

ANALYSIS FRAMEWORK

QUESTION		FUNCTION		ANSWER
What are the data?	→	Storage and retrieval	→	Present existing data
<i>What is currently stored in the database?</i> e.g. where are all the houses in the Toronto data set?				
What is the <u>pattern</u> in the data?	→	Constrained query	→	Find new patterns in existing data
<i>Search current data for entities with a certain characteristic</i> e.g. where are all the houses in the floodplain?				
What <u>could</u> the data be?	→	Modeling	→	Predicting new information
<i>Use a predictive model</i> e.g. if the river level rises 4 m, which houses will be flooded?				

APPROACHING A PROBLEM

- This is just a framework to help organize a systematic approach to a problem
- The trick is to know:
 - What kind of question you're asking
 - Which functions will help you answer the question
 - What the answer should look like

FUNCTIONS AS TOOLS

- There are a wide variety of tools that may be available for a particular problem
- Individual analysis functions must be used within an organized and complete analysis strategy
- Think of each function as a tool in a toolbox



QUERIES, SEARCHES AND ATTRIBUTE OPERATIONS

CONSTRAINED QUERIES

Mathematical Operators

+	addition
-	subtraction
*	multiplication
/	division
^	exponent

Relational Operators

<	less than
>	greater than
=	equals
<=	less than or equal to
>=	greater than or equal to
<>	not equal to

Boolean Operators

AND	<u>Intersection</u> of two or more sets
OR	<u>Union</u> of two or more sets
NOT	<u>Subtraction</u> of sets
XOR	<u>Combination</u> of sets (exclusive OR: A or B, but not both)

Spatial Operators

Are completely within...
 Completely contain...
 Have their center in...
 Contain the center of...
 Intersect...
 Are within distance of...

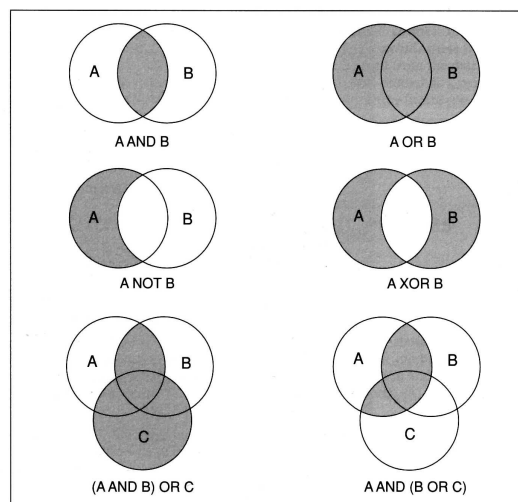


Figure 7.2. Venn diagrams showing the results of applying Boolean logic to the union and intersection of two or more sets. In each case the shaded zones are 'true'

FIELD TYPES

Short Integer:	No decimal, -32768 to 32,767
Long Integer:	No decimal, -2,147,483,648 to 2,147,483,647
Float:	Up to 7 significant digits
Double:	Up to 15 significant digits
Text:	Number of characters
Date:	no parameter to set
BLOB:	Binary large object (e.g. image data); no parameter to set