

SQL Queries

Set Operations

Sets and Multi-sets (Bags)

- Sets
 - $A = \{1, 2, 3\}$
 - Only unique elements
 - select city
from Cities_Population
 - Key attributes are sets
- Multi-sets or bags
 - $A' = \{1, 1, 2, 3, 3\}$
 - There are duplicates
 - select county
from Cities_Population
 - Attributes with duplicate values are bags

Operations on Sets and Multi-sets

- Sets

- $A = \{1,2,3\}$, $B = \{1,3,5\}$

- Union

- $A \cup B = \{1,2,3,5\}$

- Intersection

- $A \cap B = \{1,3\}$

- Difference

- $A - B = \{2\}$

- $B - A = \{5\}$

- Multi-sets or bags

- $A' = \{1,1,2,3,3\}$

- $B' = \{1,2,2,3,4\}$

- Union

- $A' \cup B' = \{1,1,1,2,2,2,3,3,3,4\}$

- Intersection

- $A' \cap B' = \{1,2,3\}$

- Difference

- $A' - B' = \{1,3\}$

- $B' - A' = \{2,4\}$

SQL Multi-sets

- SQL works with multi-sets or bags
- SQL does not eliminate duplicates by default
- select county
from Cities_Population
- Transform a multi-set to a set
- select **DISTINCT** county
from Cities_Population
- Do not apply on keys because they are already sets!
- DISTINCT is an expensive operation that can increase query runtime quite significantly

SQL Set Operations

- Set
 - UNION
 - INTERSECT
 - EXCEPT
- A UNION B
is equivalent to
DISTINCT A
UNION ALL
DISTINCT B
- Multi-set
 - UNION ALL
 - Not supported
 - INTERSECT ALL
 - EXCEPT ALL

SQL Set Operations Requirement

- The schemas of the operands have to be exactly the same, including the name and the order of the attributes
- Use renaming with AS on the SELECT

UNION

- select maker
from product
where type = 'pc'
union
select maker
from product
where type = 'laptop'

- select maker
from product
where type = 'pc'
union all
select maker
from product
where type = 'laptop'
- select maker
from product
where type = 'pc' or type = 'laptop'

INTERSECT

- select maker
from product
where type = 'pc'
intersect
select maker
from product
where type = 'laptop'

- **This does not
produce the correct
result anymore!**
 - **select maker
from product
where type = 'pc' and
type = 'laptop'**

EXCEPT

- select maker
from product
where type = 'pc'
except
select maker
from product
where type = 'laptop'

- select maker
from product
where type = 'laptop'
except
select maker
from product
where type = 'pc'
- **Incorrect!**
 - select maker
from product
where type = 'laptop' and type <> 'pc'

Multiple Attributes

```
select model, (speed+ram+hd)/price as score
```

```
from pc
```

```
union all
```

```
select model, (speed+ram+hd+screen)/price as score
```

```
from laptop
```

```
order by score desc
```

Examples

- Computers
- TPCH