

e-Fólio B

U.C. 21103

Sistemas de Gestão de Bases de Dados

2025-2026

Resolução e Critérios de Correção

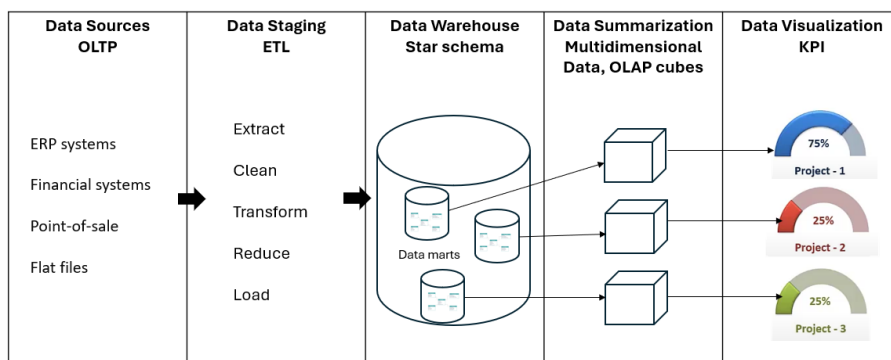
INSTRUÇÕES

- 1) O e-fólio é constituído por 5 perguntas. A cotação global é de 5 valores.
- 2) O e-fólio deve ser entregue num único ficheiro PDF, não zipado, com fundo branco, com perguntas numeradas e sem necessidade de rodar o texto para o ler. Cada pergunta com uma ou mais páginas, deve ser iniciada numa nova página. Penalização de 10% a 100%.
- 3) Não são aceites e-fólios manuscritos, i.e., tem penalização de 100%.
- 4) O nome do ficheiro: <nome estudante> + “eFolioB”.
- 5) Na primeira página do e-fólio deve constar o nome completo do estudante bem como o seu número. Penalização de 10% a 100%.
- 6) Durante a realização do e-fólio, os estudantes devem concentrar-se na resolução do seu trabalho individual, não sendo permitida a colocação de perguntas ao professor ou entre colegas.
- 7) Nesta avaliação, dada a impossibilidade de os estudantes não usarem a IAgen, use-a de forma ética e crítica, caso contrário será penalizado.
- 8) A interpretação das perguntas também faz parte da sua resolução, se encontrar alguma ambiguidade deve indicar claramente como foi resolvida.
- 9) A legibilidade, a objetividade e a clareza nas respostas serão valorizadas, pelo que, a falta destas qualidades será penalizada.
- 10) Critérios de correção gerais: todas as respostas devem ser justificadas, incluir imagens e exemplos com vista a clarificar os argumentos expostos. Devem ser utilizadas referências das páginas da bibliografia adotada e recomendada.

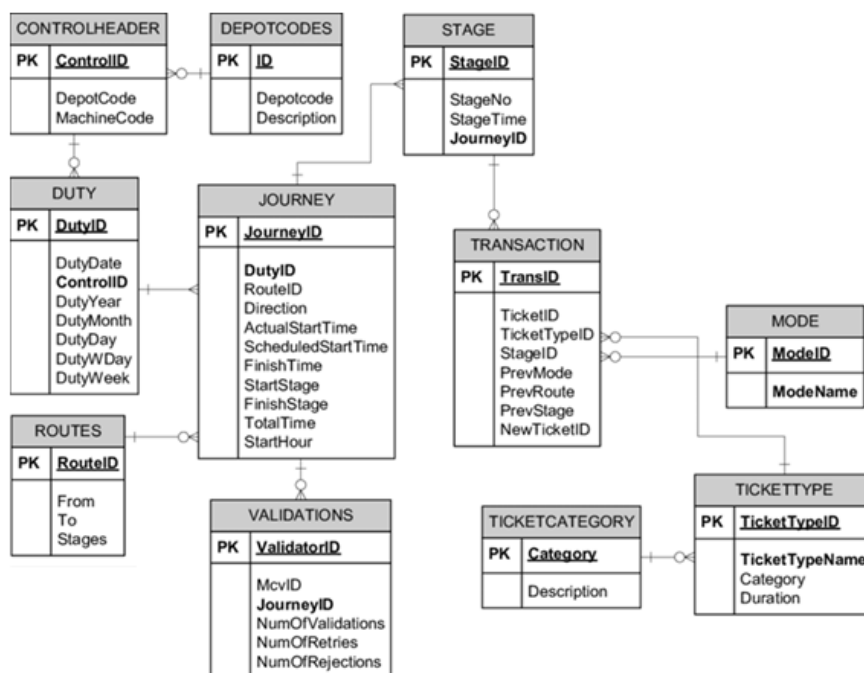
Vetor Cotações

1 2 3 4 5 pergunta
10 10 10 10 10 décimas

Um Data Warehouse (DW) é um sistema projetado para armazenar grandes volumes de dados originados de diversas fontes, com o objetivo de apoiar a tomada de decisões numa organização. Considere o modelo das 5 fases dos slides narrados.



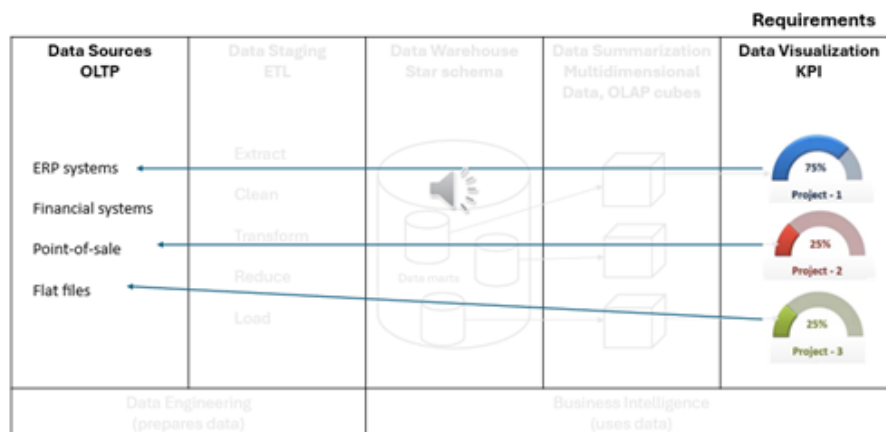
Considere a seguinte base de dados OLTP de uma empresa de transportes, para as perguntas seguintes:



□

1) (1 valor) Fontes dos dados

Considere a base de dados de uma empresa de transportes e o modelo das 5 fases dos slides narrados.



1.1) Descreva o que entende do esquema de bases de dados.

1.2) Quais as “event tables” que encontra no esquema OLTP da empresa de transportes?

1.3) Que requisitos podem ser pedidos e a que tabelas de eventos? Preencha a tabela.

requisitos\tabela eventos	A	B	C
requisito 1	X		
requisito 2		X	
requisito 3			X

Resposta:

1.1) Acerca do sistema da base de dados de gestão de transportes públicos.

Esta base de dados foi concebida para gerir de forma integrada todo o sistema de transportes públicos, abrangendo desde o planeamento operacional até ao controlo tarifário e análise de mobilidade urbana. O sistema organiza-se em torno de várias tabelas interligadas que cobrem diferentes aspetos da operação: Journey e Stage gerem viagens e percursos, Routes e Mode definem a rede disponível, Duty gere turnos de operadores, e Tickettype classifica os títulos de transporte. No entanto, dois componentes assumem particular importância no funcionamento quotidiano: as tabelas Validations e Transactions, que embora complementares, têm objetivos distintos, mas igualmente cruciais para o sucesso operacional do sistema.

A tabela Validations funciona como o guardião do sistema, registando cada interação entre passageiros e validadores. Cada passageiro possui um McvID, i.e., um identificador do cartão/título de transporte ("Machine Card Validation ID"). O seu objetivo fundamental é controlar o acesso ao sistema de transportes, distinguindo entre validações bem-sucedidas, tentativas repetidas e rejeições.

A tabela Transactions acompanha cada título de transporte ao longo de toda a jornada do passageiro, funcionando como um diário de viagem completo. Os seus objetivos principais incluem rastrear a utilização de cada bilhete, gerir transbordos entre diferentes modos de transporte através dos campos PrevMode, PrevRoute e PrevStage, e garantir o controlo tarifário verificando se o tipo de bilhete usado é adequado ao percurso realizado. Esta tabela permite também analisar padrões de mobilidade urbana, identificando rotas mais utilizadas e otimizando a oferta de serviços, e gerir o ciclo de

vida dos títulos através do campo NewTicketID que regista recargas e substituições.

1.2) As “event tables” OLTP da empresa de transportes são:

- Validations
- Transactions

Visto que são as únicas tabelas apenas com cardinalidade N, i.e., apenas com Foreign Keys.

1.3) Requisitos vs Eventos

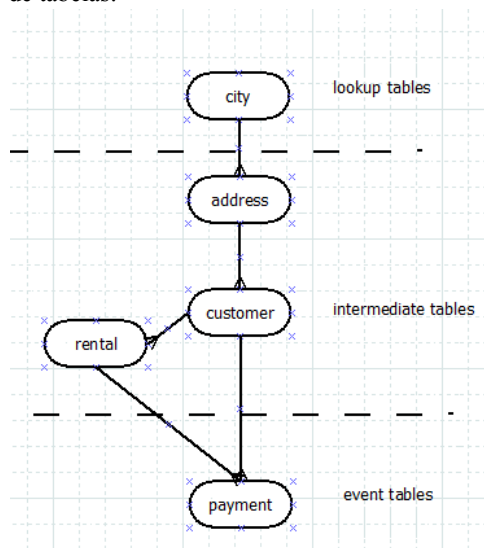
requisitos\tabela eventos	Validations	Transactions
controlar o acesso ao sistema de transportes, distinguindo entre validações bem-sucedidas, tentativas repetidas e rejeições	X	
rastrear a utilização de cada bilhete, gerir transbordos entre diferentes modos de transporte		X

Critérios de correção:

- 1.1) 3 décimas, descrição do sistemas de gestão de transportes públicos
- 1.2) 3 décimas, duas tabelas de eventos; Validations and Transactions
- 1.2) 4 décimas, requisitos versus eventos
- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%

2) (1 valor) Extração dos dados

Considere a base de dados de uma empresa de transportes. Considere os seguintes tipos de tabelas:



- Lookup tables: são tabelas com apenas cardinalidade igual a um, i.e., apenas com PK (primary keys);

- Intermediate: são para tabelas com cardinalidade 1 e N, i.e., com PKs e FKs (foreign keys);

- Event tables: são para tabelas apenas com cardinalidade N, i.e., apenas com FK.

Considere ainda, as seguintes formas desnormalizadas (FD):

- 1FD – constituída por árvores e poli-árvores separadas;
- 2FD – constituída por várias árvores separadas, com a replicação das tabelas intermédias e de lookup que forem necessárias para evitar caminho múltiplos; esta FD é equivalente ao esquema em estrela ou ao esquema floco-de-neve.
- 3FD – o processo de desnormalização termina com a junção de todas as tabelas da árvore com vista a uma rápida leitura dos dados.

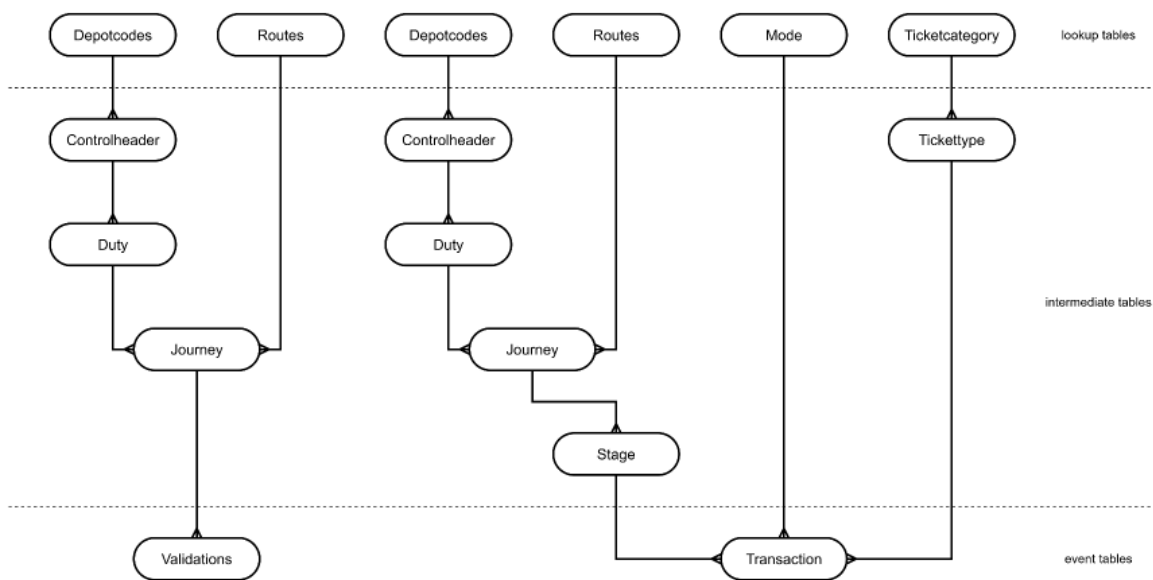
2.1) Represente graficamente os dados na 2FD da seguinte forma: as ligações de 1:N, a tabela com uma única linha é desenhada em cima e a tabela com várias linhas é desenhada por baixo.

2.2) Depois de representar as tabelas classifique-as segundo a tipologia indicada (lookup, intermédia, eventos) como na figura.

Resposta:

2.1) Representação gráfica dos dados na 2FD

2.2) com classificação da tipologia (lookup, intermédia, eventos)



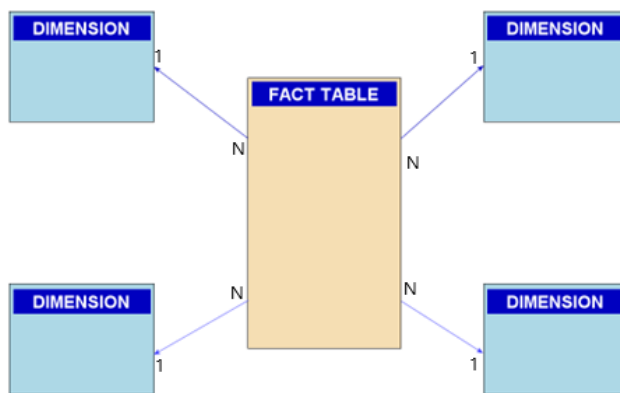
Critérios de correção:

- 2.1) 7 décimas, representação da base dados na 2FD
- 2.2) 3 décimas, tipologia das tabelas
- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%

3) (1 valor) Data warehouse: star schema

Considere novamente a base de dados da empresa de transportes, em particular a tabela de eventos com a informação de Transaction.

3.1) Crie um Data Mart, com um esquema em estrela (ou floco de neve), com os dados de Transaction com pelo menos três dimensões como o da figura.



3.2) Considere os seguintes tipos de atributos da tabela de eventos:

- Aditivos: são atributos que podem ser agregados (somados) por todas as dimensões, ex: valor da venda (usar Sum() sempre)
- Semi-aditivos: são atributos que podem ser agregados (somados) por algumas as dimensões, ex: quantidade (usar Sum() em condições particulares)
- Não-aditivos: são atributos que não podem ser agregados (somados), ex: preço unitário (usar Average() por exemplo)
- Sem factos: só existem identificadores (usar a função Count() dos identificadores).

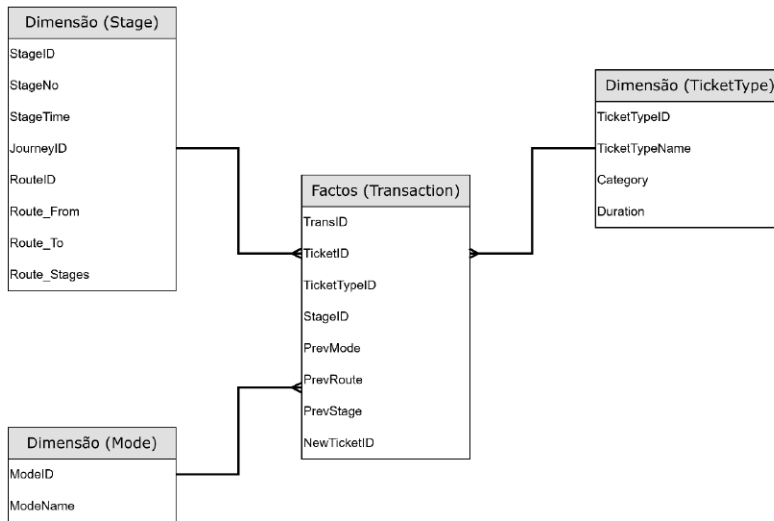
Para a tabela de factos encontrada defina os atributos aditivos, semi-aditivos, não-aditivos e sem factos. Justifique a resposta.

3.3) Diga quais das seguintes perguntas são passíveis de serem realizadas (responda verdadeiro ou falso). Justifique as respostas.

- a) Quais as transações realizadas no Metro?
- b) Quantas transações usaram o bilhete TKxxx?
- c) Quais os estágios/estações de destino das viagens de comboio?
- d) Qual o valor pago em cada transação?
- e) Qual o nome do passageiro que realizou a transação TRxxx?
- f) A que horas foi realizada cada transação?

Resposta:

3.1) Estrela (ou floco de neve), com os dados de Transaction com pelo menos três dimensões



3.2) Tipo de atributos da tabela de factos Transaction: Sem factos, só existem identificadores (usar a função Count() dos identificadores).

3.3) Perguntas possíveis (verdadeiras ou falsas)

- | | | |
|---|------------|--------------------------------------|
| a) Transações no Metro | VERDADEIRO | atributo PrevMode existe |
| b) Quantas transações do bilhete TKxxx | VERDADEIRO | atributo TicketID existe |
| c) Estágios/estações destino de comboio | VERDADEIRO | atributos PrevMode e StageID existem |
| d) Valor pago | FALSO | sem atributos monetários |
| e) Nome do passageiro | FALSO | sem atributos pessoais |
| f) Hora da transação | FALSO | sem atributos temporais |

CrITÉRIOS de correção:

- 3.1) 4 décimas, representação em estrela
- 3.2) 3 décimas, tipologia dos atributos: sem factos
- 3.3) 3 décimas, perguntas possíveis: 3 verdadeiras e 3 falsas
- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%

4) (1 valor) Sumarização de dados: operações OLAP

Considere os seguintes dados da tabela de factos com a informação de Transaction.

TransID	TicketID	TicketTypeID	StageID
TRX001	TKT20241220001	BilheteSimples	Rossio
TRX002	TKT20241220002	BilheteDiario	Cais_Sodre
TRX003	TKT20241220003	PasseMensal	Oriente
TRX004	TKT20241220004	BilheteSimples	Belem
TRX005	TKT20241220005	PasseEstudante	Alameda
TRX006	TKT20241220006	BilheteSimples	Carnide
TRX007	TKT20241220002	BilheteDiario	Terreiro_Paco
TRX008	TKT20241220007	RecargaCartao	Entrecampos
TRX009	TKT20241220009	PasseMensal	Cascais
TRX010	TKT20241220010	BilheteSimples	Aeroporto

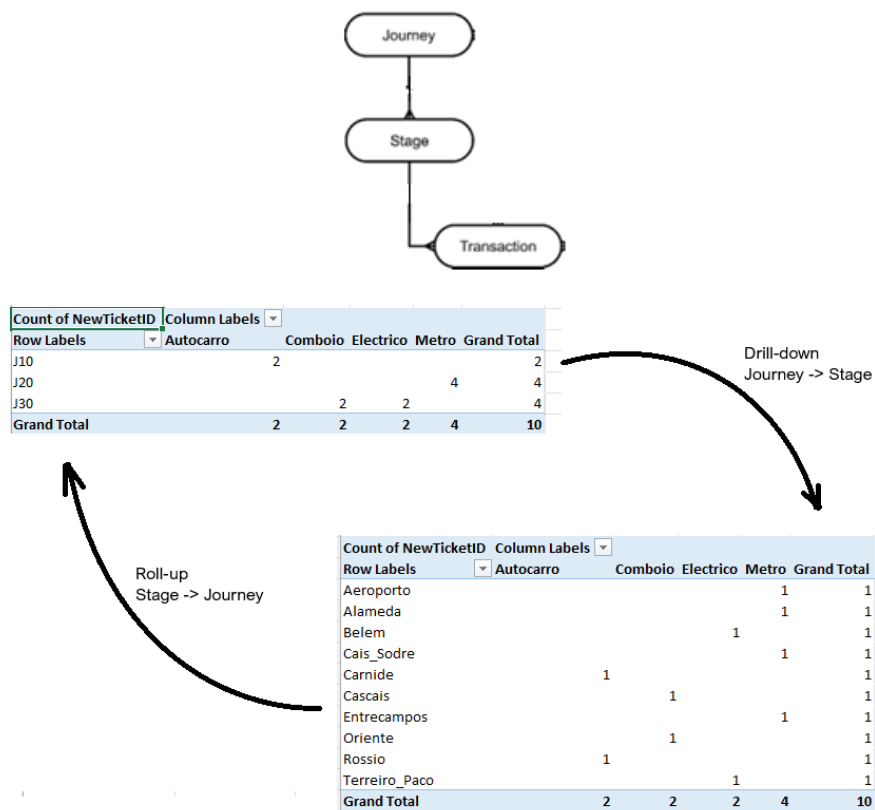
PrevMode	PrevRoute	PrevStage	NewTicketID
Autocarro	Linha28	Martim_Moniz	TKT20241220001
Metro	LinhaVerde	Baixa_Chiado	TKT20241220002
Comboio	LinhaAzul	Santa_Apolonia	TKT20241220003
Electrico	Linha15	Algés	TKT20241220004
Metro	LinhaVermelha	Saldanha	TKT20241220005
Autocarro	Linha750	Amadora	TKT20241220006
Electrico	Linha11	Cais_Sodre	TKT20241220002
Metro	LinhaAmarela	Campo_Pequeno	TKT20241220008
Comboio	LinhaCascais	Oeiras	TKT20241220009
Metro	LinhaVermelha	Oriente	TKT20241220010

4.1) Explique e exemplifique em Excel o que entende por Roll-up e Drill-down,

4.2) Explique e exemplifique em Excel o que entende por Slide e Dice

Resposta:

4.1) O Roll-up e o Drill-down consideram a hierarquia dos dados. Vamos considerar a hierarquia Journey -< Stage -< Transaction.



4.2) Slide e Dice é baseado em filtros. Slice corresponde a filtro único e Dice é a uma conjunção de filtros.

TransID	Journey	StageID	PrevMode	PrevRoute	PrevStage
TRX001	J10	Rossio	Autocarro	Linha28	Martim_Moniz
TRX002	J20	Cais_Sodre	Metro	LinhaVerde	Baixa_Chiado
TRX003	J30	Oriente	Comboio	LinhaAzul	Santa_Apolonia
TRX004	J30	Belem	Electrico	Linha15	Algés
TRX005	J20	Alameda	Metro	LinhaVermelha	Saldanha
TRX006	J10	Carnide	Autocarro	Linha750	Amadora
TRX007	J30	Terreiro_Paco	Electrico	Linha11	Cais_Sodre
TRX008	J20	Entrecampos	Metro	LinhaAmarela	Campo_Pequeno
TRX009	J30	Cascais	Comboio	LinhaCascais	Oeiras
TRX010	J20	Aeroporto	Metro	LinhaVermelha	Oriente

TransID	Journey	StageID	PrevMode	PrevRoute	PrevStage
TRX002	J20	Cais_Sodre	Metro	LinhaVerde	Baixa_Chiado
TRX005	J20	Alameda	Metro	LinhaVermelha	Saldanha
TRX008	J20	Entrecampos	Metro	LinhaAmarela	Campo_Pequeno
TRX010	J20	Aeroporto	Metro	LinhaVermelha	Oriente

TransID	Journey	StageID	PrevMode	PrevRoute	PrevStage
TRX005	J20	Alameda	Metro	LinhaVermelha	Saldanha
TRX010	J20	Aeroporto	Metro	LinhaVermelha	Oriente

Crítérios de correção:

- 4.1) 5 décimas, roll-up e drill-down com exemplos
- 4.2) 5 décimas, slice and dice com exemplos
- sem exemplos penalização de 4 décimas e 5 décimas
- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%

5) (1 valor) Visualização dados e KPI

Considere a seguinte informação dos slides narrados:

- KPI: key performance indicators
- KPI = metric + target
- Alerts: what is happening now?



5.1) O que entende por métrica e KPI?

5.2) Para a tabela de eventos Validations encontre uma métrica e um KPI.

Resposta:

5.1) Métrica e KPI

Métrica: é uma medida quantificável utilizada para acompanhar e avaliar o desempenho de um processo, atividade ou sistema.

KPI: é um indicador-chave de desempenho que resulta da combinação de uma métrica com um objetivo ou meta específica.

5.2) Para a tabela Validations temos com métrica e um KPI:

Métrica: Taxa de Rejeição de Validações, percentagem de validações rejeitadas em relação ao total de tentativas de validação.

KPI: Taxa de Rejeição de Validações com uma Meta (Target)


Métrica: Taxa de Rejeição = 3%


Target (Meta): $\leq 2\%$

Status: ALERTA - Acima da meta aceitável

Interpretação:

 Crítico: $> 5\%$ (requer intervenção imediata)

 Atenção: $2\% \dots 5\%$ (monitorizar e investigar)

 Bom: $\leq 2\%$ (dentro do esperado)

No exemplo: 3% está na zona amarela, indicando necessidade de investigação

Critérios de correção:

- 5.1) 3 décimas, definições de métrica e KPI ($KPI = \text{métrica} + \text{meta}$)

- 5.2) 7 décimas, exemplos para Validations, ex.: Taxa de Rejeição de Validações

- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%