

Тема. Підготовка наборів даних для перевірки твердження/гіпотези Аналіз даних та представлення результату дослідження

Після цього заняття потрібно вміти:

- пояснювати призначення функцій і засобів табличного процесора для опрацювання наборів однотипних об'єктів;
- пояснювати значення та виконувати обчислення основних статистичних характеристик вибірки;

Повторюємо

Пригадайте алгоритм для:

- Побудови графіка функції за допомогою табличного процесора
- Створення проміжних підсумків за даними умовами

Опрацюйте інформацію

Статистика – це наука, що властивими їй методами вивчає кількісну сторону масових суспільних явищ і процесів у нерозривному зв'язку з їх якісною стороною та дає числові вираження закономірностей суспільного розвитку в конкретних умовах місця і часу.

Так, наприклад, для аналізу тенденцій змінення маси учнів 10-х класів України за останні роки, ризику серцевих захворювань людей певного віку на планеті, популярності продуктів харчування серед населення певного регіону потрібно проаналізувати сотні тисяч або навіть мільйони даних.

Зрозуміло, що провести зважування, вивчити історії хвороб, провести анкетування сотень тисяч або навіть мільйонів людей практично неможливо. Тому для аналізу створюють певну **вибірку об'єктів дослідження**, тобто з усієї множини об'єктів дослідження відбирають певну кількість і на ній проводять дослідження.

Дані для дослідження називають **статистичними даними (вибіркою)**

Властивості вибірки

- **Масовість** (що більше така вибірка, то точніше буде проведено аналіз і зроблено відповідні висновки).
- **Репрезентативність** (франц *representatif* — показовий, характерний, типовий)

Дані, отримані з дослідженої вибірки, найчастіше заносять у таблицю. Така форма подання даних з вибірки зручна для їх аналізу та прогнозів. Дані з кожного рядка і стовпця таблиці вибірки утворюють **ряди даних**.

Розглянемо **приклад**.

Команда учнівства України бере участь у міжнародних олімпіадах з інформатики починаючи з 1992 року. У таблиці подано результати її виступів з 2005 по 2017 рік. Тут **вибіркою** є вказані в таблиці **роки**, а **рядами даних** — **загальна кількість медалей** у ці роки, а також **кількість золотих, срібних і бронзових медалей** у вказані роки

**Результати виступу команди учнівства України
на міжнародних олімпіадах з інформатики за 2005–2017 роки**

| <i>Рік</i> | <i>Кількість медалей</i> | <i>Золоті</i> | <i>Срібні</i> | <i>Бронзові</i> |
|------------|--------------------------|---------------|---------------|-----------------|
| 2005 | 4 | 2 | 1 | 1 |
| 2006 | 4 | 1 | 2 | 1 |
| 2007 | 4 | 1 | 2 | 1 |
| 2008 | 3 | 0 | 1 | 2 |
| 2009 | 4 | 1 | 1 | 2 |
| 2010 | 3 | 0 | 1 | 2 |
| 2011 | 3 | 0 | 1 | 2 |
| 2012 | 4 | 1 | 1 | 2 |
| 2013 | 4 | 0 | 1 | 3 |
| 2014 | 3 | 0 | 1 | 2 |
| 2015 | 4 | 0 | 3 | 1 |
| 2017 | 3 | 1 | 2 | 0 |

Статистичні характеристики ряду даних:

| № з\п | Характеристика | Означення | Функція MS Excel |
|-------|------------------------------|--|--|
| 1. | <i>Середнє арифметичне</i> | це частка від ділення суми цих чисел на їх кількість (обсяг вибірки). | <i>AVERAGE</i> <i>СРЗНАЧ</i> |
| 2. | <i>Обсяг вибірки</i> | кількість елементів у вибірці. | <i>COUNT</i> |
| 3. | <i>Стандартне відхилення</i> | найпоширеніший показник розсіювання статистичних даних відносно середнього значення вибірки | <i>STDEV.P</i> <i>СТАНДОТКЛОН.В</i> |
| 4. | <i>Мода</i> | число, яке найчастіше зустрічається в ряді даних. | <i>MODE.SNGL</i> <i>МОДА.НСК</i> |
| 5. | <i>Медіана</i> | це величина, що розташована в середині ряду величин, розташованих у зростаючому або спадному порядку | <i>MEDIAN</i> |

Робота за комп'ютером

Завдання 1

Є ряд даних: 56,5; 27,9; 32,4; 59,23; 44,87; 41; 39,7; 49,98. Обчисліть засобами табличного процесора середнє арифметичне, стандартне відхилення, моду й медіану для цього ряду даних.

Завдання 2

Завантажте [електронну таблицю](#) . Виконайте завдання в даній таблиці згідно [відео](#) або за планом:

- 1) Розділіть дані за стовпцями
- 2) Побудуйте графік температур за місяцями
- 3) Сформуйте стовпець з номером року (функція YEAR)
- 4) Створіть проміжні підсумки за стовпцем року з обчисленням середньої температури за рік
- 5) Побудуйте графік температур за роками
- 6) Сформуйте стовпець з номером десятиліття (формула на основі функції LEFT, щоб отримати 3 країни ліві символи з номера року =LEFT(B2;3)

- 7) Створіть проміжні підсумки за стовпцем десятиліття з обчисленням середньої температури за десятиліття
- 8) Побудуйте графік температур за десятиліттями
- 9) Порівняйте графіки і зробіть висновок

Зверніть увагу

Роботу бажано виконувати відразу в Google таблицях або після виконання у MS Office розмістити на своєму Google диску. Посилання (доступ) на виконані роботи надати вчителю на HUMAN або на електронну пошту nataliartemiuk.55@gmail.com

Джерело

[Дистосвіта](#)