

## Questão 12

### 12.1.1

Para gerar a tabela com as medidas dos quatro analitos, utilizamos a consulta do exercício 06 como *select*, e mantivemos inicialmente cinco analitos, para facilitar na remoção de tuplas nulas.

### 12.1.2

Para a remoção de nulos, fizemos cinco *updates* baseados nas definições abaixo.

$$colesterol\_total = ldl\_colesterol + hdl\_colesterol + vldl\_colesterol$$

$$colesterol\_nao\_hdl = colesterol\_total - hdl\_colesterol$$

- Calculando o 'colesterol nao-hdl, soro' a partir dos analitos 'colesterol total' e 'hdl colesterol'.
- Calculando o 'vldl colesterol' a partir dos analitos 'colesterol nao-hdl, soro' e 'ldl colesterol'.
- Calculando o 'colesterol nao-hdl, soro' a partir dos analitos 'vldl colesterol' e 'ldl colesterol'.
- Calculando o 'hdl colesterol' a partir dos analitos 'colesterol total' e 'colesterol nao-hdl, soro'.
- Calculando o 'ldl colesterol' a partir dos analitos 'colesterol nao-hdl, soro' e 'vldl colesterol'.

Por fim, realizamos uma limpeza para remover algumas tuplas que estão com dados incoerentes, ou seja, não atendem as igualdades acima, e removemos a coluna 'colesterol nao-hdl, soro', para ficarmos com somente 4 analitos.

Antes desses *updates* a tabela continha cerca de 2600 tuplas com pelo menos um analito nulo, e após eles, cerca de 90 tuplas continham algum valor nulo.

### 12.2

A tabela possui quatro atributos, sendo que um deles, o 'colesterol total', é a soma dos outros três atributos, portanto a tabela só possui três graus de liberdade. Portanto a máxima dimensão fractal possível é 3.

### 12.3

Utilizando como base a *query* do material de aula, e definimos como função de distância uma distância euclidiana considerando cada tupla como um ponto, e cada analito da tupla como uma dimensão.

Rodamos a *query* e obtivemos como resultado uma dimensão de correlação de 1.

	123 slope	123 intercept
1	1.0031478327	-4.0903425561

Figura 47 – Resultado da Questão 12.