

Educação Profissional Paulista

Técnico em
**Ciência de
Dados**

Visualização de dados

Práticas de design, formatação e apresentação dos dados

Aula 2

Código da aula: [DADOS]ANO1C3B4S29A2

Visualização de
dados

Mapa da Unidade 3 Componente 4

semana

26

Cores e elementos
visuais

semana

29

Você está aqui!

Práticas de design,
formatação e
apresentação dos
dados

semana

30

Relatório

**Visualização de
dados**

**Mapa da
Unidade 3
Componente 4**

Você está aqui!

Práticas de design, formatação
e apresentação dos dados

29

Aula 2

Código da aula:
[DADOS]ANO1C3B4S29A2



Objetivos da aula

- Conhecer como projetar, formatar e apresentar visualizações de dados eficazes de modo que sejam facilmente compreensíveis.



Recursos didáticos

- Recurso audiovisual para exibição de vídeos e imagens;
- Acesso ao laboratório de informática e/ou à internet.



Duração da aula

50 minutos.



Competências técnicas

- Criar e entender representações visuais.



Competências socioemocionais

- Desenvolver a capacidade de trabalhar em equipe para criar visualizações colaborativas e promover a comunicação eficaz ao apresentar e discutir os resultados.

Construindo
o **conceito**

Formatação de dados: padronização de unidades e escalas

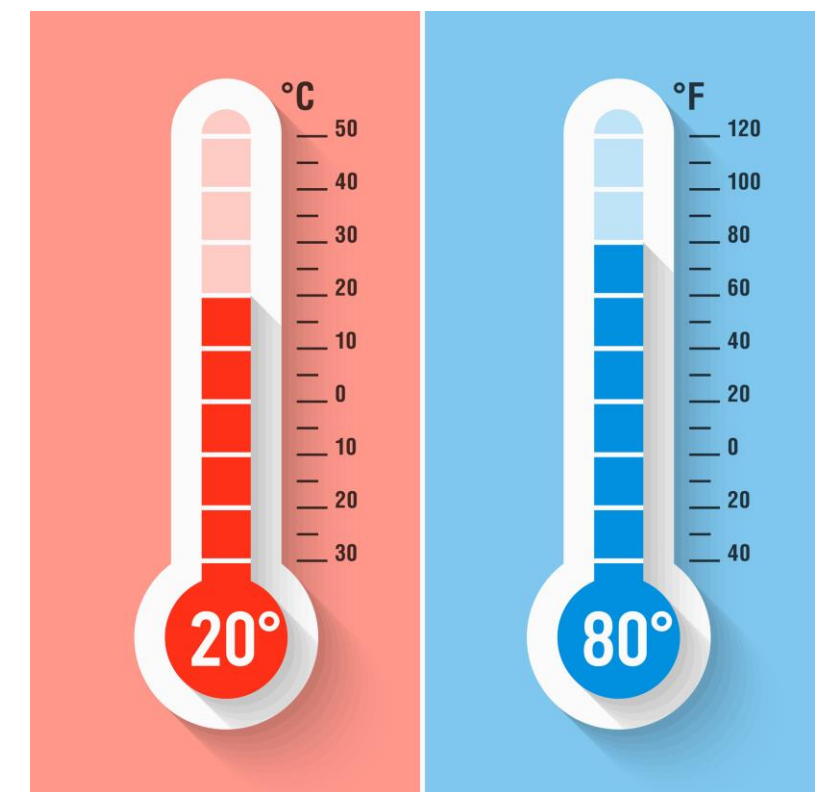
Quando estamos formatando dados, uma das coisas importantes a fazer é padronizar unidades e escalas.

Imagine que você está comparando temperaturas de vários países: se algumas estão em **Celsius** e outras em **Fahrenheit**, fica confuso, certo?



Tome nota

Padronizar significa escolher um único tipo de unidade, como Celsius, para que todos os dados estejam no mesmo "idioma". Isso ajuda a tornar tudo mais claro e as comparações, mais justas.



© Getty Images

Construindo o conceito

Formatação de dados: padronização de unidades e escalas

Quando falamos em padronizar escalas, estamos tentando garantir que todas as informações joguem no mesmo campo. É como garantir que todos os jogadores de um jogo sigam as mesmas regras para que ninguém tenha uma vantagem injusta.



Tome nota

Por exemplo, se você tem notas de escola que vão de 0 a 10 e outras de 0 a 100, ao padronizar, você ajusta tudo para uma escala comum, como 0 a 10. Isso faz com que seja mais fácil entender e analisar os dados, pois todos estão usando a mesma régua para medir.

X	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10
2	2	4	6	8	10	12	14	16	18	20
3	3	6	9	12	15	18	21	24	27	30
4	4	8	12	16	20	24	28	32	36	40
5	5	10	15	20	25	30	35	40	45	50
6	6	12	18	24	30	36	42	48	54	60
7	7	14	21	28	35	42	49	56	63	70
8	8	16	24	32	40	48	56	64	72	80
9	9	18	27	36	45	54	63	72	81	90
10	10	20	30	40	50	60	70	80	90	100

© Getty Images

Construindo
o **conceito**

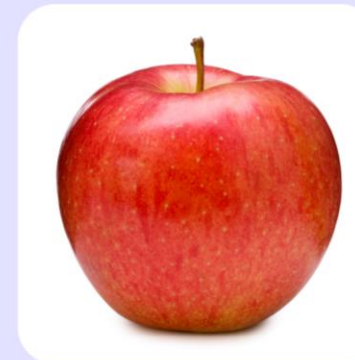
Formatação de dados: codificação de categorias

Codificação de categorias é uma maneira de transformar palavras ou etiquetas em números para que computadores possam entender e trabalhar com elas.

Esse processo é super útil, especialmente quando você está lidando com muitos dados diferentes.

Imagine que você tem uma lista de suas frutas favoritas, como maçã, banana e laranja, e você quer que um computador o ajude a organizar essas frutas.

maçã = 1



banana = 2



laranja = 3



Como os computadores são melhores com números, então você atribui a cada fruta um número e ele poderá facilmente entender e classificar sua lista.

Elaborado especialmente para o curso com imagem Getty Images.

Construindo o conceito

Formatação de dados: tratamento de datas e horas

Tratar datas e horas é como ajustar todos os relógios para que mostrem o mesmo horário, mesmo estando em diferentes lugares ou formatos.

Por exemplo, se você tem amigos em várias partes do mundo que dizem a hora que acordam, esses horários podem parecer confusos, porque cada lugar tem seu próprio fuso horário. Ao ajustar essas horas para um fuso padrão, fica muito mais fácil comparar a que horas todos acordam.



Tome nota

É o mesmo com datas em dados, você ajusta todas para um formato único (como dia/mês/ano) para que qualquer computador ou pessoa possa entender e usar facilmente.



© Getty Images

Construindo
o **conceito**

Formatação de dados: tratamento de datas e horas

As datas e horas podem vir em formatos diferentes, como "12 de março" ou "03/12". Se você está tentando organizar um evento com seus amigos e cada um usa um formato diferente, pode ficar confuso. Ao transformar todas as datas para o mesmo formato, você elimina a confusão e facilita a organização do evento.



Tome nota

Isso ajuda especialmente quando você está usando um computador para ajudar a planejar ou analisar informações, porque ele precisa de um formato consistente para processar as datas corretamente e fazer cálculos, como encontrar a diferença de dias entre duas datas.



© Getty Images

Construindo
o **conceito**

Formatação de dados: normalização ou padronização

Normalização e padronização são técnicas usadas para colocar diferentes tipos de dados na mesma escala, tornando-os mais fáceis de comparar e analisar.

Imagine que você está jogando uma partida de videogame em que você e seus amigos têm pontos que vão de 0 a 100 ou de 0 a 1.000. Será difícil comparar quem está se saindo melhor se os pontos não estiverem na mesma escala.

Normalização ajusta todos os pontos para uma escala de 0 a 100 ou de 0 a 1.000, então não importa o jogo original, todos os pontos agora são diretamente comparáveis.

Padronização é um pouco diferente, mas tem um objetivo semelhante. Ela ajusta os dados de modo que tenham uma média de zero e um desvio padrão de um.

Elaborado especialmente para o curso.

Construindo
o **conceito**

Formatação de dados: normalização ou padronização

É como ajustar as notas de uma prova para que a média da classe seja 50 com uma variação específica em torno dessa média, tornando fácil ver quem está acima ou abaixo da média.



© Getty Images



Tome nota

Ambas as técnicas ajudam a tratar os dados de maneira justa, assim, todos têm a chance de serem avaliados igualmente, seja em um jogo, numa prova escolar ou em análises de grandes quantidades de dados.



Colocando em **prática**

Estudo de caso

Tema: aplicativo de saúde

Contexto: a HealthTech Brasil é uma empresa de tecnologia focada no setor de saúde que está desenvolvendo um aplicativo inovador destinado a melhorar o gerenciamento e a experiência dos cuidados de saúde. A equipe de desenvolvimento enfrenta o desafio de inovar e adicionar rapidamente novas funcionalidades ao aplicativo, mantendo a segurança e a conformidade regulatória necessária para o manuseio de dados sensíveis de saúde.

Situação: a agilidade é essencial no desenvolvimento do aplicativo para responder rapidamente às necessidades do mercado e dos usuários. Contudo, é essencial garantir que os dados dos usuários, como informações médicas e históricos de saúde, sejam precisos, consistentes e tratados de forma segura. Isso inclui a padronização de unidades e escalas, codificação de categorias, tratamento adequado de datas e horas, além da normalização ou da padronização dos dados para análise e funcionalidade dentro do app.



15 min



Em grupo



Documento texto

Situação fictícia elaborada especialmente para o curso.



Colocando
em **prática**

Estudo de caso



15 min



Em grupo



Documento texto

1

Dada a necessidade de manter a conformidade regulatória e a integridade dos dados no aplicativo de saúde, sigam as etapas propostas:

2

Responda: como a HealthTech Brasil pode efetivamente aplicar as etapas de formatação de dados do CRISP-DM (padronização de unidades e escalas, codificação de categorias, tratamento de datas e horas, normalização ou padronização) para melhorar a precisão e a segurança dos dados dentro de seu aplicativo?

3

Descreva as ações específicas que seriam necessárias para implementar essas etapas, de forma a atender aos objetivos de negócio e regulatórios.

4

Envie a resposta do grupo para o Ambiente Virtual de Aprendizagem (AVA).



© Getty Images

O que nós
**aprendemos
hoje?**

Então ficamos assim...

- 1** Enfatizamos a necessidade de preparar e organizar os dados para que sejam fáceis de entender e analisar;
- 2** Exploramos como a codificação de categorias transforma informações não numéricas, como nomes ou etiquetas, em números que os computadores podem processar, permitindo análises mais eficientes e complexas em grandes conjuntos de dados;
- 3** Abordamos a importância de uniformizar datas e horas, ajustando todos os registros para um formato comum.

Saiba mais

Quer explorar o mundo dos dados e tirar o máximo proveito deles e não sabe por onde começar?

A Análise Exploratória de Dados (EDA) é a ferramenta que leva para o próximo nível! Com ela, você terá uma visão panorâmica e conseguirá extrair boas ideias para seus projetos.

GONZALEZ, V. Análise exploratória: primeiros passos.

Alura, 24 set. 2020. Disponível em:

<https://www.alura.com.br/artigos/analise-exploratoria>.

Acesso em 26 jul. 2024.

Referências da aula

KNAFLIC, C. N. *Storytelling com dados*: um guia sobre visualização de dados para profissionais de negócios. Rio de Janeiro: Alta Books, 2019.

Identidade visual: imagens © Getty Images

Educação Profissional Paulista

Técnico em
**Ciência de
Dados**