

CCT396, Fall 2011

Database Design and Implementation

Yuri Takhteyev
University of Toronto



This presentation is licensed under Creative Commons Attribution License, v. 3.0. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>. This presentation incorporates images from the Crystal Clear icon collection by Everaldo Coelho, available under LGPL from <http://everaldo.com/crystal/>.

Week 9

Advanced Queries

What are the names of the Diveshop's customers who paid cash for their orders?

```
select DIVECUST.Name  
from DIVECUST  
  join DIVEORDS  
    on DIVECUST.Customer_No =  
       DIVEORDS.Customer_No  
where  
  DIVEORDS.PaymentMethod = "Cash";
```

How many customers paid cash for their orders?

```
select
  count(distinct DIVECUST.Name)
from DIVECUST
  join DIVEORDS
    on DIVECUST.Customer_No =
       DIVEORDS.Customer_No
where
  DIVEORDS.PaymentMethod = "Cash";
```

How much cash did all of them paid together (in cash)?

```
select
  sum(VacationCost)
from DIVEORDS
where
  PaymentMethod = "Cash";
```

And how much money was paid through Visa?

```
select
  sum(VacationCost)
from DIVEORDS
where
  PaymentMethod = "Visa";
```

And how much money was paid through Master Card?

```
select
  sum(VacationCost)
from DIVEORDS
where
  PaymentMethod = "Master Card";
```

Can we just get all the sums for each payment method at once?

```
select  
    PaymentMethod, sum(VacationCost)  
from DIVEORDS  
group by PaymentMethod;
```


Which payment method brought in most money?

```
select  
    PaymentMethod, sum(VacationCost)  
from DIVEORDS  
group by PaymentMethod  
order by sum(VacationCost) desc;
```

Which payment method brought in most money for vacations that involved 1 or 2 people?

```
select  
  PaymentMethod, sum(VacationCost)  
from DIVEORDS  
where No_Of_People <= 2  
group by PaymentMethod  
order by sum(VacationCost) desc;
```

Note: "where" before "group by"!

Which payment method was most popular for vacations that involved 1 or 2 people?

```
select  
  PaymentMethod, sum(VacationCost)  
from DIVEORDS  
where No_Of_People <= 2  
group by PaymentMethod  
order by sum(VacationCost) desc  
limit 1;
```

Which payment method was most popular for vacations that involved 1 or 2 people and cost under \$5000?

```
select  
  PaymentMethod, sum(VacationCost)  
from DIVEORDS  
where No_Of_People <= 2 and  
VacationCost < 5000  
group by PaymentMethod  
order by sum(VacationCost) desc  
limit 1;
```

Which payment methods brought in more than \$30,000 in total?

```
select
  PaymentMethod, sum(VacationCost)
from DIVEORDS
group by PaymentMethod
having sum(VacationCost) > 30000
order by sum(VacationCost) desc;
```

Note the order!

Where vs Having

where

selects rows from the original table (after all the joins)

having by

selects rows from the aggregated table

Order of Clauses

- 7** **select ...**
- 1** **from ...**
 - join ... (several times)**
- 2** **where ...**
- 3** **group by ...**
- 4** **having ...**
- 5** **order by ...**
- 6** **limit ...**

Which payment methods brought in more than \$15,000 on average per order?

```
select
  PaymentMethod, avg(VacationCost)
from DIVEORDS
group by PaymentMethod
having avg(VacationCost) > 15000;
```


What was the average amount coming from each payment methods for orders costing over \$15,000?

```
select  
  PaymentMethod, avg(VacationCost)  
from DIVEORDS  
where VacationCost > 15000  
group by PaymentMethod;
```

What about this query?

```
select  
    PaymentMethod, avg(VacationCost)  
from DIVEORDS  
where avg(VacationCost) > 15000  
group by PaymentMethod;
```

Invalid!

And this one?

```
select
  PaymentMethod, avg(VacationCost)
from DIVEORDS
group by PaymentMethod
having VacationCost > 15000;
```

Invalid!

What payment methods were used for vacations costing *above* average?

```
select avg(VacationCost)
from DIVEORDS;
```

```
select
  PaymentMethod
from DIVEORDS
where VacationCost > 18462.8
group by PaymentMethod;
```

What payment methods were used for vacations costing *above* average?

```
set @avg_cost = (  
    select avg(VacationCost)  
    from DIVEORDS);
```

```
select  
    PaymentMethod  
from DIVEORDS  
where VacationCost > @avg_cost  
group by PaymentMethod;
```

What payment methods were used for vacations costing *above* average?

```
select
  PaymentMethod
from DIVEORDS
where VacationCost > (
  select avg(VacationCost)
  from DIVEORDS)
group by PaymentMethod;
```

An “uncorrelated” subquery

What payment methods on average brought more money than the average for all vacations?

```
select
  PaymentMethod, avg(VacationCost)
from DIVEORDS
group by PaymentMethod
having avg(VacationCost) > (
  select avg(VacationCost)
  from DIVEORDS);
```

Still "uncorrelated"

Which vacations cost more than the average for their payment method?

```
select
```

```
    PaymentMethod, VacationCost
```

```
from DIVEORDS as O
```

```
where
```

```
    VacationCost > (
```

```
        select avg(VacationCost)
```

```
        from DIVEORDS
```

```
        where
```

```
            PaymentMethod=O.PaymentMethod
```

```
);
```

This is a “correlated” query!

Which categories of marine life have more than one species? (Use BIOLIFE.)

```
select
  Category, count (*)
from BIOLIFE
group by Category
having count (*) > 1;
```

What are the smallest and the largest lengths in each of the categories that have more than 1 species?

```
select  
  Category, min (Length_cm) ,  
  max (Length_cm)  
from BIOLIFE  
group by Category  
having count (*) > 1;
```

Which category with >1 species has the largest ratio between the largest and the smallest length of species?

```
select  
  Category,  
  max (Length_cm) / min (Length_cm)  
from BIOLIFE  
group by Category  
having count (*) > 1;
```

Step 1

Which category with >1 species has the largest ratio between the largest and the smallest length of species?

```
select
  Category,
  max (Length_cm) / min (Length_cm)
  as Ratio
from BIOLIFE
group by Category
having count (*) > 1;
```

Step 2

Which category with >1 species has the largest ratio between the largest and the smallest length of species?

```
select * from (select
  Category,
  max (Length_cm) / min (Length_cm)
  as Ratio
from BIOLIFE
group by Category
having count (*) > 1) as Categories
order by Categories.Ratio desc;
```

Step 3

Questions?