

ESTATÍSTICAS PARA ESTIMATIVA DE CUSTO

Sérgio Mergen

Estatísticas para estimativa de custo

- Através de estatísticas, é possível obter estimativas a respeito de quantos registros são retornados por cada operação de um plano de execução
- Isso pode ajudar o otimizador a tomar decisões, como
 - Quais seleções fazer primeiro
 - A escolha da primeira seleção é importante, pois ela pode se valer de um índice
 - Quais junções fazer primeiro
 - Importante fazer primeiro junções que tragam menos registros
 - Quais algoritmos usar
 - O custo de cada algoritmo depende do número de registros. Sabendo esse número, dá para saber qual algoritmo tende a se sair melhor

Estatísticas para estimativa de custo

- Estatísticas usadas
 - $n(r)$: número de tuplas da relação r .
 - $v(A, r)$: cardinalidade do atributo A
 - número de valores distintos que aparecem em r para o atributo A
 - $\min(A, r)$ = menor valor em r para o atributo A
 - $\max(A, r)$ = maior valor em r para o atributo A

Estatísticas de exemplo

- Número de registros

n(depto)	=	50
n(func)	=	10.000
n(proj)	=	2.000

- Cardinalidades

v(idDepto, depto)	=	50
v(idFunc, func)	=	10.000
v(idDepto, func)	=	40
v(idProj, proj)	=	2.000
v(idLider, proj)	=	250
v(custo, proj)	=	50

- Máximos e mínimos

min(custo, proj)	=	50.000
max(custo,proj)	=	550.000

depto (idDepto, nome)

func (idFunc, nome, idDepto, salario)
idDepto referencia depto

proj (idProj, nome, custo, idLider)
idLider referencia func

ESTIMATIVAS PARA SELEÇÕES

Seletividade (probabilidade)

- Cada filtro em uma consulta é considerado como um evento independente

- Ex.

```
SELECT * FROM PROJETO  
WHERE custo > 3 AND duração = 4
```

- Um registro ter **custo > 3** forma um evento A
- Um registro ter **duração = 4** forma outro evento B

Seletividade (probabilidade)

- A seletividade de um filtro A está relacionada à probabilidade $P(A)$ de que o evento associado ao filtro de fato aconteça
- Considere que
 - a tabela **proj** tenha 2.000 registros
 - E existe 10% de chance de que um registro tenha custo > 3
 - $P(A) = 0.1$ (seletividade de custo > 3)
 - Ao aplicar esse filtro, a estimativa de retorno seria de 10% dos 2.000 registros
 - $2.000 * 0.1 = 200$ registros
- Diz-se que um valor mais alto de probabilidade significa que um filtro é pouco seletivo

Estimativa com filtros simples

- **WHERE tab.A = valor**
 - seletividade do filtro = $1 / v(A, \text{tab})$
- Exemplo:

```
SELECT *  
FROM proj  
WHERE custo = 100.000
```

n (proj)	= 2.000
v (custo, proj)	= 50

$$P(A) = \frac{1}{50} \quad [\text{custo} = 100.000]$$

$$\text{Estimativa} = 2.000 * \frac{1}{50} = 40$$

Estimativa com filtros simples

- **WHERE tab.A = valor**
 - Na falta de informações estatísticas, deve-se tentar adivinhar um valor de seletividade.
 - Ex. Seletividade = $\frac{1}{2}$.
- Exemplo:

```
SELECT *  
FROM proj  
WHERE ano = 2014
```

n (proj)	= 2.000
v (ano, proj)	= ???

$$P(A) = \frac{1}{2} \quad [\text{ano} = 2014]$$

$$\text{Estimativa} = 2.000 * \frac{1}{2} = 1000$$

Estimativa com filtros simples

- **WHERE tab.A <= v**
- Se $\min(A, \text{tab})$ e $\max(A, \text{tab})$ estiverem presentes no catálogo de estatísticas
 - se $v < \min(A, \text{tab})$
seletividade = 0
 - senão
$$\text{seletividade} = \frac{v - \min(A, \text{tab}) + 1}{\max(A, \text{tab}) - \min(A, \text{tab}) + 1}$$
- Na falta de informações estatísticas
 - seletividade = $\frac{1}{2}$.

Estimativa com filtros simples

- Exemplo:

```
SELECT *  
FROM proj  
WHERE custo <= 10.000
```

$P(A) = 0$ [custo <= 10.000]
estimativa = $2.000 * 0 = 0$

n (proj)	=	2.000
min(custo, proj)	=	50.000
max(custo,proj)	=	550.000

Estimativa com filtros simples

- Exemplo:

n (proj)	=	2.000
min(custo, proj)	=	50.000
max(custo,proj)	=	550.000

SELECT *
FROM proj
WHERE custo <= 100.000

$$P(A) = \frac{100k - 50k + 1}{550k - 50k + 1} \quad [\text{custo} \leq 100.000]$$

$$P(A) \approx 0,1$$

$$\text{estimativa} \approx 2.000 * 0,1 \approx 200$$

Estimativa com filtros simples

- Exemplo:

n (proj)	=	2.000
min(ano, proj)	=	?
max(ano,proj)	=	?

```
SELECT *  
FROM proj  
WHERE ano < 2014
```

$$P(A) = \frac{1}{2}$$

$$\text{estimativa} = 2.000 * \frac{1}{2} = 1000$$

Estimativa com negação

- A probabilidade de um evento A não aconteça (\bar{A}) é equivalente ao complemento da probabilidade de que ele aconteça
 - $P(\bar{A}) = 1 - P(A)$

Estimativa com negação

- Exemplo: Quantas tuplas satisfazem $\text{codProj} \neq 40$?

$$P(A) = \frac{1}{2000} \quad [\text{codProj} = 40]$$

$$P(\bar{A}) = 1 - \frac{1}{2000} = \frac{1999}{2000} \quad [\text{codProj} \neq 40]$$

$$\textbf{Estimativa} = 2.000 * \frac{1999}{2000} = 1999$$

- Exemplo: Quantas tuplas satisfazem $\text{custo} > 100.000$?

$$P(A) = \frac{200}{2000} \quad [\text{custo} \leq 100.000]$$

$$P(\bar{A}) = 1 - \frac{200}{2000} = \frac{1.800}{2000} \quad [\text{custo} > 100.000]$$

$$\textbf{Estimativa} = 2.000 * \frac{1800}{2000} = 1800$$

Estimativa com filtros compostos

- **Conjunção de dois filtros:** $\sigma_{\theta 1} \wedge \sigma_{\theta 2}$
- Os registros que satisfazem múltiplos filtros conjuntivos podem ser vistos como a interseção (\cap) de eventos independentes
- Pela teoria da probabilidade, a probabilidade de intersecção de dois eventos independentes A e B é dada pela seguinte fórmula

$$P(A \cap B) = P(A) \cdot P(B)$$

Estimativa com filtros compostos

SELECT *

FROM proj

WHERE custo <= 100.000 **AND** ano = 2014

$P(A) = 0,1$ [custo <= 100.000]

$P(B) = 0,5$ [ano = 2014]

$P(A \cap B) = 0,1 * 0,5 = 0,05$

Estimativa = $2.000 * 0,05 = 100$

Estimativa com filtros compostos

- **Conjunção de múltiplos filtros:** $\sigma_{\theta 1} \wedge \sigma_{\theta 2} \dots \wedge \sigma_{\theta n}$
- A conjunção de múltiplos eventos (mais do que dois) é dada pelo produto de todas as probabilidades individuais

$$P(A \cap B \cap C \cap D) = P(A) \cdot P(B) \cdot P(C) \cdot P(D)$$

Estimativa com filtros compostos

- **Disjunção de dois filtros** $\sigma_{\theta 1} \vee \sigma_{\theta 2}$
- Os registros que satisfazem pelo menos um dos filtros podem ser vistos como a união (U) de eventos independentes
- Pela teoria da probabilidade, a probabilidade da união de dois eventos independentes A e B é dada por uma fórmula que soma probabilidades individuais e descarta as intersecções

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Estimativa com filtros compostos

SELECT *

FROM proj

WHERE custo <= 100.000 **OR** ano = 2014

$$P(A) = 0,1 \quad [\text{custo} \leq 100.000]$$

$$P(B) = 0,5 \quad [\text{ano} = 2014]$$

$$P(A \cap B) = 0,05$$

$$P(A \cup B) = 0,1 + 0,5 - 0,05 = 0,55$$

$$\text{Estimativa} \rightarrow 2.000 * 0,55 \rightarrow 1.100$$

Estimativa com filtros compostos

- **Disjunção de dois filtros** $\sigma_{\theta 1} \vee \sigma_{\theta 2}$
- Fórmula geral
 - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Em alguns casos, os filtros são mutuamente exclusivos
 - Ex. custo = 2000 OR custo = 3000
- Nesses casos, $P(A \cap B)$ é zero. Isso faz com que fórmula seja simplificada
 - $P(A \cup B) = P(A) + P(B)$

Estimativa com filtros compostos

SELECT *

FROM proj

WHERE custo = 100.000 **OR** custo = 200.000

$$P(A) = \frac{1}{40} \quad [\text{custo} = 100.000]$$

$$P(B) = \frac{1}{40} \quad [\text{custo} = 200.000]$$

$$P(A \cup B) = \frac{1}{40} + \frac{1}{40} = \frac{1}{20}$$

$$\text{Estimativa} = 2.000 * \frac{1}{20} = 100$$

Estimativa com filtros compostos

- **Disjunção de múltiplos filtros** $\sigma_{\theta 1} \vee \sigma_{\theta 2} \dots \vee \sigma_{\theta n}$
- A disjunção de múltiplos eventos (mais do que dois) é dada por uma fórmula que soma ou descarta interseções
- Exemplos para três eventos
- $$P(A \cup B \cup C) = P(A) + P(B) + P(C) \\ - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ + P(A \cap B \cap C)$$

A não ser que os eventos sejam mutuamente exclusivos, a fórmula pode se tornar complexa para um número grande de eventos

Estimativa com filtros compostos

- **Disjunção de múltiplos filtros** $\sigma_{\theta 1} \vee \sigma_{\theta 2} \dots \vee \sigma_{\theta n}$
- Outra forma de encontrar a união de eventos independentes é realizando a seguinte transformação de DeMorgan
 - $(A \vee B \vee C) = \neg (\neg A \wedge \neg B \wedge \neg C)$
- Sendo assim, nós temos
 - $P(A \cup B \cup C) = \overline{P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C})}$
 - $= 1 - P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C})$

Estimativa com filtros compostos

SELECT *
FROM proj

WHERE custo <= 100.000 **OR** ano = 2014

o mesmo que

WHERE NOT (custo > 100.000 **AND** ano <> 2014)

$$P(A) = 0,1 \quad [\text{custo} \leq 100.000]$$

$$P(\bar{A}) = 1 - 0,1 = 0,9 \quad [\text{custo} > 100.000]$$

$$P(B) = 0,5 \quad [\text{ano} = 2014]$$

$$P(\bar{B}) = 1 - 0,5 = 0,5 \quad [\text{ano} \neq 2014]$$

$$P(\bar{A}) \cdot P(\bar{B}) = 0,9 * 0,5 = 0,45$$

$$P(\overline{0,45}) = 1 - 0,45 = 0,55$$

$$\text{Estimativa} = 2.000 * 0,55 \rightarrow 1.100$$

ESTIMATIVAS PARA JUNÇÕES

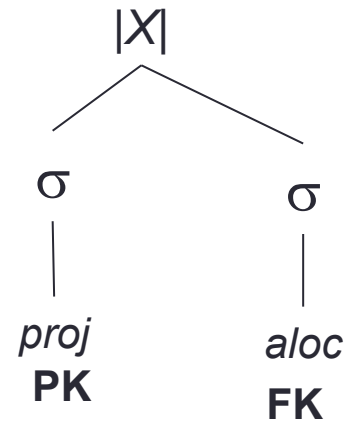
Estimativas do MySQL – Junção

- Em um plano de execução de consulta, praticamente todas operações binárias são tipos de junção
 - Assim, as estimativas referentes à junções são muito importantes
- Duas possibilidades principais
 - As colunas da junção têm relação de chave primária/estrangeira
 - Não existe critério de junção

Estimativas do MySQL – Junção

Quando as colunas da junção têm relação de chave primária/estrangeira

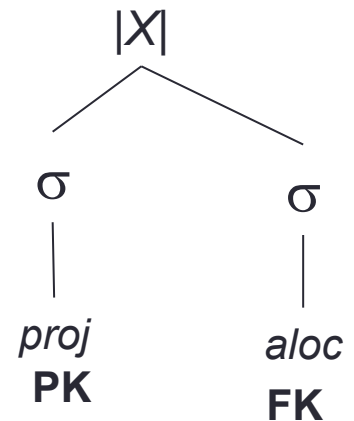
- Considerando que
 - PK: tabela onde está a chave primária
 - FK: tabela onde está a chave estrangeira:
 - $n(PK)$: número de registros que chegam da tabela onde está a PK
 - $n(FK)$: número de registros que chegam da tabela onde está a FK



Estimativas do MySQL – Junção

Quando as colunas da junção têm relação de chave primária/estrangeira

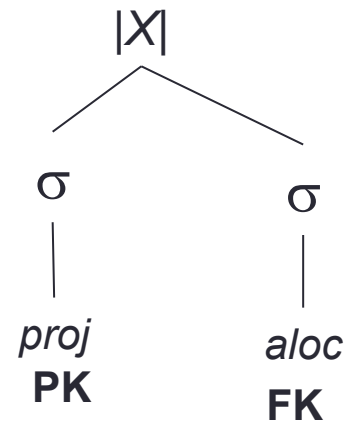
- Considerando que
 - $\text{sel}(\text{PK})$: seletividade obtida do lado PK
 - $\text{sel}(\text{FK})$: seletividade obtida do lado FK



Estimativas do MySQL – Junção

Quando as colunas da junção têm relação de chave primária/estrangeira

- O cálculo é
 - $n(\text{FK}) * \text{sel}(\text{PK})$

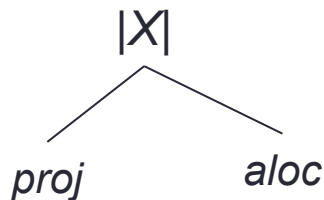


O máximo de registros gerados é equivalente ao tamanho da tabela que contém a FK

Estimativas do MySQL – Junção

- Exemplo
 - Junção por chave primária/chave estrangeira **sem filtros**

```
SELECT *  
FROM proj NATURAL JOIN aloc
```



```
aloc (idFunc, idProj, comissao)  
      idproj referencia proj  
      idfunc referencia pfunc
```

```
proj (idProj, nomeP, custo)
```

```
func(idFunc, nome, salario)
```

Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5

Estimativas do MySQL – Junção

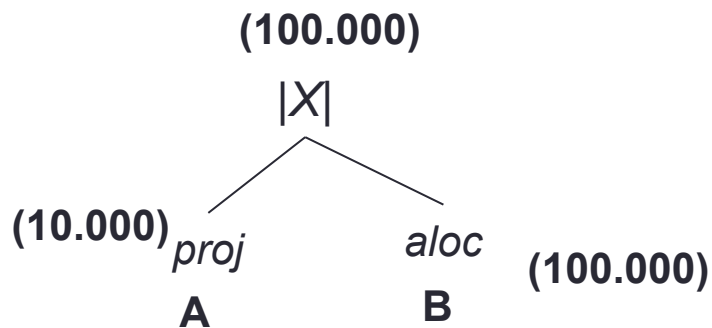
- Cálculo da junção

- = $n(\text{FK}) * \text{sel}(\text{PK})$
- = $100.000 * 1$
- = **100.000**

```
aloc (idFunc, idProj, comissao)
      idproj referencia proj
      idfunc referencia pfunc
```

```
proj (idProj, nomeP, custo)
```

```
func(idFunc, nome, salario)
```



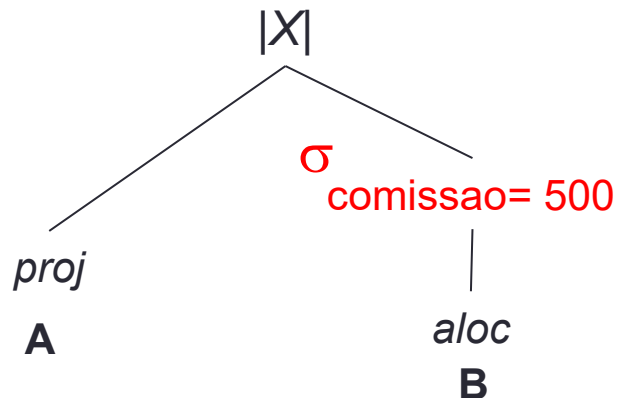
Estatísticas

$n(\text{func})$	=	500
$n(\text{proj})$	=	10.000
$n(\text{aloc})$	=	100.000
$v(\text{salario}, \text{func})$	=	50
$v(\text{custo}, \text{proj})$	=	10
$v(\text{comissao}, \text{aloc})$	=	5

Estimativas do MySQL – Junção

- Exemplo
 - Junção por chave primária/chave estrangeira **com filtro na tabela onde está a chave estrangeira**

```
SELECT *  
FROM proj NATURAL JOIN aloc  
WHERE comissao= 500
```



aloc (idFunc, idProj, **comissao**)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, custo)

func(idFunc, nome, salario)

Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5

Estimativas do MySQL – Junção

- Cálculo da junção

- = $n(\text{FK}) * \text{sel}(\text{PK})$
- = $20.000 * 1$
- = **20.000**

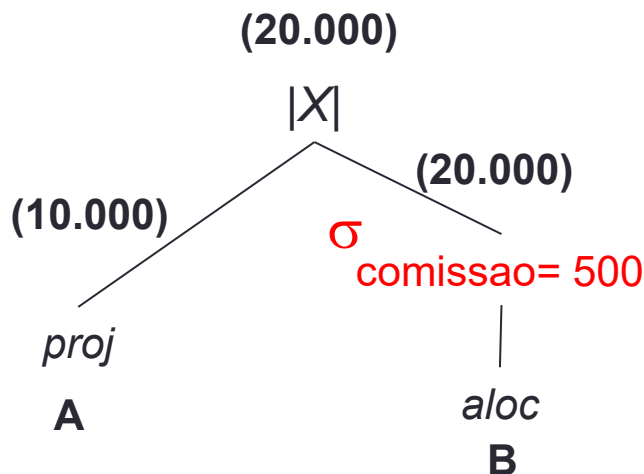
aloc (idFunc, idProj, **comissao**)

idproj referencia proj

idfunc referencia pfunc

proj (idProj, nomeP, custo)

func(idFunc, nome, salario)



Estatísticas

$n(\text{func})$	=	500
$n(\text{proj})$	=	10.000
$n(\text{aloc})$	=	100.000
$v(\text{salario}, \text{func})$	=	50
$v(\text{custo}, \text{proj})$	=	10
$v(\text{comissao}, \text{aloc})$	=	5

Estimativas do MySQL – Junção

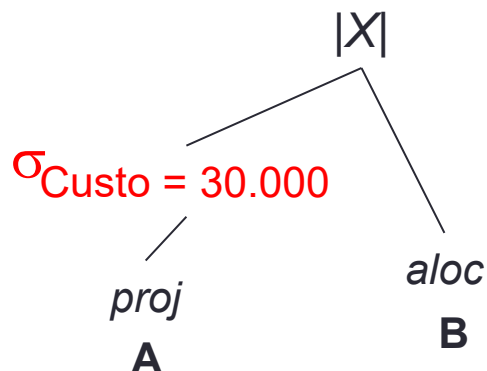
- Exemplo
 - Junção por chave primária/chave estrangeira **com filtro na tabela onde está a chave primária**

```
SELECT *  
FROM proj NATURAL JOIN aloc  
WHERE custo= 30.000
```

aloc (idFunc, idProj, comissao)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, **custo**)

func(idFunc, nome, salario)



Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5

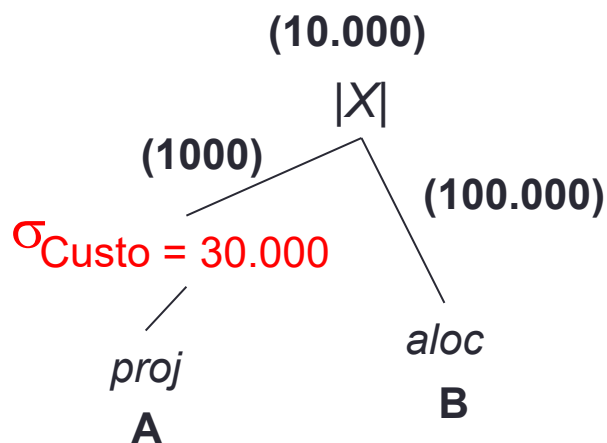
Estimativas do MySQL – Junção

- Cálculo da junção
 - = $n(\text{FK}) * \text{sel}(\text{PK})$
 - = $100.000 * 1/10$
 - = **10.000**

aloc (idFunc, idProj, comissao)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, **custo**)

func(idFunc, nome, salario)



Estatísticas

$n(\text{func})$	=	500
$n(\text{proj})$	=	10.000
$n(\text{aloc})$	=	100.000
$v(\text{salario}, \text{func})$	=	50
$v(\text{custo}, \text{proj})$	=	10
$v(\text{comissao}, \text{aloc})$	=	5

Estimativas do MySQL – Junção

- Exemplo
 - Junção por chave primária/chave estrangeira **com filtro nas duas tabelas**

```
SELECT *  
FROM proj NATURAL JOIN aloc  
WHERE custo= 30.000 and  
comissão = 500
```

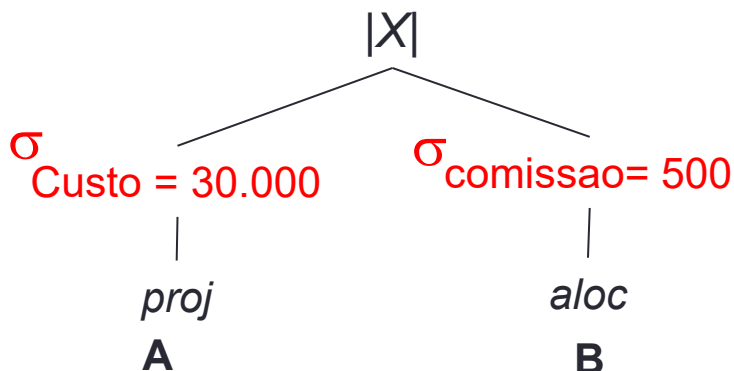
aloc (idFunc, idProj, **comissao**)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, **custo**)

func(idFunc, nome, salario)

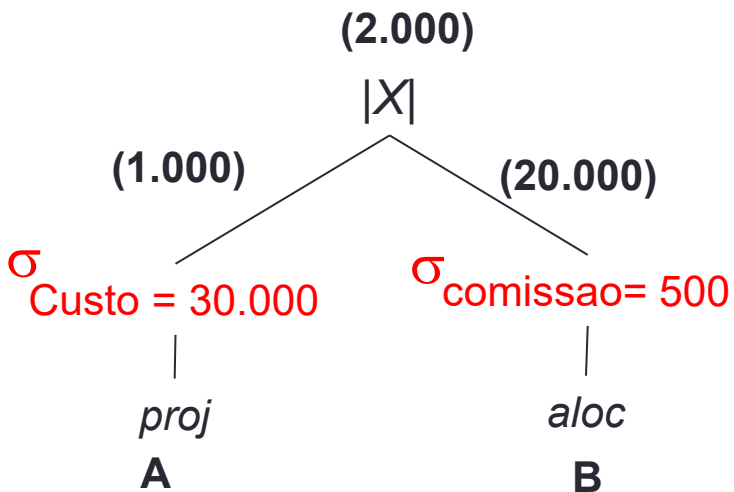
Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5



Estimativas do MySQL – Junção

- Cálculo da junção
 - $= n(\text{FK}) * \text{sel}(\text{PK})$
 - $= 20.000 * 1/10$
 - **= 2.000**



Estatísticas

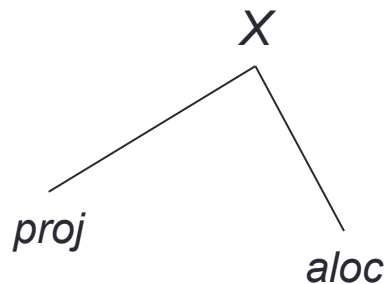
n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloca)	=	5

Estimativas do MySQL – Junção

- Em um plano de execução de consulta, praticamente todas operações binárias são tipos de junção
 - Assim, as estimativas referentes à junções são muito importantes
- Duas possibilidades principais
 - As colunas da junção têm relação de chave primária/estrangeira
 - Não existe critério de junção

Estimativas do MySQL – Junção

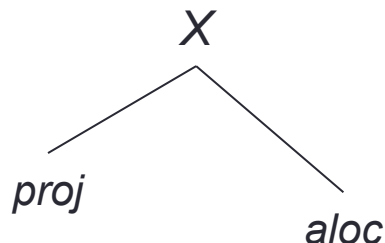
- Nesse caso, cada registro de um lado cruza com todos os registros do outro lado
- A estimativa é simplesmente o produto da quantidade de registros que chega em cada lado da junção
 - $n(\text{tab1}) * n(\text{tab2})$



Estimativas do MySQL – Junção

- Exemplo
 - Cruzamento sem nenhum critério de junção

```
SELECT *  
FROM proj CROSS JOIN aloc
```



```
aloc (idFunc, idProj, comissao)  
      idproj referencia proj  
      idfunc referencia pfunc
```

```
proj (idProj, nomeP, custo)
```

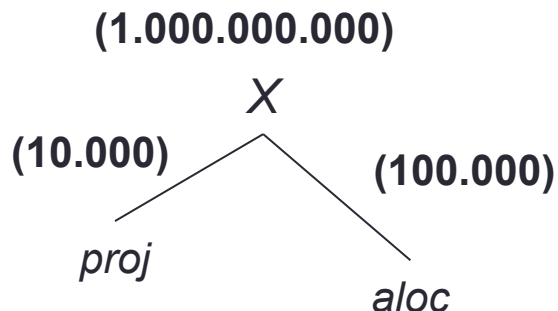
```
func(idFunc, nome, salario)
```

Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5

Estimativas do MySQL – Junção

- Cálculo da junção
 - $= n(\text{proj}) * n(\text{aloc})$
 - $= 10.000 * 100.000$
 - **$= 1.000.000.000$**



aloc (idFunc, idProj, comissao)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, custo)

func(idFunc, nome, salario)

Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloca)	=	5

Estimativa com composição de junções

- Vimos como estimar o resultado de uma junção
 - Mas uma expressão geralmente possui várias junções
- Como estimar o resultado de uma expressão complexa?
 - Deve-se estimar o resultado das operações de nível mais baixo primeiro
 - O resultado é então usado para estimar o resultado das operações de nível mais alto

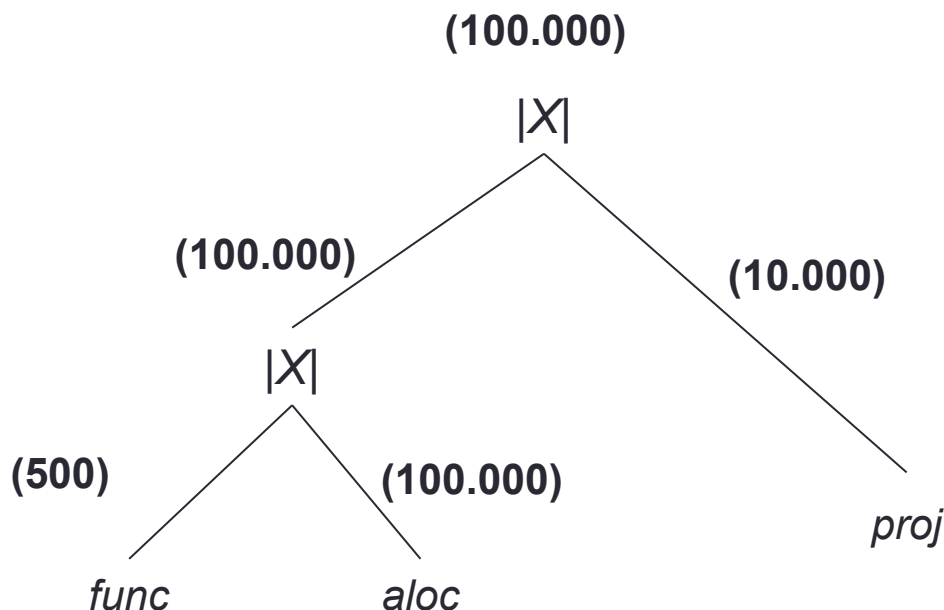
Estimativa com composição de junções

- Ex.

aloc (idFunc, idProj, comissao)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, custo)

func(idFunc, nome, salario)



Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5

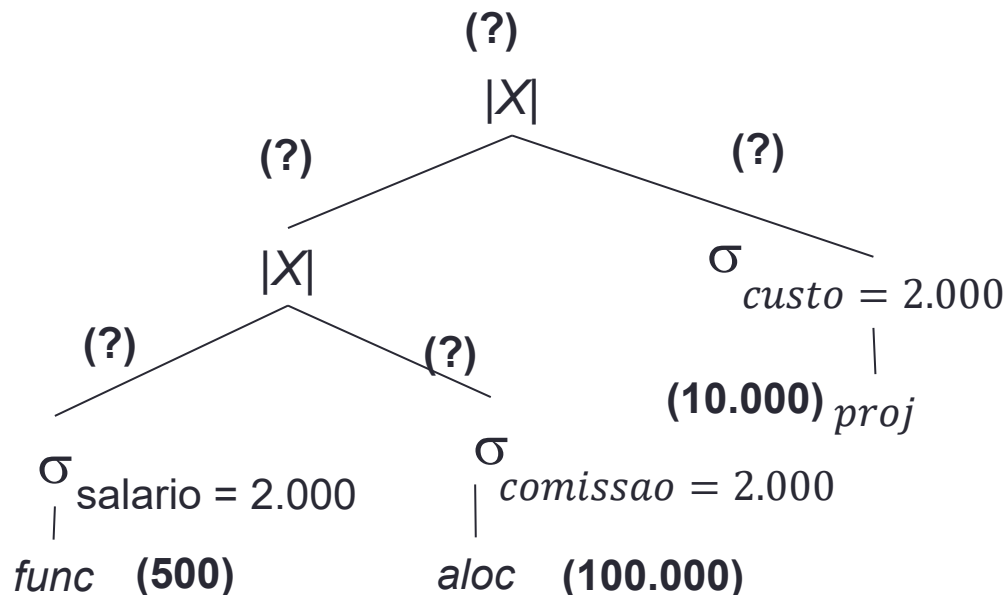
Estimativa com composição de junções

- Como estimar nesse caso?

aloc (idFunc, idProj, comissao)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, custo)

func(idFunc, nome, salario)

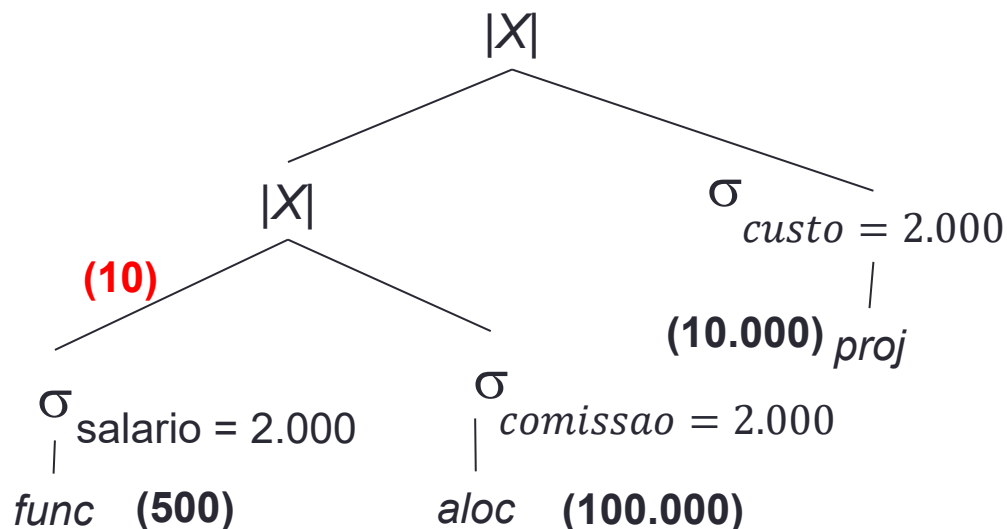


Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5

Estimativa com composição de junções

- Estima-se que 10 funcionários ganhem 2.000
 - = $n(\text{func}) / v(\text{salario}, \text{func})$
 - = $500 / 50$
 - = 10



`aloc (idFunc, idProj, comissao)`
idproj referencia proj
idfunc referencia pfunc

`proj (idProj, nomeP, custo)`

`func(idFunc, nome, salario)`

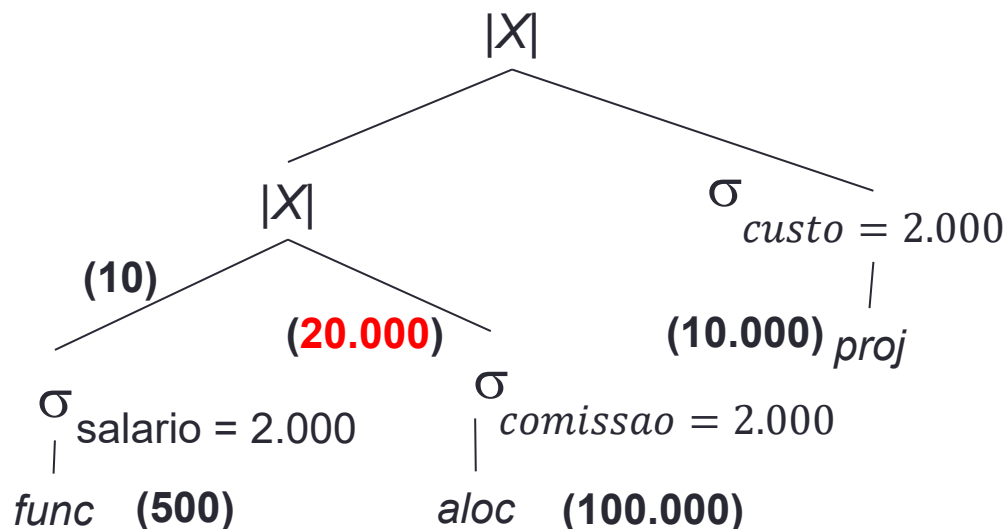
Estatísticas

$n(\text{func})$	=	500
$n(\text{proj})$	=	10.000
$n(\text{aloc})$	=	100.000
$v(\text{salario}, \text{func})$	=	50
$v(\text{custo}, \text{proj})$	=	10
$v(\text{comissao}, \text{aloc})$	=	5

Estimativa com composição de junções

- Estima-se que 20.000 alocações tenham comissão = 2.000

- $= n(\text{aloc}) / v(\text{comissao}, \text{aloc})$
- $= 100.000 / 5$
- $= 20.000$



$\text{aloc}(\underline{\text{idFunc}}, \underline{\text{idProj}}, \text{comissao})$
idproj referencia proj
idfunc referencia pfunc

$\text{proj}(\underline{\text{idProj}}, \text{nomeP}, \text{custo})$

$\text{func}(\text{idFunc}, \text{nome}, \text{salario})$

Estatísticas

$n(\text{func})$	=	500
$n(\text{proj})$	=	10.000
$n(\text{aloc})$	=	100.000
$v(\text{salario}, \text{func})$	=	50
$v(\text{custo}, \text{proj})$	=	10
$v(\text{comissao}, \text{aloc})$	=	5

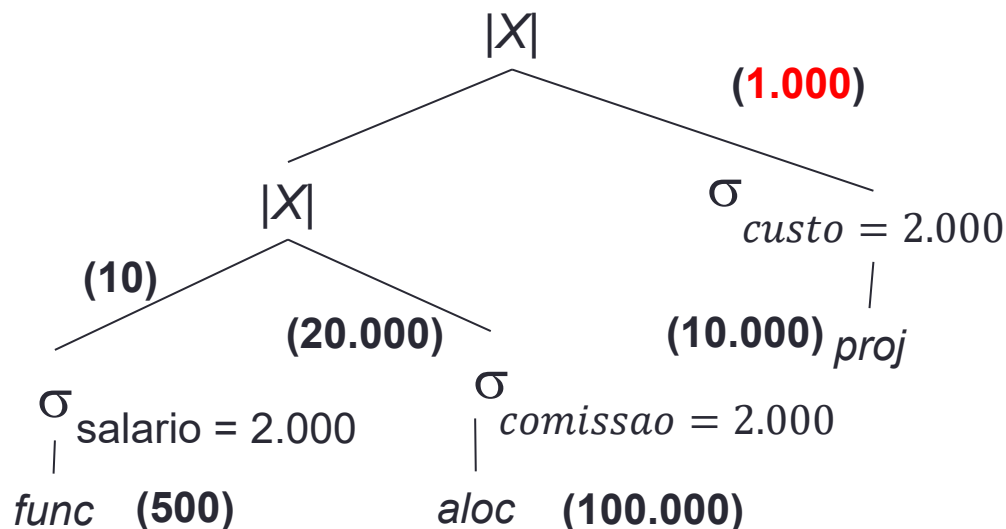
Estimativa com composição de junções

- Estima-se que 1.000 projetos tenham custo = 2.000
 - = $n(\text{proj}) / v(\text{custo}, \text{proj})$
 - = $10.000 / 10$
 - = 1.000

aloc (idFunc, idProj, comissao)
idproj referencia proj
idfunc referencia pfunc

proj (idProj, nomeP, custo)

func(idFunc, nome, salario)

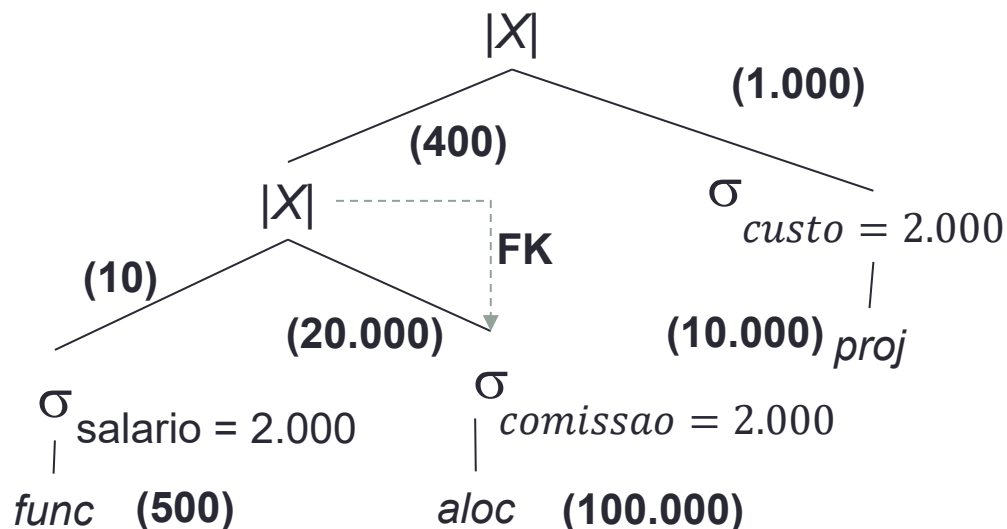


Estatísticas

n(func)	=	500
n(proj)	=	10.000
n(aloc)	=	100.000
v(salario, func)	=	50
v(custo, proj)	=	10
v(comissao, aloc)	=	5

Estimativa com composição de junções

- Junção entre (func,aloc)
 - = $n(\text{FK}) * \text{sel}(\text{PK})$
 - = $20.000 * 1/50$
 - = **400**

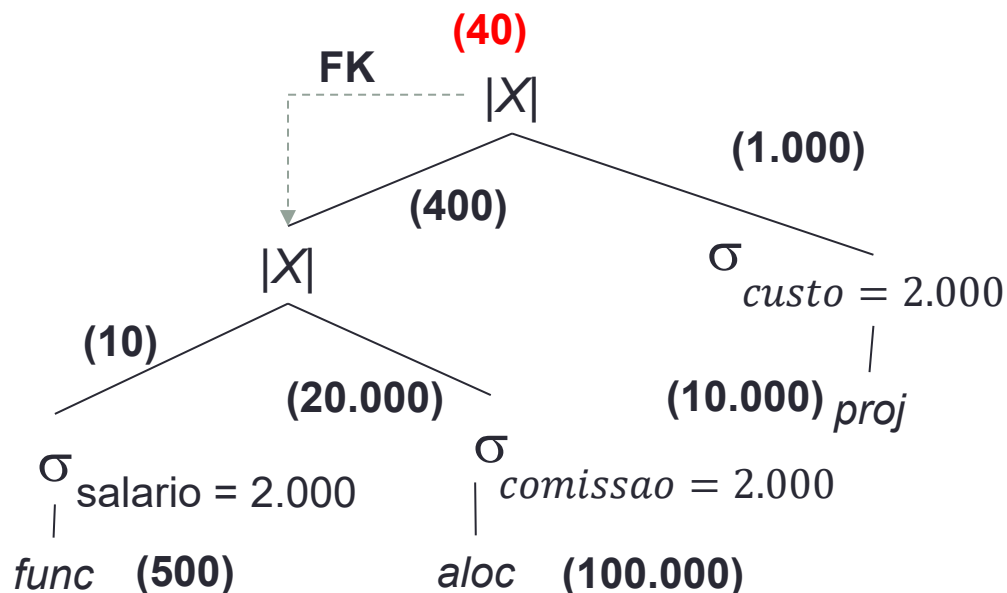


Estatísticas

$n(\text{func})$	=	500
$n(\text{proj})$	=	10.000
$n(\text{aloc})$	=	100.000
$v(\text{salario}, \text{func})$	=	50
$v(\text{custo}, \text{proj})$	=	10
$v(\text{comissao}, \text{aloc})$	=	5

Estimativa com composição de junções

- Junção entre ((func,aloc), proj)
 - $n(\text{FK}) * \text{sel}(\text{PK})$
 - $= 400 * 1/10$
 - **$= 40$**



Estatísticas

$n(\text{func})$	=	500
$n(\text{proj})$	=	10.000
$n(\text{aloc})$	=	100.000
$v(\text{salario}, \text{func})$	=	50
$v(\text{custo}, \text{proj})$	=	10
$v(\text{comissao}, \text{aloc})$	=	5

Considerações Finais

- A estimativa de quantidade de registros retornados é realizada pelo próprio SGBD
- No entanto, é importante ter noção de como a estimativa é realizada
 - Para escolha de que atributos indexar
 - com base na seletividade
 - Para uso de índices mais detalhados
 - como histogramas
 - Para preparar melhor o ambiente de execução
 - Com base no volume de dados que deve ser trabalhado

Atividade Individual

- Com base no banco de dados disponibilizado
 - Execute o script de importação
 - Descubra as estatísticas referentes a cada tabela
 - Estime a quantidade de registros que cada consulta deverá retornar
 - Usando as estatísticas descobertas
 - Compare a estimativa com a quantidade que é realmente retornada para cada consulta

Atividade Individual

- Complete com as estatísticas descobertas

N (func)	
N (depto)	
N (proj)	
N (aloc)	
V(salario,func)	
Min(salario, func)	
Max(salario, func)	
V(custo, proj)	
Min(custo, proj)	
Max(custo, proj)	
V(nomeDepto,depto)	

Atividade Individual

- Consultas a serem executadas:
 1. **SELECT * FROM proj NATURAL JOIN func**
 2. **SELECT * FROM func NATURAL JOIN aloc**
WHERE func.salario <= 10000
 3. **SELECT * FROM proj NATURAL JOIN aloc NATURAL JOIN func**
WHERE proj.custo >= 0 AND func.salario = 10000
 4. **SELECT * FROM proj NATURAL JOIN aloc NATURAL JOIN func**
WHERE proj.custo = 200000
AND (func.salario = 5000 OR func.salario = 8000)
 5. **SELECT * FROM proj NATURAL JOIN aloc**
NATURAL JOIN func NATURAL JOIN depto
WHERE depto.nomeDepto = 'Depto:1'
OR (proj.custo = 200000 AND func.salario = 8000)

Atividade Individual

- Modelo relacional:

Func (idFunc, nome, salario, idDepto)

idDepto referencia depto

Depto (idDepto, nomeDepto)

Proj (idProj, nomeP, custo, duracao)

Aloc (idProj, idFunc)

idProj referencia Proj

idFunc referencia Func