SYMBOLIC DATA MINING

- **Frequent**: If supp(X) ≥ min-supp, then X is frequent
- **Closed-Itemset**: An itemset, doesn't have a <u>superset</u> with same supp
- **Generator**: An itemset, doesn't have a <u>subset</u> with same supp

	A	B	C	D	E
1	X	X		X	X
2	X		X		
3	X	X	X		X
4		X	X		X
5	X	X	X		X

Itemset: A, AB, BCE...

Support(A) = 4

Support(AB) = 3

Support(BCE) = 3

Min-Support: Threshold

<u>One Size Larger Superset:</u>

| SAME | X |

| SAME | Y | +

⇓

| SAME | X | Y |

APRIORI (breadth-first)(levelwise)

	A	B	C	D	E
1	X	X		X	X
2	X		X		
3	X	X	X		X
4		X	X		X
5	X	X	X		X

min-supp = 2

FC_1

A	4
B	4
C	4
~~D~~	~~1~~
E	4

→ F_1

A	4
B	4
C	4
E	4

FC_2

AB	3
AC	3
AE	3
BC	3
BE	4
CE	3

→ F_2

AB	3
AC	3
AE	3
BC	3
BE	4
CE	3

FC_3

ABC	2
ABE	3
ACE	2
BCE	3

→ F_3

ABC	2
ABE	3
ACE	2
BCE	3

FC_4

ABCE	2

→ F_4

ABCE	2

FC_5

∅ stop!

$|F|_5 = 15$

APRIORI-CLOSE (breadth-first) (levelwise)

	A	B	C	D	E
1	X	X		X	X
2	X		X		
3	X	X	X		X
4		X	X		X
5	X	X	X		X

min-supp = 3

FC_1

A	4
B	4
C	4
D	1
E	4

F_1

itemset	support	closed?
A	4	yes
B	4	~~yes~~ no
C	4	yes
E	4	~~yes~~ no

FC_2

A B	3
A C	3
A E	3
B C	3
B E	4
C E	3

F_2

itemset	support	closed?
AB	3	~~yes~~ no
AC	3	yes
AE	3	~~yes~~ no
BC	3	~~yes~~ no
BE	4	yes
CE	3	~~yes~~ no

FC_3

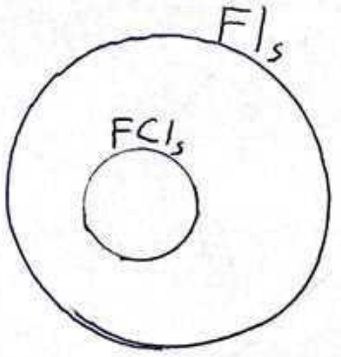
~~ABC~~	~~2~~
ABE	3
~~ACE~~	~~2~~
BCE	3

F_3

itemset	support	closed?
ABE	3	yes
BCE	3	yes

FC_4

$\emptyset$ stop!



$FI_5 = 12$

$FCI_5 = 6$

ECLAT (vertical) (depth-first)

1 → ABDE
2 → AC
3 → ABCE
4 → BCE
5 → ABCE

A	B	C	D	E
1	1	2	1	1
2	3	3		3
3	4	4		4
5	5	5		5

	A	B	C	D	E
1	X	X		X	X
2	X		X		
3	X	X	X		X
4		X	X		X
5	X	X	X		X

min-supp=3