

# Benchmarks

Profa. Dra. Cristina Dutra de Aguiar Ciferri

Prof. Dr. Ricardo Rodrigues Ciferri

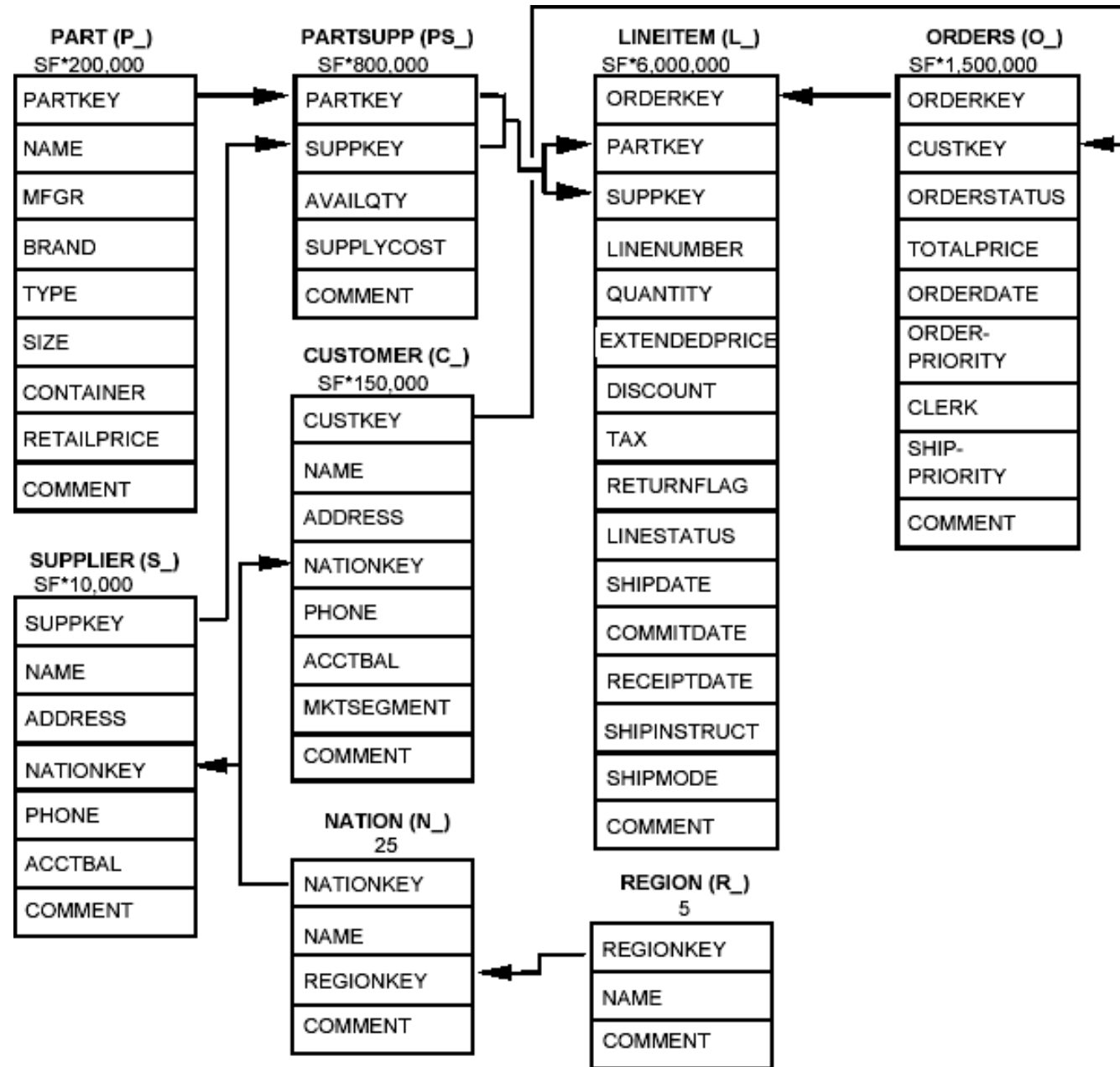
# Benchmarks

- Usados para avaliar o desempenho de aplicações de banco de dados
- Funcionalidades oferecidas
  - esquema de dados
  - diferentes tipos de consulta
  - diferentes operações de inserção, remoção, atualização
- Aspectos considerados
  - seletividade
  - volume de dados

# Benchmark TPC-H

- Características
  - voltado à tomada de decisão
    - DW de uma aplicação que representa dados históricos relativos a compras e vendas de uma corporação
  - especifica como dados sintéticos para DW devem ser gerados
- Especificação completa
  - <http://www.tpc.org/tpch/>
  - Poess, M. and Floyd, C. (2000) “New TPC Benchmarks for Decision Support and Web Commerce”, SIGMOD Record, 29 (4): 64-71.

# TPC-Q

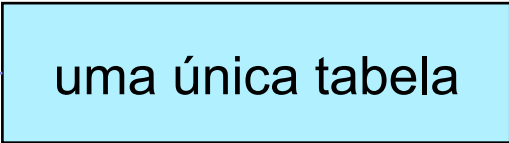


# Consultas

- Define diferentes tipos de consulta
- Exemplo: Q1
  - Lista diversas informações de interesse relacionadas ao negócio.

# Consulta Q1

```
select
  l_returnflag,
  l_linestatus,
  sum(l_quantity) as sum_qty,
  sum(l_extendedprice) as sum_base_price,
  sum(l_extendedprice*(1-l_discount)) as sum_disc_price,
  sum(l_extendedprice*(1-l_discount)*(1+l_tax)) as sum_charge,
  avg(l_quantity) as avg_qty,
  avg(l_extendedprice) as avg_price,
  avg(l_discount) as avg_disc,
  count(*) as count_order
from
  lineitem
where
  l_shipdate <= date '1998-12-01' - interval '[DELTA]' day (3)
group by
  l_returnflag,
  l_linestatus
order by
  l_returnflag,
  l_linestatus;
```




# Consulta Q3

- Especificação
  - Esta consulta recupera os 10 pedidos ainda não enviados que possuem os valores mais elevados.

# Consulta Q3

```
select
  l_orderkey,
  sum(l_extendedprice*(1-l_discount)) as revenue,
  o_orderdate,
  o_shippriority
from
  customer,
  orders,
  lineitem
where
  c_mktsegment = '[SEGMENT]'
  and c_custkey = o_custkey
  and l_orderkey = o_orderkey
  and o_orderdate < date '[DATE]'
  and l_shipdate > date '[DATE]'
group by
  l_orderkey,
  o_orderdate,
  o_shippriority
order by
  revenue desc,
  o_orderdate;
```

junção de tabelas  
- esquema estrela -





# Consulta Q9

- Especificação
  - Obtém as receitas dos fornecedores agrupadas por nação e ano, realizando a junção das tabelas de fatos LineItem e PartSupp por intermédio das tabelas de dimensão compartilhadas Customer e Supplier.

# Consulta Q9

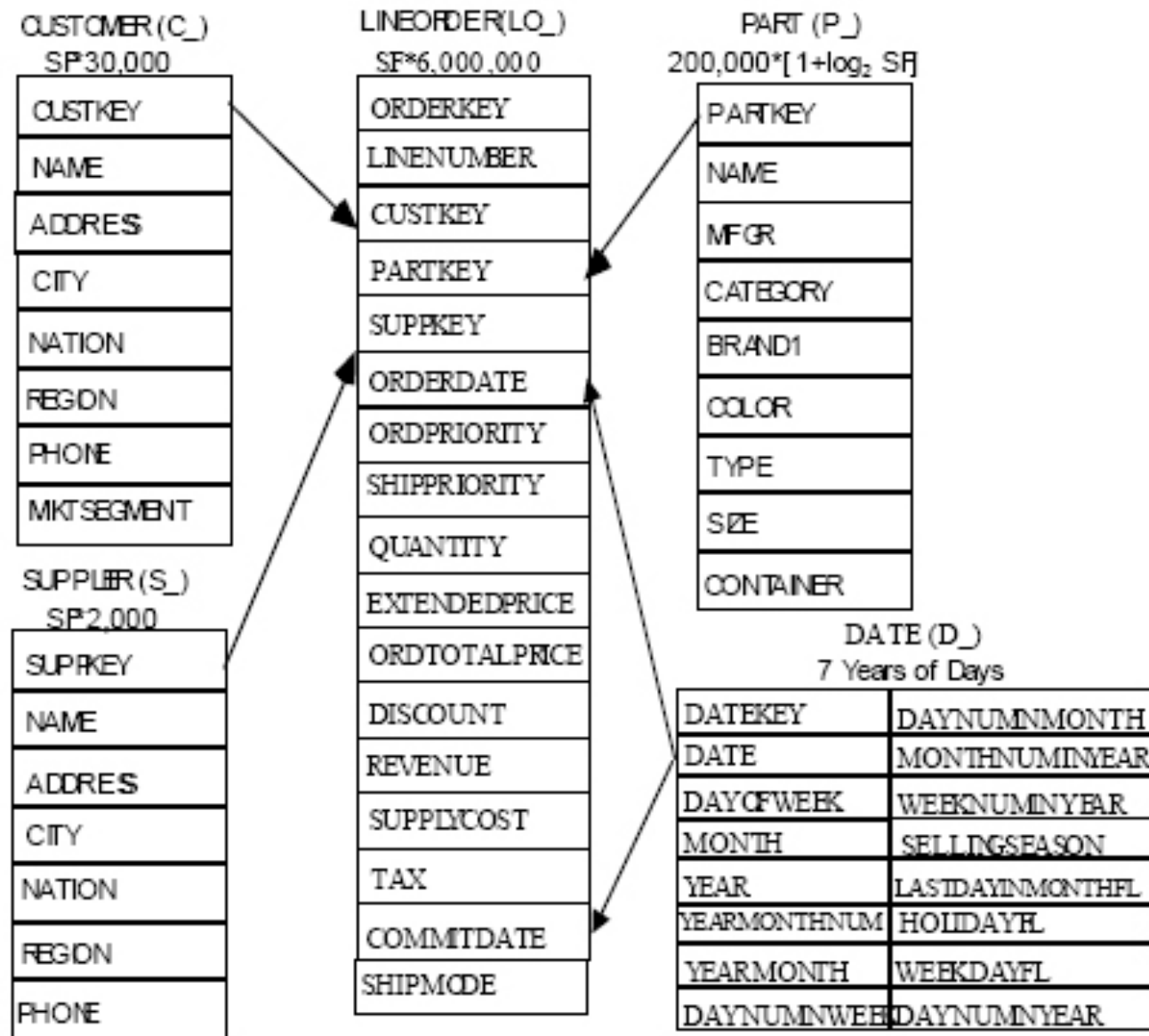
```
select
  nation,
  o_year,
  sum(amount) as sum_profit
from (
  select
    n_name as nation,
    extract(year from o_orderdate) as o_year,
    l_extendedprice * (1 - l_discount) - ps_supplycost * l_quantity as amount
  from
    part,
    supplier,
    lineitem,
    partsupp,
    orders,
    nation
  where
    s_suppkey = l_suppkey
    and ps_suppkey = l_suppkey
    and ps_partkey = l_partkey
    and p_partkey = l_partkey
    and o_orderkey = l_orderkey
    and s_nationkey = n_nationkey
    and p_name like '%[COLOR]%'
) as profit
group by
  nation,
  o_year
order by
  nation,
  o_year desc;
```

← junção de tabelas  
- floco de neve -

# Benchmark SSB

- Baseado no benchmark TPC-H
  - Foram juntadas as tabelas de fatos Lineitem e Order, gerando a tabela de fatos Lineorder
  - Foi excluída a tabela de fatos Partsupp
  - Foi adicionada a tabela de dimensão Date
- Especificação completa
  - <http://www.cs.umb.edu/~poneil/StarSchemaB.pdf>

# TPC-S



**LINEORDER Table Layout SF\*6,000,000**

LO\_ORDERKEY numeric (int up to SF 300) first 8 of each 32 keys populated

LO\_LINENUMBER numeric 1-7

LO\_CUSTKEY numeric identifier FK to C\_CUSTKEY

LO\_PARTKEY identifier FK to P\_PARTKEY

LO\_SUPPKEY numeric identifier FK to S\_SUPPKEY

LO\_ORDERDATE identifier FK to D\_DATEKEY

LO\_ORDERPRIORITY fixed text, size 15 (See pg 91: 5 Priorities: 1-URGENT, etc.)

LO\_SHIPPRIORITY fixed text, size 1

LO\_QUANTITY numeric 1-50 (for PART)

LO\_EXTENDEDPRICE numeric  $\leq 55,450$  (for PART)

LO\_ORDTOTALPRICE numeric  $\leq 388,000$  (ORDER)

LO\_DISCOUNT numeric 0-10 (for PART, percent)

LO\_REVENUE numeric (for PART:  
(lo\_extendedprice\*(100-lo\_discnt))/100)

LO\_SUPPLYCOST numeric (for PART)

LO\_TAX numeric 0-8 (for PART)

LO\_COMMITDATE FK to D\_DATEKEY

LO\_SHIPMODE fixed text, size 10 (See pg. 91: 7 Modes: REG AIR, AIR, etc.)

Compound Primary Key: LO\_ORDERKEY,  
LO\_LINENUMBER

# Consultas

- Define diferentes tipos de consulta
- Exemplo: Q1
  - quantifica a renda a partir da eliminação de certos descontos da empresa, dada uma porcentagem de produtos enviados em um determinado ano

# Consulta Q1.1

Q1.1 YEAR = 1993, DISCOUNT = 2, QUANTITY = 25, so predicates are d\_year = 1993, lo\_quantity < 25, lo\_discount between 1 and 3.

```
select sum(lo_extendedprice*lo_discount) as revenue
from lineorder, date
where lo_orderdate = d_datekey
and d_year = 1993
and lo_discount between 1 and 3
and lo_quantity < 25;
```

$FF = (1/7)*0.5*(3/11) = 0.0194805$ . Number of lineorder rows selected, for SF = 1, is  $0.0194805*6,000,000 \approx 116,883$ .

# Consulta Q1.2

Q1.2 d\_yearmonthnum = 199401, lo\_quantity between 26 and 35, lo\_discount between 4 and 6.

```
select sum(lo_extendedprice*lo_discount) as revenue
from lineorder, date
where lo_orderdate = d_datekey
and d_yearmonthnum = 199401
and lo_discount between 4 and 6
and lo_quantity between 26 and 35;
```

$FF = (1/84)*(3/11)*0.2 = 0.00064935$ . Number of lineorder rows selected, for SF = 1:  
 $0.00064935*6,000,000 \approx 3896$ .



# Consulta Q1.3

Q1.3 d\_weeknuminyear = 6 and d\_year = 1994,  
lo\_quantity between 36 and 40, lo\_discount between 5  
and 7.

```
select sum(lo_extendedprice*lo_discount) as revenue
from lineorder, date
where lo_orderdate = d_datekey
and d_weeknuminyear = 6
and d_year = 1994
and lo_discount between 5 and 7
and lo_quantity between 26 and 35;
```

$FF = (1/364)*(3/11)*0.1 = .000075$ . Number of li-  
neorder rows selected, for  $SF = 1$ , is  
 $.000075*6,000,000 \approx 450$ .

# Medidas de Desempenho

- Os relatórios devem conter
  - planos de consulta
  - número de linhas acessadas
  - tempo execução da consulta
  - tempo da CPU utilizada
  - uso de entrada e saída