




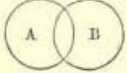
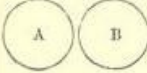
Introduction to data science & artificial intelligence (IF7100)

Arthur Charpentier

#412 Sets

été 2020

Sets

(i) Diagrammatic	(ii) Common Logic	(iii) Quantified	(iv) Symbolic
	$\left. \begin{array}{l} \text{All } A \text{ is } B \\ \text{All } B \text{ is } A \end{array} \right\}$	All A is all B	$\left. \begin{array}{l} AB = 0 \\ \bar{A}\bar{B} = 0 \end{array} \right\}$
	$\left. \begin{array}{l} \text{All } A \text{ is } B \\ \text{Some } B \text{ is not } A \end{array} \right\}$	All A is some B	$\left. \begin{array}{l} AB = 0 \\ \bar{A}B = v \end{array} \right\}$
	$\left. \begin{array}{l} \text{All } B \text{ is } A \\ \text{Some } A \text{ is not } B \end{array} \right\}$	Some A is all B	$\left. \begin{array}{l} \bar{A}B = 0 \\ A\bar{B} = v \end{array} \right\}$
	$\left. \begin{array}{l} \text{Some } A \text{ is } B \\ \text{Some } A \text{ is not } B \\ \text{Some } B \text{ is not } A \end{array} \right\}$	Some A is some B	$\left. \begin{array}{l} AB = v \\ A\bar{B} = v \\ \bar{A}B = v \end{array} \right\}$
	No A is B	No A is any B	$AB = 0$

John Venn, *Symbolic Logic*, 1881.

Intersection & Union

the **intersection** of two sets A and B , denoted by $A \cap B$, is the set containing all elements of A that also belong to B

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

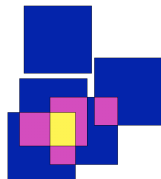
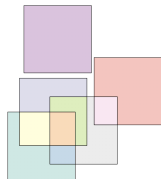
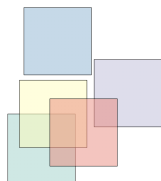
E.g. $\{1, 2, 3\} \cap \{2, 3, 4\} = \{2, 3\}$

the **union** of two sets A and B , denoted by $A \cup B$, is the set containing all elements of A and B

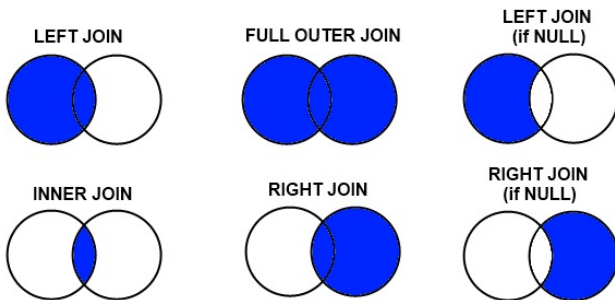
$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

E.g. $\{1, 2, 3\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\}$

- ▶ important in probabilities
- ▶ important when merging datasets



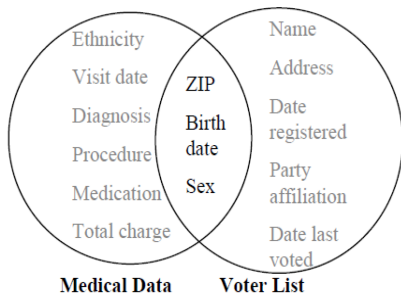
Merging Datasets



via [Merge and Join DataFrames with Pandas](#)

k-anonymity

See Protecting privacy when disclosing information: *k*-anonymity and its enforcement through generalization and suppression



via Privacy Preserving Data Publishing with Multiple Sensitive Attributes based on Overlapped Slicing

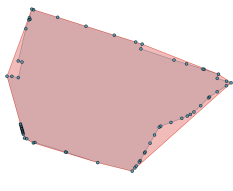
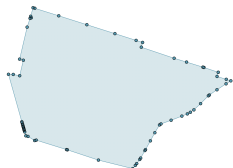
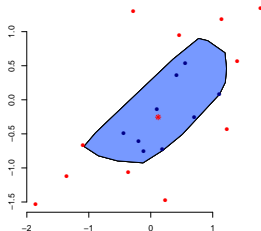
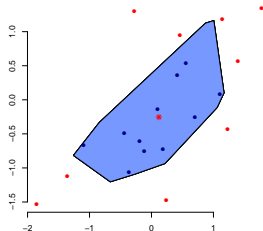
Convexity

A set is **convex** if, with any two points, it contains the whole line segment that joins them,

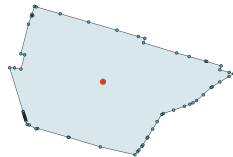
$$(1 - t)\mathbf{x} + t\mathbf{y} \in C,$$

for all \mathbf{x} and $\mathbf{y} \in C$, and $t \in (0, 1)$.

The **convex hull** of a shape is the smallest convex set that contains it



Centroid



the **centroid** (geometric center) of a plane figure is the arithmetic mean position of all the points in the figure

