Academia SQL + Teradata

Antes de realizarmos transformações nos dados, é importante estabelecermos processos que apontem erros de qualidade nos dados que estamos trabalhado, dessa forma é possível ter clareza das inconsistências comuns e assim criar formas de melhorar a qualidade dos dados. Pensando nisso, a primeira atividade planejada é criarmos em cada tabela colunas adicionais reportando o tipo de inconsistência que encontrarmos. Obtendo os dados Dados previamente carregados no database “marcos.nagato@capgemini.com” no ambiente Lab da Teradata com os nomes de tabela airports, planes e flights. Criar as estruturas no database do seu próprio usuário, e copiar os dados das 3 tabelas. As permissões apropriadas foram previamente atribuídas. Airports Dataset

*--Criar as estruturas no database do seu próprio usuário, e copiar os dados das 3 tabelas.*

*--As permissões apropriadas foram previamente atribuídas.*

*--Entra no banco de dados*

**database** pollyana.ferreira@capgemini.com**;**

*--Visualiza a tabela que quer copiar*

**show** **table** "marcos.nagato@capgemini.com".airports**;**

*--Copia os dados do Show table altera o email e executa pra criar uma copia do banco*

**CREATE** **SET** **TABLE** "pollyana.ferreira@capgemini.com".airports ,**NO** **FALLBACK** ,

**NO** **BEFORE** **JOURNAL**,

**NO** **AFTER** **JOURNAL**,

**CHECKSUM** = **DEFAULT**,

**DEFAULT** **MERGEBLOCKRATIO**

(

faa VARCHAR(5) CHARACTER **SET** LATIN **NOT** CASESPECIFIC,

**name** VARCHAR(51) CHARACTER **SET** LATIN **NOT** CASESPECIFIC,

lat FLOAT,

lon FLOAT,

alt INTEGER,

tz INTEGER COMPRESS (5 ,6 ,7 ,8 ,-11 ,-10 ,-9 ,-8 ,-7 ,-6 ,-5 ,-4 ),

dst CHAR(1) CHARACTER **SET** LATIN **NOT** CASESPECIFIC)

**PRIMARY** **INDEX** ( faa )**;**

*--copia os dados da tabela airports*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".airports **SELECT** \* **FROM** "marcos.nagato@capgemini.com".airports**;**

**select** \* **from** airports**;**

*---copia os dados da tabela flights*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".flights **SELECT** \* **FROM** "marcos.nagato@capgemini.com".flights**;**

**select** \* **from** flights**;**

*--copia os dados da tabela planes*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".planes **SELECT** \* **FROM** "marcos.nagato@capgemini.com".planes**;**

**select** \* **from** planes**;**

*--Criar as estruturas no database do seu próprio usuário, e copiar os dados das 3 tabelas.*

*--As permissões apropriadas foram previamente atribuídas.*

*--Entra no banco de dados*

**database** pollyana.ferreira@capgemini.com**;**

*--Visualiza a tabela que quer copiar*

**show** **table** "marcos.nagato@capgemini.com".airports**;**

*--Copia os dados do Show table altera o email e executa pra criar uma copia do banco*

**CREATE** **SET** **TABLE** "pollyana.ferreira@capgemini.com".airports ,**NO** **FALLBACK** ,

**NO** **BEFORE** **JOURNAL**,

**NO** **AFTER** **JOURNAL**,

**CHECKSUM** = **DEFAULT**,

**DEFAULT** **MERGEBLOCKRATIO**

(

faa VARCHAR(5) CHARACTER **SET** LATIN **NOT** CASESPECIFIC,

**name** VARCHAR(51) CHARACTER **SET** LATIN **NOT** CASESPECIFIC,

lat FLOAT,

lon FLOAT,

alt INTEGER,

tz INTEGER COMPRESS (5 ,6 ,7 ,8 ,-11 ,-10 ,-9 ,-8 ,-7 ,-6 ,-5 ,-4 ),

dst CHAR(1) CHARACTER **SET** LATIN **NOT** CASESPECIFIC)

**PRIMARY** **INDEX** ( faa )**;**

*--copia os dados da tabela airports*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".airports **SELECT** \* **FROM** "marcos.nagato@capgemini.com".airports**;**

**select** \* **from** airports**;**

*---copia os dados da tabela flights*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".flights **SELECT** \* **FROM** "marcos.nagato@capgemini.com".flights**;**

**select** \* **from** flights**;**

*--copia os dados da tabela planes*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".planes **SELECT** \* **FROM** "marcos.nagato@capgemini.com".planes**;**

**select** \* **from** planes**;**

*--Criar as estruturas no database do seu próprio usuário, e copiar os dados das 3 tabelas.*

*--As permissões apropriadas foram previamente atribuídas.*

*--Entra no banco de dados*

**database** pollyana.ferreira@capgemini.com**;**

*--Visualiza a tabela que quer copiar*

**show** **table** "marcos.nagato@capgemini.com".airports**;**

*--Copia os dados do Show table altera o email e executa pra criar uma copia do banco*

**CREATE** **SET** **TABLE** "pollyana.ferreira@capgemini.com".airports ,**NO** **FALLBACK** ,

**NO** **BEFORE** **JOURNAL**,

**NO** **AFTER** **JOURNAL**,

**CHECKSUM** = **DEFAULT**,

**DEFAULT** **MERGEBLOCKRATIO**

(

faa VARCHAR(5) CHARACTER **SET** LATIN **NOT** CASESPECIFIC,

**name** VARCHAR(51) CHARACTER **SET** LATIN **NOT** CASESPECIFIC,

lat FLOAT,

lon FLOAT,

alt INTEGER,

tz INTEGER COMPRESS (5 ,6 ,7 ,8 ,-11 ,-10 ,-9 ,-8 ,-7 ,-6 ,-5 ,-4 ),

dst CHAR(1) CHARACTER **SET** LATIN **NOT** CASESPECIFIC)

**PRIMARY** **INDEX** ( faa )**;**

*--copia os dados da tabela airports*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".airports **SELECT** \* **FROM** "marcos.nagato@capgemini.com".airports**;**

**select** \* **from** airports**;**

*---copia os dados da tabela flights*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".flights **SELECT** \* **FROM** "marcos.nagato@capgemini.com".flights**;**

**select** \* **from** flights**;**

*--copia os dados da tabela planes*

**INSERT** **INTO** "pollyana.ferreira@capgemini.com".planes **SELECT** \* **FROM** "marcos.nagato@capgemini.com".planes**;**

**select** \* **from** planes**;**

*----------------------Airpots-----------------------------------------*

*--Tarefas*

*--1. Crie a coluna qa\_faa e aponte inconsistências da coluna faa de acordo com as regras abaixo:*

**ALTER** **TABLE** airports

**ADD** qa\_faa VARCHAR(5) CHARACTER **SET** LATIN **NOT** CASESPECIFIC

**;**

*--M : Indica que está com dado faltante.*

**insert** **into** airports(faa, **name**, lat, lon, alt, tz, dst)

**values**("124jjhd", "Whiteman Airport", 34.26, -118.41, 1.003, -8, "A")

**;**

**update** airports **set** qa\_faa = 'M'

**where** (faa **is** **null**

**or** trim(faa) = ''

**or** length(trim(faa)) = 0

)**;**

*--F : Indica que não respeita o formato de 3-5 caracteres alfanuméricos.*

**update** airports **set** qa\_faa = 'F'

**where** (faa **is** **not** **null**

**and** trim(faa) <> ''

**and** (length(trim(faa)) < 3

**or** length(trim(faa)) > 5)

)**;**

*--2. Crie a coluna qa\_name e aponte inconsistências da coluna name de acordo com as regras abaixo:*

**ALTER** **TABLE** airports

**ADD** qa\_name VARCHAR(51) CHARACTER **SET** LATIN **NOT** CASESPECIFIC

**;**

*--M : Indica que está com dado faltante.*

**update** airports **set** qa\_name = 'M'

**where** (**name** **is** **null**

**or** trim(**name**) = ''

**or** length(trim(**name**)) = 0

)**;**

*--3. Crie a coluna qa\_lat e aponte inconsistências da coluna lat de acordo com as regras abaixo:*

**ALTER** **TABLE** airports

**ADD** qa\_lat char(1)

**;**

*--M : Indica que está com dado faltante.*

**update** airports **set** qa\_lat = **null;**

**update** airports **set** qa\_lat = 'M'

**where** (lat **is** **null** **or** trim(lat) = '' **or** length(trim(lat)) = 0)

**;**

*--I : Indica que o valor excede o intervalo [−180, 180].*

**update** airports **set** qa\_lat = 'I'

**where** lat **not** between -180 and 180

**;**

*--A : Indica que o valor é alfanumérico.*

*--anulada*

*--4. Crie a coluna qa\_lon e aponte inconsistências da coluna lon de acordo com as regras abaixo:*

**ALTER** **TABLE** airports

**ADD** qa\_lon char(1)

**;**

*--M : Indica que está com dado faltante.*

**update** airports **set** qa\_lon = **null;**

**update** airports **set** qa\_lon = 'M'

**where** (lon **is** **null** **or** trim(lon) = '' **or** length(trim(lon)) = 0)**;**

*-- I : Indica que o valor excede o intervalo [−180, 180].*

**update** airports **set** qa\_lon = 'I'

**where** lon **not** between -180 and 180

**;**

*--A : Indica que o valor é alfanumérico.*

*--anulada*

*--5. Crie a coluna qa\_alt e aponte inconsistências da coluna alt de acordo com as regras abaixo:*

**ALTER** **TABLE** airports

**ADD** qa\_alt char(1)

**;**

*--M : Indica que está com dado faltante.*

**update** airports **set** qa\_alt = 'M'

**where** (alt **is** **null** **or** trim(alt) = '' **or** length(trim(alt)) = 0)**;**

*--I : Indica que o valor excede o intervalo [-100,+∞].*

**update** airports **set** qa\_alt = **null;**

**update** airports **set** qa\_alt = 'I'

**where** alt **not** between > -100**;**

*--A : Indica que o valor é alfanumérico.*

*--anulada*

*--6. Crie a coluna qa\_tz e aponte inconsistências da coluna tz de acordo com as regras abaixo.*

**alter** **table** airports **drop** qa\_tz**;**

**ALTER** **TABLE** airports

**ADD** qa\_tz char(1)

**;**

*--M : Indica que está com dado faltante.*

**update** airports **set** qa\_tz = **null;**

**update** airports **set** qa\_tz = 'M'

**where** (tz **is** **null** **or** trim(tz) = '' **or** length(trim(tz)) = 0)**;**

*--I : I : Indica que o valor excede o intervalo [−11,+14].*

**update** airports **set** qa\_tz = 'I'

**where** tz **not** between -11 and 14

**;**

*--A : Indica que o valor é alfanumérico.*

*--anulada*

*--*

*--7. Crie a coluna qa\_dst e aponte inconsistências da coluna dst de acordo com as regras abaixo:*

**ALTER** **TABLE** airports

**ADD** qa\_dst char(1)

**;**

*--M : Indica que está com dado faltante.*

**update** airports

**set** qa\_dst = 'M'

**where** to\_number (dst) >= 0 **and** to\_number (dst) <= 9**;**

*--C : Indica que o valor não pertence a nenhuma das categorias esperadas: E, A, S, O, Z, N, U.*

**update** airports

**set** qa\_dst = 'C'

**Where** dst **NOT** **IN**('E', 'A', 'S', 'O', 'Z', 'N', 'U')

**;**

*N : Indica que o valor é numérico.*

**update** airports

**set** qa\_dst = 'N'

**Where** dst **NOT** **IN**('E', 'A', 'S', 'O', 'Z', 'N', 'U')

**;**

**select** \* **from** airports**;**