

INF010

RELATÓRIO DE OTIMIZAÇÃO DE CONSULTAS
INSTITUTO FEDERAL DE CIÊNCIA E TECNOLOGIA DA BAHIA

João Manuel dos Santos Lapa
Tecnologia em análise e desenvolvimento de sistemas

Salvador, Bahia
28 de abril de 2021

Municípios que não pertencem à região Norte:

```
select
    m.*
from
    municipio as m
join estado as e on
    e.codestado = m.codestado
where
    e.codregiao <> 1;
```

A consulta executada pelo SGBD Postgres primeiro faz uma seleção sobre a tabela estado, filtrando resultados conforme a condição especificada na cláusula “where”, aplicando em seguida uma operação de junção com a tabela município, respeitando o predicado especificado na cláusula “on” da junção. Neste cenário, nenhuma sinalização direta ao SGBD que pudesse melhorar a performance foi identificada.

Municípios homônimos:

```
select
    m.*
from
    municipio as m
join (
    select
        m.nomemunicipio,
        count(*) as quantidade
    from
        municipio as m
    group by
        m.nomemunicipio
    having
        count(*) > 1
) as tb on
    tb.nomemunicipio = m.nomemunicipio
order by
    m.nomemunicipio
```

A consulta executada pelo SGBD Postgres primeiro faz uma agregação de contagem sobre a tabela município através do nome, selecionando aqueles caso a quantidade seja superior a 1. Após, realiza um join sobre a tabela de municípios e o resultado da agregação, utilizando a igualdade entre nomemunicipio como predicado, seguindo por último por uma operação de ordenação. Neste cenário, nenhuma sinalização direta ao SGBD que pudesse melhorar a performance foi identificada.

Média de municípios por estado e por região:

```
select
    sum(tb.media_regiao) as media_municipios_por_regiao,
    sum(tb.media_estado) as media_municipios_por_estado
from (
    (select
        avg(tb.count)::decimal(8,2) as media_regiao,
        0::decimal(8,2) as media_estado
    from (
        select
            e.codregiao,
            count(m.*)
        from
            estado as e
        join municipio as m on
            e.codestado = m.codestado
        group by
            e.codregiao
    ) as tb)
    union all
    (select
        0::decimal(8,2) as media_regiao,
        avg(tb.count)::decimal(8,2) as media_estado
    from (
        select
            e.codestado,
            count(m.*)
        from
            estado as e
        join municipio as m on
            e.codestado = m.codestado
        group by
            e.codestado
    ) as tb)
) as tb
```

A consulta executada pelo SGBD Postgres tanto no ramo que calcula a média por estado, quanto na que calcula a média por região, faz primeiro uma junção entre as tabelas município e estado, aplicando em seguida uma agregação, sobre o resultado da junção, de forma a obter o total de municípios por estado. A seguir, uma nova agregação é feita, sendo que em um dos ramos é obtida a média da contagem d municípios sobre o código da região, e no outro ramo é obtida a média da contagem sobre o código do estado. Em ambos os casos, são projetados os resultados obtidos, bem como o valor zero para as médias calculadas pelo ramo oposto. Por fim, é executada uma união entre os 2 conjuntos de dados e posterior soma dos campos comuns, de forma a obter em um único conjunto as médias estaduais e regionais.

Sigla dos estados com as respectivas quantidades de municípios:

```
select
    e.siglaestado,
    tb.qtd_municipios
from
    estado as e
join (
    select
        m.codestado,
        count(m.*) as qtd_municipios
    from
        municipio as m
    group by
        m.codestado
) as tb on
    tb.codestado = e.codestado;
```

A consulta executada pelo SGBD Postgres primeiro realiza uma agregação de contagem na tabela municipio através da coluna codestado. Em seguida, é realizado um join entre a tabela estado e a “tabela temporária” resultante da primeira query, de forma a projetar juntamente com a quantidade de municípios por estado, a sigla de cada estado. Não foram observadas possibilidades de melhoria.

Município com as pessoas mais idosas:

```
select
    i.codmunicipio
from
    indice as i
order by
    i.idh_longevidade desc
limit 1
```

A consulta executada pelo SGBD Postgres primeiro realiza uma ordenação decrescente na tabela indice através da coluna idh_longevidade, em seguida a projeção da coluna codmunicipio e logo após uma seleção do primeiro registro. Não foram observadas possibilidades de melhoria.

Ano em que Salvador obteve o melhor índice de Instrução:

```
select
    i.ano
from
    indice as i
where
    i.codmunicipio = (
        select
            m.codmunicipio
        from
            municipio as m
        where
            m.nomemunicipio like 'Salvador%'
            and m.codestado = 5
    )
order by
    i.idh_educacao desc
limit 1;
```

A consulta executada pelo SGBD Postgres primeiro realiza uma seleção sobre a tabela municipio com predicados de igualdade sobre as colunas nomemunicipio e codestado. Em seguida é feita nova seleção, desta vez sobre a tabela indice, com predicado de igualdade entre a colunas codmunicipio e o resultado da primeira query. Sobre o resultado da seleção é feita uma ordenação decrescente através da coluna idh_educacao, uma projeção da coluna i e por último, seleção do primeiro registro resultante da ordenação. Não foram observadas possibilidades de melhoria.

Qual o município com a melhor distribuição de renda:

```
select
    i.codmunicipio
from
    indice as i
order by
    i.idh_renda desc
limit 1;
```

A consulta executada pelo SGBD Postgres primeiro realiza uma ordenação decrescente na tabela indice através da coluna idh_renda em seguida a projeção da coluna codmunicipio e logo após uma seleção do primeiro registro. Não foram observadas possibilidades de melhoria.

Quais estados possuem municípios com IDH geral maior que 0,8:

```
select distinct
    m.codestado
from
    municipio as m
where
    m.codmunicipio in (
        select
            i.codmunicipio
        from
            indice as i
        where
            i.idh_geral > 0.8
    )
```

A consulta executada pelo SGBD Postgres primeiro seleciona todos os registros da tabela indice os quais o valor da coluna idh_geral seja maior que 0,8. Em seguida faz uma seleção sobre a tabela município onde os valores da coluna codmunicipio estejam contidos no resultado da primeira consulta (semi join). Por fim, é feita projeção da coluna codestado e agrupamento dos valores, de forma obter apenas valores distintos. Não foram observadas possibilidades de melhoria.

Qual o maior IDH de educação por estado:

```
select
    m.codestado,
    max(i.idh_educacao) max_idh_educacao
from
    indice as i
join municipio as m on
    m.codmunicipio = i.codmunicipio
group by
    m.codestado
```

A consulta executada pelo SGBD Postgres primeiro realiza um join entre as tabelas município e indice. Em seguida, realiza uma agregação de valor máximo do campo idh_educacao da tabela indice a partir do campo codestado da tabela município. Por fim, é feita projeção dos campos codestado e max_idh_educacao. Não foram observadas possibilidades de melhoria.

Relatório de Todos IDHs da Bahia de 91 e 2000, inclusive com a diferença entre os mesmos:

```
select
    tb2.*,
    idh_geral_2000 - idh_geral_1991 as idh_geral_variacao,
    idh_educacao_2000 - idh_educacao_1991 as idh_educacao_variacao,
    idh_longevidade_2000 - idh_longevidade_1991 as idh_longevidade_variacao,
    idh_renda_2000 - idh_renda_1991 as idh_renda_variacao
from (
    select
        sum(idh_geral_1991) as idh_geral_1991,
        sum(idh_educacao_1991) as idh_educacao_1991,
        sum(idh_longevidade_1991) as idh_longevidade_1991,
        sum(idh_renda_1991) as idh_renda_1991,
        sum(idh_geral_2000) as idh_geral_2000,
        sum(idh_educacao_2000) as idh_educacao_2000,
        sum(idh_longevidade_2000) as idh_longevidade_2000,
        sum(idh_renda_2000) as idh_renda_2000
    from (
        (select
            avg(i.idh_geral)::decimal(8,3) as idh_geral_1991,
            avg(i.idh_educacao)::decimal(8,3) as idh_educacao_1991,
            avg(i.idh_longevidade)::decimal(8,3) as idh_longevidade_1991,
            avg(i.idh_renda)::decimal(8,3) as idh_renda_1991,
            0 as idh_geral_2000,
            0 as idh_educacao_2000,
            0 as idh_longevidade_2000,
            0 as idh_renda_2000
        from
            indice as i
        join municipio as m on
            m.codmunicipio = i.codmunicipio
        where
            m.codestado = 5
            and ano = '1991')
        union
        (select
            0 as idh_geral_1991,
            0 as idh_educacao_1991,
            0 as idh_longevidade_1991,
            0 as idh_renda_1991,
            avg(i.idh_geral)::decimal(8,3) as idh_geral_2000,
            avg(i.idh_educacao)::decimal(8,3) as idh_educacao_2000,
            avg(i.idh_longevidade)::decimal(8,3) as idh_longevidade_2000,
            avg(i.idh_renda)::decimal(8,3) as idh_renda_2000
        from
```

```

        indice as i
    join municipio as m on
        m.codmunicipio = i.codmunicipio
    where
        m.codestado = 5
        and ano = '2000')
) as tb1
) as tb2

```

A consulta executada pelo SGBD Postgres inicia a partir da união de 2 subconsultas. Estas subconsultas são as consultas dos idhs médios do estado da Bahia em 2 períodos. Em cada subconsulta, é efetuado um join entre as tabelas indice e município, seguido de agregação de média de todos os idhs dos anos 1991 e 2000 respectivamente. Após a união, o sgbd executa, sobre o conjunto de dados resultante, nova agregação, desta vez de soma de todas as colunas. Por fim, o sgbd executa a projeção dos valores deste último conjunto de dados resultante juntamente com os campos de variação do idh, resultantes de uma subtração dos idhs do ano 1991 dos ids do ano 2000. Nenhuma melhoria foi observada em relação à query montada.

Relatório comparativo entre as médias dos IDHs de SC e AL, de 2000 e 91:

```

select
    tb2.*,
    abs(idh_geral_sc - idh_geral_al) as idh_geral_variacao,
    abs(idh_educacao_sc - idh_educacao_al) as idh_educacao_variacao,
    abs(idh_longevidade_sc - idh_longevidade_al) as idh_longevidade_variacao,
    abs(idh_renda_sc - idh_renda_al) as idh_renda_variacao
from (
    select
        ano,
        sum(idh_geral_al) as idh_geral_al,
        sum(idh_educacao_al) as idh_educacao_al,
        sum(idh_longevidade_al) as idh_longevidade_al,
        sum(idh_renda_al) as idh_renda_al,
        sum(idh_geral_sc) as idh_geral_sc,
        sum(idh_educacao_sc) as idh_educacao_sc,
        sum(idh_longevidade_sc) as idh_longevidade_sc,
        sum(idh_renda_sc) as idh_renda_sc
    from (
        (select
            i.ano,
            avg(i.idh_geral)::decimal(8,3) as idh_geral_al,
            avg(i.idh_educacao)::decimal(8,3) as idh_educacao_al,
            avg(i.idh_longevidade)::decimal(8,3) as idh_longevidade_al,
            avg(i.idh_renda)::decimal(8,3) as idh_renda_al,
            0 as idh_geral_sc,

```



```

        0 as idh_educacao_sc,
        0 as idh_longevidade_sc,
        0 as idh_renda_sc
    from
        indice as i
    join municipio as m on
        m.codmunicipio = i.codmunicipio
    where
        m.codestado = 2
    group by
        i.ano)
union all
(select
    i.ano,
    0 as idh_geral_al,
    0 as idh_educacao_al,
    0 as idh_longevidade_al,
    0 as idh_renda_al,
    avg(i.idh_geral)::decimal(8,3) as idh_geral_sc,
    avg(i.idh_educacao)::decimal(8,3) as idh_educacao_sc,
    avg(i.idh_longevidade)::decimal(8,3) as idh_longevidade_sc,
    avg(i.idh_renda)::decimal(8,3) as idh_renda_sc
    from
        indice as i
    join municipio as m on
        m.codmunicipio = i.codmunicipio
    where
        m.codestado = 24
    group by
        i.ano)
) as tb1
group by
    ano
) as tb2

```

A consulta executada pelo SGBD Postgres inicia a partir da união de 2 subconsultas. Estas subconsultas são as consultas dos idhs médios do estados de Alagoas e Santa Catarina agrupados por ano (1991 e 2000). Em cada subconsulta é feita uma filtragem de municípios dos estados AL e SC, respectivamente, depois é efetuado um join entre as tabelas indice e município, seguido de agregação de média de todos os idhs agrupados por ano. Após a união, o sgbd executa, sobre o conjunto de dados resultante, nova agregação, desta vez de soma de todas as colunas. Por fim, o sgbd executa a projeção dos valores deste último conjunto de dados resultante juntamente com os campos de variação do idh, resultantes de uma subtração dos idhs de Al dos ids de SC. Nenhuma melhoria foi observada em relação à query montada.