# 5.4. Mining Diverse Patterns: Mining Negative Correlations
## In-Lecture Question

# Question

Consider the support-base and null-invariant definitions for negative patterns. Let $\epsilon$ = 0.01. In a database containing $10^9$ total transactions, $10^6$ transactions contain item A and $10^4$ transactions contain item B. $10^2$ transactions contain both A and B. Which of the following is true?

- ❑ {A, B} is a negative pattern only by the support-based definition.
- ❑ {A, B} is a negative pattern only by the null-invariant definition.
- ❑ {A, B} is a negative pattern by both definitions.
- ❑ {A, B} is a negative pattern by neither definition.
- ❑ More information is needed to determine whether {A, B} is a negative pattern.

# Question

Consider the support-base and null-invariant definitions for negative patterns. Let $\epsilon = 0.01$. In a database containing $10^9$ total transactions, $10^6$ transactions contain item A and $10^4$ transactions contain item B. $10^2$ transactions contain both A and B. Which of the following is true?

- ❑ {A, B} is a negative pattern only by the support-based definition.

- ❑ {A, B} is a negative pattern only by the null-invariant definition.

- ❑ {A, B} is a negative pattern by both definitions.

- ❑ {A, B} is a negative pattern by neither definition.

- ❑ More information is needed to determine whether {A, B} is a negative pattern.

- ❑ **Answer**: {A, B} is a negative pattern only by the null-invariant definition.

- ❑ **Explanation**: $sup(A) = 10^{-3}$, $sup(B) = 10^{-5}$, $sup(A \cup B) = 10^{-7} > sup(A) \times sup(B) = 10^{-8}$
  => {A, B} is not a negative pattern by the support-based definition.
  $P(B|A) = P(A \cup B)/P(A) = 10^2/10^6 = 10^{-4}$, $P(A|B) = P(A \cup B)/P(B) = 10^2/10^4 = 10^{-2}$
  $(P(B|A) + P(A|B))/2 = (10^{-4} + 10^{-2})/2 = 0.00505 < 0.01$
  => {A, B} is a negative pattern by the null-invariant definition.