

PROCESSAMENTO DE CONSULTAS E OTIMIZAÇÃO

IFC - Prof.MSc. Leila Lisiane Rossi

PROCESSAMENTO DE CONSULTAS

- Uma linguagem de consulta expressa em uma linguagem de consulta de alto nível, tal como SQL, deve primeiro passar por uma **análise léxica**, uma **análise sintática** e **ser validada**.

PROCESSAMENTO DE CONSULTAS

- A análise léxica
- A análise sintática

PROCESSAMENTO DE CONSULTAS

- A consulta também deve ser **validada** por meio da verificação de que todos os atributos e nomes de relacionamentos são válidos, e se são nomes com significados semânticos no esquema do banco de dados específico que está sendo consultado;
- **Otimização de consulta**

OTIMIZAÇÃO DE CONSULTA

- O SGBD deve planejar uma estratégia de execução. Em geral, uma consulta possui muitas estratégias de execução possíveis e o processo de escolha de uma estratégia adequada para o processamento de uma consulta é chamado **Otimização de Consulta**

OTIMIZAÇÃO DE CONSULTA

- A finalidade do otimizador de um banco de dados é livrar os usuários de suas complexidades e das exigências necessárias para se conseguir consultas de forma eficiente.

OTIMIZAÇÃO DE CONSULTA

- Usar técnicas para reformular consultas em outras que desempenham a mesma funcionalidade, mas com tempo de resposta menor que a consulta original é uma das principais características de um otimizador de consultas

OTIMIZAÇÃO DE CONSULTA

- O otimizador faz com que reduza o esforço manual e repetitivo de identificar e corrigir comandos de SQL realizados de maneira incorreta, causando em baixos desempenhos. Isso é uma vantagem do otimizador evitando-se assim aquisição de novos *hardwares* muito mais poderosos para se alcançar ainda mais quanto ao desempenho

OTIMIZAÇÃO DE CONSULTA

- O módulo otimizador de consulta tem a função de produzir um **plano de execução** e o gerador de código **gera o código** que executa aquele plano. O processador tem um tempo de execução do banco de dados tem a função de **executar o código de consulta**, quer seja no modo interpretado, quer seja no modo compilado, a fim de obter o resultado da consulta.

OTIMIZAÇÃO DE CONSULTA

- O nome mais adequado para **Otimização** de Consulta poderia ser **Planejamento de uma estratégia de execução**

TRANSFORMAÇÕES DE CONSULTAS SQL PARA ÁLGEBRA

- A consulta SQL é decomposta em pequenas unidades chamadas de **blocos de consulta**, e estas são otimizadas um bloco por vez. Os blocos aninhados são tratados como uma chamada de uma sub-rotina, feita uma vez por uma tupla mais externa. São esses blocos que serão convertidos pela álgebra relacional

TRANSFORMAÇÕES DE CONSULTAS SQL PARA ÁLGEBRA

- Traduzir a consulta em uma expressão equivalente da álgebra relacional e representá-la em uma estrutura de dados conhecida como árvore de consulta
- O otimizador decide como avaliar a consulta transformada

OTIMIZAÇÃO DE CONSULTA

- É uma tarefa importante em sistemas relacionais. É necessário compreender otimização para entender o impacto de desempenho de um projeto de banco de dados (relações, índices) em uma carga de trabalho (conjunto de consultas).

Consulta em linguagem de alto nível

ANÁLISE LÉXICA,
ANÁLISE SINTÁTICA
E VALIDAÇÃO

Forma intermediária da consulta

OTIMIZADOR DE CONSULTA

Plano de execução

GERADOR DE CÓDIGO
DE CONSULTA

Código para executar a consulta

PROCESSADOR EM TEMPO
DE EXECUÇÃO DO
BANCO DE DADOS

Resultado da consulta

ALGORITMOS DE SELEÇÃO

- Pesquisa Linear
- Pesquisa Binária
- Usando Índice Primário
- Usando Chave Hash
- Combinado com o Índice Primário
- Usando Índice de Agrupamento
- Usando Índice Secundário

SELEÇÃO

- Conjuntiva – AND – Exemplo: Candidato Cargo de São Paulo e Solteiro
- Disjuntiva – OR
- Índice
- Seletividade

OTIMIZAÇÃO DE CONSULTA

- **Utilização de Heurísticas na Otimização de Consultas**
 - O analisador sintático de uma consulta de alto nível primeiro gera uma representação interna inicial, que depois é **otimizada** de acordo com **regras de heurística**. Na sequência, um plano de execução de consulta é gerado para executar grupos de operações com base nos caminhos de acesso disponíveis para os arquivos envolvidos na consulta

OTIMIZAÇÃO DE CONSULTA

- **Utilização de Heurísticas na Otimização de Consultas**
 - Informática - Método de investigação baseado na aproximação progressiva de um dado problema
 - Uma das principais regras heurísticas é aplicar as operações **SELECT** e **PROJECT** antes de aplicar o JOIN ou outras operações binárias

OTIMIZAÇÃO DE CONSULTA

- **Notação de Árvores de Consulta e de Grafos de Consulta**
 - Uma árvore de consulta é uma estrutura de dados de árvore que corresponde a uma expressão da **álgebra relacional**. Ela representa as relações de entrada de uma consulta como nós folhas da árvore e representa as operações da álgebra relacional como nós internos.

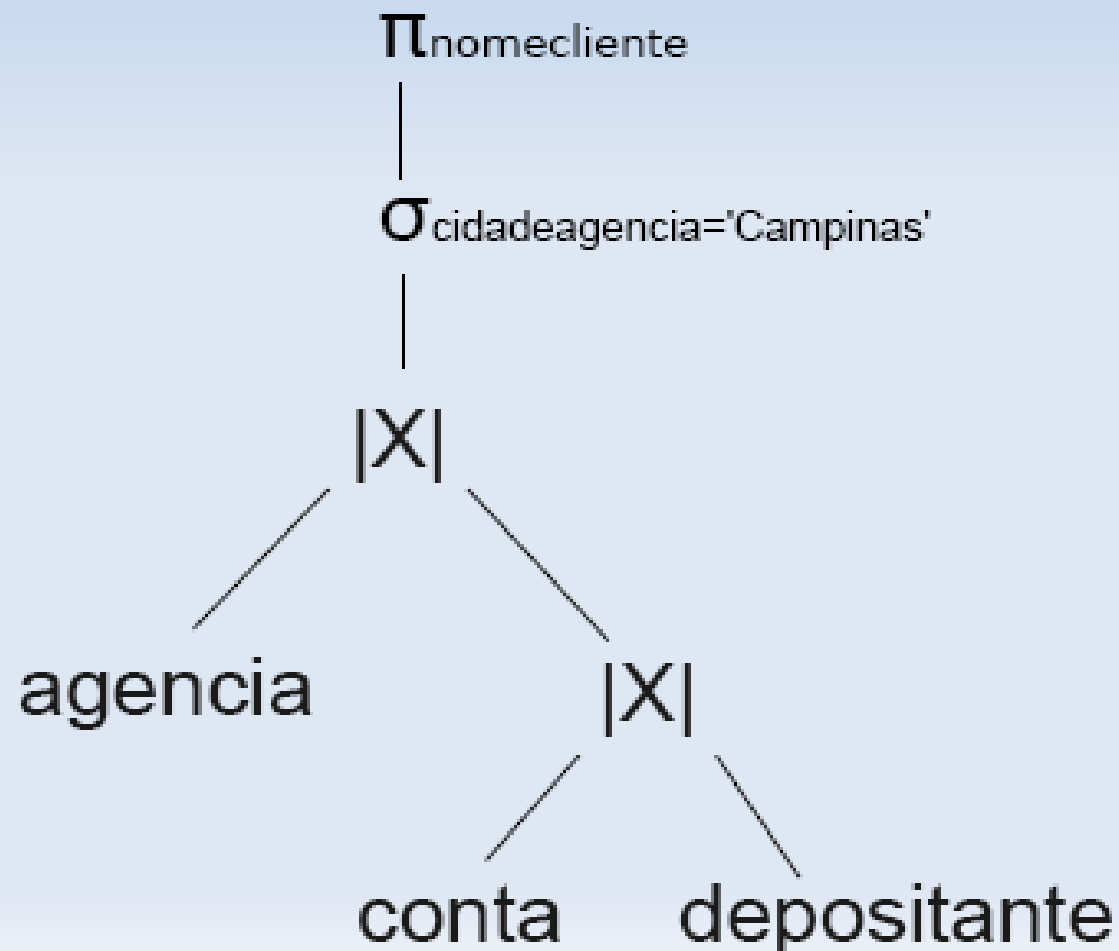
OTIMIZAÇÃO DE CONSULTA

- Uma execução de árvore de consulta consiste na execução de uma operação de nó interno sempre que seus operandos estiverem disponíveis e depois da substituição do nó interno pela relação que resulta da execução da operação. A execução termina quando o nó raiz é executado e produz a relação de resultado da consulta.

OTIMIZAÇÃO DE CONSULTA

- **Árvore de Consulta**
 - Estrutura que representa uma expressão em álgebra relacional
 - **Nós-Folhas – Relações**
 - **Nós-Internos – Operações da Álgebra**

$\pi_{\text{nomecliente}}(\sigma_{\text{cidadeagencia}=\text{'Campinas'}}(\text{agencia } |X| (\text{conta } |X| \text{ depositante})))$



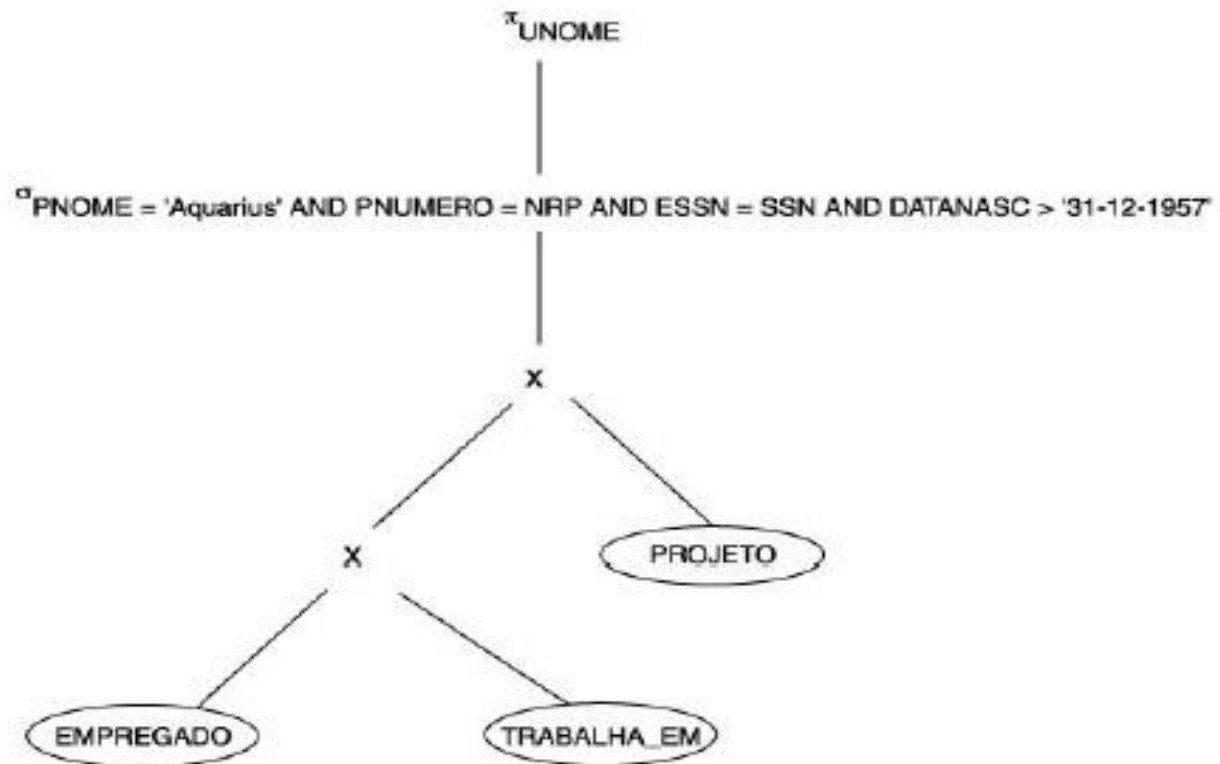
OTIMIZAÇÃO DE CONSULTA

- Otimização Heurística de Árvores de Consulta
 - Ex: Encontre os últimos nomes dos empregados nascidos após 1957 que trabalham no projeto “Aquarius”. Em SQL:
 - **SELECT Unome FROM Empregado, Trabalha_em, Projeto WHERE P_nome = ‘Aquarius’ AND Pnumero = NRP AND ESSN = SSN AND Data_nasc > ‘31-12-1957’;**

OTIMIZAÇÃO DE CONSULTA

- Árvore Canônica

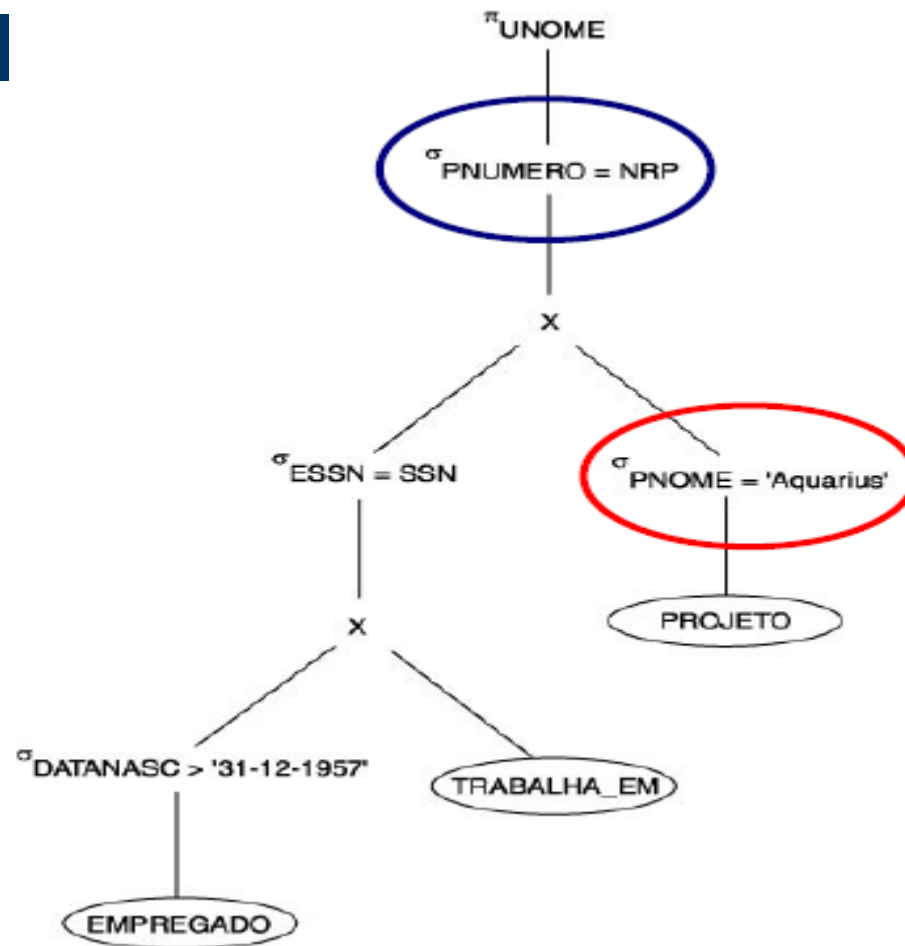
OTIMIZAÇÃO DE CONSULTA



OTIMIZAÇÃO DE CONSULTA

- Operações de Seleção para baixo

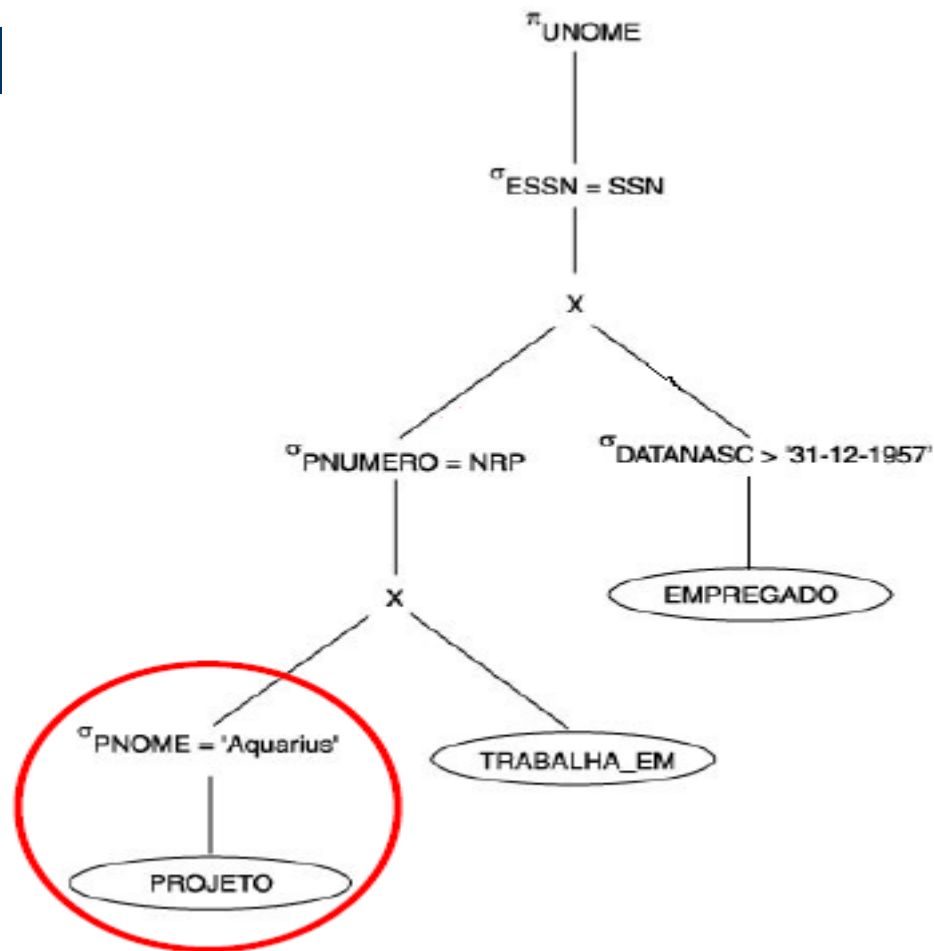
OTIMIZAÇÃO DE CONSULTA



OTIMIZAÇÃO DE CONSULTA

- **Prioridade Seleção mais Restritiva**
 - Seletividade
 - Seleção em direção às folhas

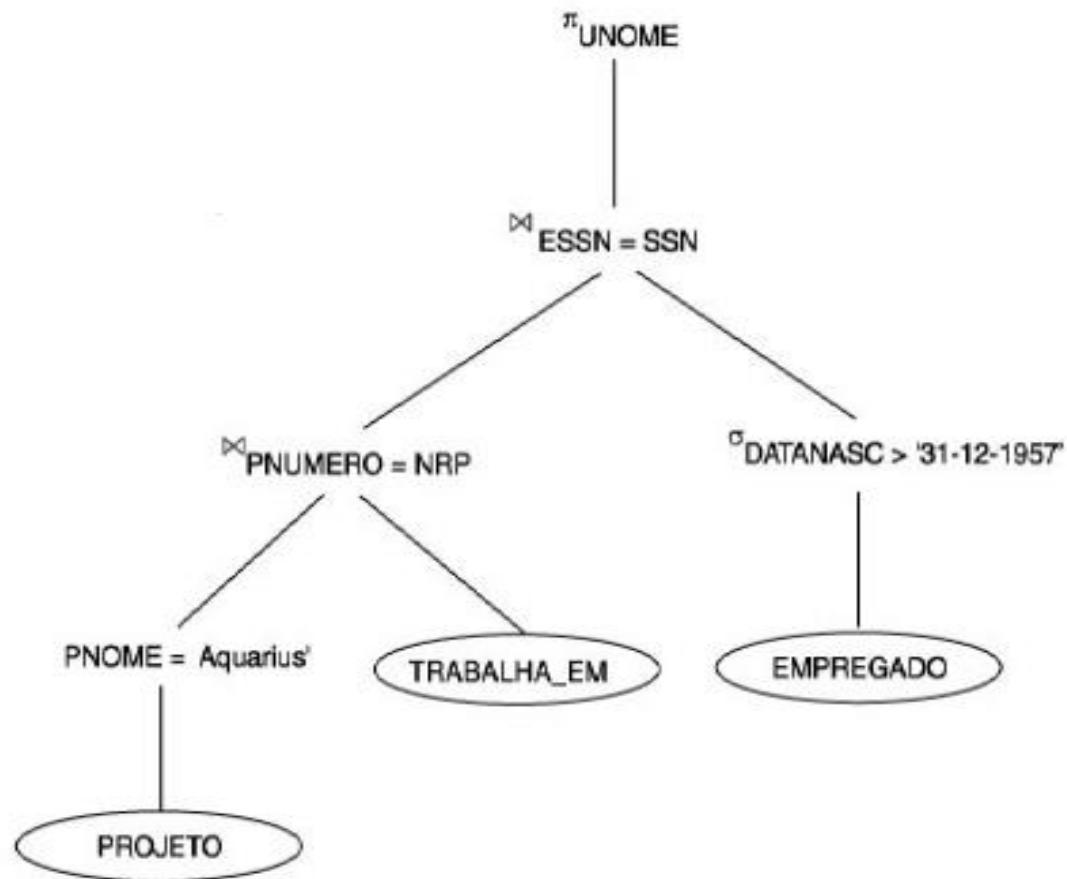
OTIMIZAÇÃO DE CONSULTA



OTIMIZAÇÃO DE CONSULTA

- Transformação dos Produtos Cartesianos em Junção

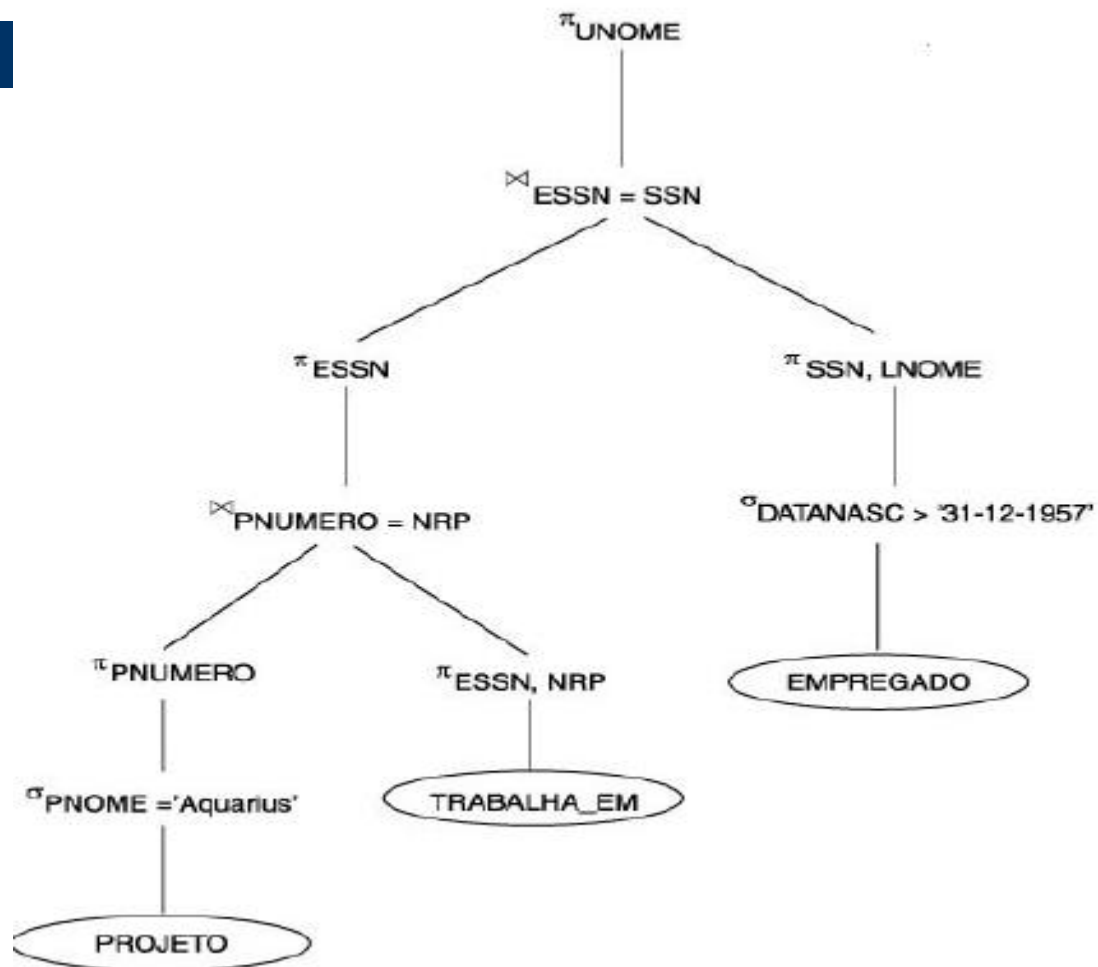
OTIMIZAÇÃO DE CONSULTA



OTIMIZAÇÃO DE CONSULTA

- Operações de Projeção

OTIMIZAÇÃO DE CONSULTA



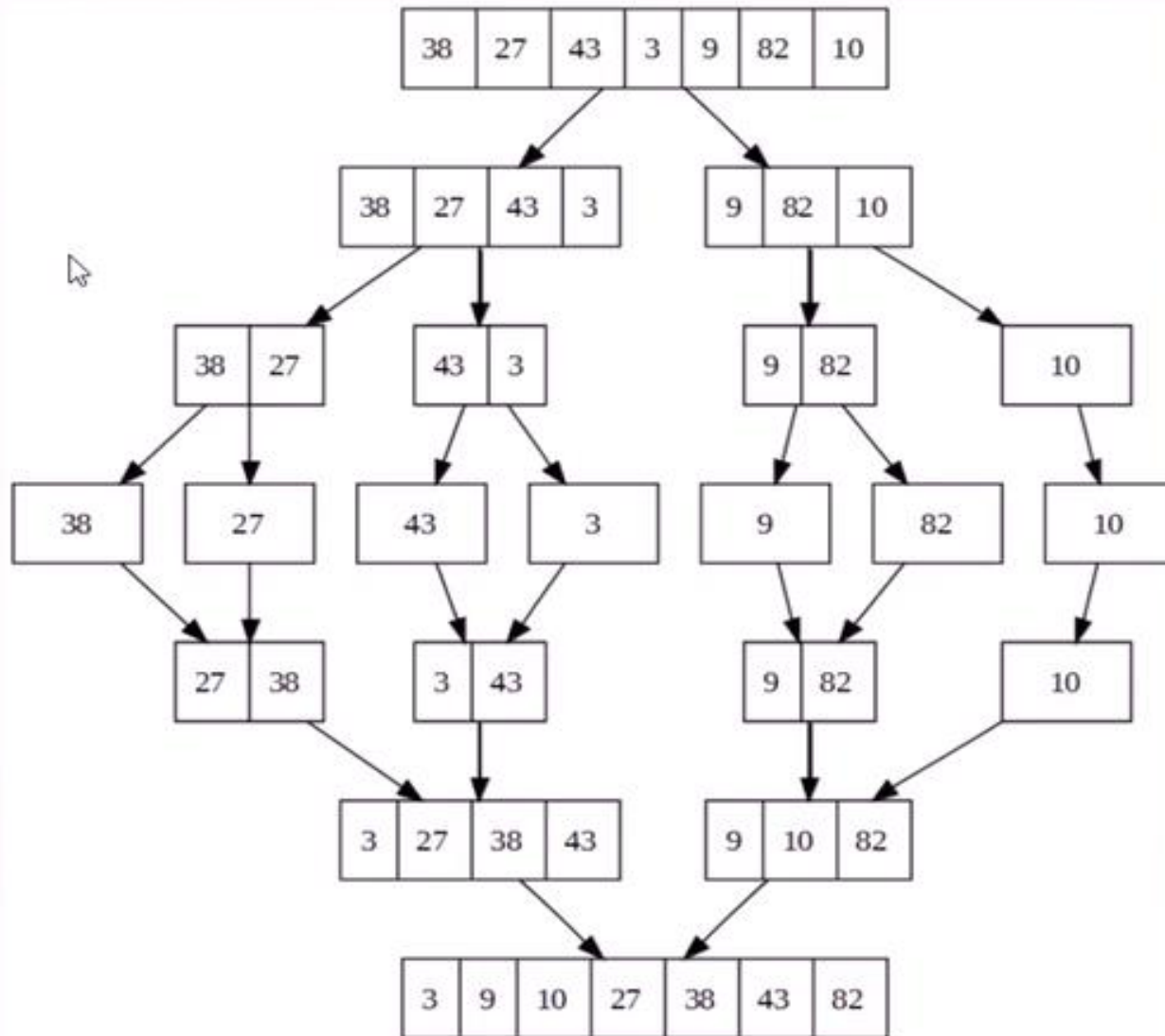
OTIMIZAÇÃO DE CONSULTA

- **Conversão de Árvores de Consulta em Planos de Execução de Consulta**
 - Um plano de execução para uma expressão de álgebra relacional representada como uma árvore de consulta, inclui informações sobre os **métodos de acesso** disponíveis para cada relação, bem como os **algoritmos** a serem utilizados na computação dos operadores relacionais representados na árvore

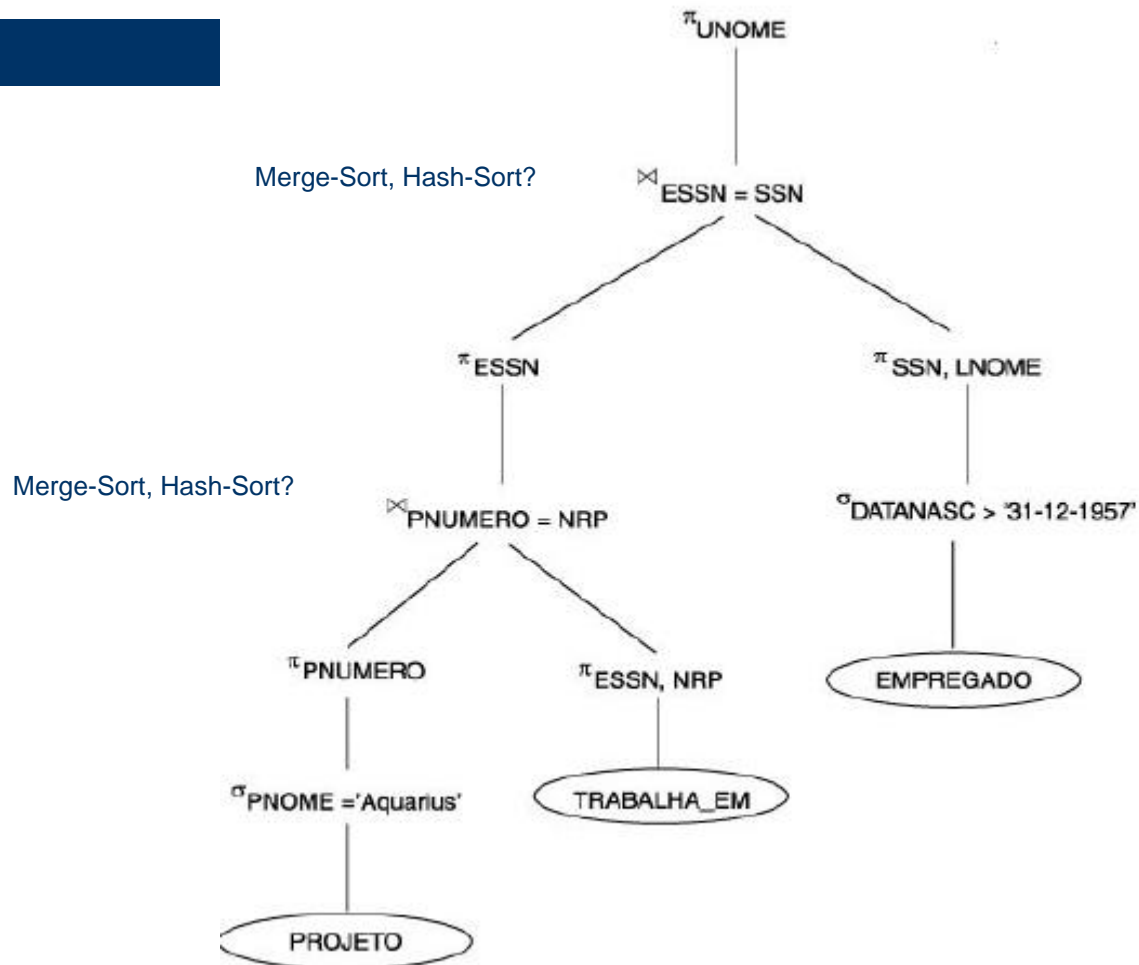
OTIMIZAÇÃO DE CONSULTA

- Ordenação Externa
 - **Sort-Merge** (Select)
 - Junção **Hash**, Junção **Sort-Merge** (Junção/Select mais complexo)
 - Variações **Sort-Merge** (Project / Conjuntos)
 - <https://www.youtube.com/watch?v=JSceec-wEyw>

Merge Sort Tradicional



OTIMIZAÇÃO DE CONSULTA



OTIMIZAÇÃO DE CONSULTA

- **Utilização de Seletividade e Estimativa de Custo na Otimização de Consultas**
 - Um otimizador de consultas não deve depender somente de regras heurísticas.

OTIMIZAÇÃO DE CONSULTA

- **Utilização de Seletividade e Estimativa de Custo na Otimização de Consultas**
 - Custo de Armazenamento
 - Custo de Computação
 - Custo de Acesso ao Armazenamento Secundário
 - Custo do uso de Memória
 - Custo de Comunicação

OTIMIZAÇÃO DE CONSULTA













- O PostgreSQL usa os seguintes comandos para a otimização de consultas:
 - EXPLAIN
 - ANALYZE
 - VACUUM



OTIMIZAÇÃO DE CONSULTA

- EXPLAIN
 - O comando EXPLAIN exhibe o Plano de Execução escolhido pelo Otimizador

FileEditQueryFavouritesMacrosViewHelp



astronomia on postgres@localhost:5432

SQL Editor

Graphical Query Builder

explain select * from wise

Delete

Delete All

Scratch pad

Output pane

Data Output

Explain

Messages










History

	QUERY PLAN
	text
1	Seq Scan on wise (cost=0.00..601.39 rows=10939 width=379)

OK.

UnixLn 1 Col 9 Ch 9

1 row.10 ms



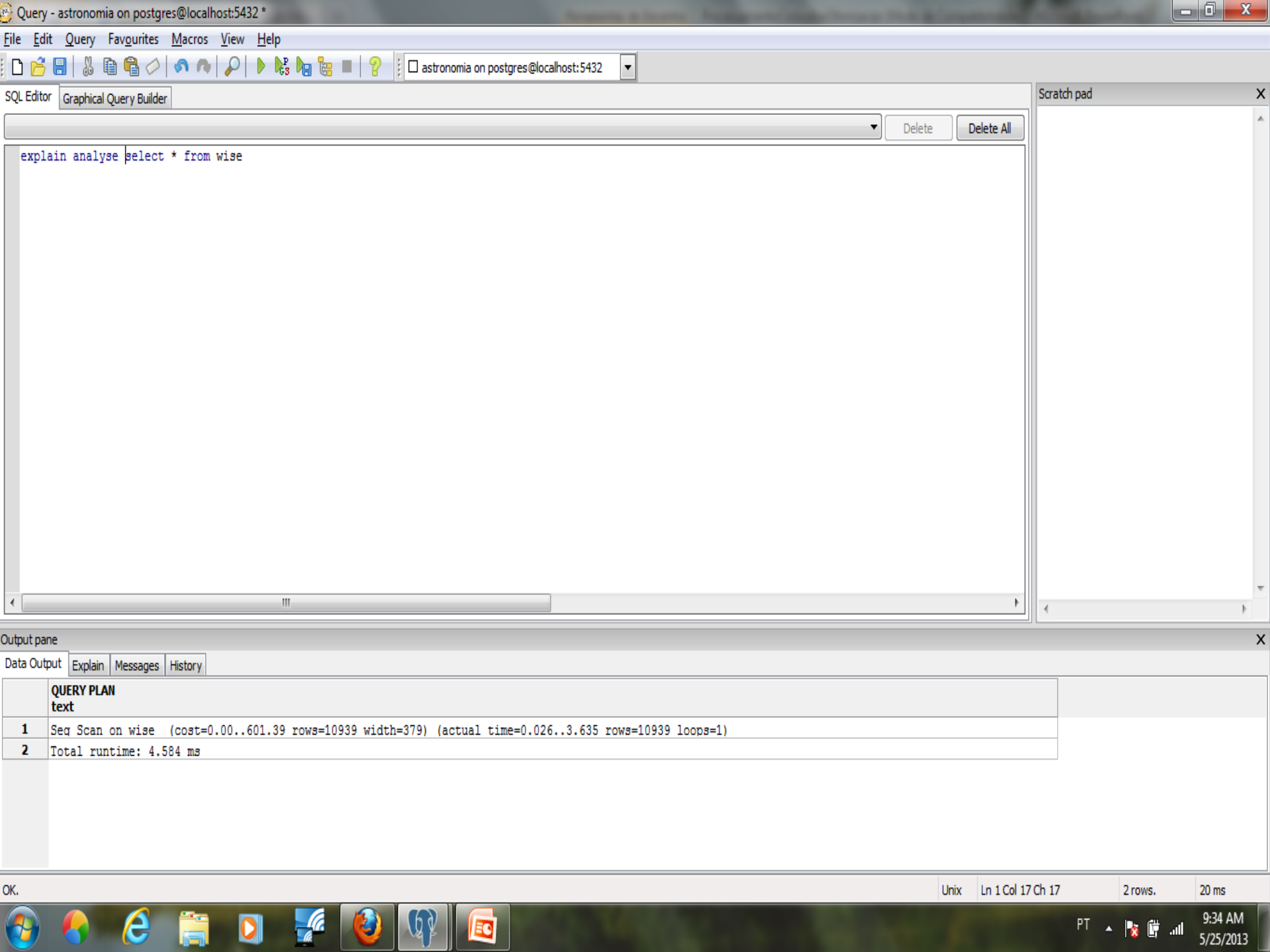
PT9:31 AM5/25/2013

OTIMIZAÇÃO DE CONSULTA

- ANALYSE
 - O comando analyse coleta estatísticas sobre o conteúdo das tabelas do banco de dados e armazena os resultados na tabela do sistema pg_statistic

OTIMIZAÇÃO DE CONSULTA

- ANALYSE
 - Com base nestas estatísticas o planejador de comandos determina o plano de execução mais eficiente para os comandos. Estas estatísticas devem ser atualizadas com frequência para não comprometer o desempenho do banco de dados por uma escolha errada do plano de comandos



Delete

Delete All

```
explain analyse select * from wise
```

QUERY PLAN
text

- | | |
|---|---|
| 1 | Seq Scan on wise (cost=0.00..601.39 rows=10939 width=379) (actual time=0.026..3.635 rows=10939 loops=1) |
| 2 | Total runtime: 4.584 ms |

Delete

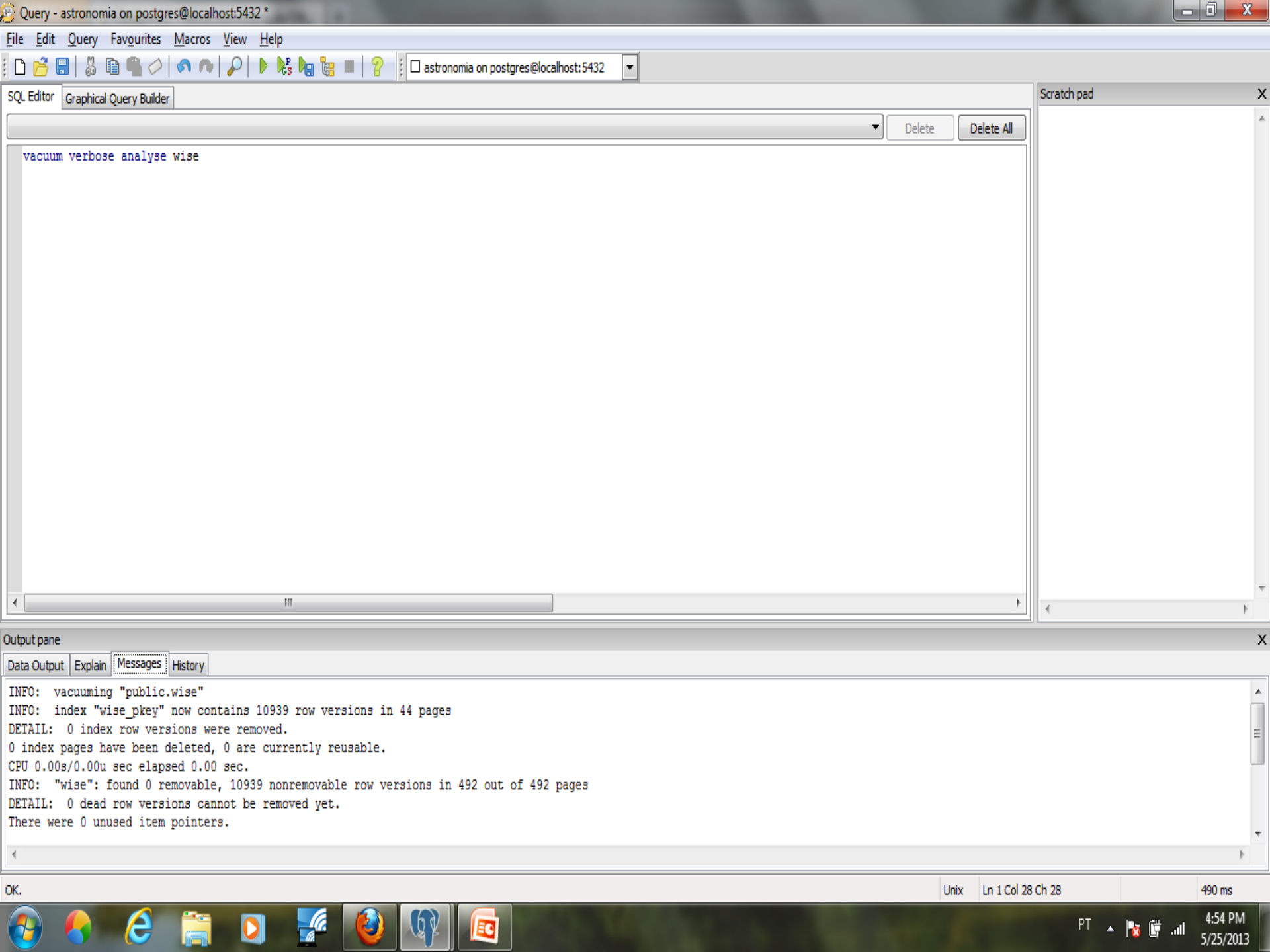
Delete All

```
select * from pg_statistic
```

	starelid oid	staatnum smallint	stainherit boolean	stanullfrac real	stawidth integer	stadistinct real	stakind1 smallint	stakind2 smallint	stakind3 smallint	stakind4 smallint	staop1 oid	staop2 oid	staop3 oid	staop4 oid	stanumbers1 real[]	stanumbers2 real[]	stanumbers3 real[]	stanumbers4 real[]	stavalues1 anyarray	stavalues2 anyarray	stavalues3 anyarray
1	1259	16	f	0	1	5	1	3	0	0	92	631	0	0	{0.347607,0	{0.111208}			{i,r,v,t,		
2	1260	1	f	0	64	-1	0	0	0	0	0	0	0	0							
3	1260	2	f	0	1	-1	0	0	0	0	0	0	0	0							
4	1260	3	f	0	1	-1	0	0	0	0	0	0	0	0							
5	1260	4	f	0	1	-1	0	0	0	0	0	0	0	0							
6	1260	5	f	0	1	-1	0	0	0	0	0	0	0	0							

OTIMIZAÇÃO DE CONSULTA

- VACUUM
 - O comando VACUUM recupera espaço em disco e otimiza o desempenho do banco e previne contra perda de dados muito antigos
 - FULL – Limpeza completa
 - VERBOSE – Relatório detalhado da atividade de limpeza de cada tabela



```
vacuum verbose analyse wise
```

```
INFO: vacuuming "public.wise"
INFO: index "wise_pkey" now contains 10939 row versions in 44 pages
DETAIL: 0 index row versions were removed.
0 index pages have been deleted, 0 are currently reusable.
CPU 0.00s/0.00u sec elapsed 0.00 sec.
INFO: "wise": found 0 removable, 10939 nonremovable row versions in 492 out of 492 pages
DETAIL: 0 dead row versions cannot be removed yet.
There were 0 unused item pointers.
```


TRABALHO I

- Processamento e Otimização de Consultas
 - Vídeo

REFERÊNCIAS

- ELMASRI, NAVATHE – Sistemas de Banco de Dados – São Paulo - Addison Wesley, 2011
- KORTH, SILBERSCHATZ - Sistemas de Banco de Dados – São Paulo Makron Books, 2006
- NETO, Acácio Feliciano e outros – Engenharia da Informação – São Paulo – McGraw-Hill Ltda, 1988
- <http://imasters.com.br/artigo/22264/banco-de-dados/10-tecnicas-para-otimizacao-de-instrucoes-sql/>
- <http://wp.sigmod.org/?searchterm=query+optimization&p=1075>
- <https://www.youtube.com/watch?v=cqXeTlvfyk4&t=3622s>
- <https://www.youtube.com/watch?v=nmhjrl-aW5o>
- https://wiki.postgresql.org/wiki/Introduction_to_VACUUM,_ANALYZE,_EXPLAIN,_and_COUNT
- <https://www.youtube.com/watch?v=JSceec-wEyw>