



Trabalho em grupo

Análise é definida com um processo de separação para melhor entender o todo.

A análise de dados é um processo importante para melhor entendê-los.

O WizRule é um programa desenvolvido pra examinar e descrever conjuntos de dados, detectando erros e exceções entre eles, e é útil na detecção de desvios por meio da extração de regras tipo $SE \ x \ ENTÃO \ y$. Relatórios de desvios são particularmente úteis em auditoria e detecção de fraudes.

WizRule pode ajudar a entender as relações entre os campos que aparentemente não tem nenhuma relação, contribuindo assim para melhor entendimento sobre a compreensão dos dados analisados. Além disso o WizRule pode (a) revelar fórmulas matemáticas que relacionam os campos; (b) revelar irregularidades de ortografia; e (c) exibir regras nos modos texto e gráfico.

Observe que para realizar uma análise é fundamental a compreensão dos dados (o que eles representam no todo e o significado de cada campo).

- 1) Utilize o WizRule sobre uma base de dados vistas em sala de aula (laboratório) e indique:
 - a. as principais regras que apontem “desvios” nos dados; procure descrever os seus significados numa frase que possa ser interpretada por qualquer pessoa que não seja um especialista na base de dados que está sendo analisada (por exemplo, a regra “SE (data_nasc < 2000) E (sexo = “1”) ENTÃO (registro ≠ vazio)” tem como exceção os registros 31 e 79, deve ser escrita como “verificar os registros 31 e 79 pois os funcionários tem mais de 16 anos, do sexo masculino e não tem registro de trabalho)
 - b. para cada regra/exceção, apresente uma possível solução (por exemplo, no caso acima, pode ser: “verificar o registro do funcionário de matrícula XXX” – o número da matrícula foi feita na base de dados! – ou “corrigir o registro do funcionário de matrícula XXX”)
- 2) Faça a questão anterior sobre uma base de dados de sua escolha. Observe que a base de dados deve ser no formato de tabela! Caso os dados estejam dispersos em mais de uma tabela, deve ser feito antes um pré-processamento para “juntar” os campos adequados das diversas tabelas que contém os dados desejados.

Responda:

- 3) Qual a diferença entre análise de fatos e análise de dados? Cite um ou mais exemplos.
- 4) Apresente uma decisão tomada com base na (i) razão; (ii) intuição; (iii) emoção; e (iv) conhecimento.
- 5) Que dados podem ser obtidos (i) nos sites da Internet; (ii) nas redes sociais; (iii) num sistema de sua empresa ou escola.
- 6) Conceitue mostrando as principais diferenças: (i) dados; (ii) informação; e (iii) conhecimento.
- 7) Dê um ou mais exemplos de conhecimento tácito. Descreva-o explicitamente.
- 8) De que forma podem ser representados/descritos (i) os dados do censo demográfico; (ii) o percurso entre 2 localidades; (iii) organização da diretoria de uma empresa. Apresente uma forma de representação alternativa para cada uma e descreva a principal vantagem que diferencia a melhor em relação à outra.
- 9) Defina com suas palavras os 5 V's (volume, velocidade, variabilidade, valor e veracidade) que caracterizam o ambiente *bigdata*.
- 10) Defina com suas palavras as principais características das análises (i) descritiva; (ii) diagnóstica; (iii) preditiva; e (iv) prescritiva.