

Практична робота

Статистична обробка медичних даних

Мета роботи – навчитися проводити автоматизовану статистичну обробку масивів даних.

Завдання

1. Провести статистичну обробку даних завантаженого на минулій практиці датасету.
2. Визначити:
 - 2.1 Осяг вибірки ($=\text{СЧЁТ}(\text{С649}:\text{С668})$).
 - 2.2 Мінімальне ($=\text{МИН}(\text{С649}:\text{С668})$) та максимальне ($=\text{МАКС}(\text{С649}:\text{С668})$) значення .
 - 2.3 Розмах – різниця між мінімальним та максимальним значенням.
 - 2.3 Середні значення за параметрами:

Середнє арифметичне ($=\text{СРЗНАЧ}(\text{С649}:\text{С668})$)– це сума всіх значень за параметром, поділена на кількість екземплярів.

Середнє геометричне ($=\text{СРГЕОМ}(\text{С649}:\text{С668})$) – це середнє значення за добутком значень параметра за примірниками.

Середнє гармонійне ($=\text{СРГАРМ}(\text{С649}:\text{С668})$) розраховується як окрема кількість екземплярів на суму зворотних значень параметра для кожного екземпляра.
 - 2.4 Мода ($=\text{МОДА}(\text{С649}:\text{С668})$) – це число, яке має максимальну частоту у всьому наборі даних, або, іншими словами, мода – це число, яке зустрічається максимальну кількість разів. Дані можуть мати одну або кілька мод.
 - 2.5 Медіана ($=\text{МЕДИАНА}(\text{С649}:\text{С668})$) – це точка, яка поділяє всі дані на дві рівні половини. Одна половина даних менше медіани, а інша – більше. Медіана розраховується шляхом розташування даних у порядку неспадання. Якщо кількість спостережень непарна, медіана визначається за середнім спостереженням у відсортованому вигляді. Якщо кількість спостережень парна,

медіана визначається середнім значенням між двома середніми спостереженнями у відсортованому вигляді. Медіана – це центральна точка або 50-й перцентиль даних.

2.6 Дисперсія (=ДИСП(C649:C668)) – показує, наскільки віддалене значення від середнього. Високе значення дисперсії означає, що значення розподілені далеко від середнього і дані значно варіюються. Відповідно низька дисперсія означає низьку варіативність.

2.7 Середньоквадратичне відхилення (=СТАНДОТКЛОН(C649:C668)) визначає частку дисперсії даних. Воно відповідає тій частині даних, які не збігаються із середнім значенням. Визначається як корінь квадратний від дисперсії.

3. Зробити висновки по роботі.