação ao attrition?
- Funcionários com mais promoções, treinamentos ou envolvimento permanecem mais?

Essa análise pode apoiar decisões estratégicas de RH como:

- Reforçar programas de retenção e reconhecimento;
- Ajustar cargas de trabalho ou treinamento;
- Revisar critérios de promoção ou gestão de desempenho.



 **Nome:** base_governamental



Clique aqui para acessar a pasta com a base

Segmento: Políticas Públicas / Agricultura / Economia

Tipo de análise: Variação de preços mensais por produto e estado

Objetivo: Analisar a variação de preços ao longo do ano e entender seus impactos regionais

Entendimento do Negócio

Essa base representa os preços mensais de produtos agrícolas e alimentícios no Brasil, divulgados por órgãos oficiais, em diferentes níveis de comercialização (produtor, atacado, varejo) e por Unidade Federativa (U.F.).

Você atuará como um analista de dados para uma secretaria de políticas públicas, com a missão de monitorar inflação setorial, flutuações regionais e apoiar decisões estratégicas sobre produção, subsídios ou políticas de abastecimento.

Algumas perguntas que podem orientar sua análise:

- Quais produtos apresentaram maior variação de preço ao longo do ano?
- Há sazonalidade clara em certos produtos (ex: frutas, hortaliças)?
- Em quais estados há mais instabilidade de preços?
- Há diferença entre os níveis de comercialização (produtor vs. atacado)?
- Existe concentração de aumentos em algum mês ou período?

Essa análise pode ajudar gestores públicos a antecipar impactos na cadeia produtiva, sugerir políticas regionais específicas e monitorar o custo de vida.



 **Nome:** base_streaming



Clique aqui para acessar a pasta com a base

Segmento: Mídia e Entretenimento

Tipo de análise: Catálogo de filmes e séries por plataforma

Objetivo: Analisar padrões de conteúdo, popularidade e estratégias de catálogo entre plataformas de streaming

Entendimento do Negócio

Essa base reúne informações sobre o catálogo de filmes disponíveis em diferentes plataformas de streaming (Netflix, Prime, Max, Apple TV). Estão incluídos dados como gênero, ano de lançamento, avaliações no IMDb e número de votos. Você atuará como um analista de dados de uma consultoria de mídia e entretenimento, responsável por gerar insights sobre o comportamento do mercado de streaming, analisando preferências de público, curadoria por plataforma e tendências de conteúdo.

Algumas perguntas que podem nortear sua análise:

- Quais plataformas concentram os filmes mais bem avaliados?
- Há diferença de foco entre gêneros por plataforma?
- Os títulos mais votados estão presentes em múltiplas plataformas ou são exclusivos?
- Qual o perfil médio dos títulos por plataforma (ano, gênero, nota)?
- Qual plataforma aposta mais em clássicos, e qual foca em lançamentos?

A ideia é gerar insights que possam ser usados por equipes de aquisição de conteúdo, marketing ou recomendação de algoritmos.



 **Nome:** base_credit_score_data_science



Clique aqui para acessar a pasta com a base

Segmento: Crédito / Análise de Risco

Tipo de análise: Modelagem de risco de crédito (Score de inadimplência)

Objetivo: Construir um modelo preditivo de inadimplência com base em atributos transacionais e cadastrais

Entendimento do Negócio

Essa base simula o comportamento de clientes de cartão de crédito, com centenas de atributos transacionais e cadastrais, além do rótulo `bad_flag`, que indica se o cliente se tornou inadimplente.

Você atuará como cientista de dados de uma instituição financeira, com a missão de construir um modelo de classificação para prever o risco de inadimplência, utilizando dados históricos de comportamento.

Observação importante:

Este projeto exige um nível intermediário a avançado de Python, especialmente com bibliotecas como pandas, sklearn e matplotlib. A entrega envolverá treinamento, validação e avaliação de modelos preditivos, utilizando algoritmos como Regressão Logística, Random Forest, XGBoost ou LightGBM.

Perguntas-chave para guiar sua análise:

- Quais variáveis melhor se correlacionam com a inadimplência?
- Como tratar os dados faltantes, outliers e escalas?
- Qual modelo apresenta melhor desempenho preditivo?
- Como avaliar se o modelo é robusto (AUC, KS, matriz de confusão)?

Além disso, o dataset está dividido em duas partes:

- Base de treino (com `bad_flag` conhecido)
- Base de validação (para simular o ambiente real de aplicação do modelo)



Todos os alunos devem entregar um notebook .ipynb estruturado com análise completa, narrativa clara e visualizações relevantes. O foco será sempre resolver um problema de negócio por meio dos dados, com entregas que simulem uma atuação real de um profissional da área.

Entregáveis Comuns a Todos os Projetos

1. Notebook em Python estruturado com:

- Título e Introdução
- Objetivo do projeto
- Análise Exploratória dos Dados (EDA)
- Visualizações com matplotlib, seaborn ou plotly
- Insights principais
- Conclusão com sugestões ou recomendações

2. Organização e Documentação:

- Códigos comentados
- Uso de funções quando necessário
- Evitar notebooks extensos e desorganizados
- Títulos e subtítulos que guiem a leitura

3. Apresentação Final (Pitch Analítico):

- Slides ou narrativa no próprio notebook
- Foco na clareza dos insights
- Deve simular uma entrega para a área de negócio ou gestão



Projeto Financeiro (Análise de Adimplentes)

- Foco: Segmentar clientes e propor ações de engajamento
- Expectativa: Análise comportamental e sugestões para o time de CRM
- Técnicas esperadas: agrupamentos, análise descritiva, visualizações comparativas

Projeto Comercial (E-commerce)

- Foco: Compreensão do comportamento de compra e avaliações
- Expectativa: Insights por categoria, faixa etária, sentimento
- Técnicas esperadas: análises de correlação, agrupamentos por gênero e categoria

Projeto RH (Attrition/People Analytics)

- Foco: Identificar padrões na rotatividade de colaboradores
- Expectativa: Insights com viés de retenção e clima organizacional
- Técnicas esperadas: comparação entre grupos, destaque de variáveis-chave

Projeto Governamental (Preços Agrícolas)

- Foco: Monitoramento de variação de preços por produto e estado
- Expectativa: Gráficos temporais, sazonalidade, observações por região
- Técnicas esperadas: análise de séries temporais simples, gráficos de linha



Projeto Streaming (Catálogo e Popularidade)

- Foco: Comparação entre catálogos das plataformas de streaming
- Expectativa: Diferenças entre plataformas, ranking de filmes, exclusividades
- Técnicas esperadas: análise de frequências, rankings, análise por gênero e ano

Projeto Crédito / Data Science (Score de Inadimplência)

- Foco: Construção de modelo preditivo de inadimplência
- Expectativa: Treinamento e validação de modelos supervisionados
- Técnicas esperadas: train_test_split, regressão logística, árvore, Random Forest, XGBoost, avaliação de AUC, KS, matriz de confusão
- ⚠ **Observação: Projeto mais técnico, exige domínio de modelagem com sklearn**

⚠ Observações Técnicas

- Evite o uso de notebooks desorganizados ou códigos muito longos em uma só célula
- Trate valores ausentes, padronize colunas e explique decisões de limpeza
- Use markdown para se comunicar dentro do notebook — isso também é storytelling
- **A clareza dos seus insights será mais importante que a complexidade técnica**

Entrega e Apresentação


Apresentação Final (Pitch)

A apresentação será no dia 07/06, no formato de pitch analítico, com tempo máximo de 5 minutos por aluno. O pitch é uma apresentação rápida e objetiva que simula a entrega de um projeto para uma equipe de negócios ou gestão. O foco não é explicar cada linha de código, mas sim:

- Qual era o problema ou desafio proposto?
- Quais dados você analisou e o que descobriu?
- Quais os principais insights ou recomendações?
- Como isso gera valor para a área de negócio?

 **Lembre-se: menos é mais. Apresente com clareza, segurança e foco no impacto.**

Envio do Projeto

- Data limite para envio do notebook: até 07/06 (antes da apresentação)
- Formato: .ipynb (notebook Jupyter ou Google Colab)
- Envio por e-mail para:
-  nayara.valevskii@gmail.com
- Quem tiver domínio, pode enviar apenas o link do repositório do GITHUB (Fica como desafio).

Entrega e Apresentação

Feedback Antecipado (opcional)

Se quiser orientações ou sugestões antes da entrega, envie seu notebook até 03/06.

No dia 04/06, você receberá um retorno com dicas, sugestões e possíveis melhorias para fortalecer seu projeto.

Essa etapa é opcional, mas recomendada se você quiser evoluir seu projeto com apoio direto.

Ajustes Finais

Os dias 05 e 06/06 estão reservados para quem quiser fazer ajustes após o feedback.

Se possível, abrirei agendas individuais de dúvidas na semana da apresentação (informações serão divulgadas posteriormente).

Um Projeto de Portfólio, Não de Nota

Esse projeto não será avaliado com nota, mas tem peso direto na sua carreira.

Ele representa uma entrega real, estruturada, com dados e storytelling — exatamente como se espera de um analista ou cientista de dados no mercado.

Leve a sério. Essa entrega é sua chance de colocar em prática o que aprendeu, exercitar sua análise e ter um case autêntico no seu portfólio profissional.

"Você não controla os dados que o mundo te dá. Mas pode escolher como interpretá-los e é aí que mora o poder da análise."

— Inspirado em Leonard Mlodinow, O Andar do Bêbado

**E aí,
o que achou?**

Que tal compartilhar o andamento do seu projeto no LinkedIn e me marcar? Lembrem-se, quem não é visto, não é lembrado. Conecta aí!

