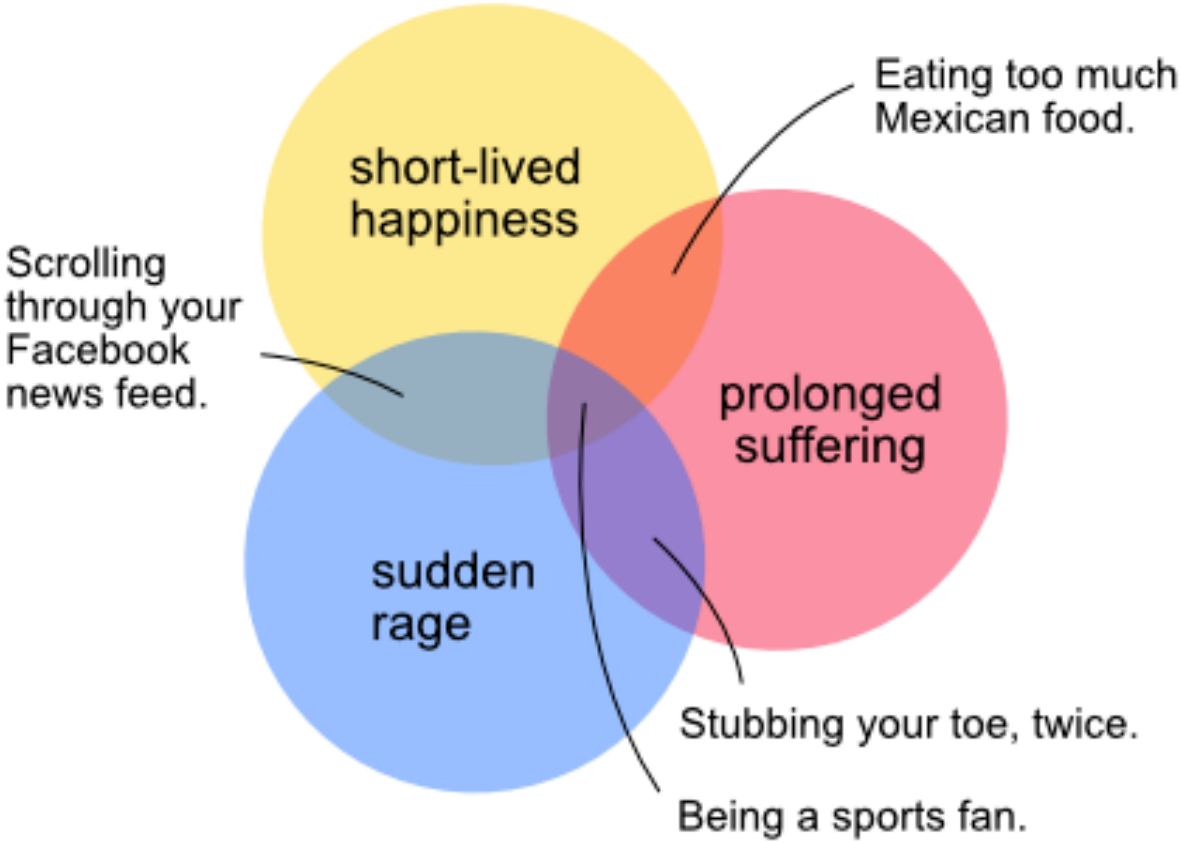
Set theory

How data can be related

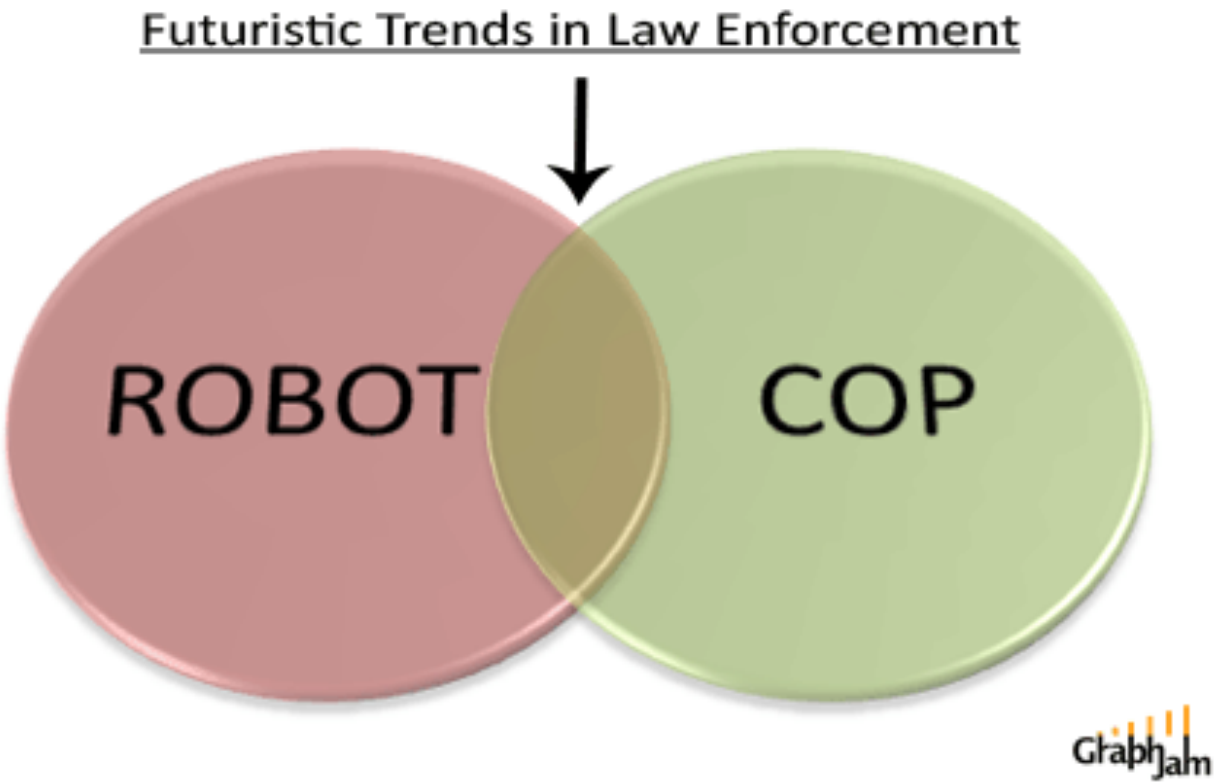
Set theory

- A set is represented by a circle
- Each set represents a group of entities that have a particular attribute.
- Intersecting sets overlap when they have attributes in common.
- Sets that have no commonality do not overlap.

Venn Diagram of Emotions

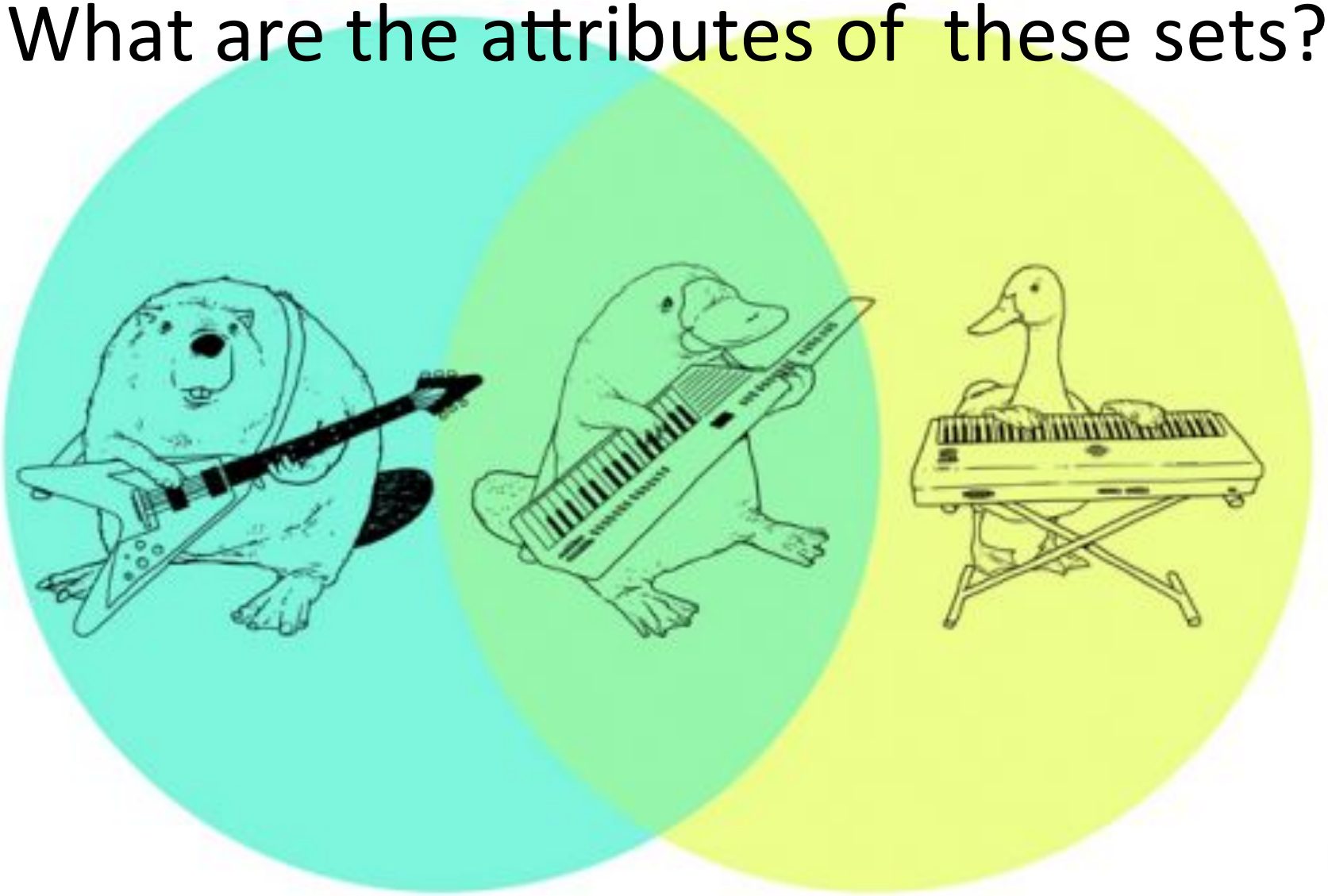


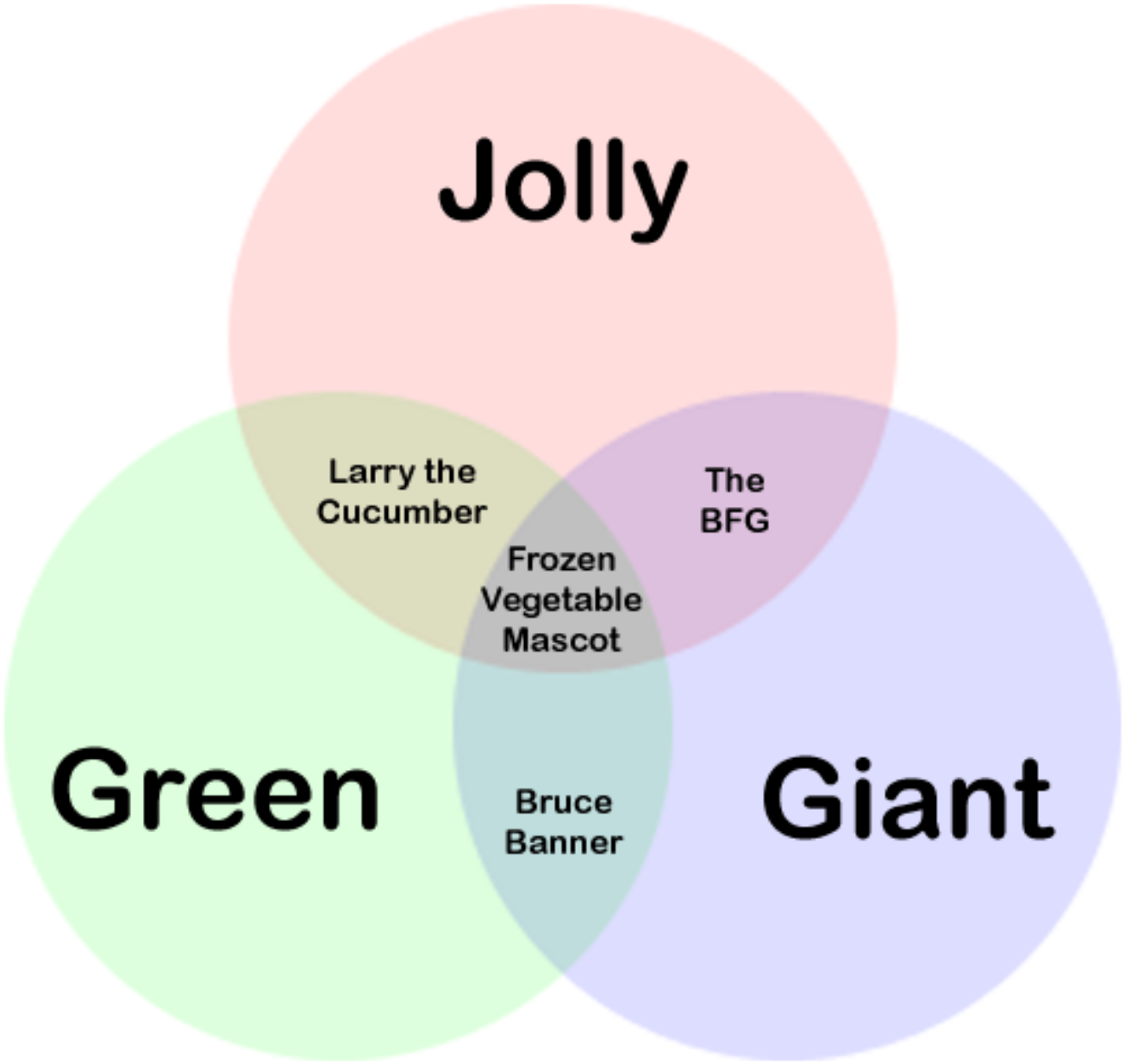
Intersection



http://graphjam.files.wordpress.com/2008/03/robo-cop-venn-diagram-2.gif

What are the attributes of these sets?

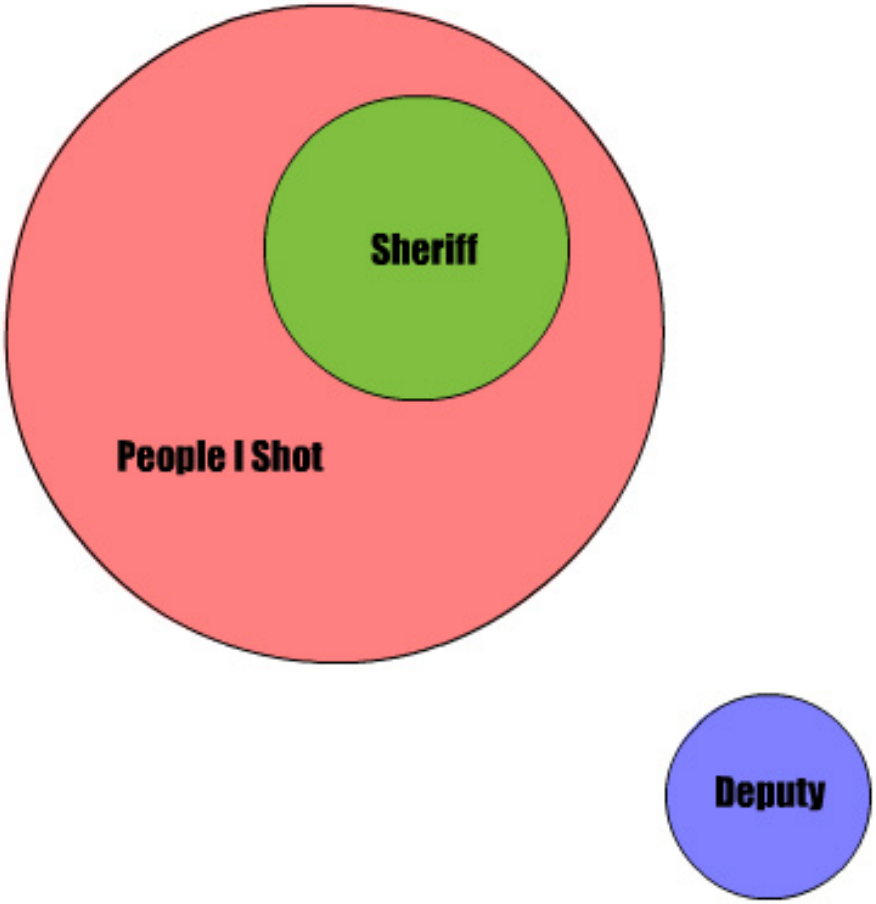



Jolly
Larry the Cucumber
The BFG
Frozen Vegetable Mascot
Green
Bruce Banner
Giant

Non-intersecting sets

- Have no attributes in common.
- Do not overlap.

Summary of Contentions by Bob Marley



Handling sets

- In general, where there are sets of data, some will overlap and some will not.
- It is useful to find:
  - The intersection of the sets
  - The union of sets
  - Differences between sets
  - Subsets
- SQL allows us to manipulate these relationships.

Questions we might ask

- Return the list of cops that aren't robots.
- Return the list of animals that have beaks and play six-stringed instruments.
- Return the list of things that are Jolly and Green, but not Giant.
- Return the list of law enforcement officers that Bob shot.

Assumptions we might make

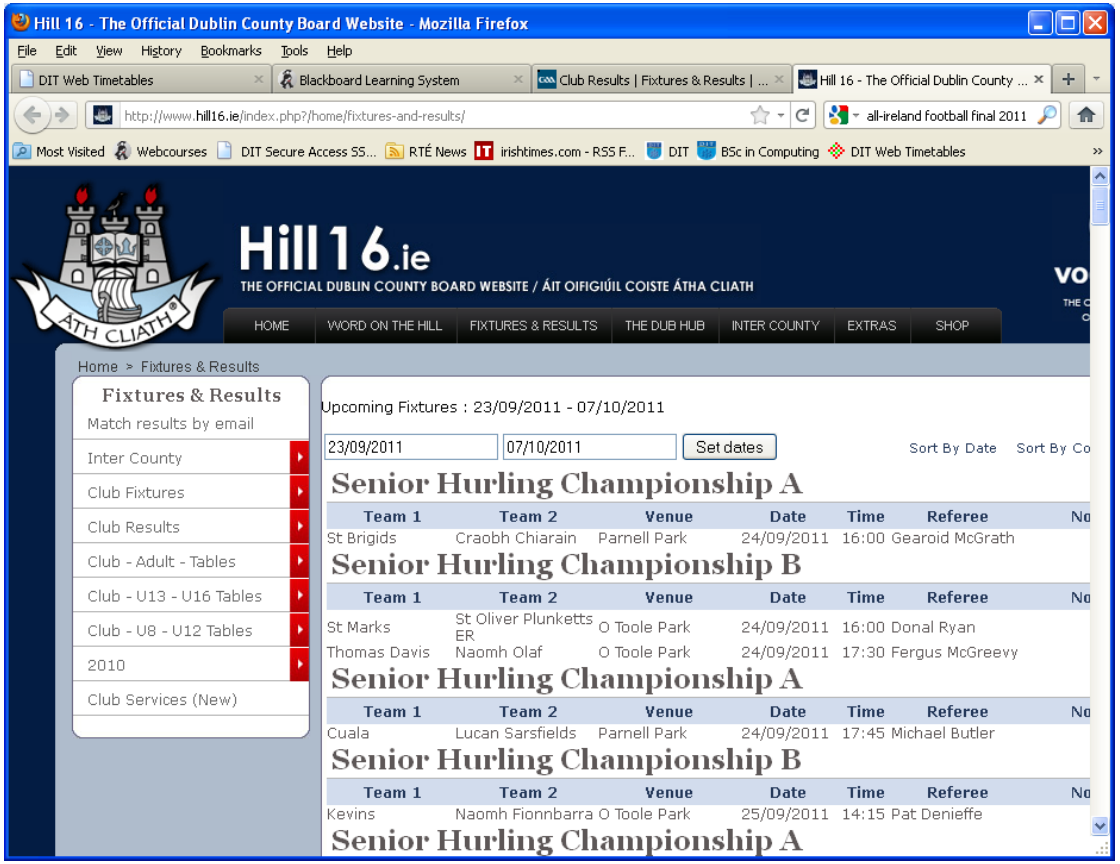
- There are no sheriffs that Bob didn't shoot.
  - There is nothing in the set of sheriffs that is not in the set of people Bob shot.
  - Sheriffs are a subset of the people that Bob shot.

Sets

Although there are a lot of divisions between these groups, there is also a lot they have in common.

They are perfect groups to use for set relations.

Related domains



When do we use SET theory?

- SET theory
  - When the tables have similar columns and constraints.
- JOINs
  - When the tables are related, but NOT necessarily with the same columns and constraints.
  - The tables must have one or more columns with attributes in common.

Converting to SQL

- Using SET theory to formulate queries
- Set theory
- Intersection, Union, Difference

Relational algebra terms

- Projection
- Selection
- Union
- Intersect
- Minus
- Divide

Projection

- This is where we return just some ***columns*** from a table
- i.e., instead of
  - SELECT * FROM STOCK,

we select just some of the fields:

  - SELECT stock_code, stock_description FROM STOCK;

Selection

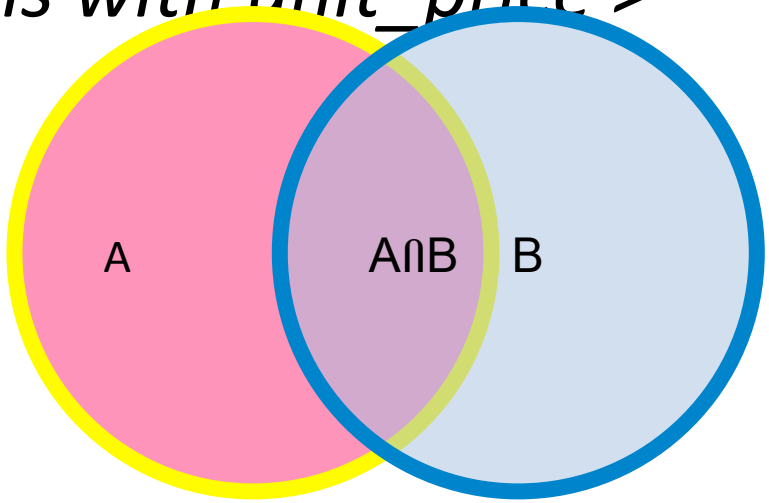
- This is where we select just some ***rows*** from a table, by filtering out the ones we don't want:
  - SELECT * FROM STOCK WHERE stock_code like 'A101';

Sets from the Stock Table

- Consider set A as all stock items supplied by supplierid 501.
  - **`select * from stock where supplier_id = 501;`**
- Consider set B as all stock items that have a unit_price of more than €200.
  - **`select * from stock where unit_price > 200;`**

Intersection in a Venn diagram

- Intersection is where the two sets overlap.
- *If A is the set of stock items supplied by SupplierId 501 and*
- *B is the set of stock items with unit_price > €200,*
- *What's A∩B?*



Using INTERSECT in SQL

- Determine the sets in separate SQL statements.
- Use the word INTERSECT between the two SQL statements.
  - Omit the ; from the first select statement.

```
select * from stock where supplier_id = 501;

select * from stock where unit_price > 200;
```

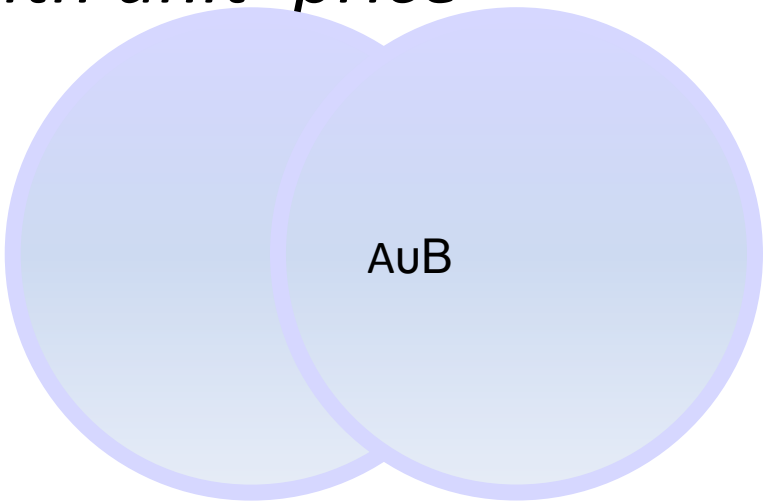
Using INTERSECT in SQL

- Determine the sets in separate SQL statements.
- Use the word INTERSECT between the two SQL statements.
  - Omit the ; from the first select statement.

```
select * from stock where supplier_id = 501
INTERSECT
select * from stock where unit_price > 200;
```

Union in a Venn diagram

- This is what it looked like in a VENN diagram
- *If A is the set of stock items supplied by SupplierId 501 and*
- *B is the set of stock items with unit price > €200,*
- *What's A∪B?*



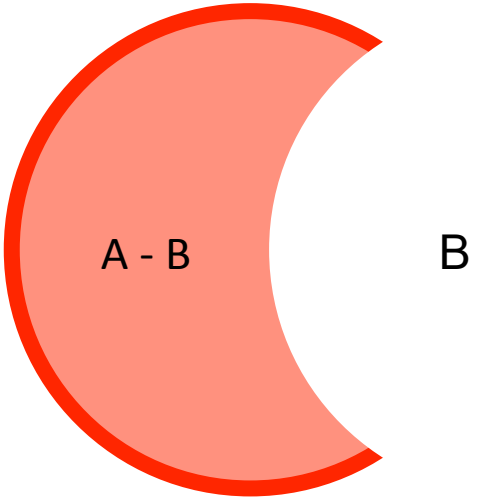
Using UNION in SQL

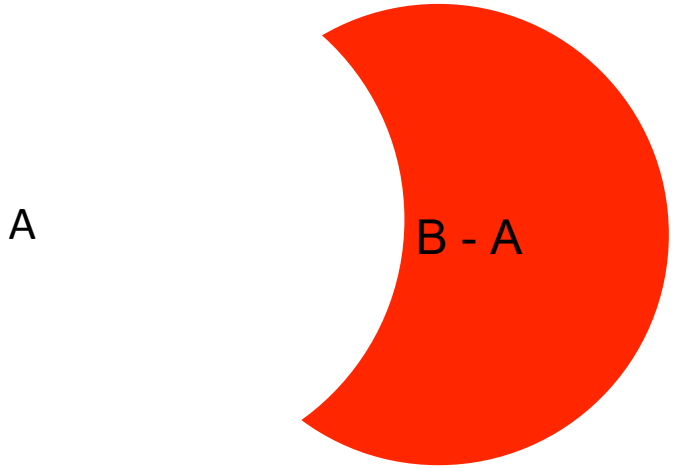
- Determine the sets in separate SQL statements.
- Use the word UNION between the two SQL statements.
  - Omit the ; from the first select statement.

```
select * from stock where supplier_id = 501
UNION
select * from stock where unit_price > 200;
```

Difference Venn diagram

- A – B
  - *What is A-B in our example?*
- B – A
  - *What is B – A in our example?*





A – B in SQL

- Determine the sets in separate SQL statements.
- Use the word MINUS between the two SQL statements.
  - Omit the ; from the first select statement.

```
select * from stock where supplier_id = 501
minus
select * from stock where unit_price > 200;
```

B - A in SQL

- Determine the sets in separate SQL statements.
- Use the word MINUS between the two SQL statements.
  - Omit the ; from the first select statement.

```
select * from stock where unit_price > 200
minus
select * from stock where supplier_id = 501;
```

- Socrative quiz on sets.
- Before question 5 do views.

Divide

- A ÷ B = rows in A that are related to every row in B.
  - This always requires a third table.
- Example:
  - Student that passed all modules.
  - Supplier who supplies all parts.
  - Consumer who eats all types of crisps.
- See later!