

e-book



Pós-Graduação



em



Business _____



Intelligence _____



& Analytics. _____

Como alcançar
resultados de
forma inteligente.



jornada do curso

MÓDULO 01_O PRINCÍPIO DE TUDO

MÓDULO 02_DADOS, INFORMAÇÕES E INTELIGÊNCIA

MÓDULO 03_OS DADOS CONTAM HISTÓRIAS

**AULA 07_ CONECTAR DADOS A RESULTADOS:
OLHAR PARA TRÁS**

**AULA 08_ CONECTAR DADOS A RESULTADOS:
OLHAR PARA FRENTE**

**AULA 09_ DADOS COMO HIPÓTESES NA NOVA ECONOMIA:
ENTRE PREVISÕES E RESULTADOS DE SUCESSO**



AULA 10_CASE #01

AULA 11_CHECKPOINT

MÓDULO 04_APRESENTAR OS DADOS E AS DECISÕES





MÓDULO 03:
OS DADOS CONTAM HISTÓRIAS

CASE #01



ou vai ouvir no



PÓS-GRADUAÇÃO EM
BUSINESS INTELLIGENCE & ANALYTICS



sumário



Clique, selecione ou escreva! Este material possui recursos interativos para enriquecer sua experiência.



M

missão da aula

10

I

Consolidar os temas estudados no módulo por meio de um case, realizando o passo a passo de tratamento e análise de dados e de validação de hipóteses.

S

S

Ã



O



[10]

interativo! 

preparando o **terreno**





revisão do módulo 03:

os dados contam histórias

aula 07



_Conectar dados a resultados:
olhar para trás

- > tomada de decisão e confiabilidade
- > analisando resultados: olhando para trás
- > técnicas de classificação e clusterização

análise de **sucesso**
gera **insight confiável**
para o seu negócio!



aula 08



_Conectar dados a resultados:
olhar para frente

- > conceitos estatísticos
- > correlação
- > regressão linear

a análise do **passado** garante uma
melhor avaliação e **previsão dos**
resultados do futuro.

aula 09



_Dados como hipóteses na nova economia:
entre previsões e resultados de sucesso

- > análises preditivas
- > teste A/B
- > teste A/B e teste de hipóteses

não adivinhe nada. **teste!**



CASE #01: como vai funcionar?

#regras #do jogo

São dois cases com **cenários** e **contextos diferentes**, mas as técnicas e as ferramentas de resolução de ambos são as mesmas.

> Case 1: o professor apresenta e soluciona.

> Case 2: o professor apresenta e a turma resolve.

Os cases estão divididos em **três partes**.
Assim, **resolveremos uma parte de cada vez**.

Importante!

As respostas para as perguntas dos cases estão disponíveis no final do seu e-book.



[10]

interativo! 

parte 01: **análise** **descritiva**





mão na massa

Finaceira Broke – parte 1

A Broke é uma empresa financeira do ramo de cartões de crédito e está passando por uma reestruturação interna, pois, nos últimos anos, o seu resultado foi abaixo do esperado.

O primeiro passo para a nova diretoria conseguir melhorar os resultados é entender qual é o perfil de consumo dos clientes que utilizam os cartões de crédito da Broke.

Para isso, a empresa disponibilizou algumas informações do seu banco de dados:

> Dados dos cartões (clientes, limites, idades etc.);

> Produtos comprados.

Para a análise do perfil dos clientes, utilize as duas bases de dados e responda:

1. Existe uma relação de idade x limite no cartão de crédito?

2. Quais são as principais categorias de produtos comprados nos cartões de crédito?



mão na massa

Pop Store – parte 1

A Pop Store contratou sua empresa de consultoria para melhorar o posicionamento e a receita. Para ajudar a alavancar as vendas, o primeiro passo é entender o perfil dos clientes e do consumo.

Para isso, a empresa disponibilizou algumas informações do seu banco de dados:

> Cadastro de clientes;

> Produtos vendidos;

> Pedidos realizados.

Faça sua análise com base nesses arquivos e responda:

1. Qual é o perfil dos clientes (onde eles moram, as suas preferências)?

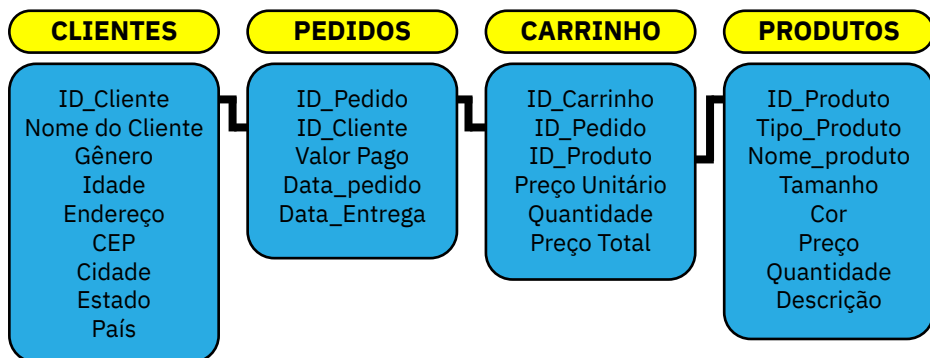
2. Quais foram os produtos mais vendidos no último mês?

3. Qual foi o comportamento das vendas e a receita nos últimos meses?



mão na massa

Pop Store – parte 1





[10]

interativo! 

parte 02:

modelo preditivo





mão na massa

Financeira Broke – parte 2

Um grande problema encontrado na Broke foi a falta de uma estratégia bem definida para a aprovação de novos cartões de crédito.

Recebemos uma base histórica com os dados cadastrais de clientes e o limite de cartão de crédito ideal que deveria ter sido disponibilizado para cada cliente.

Com base nesse dados, precisamos criar um modelo preditivo para definir o limite de cartão de crédito de novos clientes de uma maneira automatizada e padronizada.



mão na massa

Pop Store – parte 2

A equipe de comunicação da Pop Store precisa rever a estratégia de marketing. Para isso, pediu a ajuda da consultoria para entender qual é o retorno esperado, caso resolva investir diferentes valores nos canais de comunicação: podcast, televisão e TikTok.

De acordo com o arquivo, responda às perguntas a seguir:

1. Qual é o canal que traz o melhor resultado?

2. Qual será o aumento de receita esperado se forem investidos R\$ 10.000 a mais em cada mídia?

3. Qual será a receita se forem investidos:

R\$ 20.000 em podcast?

R\$ 32.000 no TikTok?

R\$ 40.000 na televisão?



[10]

interativo! 

parte 03: **A / B** **teste**





mão na massa

Financeira Broke – parte 3

A Broke sempre foi mais tradicional e entrava em contato com os seus clientes por meio de um call center para cobrar os valores em atraso. Com isso, sempre teve baixos índices de recuperação e, muitas vezes, os clientes não sabiam que estavam com saldo devedor ou esqueciam de efetuar o pagamento.

Por isso, a empresa resolveu testar o ZapWhat, um novo canal para abordar os clientes com o objetivo de melhorar a comunicação e os indicadores de inadimplência.

**Os resultados do teste estão no arquivo anexado.
Será que essa comunicação foi mais efetiva?**



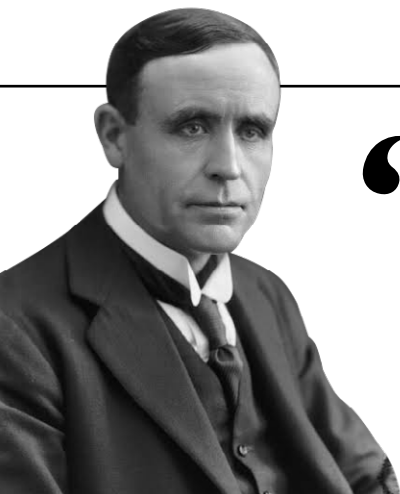
mão na massa

Pop Store – parte 3

Agora, o último passo para melhorar os resultados da Pop Store é atrair um maior número de clientes. Para isso, a loja está testando um novo *site*, com um novo *layout*.

A empresa disponibilizou os dados de performance do teste A/B que realizou com o *site* atual *versus* a nova proposta.

Com base nos dados disponibilizados, utilize o **teste de hipóteses** para avaliar se o novo *site* trará mais resultados para a loja.



“

Não ponha fé no que as estatísticas dizem até que você tenha considerado cuidadosamente o que elas não dizem.”

WILLIAM WATT



[10]

interativo! 



**CAIXA DE
DÚVIDAS**





desafio conquer #10

- > Imagine que você trabalha numa grande loja de varejo e gostaria de **prever suas vendas e demandas** com precisão. Atualmente, a empresa está enfrentando um problema causado por demandas imprevisíveis e, algumas vezes, fica sem estoque devido ao modelo atual, que está inadequado.
- > O desafio é **entender** o impacto das variáveis nas vendas semanais e, em seguida, **construir** um modelo para prever essa demanda.
- > Na base de dados, temos dados históricos de 45 lojas dessa rede, localizadas em diferentes regiões.
- > Por ser uma grande loja, suas vendas também são influenciadas por feriados importantes, e as semanas que possuem feriados têm um peso 5 vezes maior na avaliação do que semanas sem feriados (como utilizar essa informação é parte do desafio!).



[10]

interativo! 

Conquer **notes**





site

ANÁLISES PREDITIVAS: O QUE SÃO E QUAL SUA IMPORTÂNCIA?

Victor Padilha & André Carvalho, ICMC-USP

FORECAST NO EXCEL

Suporte Microsoft

MINEIRAÇÃO DE DADOS EM PYTHON – INICIANDO O ESTUDO E EXPLORAÇÃO DE DADOS

Victor Padilha & André Carvalho, ICMC-USP

+
+
+
quero

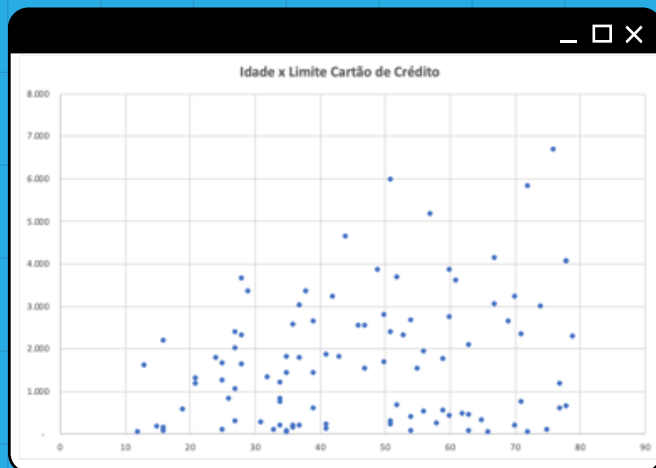


[10]

mão na massa

Respostas: Financeira Broke – parte 1

1. Existe uma relação de idade x limite no cartão de crédito?



Correl = 26,37%

2. Quais são as principais categorias de produtos comprados nos cartões de crédito?

Categoria	Silver
PETROL	8.841
CAMERA	8.752
TRAIN TICKET	7.523
AIR TICKET	6.994
FOOD	6.362

Categoria	Gold
PETROL	8.841
CAMERA	8.752
TRAIN TICKET	7.523
AIR TICKET	6.994
FOOD	6.362

Categoria	Platinum
PETROL	18.294
FOOD	17.602
CAMERA	13.645
CLOTHES	11.973
AIR TICKET	11.580



mão na massa

Respostas: Pop Store – parte 1

1. Qual é o perfil dos clientes (onde eles moram, as suas preferências)?

Distribuição por UF	%	Gênero	Média de Idade	Produtos mais Vendidos	Qtde
Australian Capital Territory	12,1%	Agender	48	Denim	527
New South Wales	13,2%	Bigender	51	Joggers	334
Northern Territory	12,5%	Female	49	Pleated	308
Queensland	13,4%	Genderfluid	48	Casual Slim Fit	306
South Australia	13,9%	Genderqueer	50	Shearling	302
Tasmania	10,4%	Male	51	Relaxed Leg	301
Victoria	12,1%	Non-binary	52	Trench Coat	299
Western Australia	12,4%	Polygender	50	Puffer	298
Total Geral	100,00%	Idade média	50	Bomber	297

2. Quais foram os produtos mais vendidos no último mês?

Produtos mais vendidos em outubro/2021:

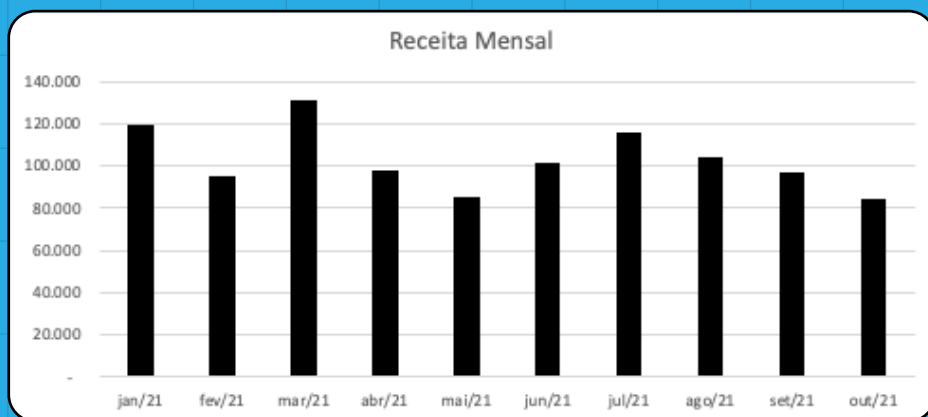
Produtos mais Vendidos	Qtde
Denim	45
Cords	37
High-Waisted	36
Joggers	34
Windbreaker	33
Mandarin Collar	31
Drawstring	30
Camp Collared	29
Shearling	29
Chambray	29
Cuban Collar	28
Pleated	28
Chinos	28



mão na massa

Respostas: Pop Store – parte 1

3. Qual foi o comportamento das vendas e a receita nos últimos meses?





mão na massa

Respostas: Financeira Broke – parte 2

Modelo preditivo

SUMÁRIO DOS RESULTADOS					
Estatística de regressão					
R múltiplo	0,99123279				
Quadrado de R	0,98254245				
Quadrado de R ajustado	0,98141615				
Erro-padrão	455,31806				
Observações	100				
ANOVA					
	gl	SQ	MQ	F	F de significância
Regressão	6	1085127208	180854535	872,367845	1,98911E-79
Residual	93	19280251,8	207314,536		
Total	99	1104407459			
	Coeficientes	Erro-padrão	Stat t	valor P	95% inferior
Interceptar	8818,98548	113,865848	77,4506637	2,9123E-86	8592,870469
Renda	1,5240988	0,02613868	58,3081684	4,9992E-75	1,472192558
Cartao_Silver	-1854,0104	119,144824	-15,560981	1,249E-27	-2090,608371
Cartao_Gold	-1008,8106	109,420907	-9,2195413	9,1643E-15	-1226,098802
Self Employed	-3434,192	116,134589	-29,570794	4,3763E-49	-3664,812266
Salaried_MNC	-1921,328	142,921691	-13,443222	1,5997E-23	-2205,142162
Salaried_Pvt	-2853,4353	143,125068	-19,936657	2,4104E-35	-3137,653353

Limite = 8.818 + 1,52 * renda – 1854 * cartão_silver - 1008 * cartão_gold - 3434* self_employed - 1921 * salaried_mnc + 2853 * salaried_pvt



mão na massa

Respostas: Pop Store – parte 2

Modelo preditivo

SUMÁRIO DOS RESULTADOS					
<i>Estatística de regressão</i>					
R múltiplo	0,94159053				
Quadrado de R	0,88659272				
Quadrado de R ajustado	0,87628297				
Erro-padrão	3615,97953				
Observações	25				
ANOVA					
	gl	SQ	MQ	F	F de significância
Regressão	2	2248836216	1124418108	85,9955355	3,99071E-11
Residual	22	287656775	13075307,9		
Total	24	2536492991			
	Coeficientes	Erro-padrão	Stat t	valor P	95% inferior
Interceptar	19670,609	9591,80564	2,05077226	0,05238742	-221,5784181
Podcast	2,01724293	0,48596642	4,15099245	0,00041727	1,009410265
TikTok	2,97089277	0,2282554	13,0156514	8,2242E-12	2,497520031

Receita = 19.670 + 2,01 * podcast + 2,97 * TikTok



mão na massa

Respostas: Pop Store – parte 2

Modelo preditivo

De acordo com o modelo final, temos que:

Qual é o canal que traz o melhor resultado?

TikTok é o canal que traz o melhor resultado, pois o seu coeficiente é o maior entre as variáveis dependentes.

Qual será o aumento de receita esperado se forem investidos R\$ 10.000 a mais em cada mídia?

**A cada R\$ 10.000 a mais, terá aumento na receita:
Podcast: R\$ 20.172
TikTok: R\$ 29.708
Televisão: R\$ 0 (pois a variável ficou de fora do modelo ótimo)**

Qual será a receita se forem investidos:

R\$ 20.000 em podcast?

R\$ 32.000 no TikTok?

R\$ 40.000 na televisão?

A receita será igual a R\$ 155.084,00.



mão na massa

Respostas: Financeira Broke – parte 3

Teste T:

Teste T: duas amostras com variâncias iguais		
	Controle	Teste
Média	0,37672006	0,45197119
Variância	0,00126536	0,00087783
Observações	40	40
Variância agrupada	0,00107159	
Hipótese de diferença de média	0	
gl	78	
Stat t	-10,280473	
P(T<=t) uni-caudal	1,8285E-16	
t crítico uni-caudal	1,66462464	
P(T<=t) bi-caudal	3,6571E-16	
t crítico bi-caudal	1,99084707	

A um nível de significância de 95%:

P-Valor = 0,000000

Alfa = 0,05

Como o **p-valor** é menor que o **alfa**, podemos concluir que o teste foi bem sucedido.



mão na massa

Respostas: Pop Store – parte 3

Teste T:

Teste T: duas amostras com variâncias iguais		
	Controle	Teste
Média	0,11592561	0,15656625
Variância	0,00206313	0,00464527
Observações	40	40
Variância agrupada	0,0033542	
Hipótese de diferença de média	0	
gl	78	
Stat t	-3,1382026	
P(T<=t) uni-caudal	0,00119971	
t crítico uni-caudal	1,66462464	
P(T<=t) bi-caudal	0,00239942	
t crítico bi-caudal	1,99084707	

A um nível de significância de 95%:

P-Valor = 0,002399

Alfa = 0,05

Como o **p-valor** é menor que o **alfa**, podemos concluir que o teste foi bem sucedido.