



Disciplina Inteligência Artificial	Departamento Ciência da Computação	Turno Manhã	Período 9º
Professor Hugo de Paula (hugo@pucminas.br)			

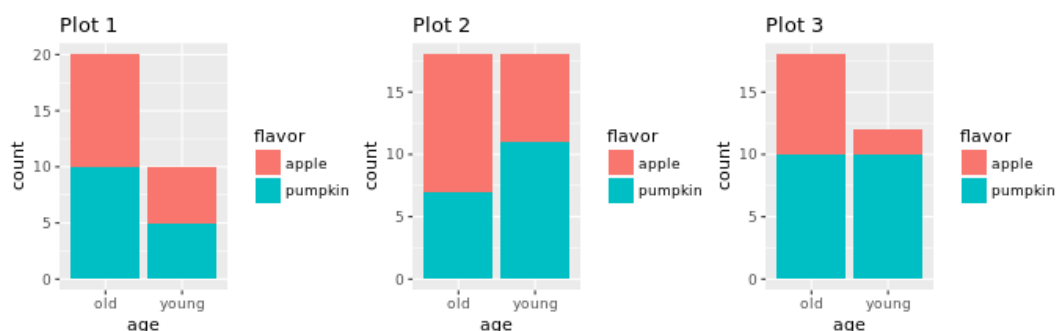
Exercício 1

1. (Adaptado de Tan, Steinbach e Kumar, 2006) Classifique os tipos de atributos a seguir baseados nos seus tipos de dados: qualitativo (polinomial ou ordinal) ou quantitativo (intervalar ou razão), binomial, discreto ou contínuo.

Em alguns casos poderá haver mais de uma interpretação. Neste caso, indique o raciocínio utilizado para cada interpretação.

Exemplo: Idade em anos. Resposta: Discreto, quantitativo, razão.

- (a) Tempo em termos de AM e PM (manhã e tarde).
 - (b) Brilho medido por um sensor de luminosidade.
 - (c) Brilho medido pelo julgamento de uma pessoa.
 - (d) Ângulo medido em graus entre 0° e 360° .
 - (e) Ouro, prata e bronze conquistados nas olimpíadas.
 - (f) Altura acima do nível do mar.
 - (g) Número de pacientes em um hospital.
 - (h) Números ISBN de livros.
 - (i) Capacidade de permitir a passagem da luz em termos dos valores: opaco, transluzente e transparente.
 - (j) Patente militar.
 - (k) Distância do centro do campus.
 - (l) Densidade de uma substância em gramas por centímetro cúbico.
2. Suponha que você perguntou 30 pessoas, algumas jovens, outras adultas, qual seu sabor preferido de torta: torta de maçã ou torta de abóbora. Os dados obtidos podem ser visualizados pelos gráficos em barra a seguir.



Matrícula:	Aluno:
------------	--------

Qual dos gráficos de barra não exibe nenhuma relação entre idade e sabor da torta?

- Reproduza dos gráficos da Questão 2 em Python. Crie um **DataFrame** que represente a tabela de dados, e utilize o **Matplotlib** para exibir os gráficos. Obs.: Os gráficos não precisam ficar idênticos, mas apresentarem a mesma informação.
- O dataset *InsectSprays*, disponível no SGA, relata o número de insetos encontrados em unidades agrícolas experimentais tratadas com diferentes inseticidas. Faça um boxplot (usando Python) e determine qual inseticida parece ser mais eficiente. Justifique sua resposta apresentando o gráfico e sua interpretação do mesmo.
- Cite duas medidas estatísticas de localização de dados numéricos. Pontue que característica dos dados deve ser considerada na escolha da melhor medida de localização.
- Utilize o dataset de Microdados do MEC, disponível no SGA, para realizar a limpeza, transformação e enriquecimento dos dados. Realize pelo menos 4 operações sobre os dados, tais como agregação de valores, substituição de dados omissos, enriquecimento dos dados, etc.

Todo o código Python desenvolvido para as atividades deve ser postado no SGA.