

Урок-БІТ «Основи статистичного аналізу даних. Ряди даних. Обчислення основних статистичних характеристик вибірки»

Мета:

➤ **навчальна:** сформувати поняття основ статистичного аналізу даних, рядів даних, обчислення основних статистичних характеристик вибірки;

➤ **розвивальна:** розвивати логічне мислення, пам'ять; формувати вміння узагальнювати;

➤ **виховна:** виховувати інформаційну культуру, формувати бережливе ставлення до обладнання комп'ютерного кабінету, виховувати вміння працювати в групі; формувати позитивне ставлення до навчання.

Обладнання та наочність: дошка, комп'ютери з підключенням до мережі Інтернет, підручник, навчальна презентація.

Програмне забезпечення: офісні програми.

Сценарій заходу

БІТ-урок включає три взаємопов'язаних елементи: бесіду, гру, творчість.

Для її успішного проведення першої частини уроку - бесіди учитель ставить перед учнями основну мету уроку. На робочому місці учні знаходять опорний конспект і конверт з роздатковим матеріалом для уроку. Колективна робота створює ту творчу лабораторію, в якій кожний учень виступає як дослідник, шукає і знаходить відповіді, запрошуючи учителя як консультанта.

За допомогою гри-естафети закріплюються отримані знання під час бесіди. Учні об'єднуються у команди. Кожна команда отримує естафетні картки, які починають заповнювати учні із останнього ряду. Кожний із учнів вписує одну назву і передає картку далі. Виграє та команда, яка заповнила картку без помилок і швидше за інших.

На третьому етапі учням пропонують продемонструвати вміння оригінально застосовувати одержані знання. Розв'язання поставленого завдання вимагає від учнів творчого підходу, знання основ предмета.

Для забезпечення швидкого та ефективного включення учнів в інтерактивну діяльність доцільно давати їм пам'ятки, що містять опис алгоритму діяльності у вигляді роздаткового матеріалу, плакатів або будь-якої наочності, використовуваної за допомогою технічних засобів навчання.

Хід уроку

I. Організаційний етап

- привітання
- перевірка присутніх
- перевірка готовності учнів до уроку

II. Актуалізація опорних знань. Мотивацій навчальної діяльності у вигляді бесіди.

Уч. – Практично всім відоме слово «статистика» та з ним в кожній людині виникають свої асоціації.

Подумайте, для чого потрібна статистика? Яке місце займає статистика в сучасному світі?

Учні відповідають: для збирання інформації, для аналізу даних, без статистичних даних неможливе управління державою, розвиток економіки.

Уч. - Статистика – це тільки чорно-білі цифри, які подають у бюлетенях, збірниках чи щорічниках. Усе значно цікавіше! Кожен у цьому світі долучається до створення статистики і ти – також.

Сьогодні ми з вами спробуємо відкрити завісу над таємницею статистичного аналізу даних.

III. Вивчення нового матеріалу

Пояснення вчителя з елементами демонстрування презентації

Протягом уроку учні доповнюють опорний конспект, які вчитель роздає на початку уроку. (див додаток 1)

Ви знаєте, що методи отримання, опрацювання й аналізу даних, які характеризують масові явища, вивчає наука **статистика** (лат. *status* — стан). Отже, темою нашого уроку є «Основи статистичного аналізу

даних. Ряди даних. Обчислення основних статистичних характеристик вибірки»



ОПОРНИЙ КОНСПЕКТ

Тема уроку: _____

наука про методи збирання, опрацювання, аналізу та інтерпретації даних, що характеризують масові явища та процеси



Рис. 1. Учні записують тему уроку та назву науки

Розрізняють *три етапи аналізу даних*:

1. Збір і групування первинних статистичних даних, отриманих шляхом спостережень або вимірювань;
2. Опрацювання отриманих даних;
3. Отримання висновків за даними спостережень.



Рис. 2. Учні записують етапи аналізу даних

Нехай необхідно проаналізувати тенденції змінення маси учнів 10-х класів України за останні роки, ризик серцевих захворювань людей певного віку на планеті, популярність продуктів харчування серед населення певного регіону потрібно проаналізувати сотні тисяч або навіть мільйони даних.

Тому для аналізу створюють певну **вибірку об'єктів дослідження**, тобто з усієї множини об'єктів дослідження відбирають певну кількість і на ній проводять дослідження. Чим більше така вибірка, тим точніше буде проведено аналіз і зроблено відповідні висновки. Тобто вибірка повинна бути масовою.

Але не тільки кількість даних у вибірці визначає рівень точності аналізу і висновків. Так, у першому і другому з наведених вище прикладів доцільно вибирати людей різних регіонів і різної статі, а у третьому — людей різного віку. Кажуть, що вибірка даних має бути **репрезентативною** (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 такої таблиці утворюють

Багаторазове вимірювання та збереження при цьому відповідних результатів приводить до накопичення даних, які опрацьовують спеціальними методами, які вивчаються у статистиці. Такі дані називаються **статистичними даними**.

Статистичні дані — сукупність упорядкованих, класифікованих даних про деяке масове явище або процес.

Статистичні дані дають змогу не тільки охопити картину певного питання на даний час, а й планувати необхідні дії на майбутнє.

Велику множину об'єктів, що є предметом статистичного дослідження, називають генеральною сукупністю.

Дослідник, як правило, не має змоги оперувати всією генеральною сукупністю.

Наприклад, опитати кожного громадянина країни нереально. Натомість досліджують вибірку — деяку множину об'єктів, вибраних з генеральної сукупності, і, проаналізувавши її, роблять висновки щодо властивостей генеральної сукупності загалом.

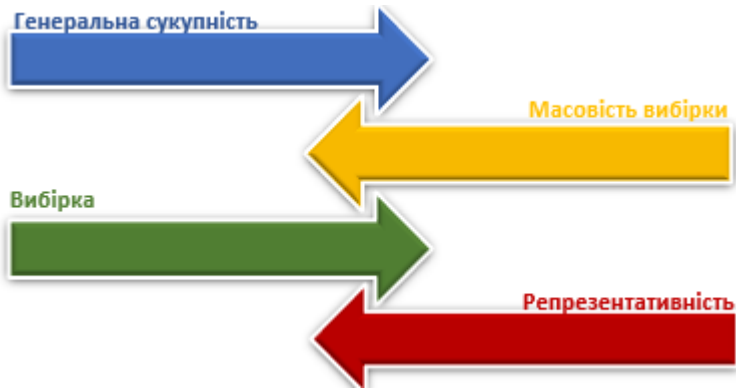


Рис. 3. Учні записують пояснення для основних понять статистики

Маючи в розпорядженні дані статистичного спостереження, що характеризують те чи інше явище, перш за все необхідно їх упорядкувати, тобто надати характер системності.

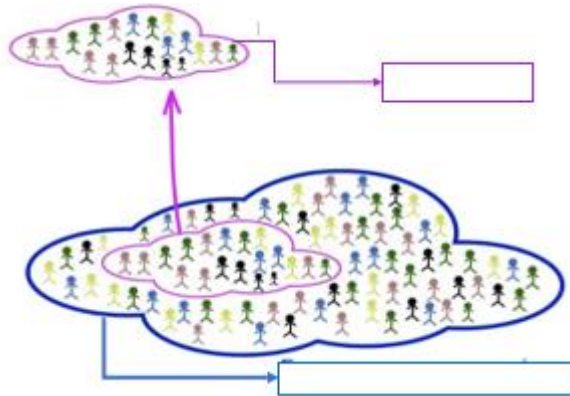


Рис.4. Вибірка та генеральна сукупність

Статистичний ряд розподілу — це впорядковані статистичні дані. Ряди розподілу одиниць сукупності за ознаками, що мають кількісний вираз, називаються *варіаційними рядами*. У таких рядах значення ознаки (варіанти) розташовані в порядку зростання або спадання.

Варіанта — це окреме значення груповальної ознаки; **частота** — число, яке показує, скільки разів трапляється кожна варіанта.

Дискретні варіаційні ряди — це такі ряди розподілу, в яких варіанта як величина кількісної ознаки може набувати тільки певного значення.

Інтервальні варіаційні ряди — такі ряди розподілу, в яких значення варіанти дано у вигляді інтервалів, тобто значення ознак можуть відрізнятися одне від одного на як завгодно малу величину.

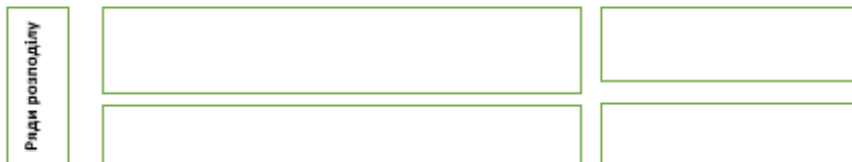


Рис. 5. Варіаційні ряди та їх характеристики

Основними статистичними характеристиками вибірки є **середнє, мода та медіана**, які ще називають мірами центральної тенденції.

Медіану можна визначити як точку на ряді розподілу (впорядкований набір значень змінної для різних спостережень).

Мода — це значення, яке найчастіше трапляється.

Стандартне відхилення - характеризує, наскільки широко розташовані значення ряду даних відносно їх середнього арифметичного

ОСНОВНІ СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ ВИБІРКИ

ОБСЯГ ВИБІРКИ	
РОЗМАХ ВИБІРКИ	
СЕРЕДНЄ ЗНАЧЕННЯ	
СТАНДАРТНЕ ВІДХИЛЕННЯ	
МОДА ВИМІРЮВАННЯ	
МЕДІАНА ВИБІРКИ	

Рис. 6. Таблиця з основними статистичними характеристиками

У табличному процесорі центральна тенденція представлена функціями з категорії **Статистичні**, та майже всі вони призначені для обчислення узагальнювальних статистичних характеристик вибірки.

Найважливішими функціями категорії Статистичні є:

- AVERAGE (число1; число2;...)— обчислює середнє значення;
- MAX (число1; число2;...) — обчислює максимальнє значення;
- MIN (число1; число2;...) — обчислює мінімальнє значення;
- MEDIAN (число1; число2;...) — повертає медіану;
- MODE (число1; число2;...) — повертає моду;
- RANK (число; посилання; порядок) (рос. РАНГ) — повертає ранг числа у списку чисел, тобто його номер у впорядкованій послідовності чисел із вказаного діапазону.

Якщо порядок=0, то ранг 1 має найбільше число, якщо порядок дорівнює будь—якому числу або 1, то ранг 1 має найменше число.

ОСНОВНІ СТАТИСТИЧНІ ФУНКЦІЇ:

Назва функції (англ.)	Призначення
AVERAGE{аргумент1;[аргумент2];...}	
MIN{аргумент1;[аргумент2];...}	
MAX{аргумент1;[аргумент2];...}	
MEDIAN{аргумент1;[аргумент2];...}	
MODE.SNGL{аргумент1;[аргумент2];...}	
STDEV.S{аргумент1;[аргумент2];...}	
RANK (аргумент; посилання; порядок)	

Рис. 7. Таблиця з основними статистичними функціями

IV. Засвоєння нових знань, формування вмінь

Кожна команда, яка об'єднує учнів на ряду, отримує естафетні картки з завданнями. Спочатку виконують завдання ті учні, які сидять на останніх партах та передають вперед.

В картці необхідно ознайомитись з вибіркою та визначити основні статистичні характеристики:

1. Обсяг вибірки
2. Розмах вибірки
3. Середнє арифметичне
4. Мода
5. Медіана

Завдання 1. Розглянемо зріст (у сантиметрах) учнів вашої паралелі класів на прикладі довільно вибраних 10 учнів. Впорядкуємо за зростанням:

168 169 172 172 172 174 175 178 178 182

Обсяг вибірки: 10 (вимірювали зріст 10 учнів).

Розмах вибірки: максимальне значення — 182, мінімальне значення — 168. Різниця $182 - 168 = 14$ елементів вибірки.

Середнє арифметичне ряду чисел: 174 см.

Мода = 172, адже зріст 172 зустрічається частіше за інші значення — 3 рази.

Медіана = 173: у нашому прикладі 10 елементів — число парне, в упорядкованому наборі цих елементів посередині, тобто п'ятий і шостий елементи мають значення 172 і 174. Середнє арифметичне цих чисел: $(172 + 174) : 2 = 173$. Якби досліджували зріст 9 учнів, посередині був би п'ятий елемент зі значенням 172 — медіана мала б саме таке значення.

Завдання 2. Завдання з ЗНО.

Упродовж чверті учень отримав 12 оцінок з алгебри. Інформацію про отримані оцінки та їхню кількість відображено у таблиці.

Оцінка	5	6	7	8	10
Кількість оцінок	2	1	3	5	1

Обсяг вибірки: 12

Розмах вибірки: 5

Середнє арифметичне: 7,25

Мода: 8

Медіана: 7,5

Впорядкуємо ряд: 5, 5, 6, 7, 7, 7, 8, 8, 8, 8, 8, 10.

Завдання 3.

Лучник здійснив 11 пострілів по мішені і набрав відповідно 8, 5, 7, 9, 6, 9, 10, 6, 9, 9, 10 очок.

Обсяг вибірки: 11

Розмах вибірки: 5

Середнє арифметичне: 8

Мода: 9

Медіана: 9

Впорядкуємо ряд: 5, 6, 6, 7, 8, 9, 9, 9, 9, 10, 10

V. Узагальнення і систематизація знань, застосування їх у різних ситуаціях

Робота за комп'ютером

- 1) Повторення правил безпечної поведінки за комп'ютером.
- 2) Інструктаж учителя.
- 3) Практична робота за комп'ютерами.

Завдання: обчислити основні статистичні характеристики засобами табличного процесора.

Учні працюють за комп'ютерами, використовуючи пам'ятки, в яких описується алгоритм виконання завдання.

Завдання 1. Витрати сімейного бюджету

Протягом деякого року сім'я витратила на опалення будинку та користування електроенергією кошти, які подано у статистичному ряді таблицею 1. Знайдіть основні статистичні характеристики ряду даних.

Таблиця 1

Місяць	01	02	03	04	05	06	07	08	09	10	11	12
Витрати	990	670	420	220	190	200	190	80	200	200	400	500

1. Обчислимо середнє арифметичне:

$$\frac{990 + 670 + 420 + 220 + 190 + 200 + 190 + 80 + 200 + 200 + 400 + 500}{12} = 355.$$

Перевіримо отримане значення за допомогою табличного процесора.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Місяць	1	2	3	4	5	6	7	8	9	10	11	12	Середнє
2	Витрати	990	670	420	220	190	200	190	80	200	200	400	500	355

Отримаємо середнє значення 355.

2. Для знаходження медіани виконаємо упорядкування даних за зростанням. Для цього виділіть діапазон $B1:M2$, на вкладці *Основне*



у списку оберіть *Налаштовуване сортування*. У вікні *Сортування* оберіть *Параметри* й у вікні *Параметри сортування* вкажіть *стовпці діапазону*. Натисніть *ОК*.

3. Зазначте 2 рядок для сортування. Перевірте, чи отримали такий вигляд таблиці (мал. 10.7):

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Місяць	8	5	7	6	9	10	4	11	3	12	2	1	Середнє
2	Витрати	80	190	190	200	200	200	220	400	420	500	670	990	355

4. Центральними значеннями вибірки ($n = 12$) є значення у стовпцях G та H .

$$\frac{200 + 220}{2} = 210.$$

Тому медіану обчислюють за формулою:

Знайдіть медіану засобами табличного процесора (=MEDIAN (B2:M2)). Перевірте, чи отримали таке саме значення. Маємо, що середнє значення витрат більше тяжіє до половини значень у правій частині відсортованого діапазону.

5. За даними витрат найчастіше для оплати послуг витратили 200 грн (6, 9, 10 місяці). Це ж значення можна отримати в табличному процесорі за формулою = MODE (B2:M2).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Місяць	8	5	7	6	9	10	4	11	3	12	2	1	Середнє
2	Витрати	80	190	190	200	200	200	220	400	420	500	670	990	355
3	Медіана													210
4	Мода													200

Завдання 2. Кількість опадів

За таблицею з відомостями про кількість опадів, яку склали кожного місяця протягом трьох років, потрібно визначити та проаналізувати показники центру та варіації розподілу опадів за роками і місяцями.

1. Відкрийте файл *Опади.xlsx*, де міститься таблиця з даними про кількість опадів.

2. Для знаходження середньої кількості опадів за кожний місяць у клітинку В6 уведіть формулу обчислення середнього значення за діапазоном В3:В5. Скопіюйте формулу у клітинки діапазону С6:М6.

3. Для знаходження рангів місяців за вологістю у клітинку В7 уведіть формулу для обчислення рангу клітинки В6 у діапазоні В6:М6 за спаданням кількості опадів. Скопіюйте формулу у клітинки діапазону С7:М7. Зверніть увагу, що посилання на діапазон має бути абсолютним, а порядок ранжування (за спаданням) визначає третій аргумент функції RANK.

4. Аналогічним чином у клітинках В8:М8 обчисліть ранг місяців за засушливістю, тобто за зростанням кількості опадів.

5. У клітинках N3:P5 обчисліть середньомісячну кількість опадів для кожного з трьох (2002-2004) років, стандартне відхилення та асиметрію вибірки щомісячних опадів для кожного року. Формули потрібно вводити тільки у клітинки N3:P3, а в діапазон N4:P5 їх можна скопіювати.

6. За результатами аналізу самостійно зробіть такі висновки:

- Які місяці є найбільш та найменш засушливими?
- Чи можна вважати розподіл кількості опадів за місяцями рівномірним? Який показник дозволяє дати відповідь на це питання?
- Які місяці — найбільш чи найменш засушливі — є рідкіснішими у цій вибірці?

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	<i>Дані про кількість опадів</i>													
		I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	Середня кількість опадів
2														
3	2002	37,2	11,4	16,5	19,5	11,7	129	57,1	43,8	85,7	86	12,5	21,2	44,30
4	2003	34,5	51,3	20,5	26,9	45,5	71,5	153	96,6	74,8	14,5	21	22,3	52,70
5	2004	8	1,2	3,8	11,9	66,3	60	50,6	14,5	79,9	74,9	56,6	9,4	36,43
6	Середнє за місяць	26,6	21,3	13,6	19,4	41,2	86,8	86,9	51,6	80,1	58,5	30,0	17,6	
7	Ранг місяця за вологістю	8	9	12	10	6	2	1	5	3	4	7	11	
8	Ранг місяця за засушливістю	5	4	1	3	7	11	12	8	10	9	6	2	
9														

VI. Підсумки уроку

Фронтальне опитування

1. Що вивчає наука статистика? Наведіть приклади.
2. Яким вимогам має відповідати вибірка об'єктів дослідження?

Поясніть кожну з них.

3. Що таке ряд даних? Наведіть приклади.
4. Що характеризує середнє арифметичне ряду даних?
5. У яких випадках середнє арифметичне більш показово характеризує весь ряд даних, а у яких — менш показово?
6. Що таке стандартне відхилення ряду даних? Що воно характеризує?
7. Що таке мода ряду даних? Що вона характеризує?
8. Що таке медіана ряду даних? Що вона характеризує?
9. У впорядкованому ряді чисел найбільше число збільшили на 5. Як зміняться значення середнього арифметичного, моди і медіани цього ряду? Відповідь поясніть.

Рефлексія

1. Під час уроку я
 - ✓ дізнався...
 - ✓ зрозумів...
 - ✓ навчився...
2. Найбільше мені сподобалося...
3. На уроках найкраще в мене виходило...
4. Я мав (-ла) труднощі з...
5. Я хотів би ще дізнатися про...

VII. Домашнє завдання

Підручник § 10.4-10.5 ст. 98-102, виконати вправи 7, 8 нас.102.

Вправа 7. На заводі протягом семигодинного робочого дня робітник виготовляв: 10, 8, 11,12, 11, 9, 7 деталей (табл. 10.3). Знайдіть моду, медіану, середнє значення. Обчислення проведіть у табличному процесорі.

Таблиця 10.3

Години	1	2	3	4	5	6	7
Кількість деталей	10	8	11	12	11	9	7

Вправа 8. Директор фірми отримує 20 000 грн на місяць, два його заступники — по 14 000 грн, а решта 17 робітників фірми — по 7500 грн на місяць. Знайдіть середнє значення, моду, медіану заробітних плат у цій фірмі. Обчислення проведіть у табличному процесорі.

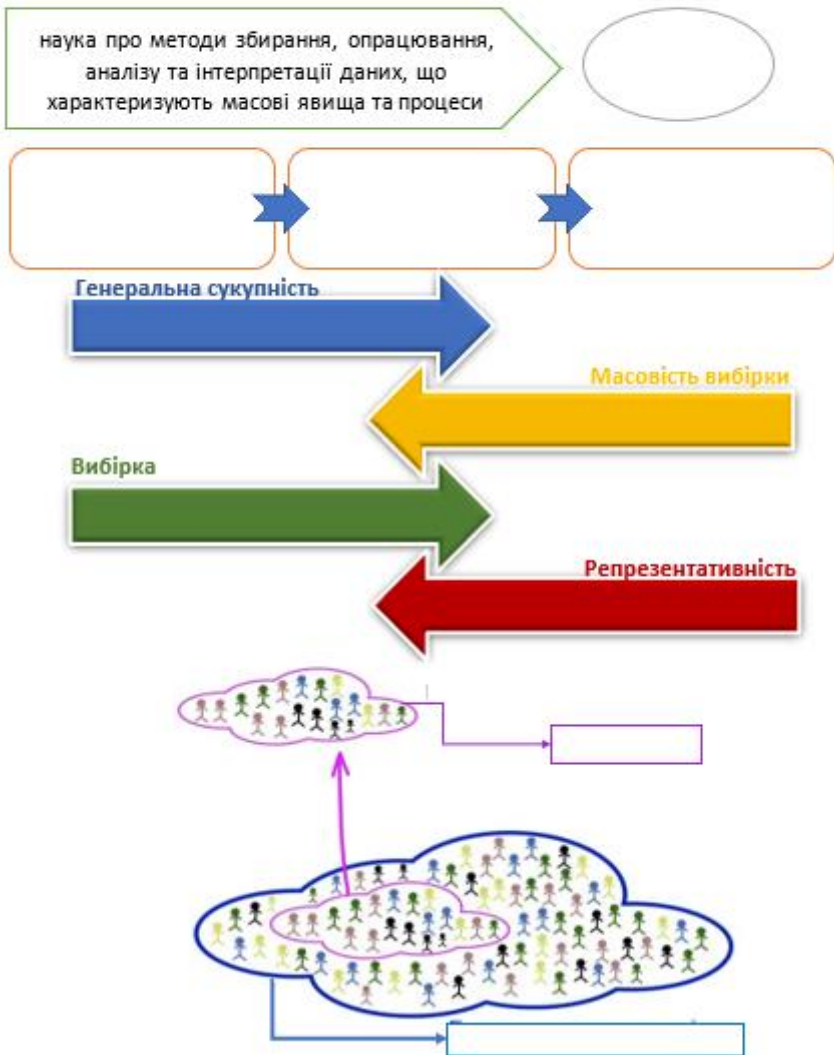
VIII. Оцінювання роботи учнів

Учні демонструють заповнені опорні конспекти (див. додаток 2) та результати виконаних завдань.



ОПОРНИЙ КОНСПЕКТ

Тема уроку: _____



Ряди розподілу

ОСНОВНІ СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ ВИБІРКИ

ОБСЯГ ВИБІРКИ	
РОЗМАХ ВИБІРКИ	
СЕРЕДНЄ ЗНАЧЕННЯ	
СТАНДАРТНЕ ВІДХИЛЕННЯ	
МОДА ВИМІРЮВАННЯ	
МЕДІАНА ВИБІРКИ	

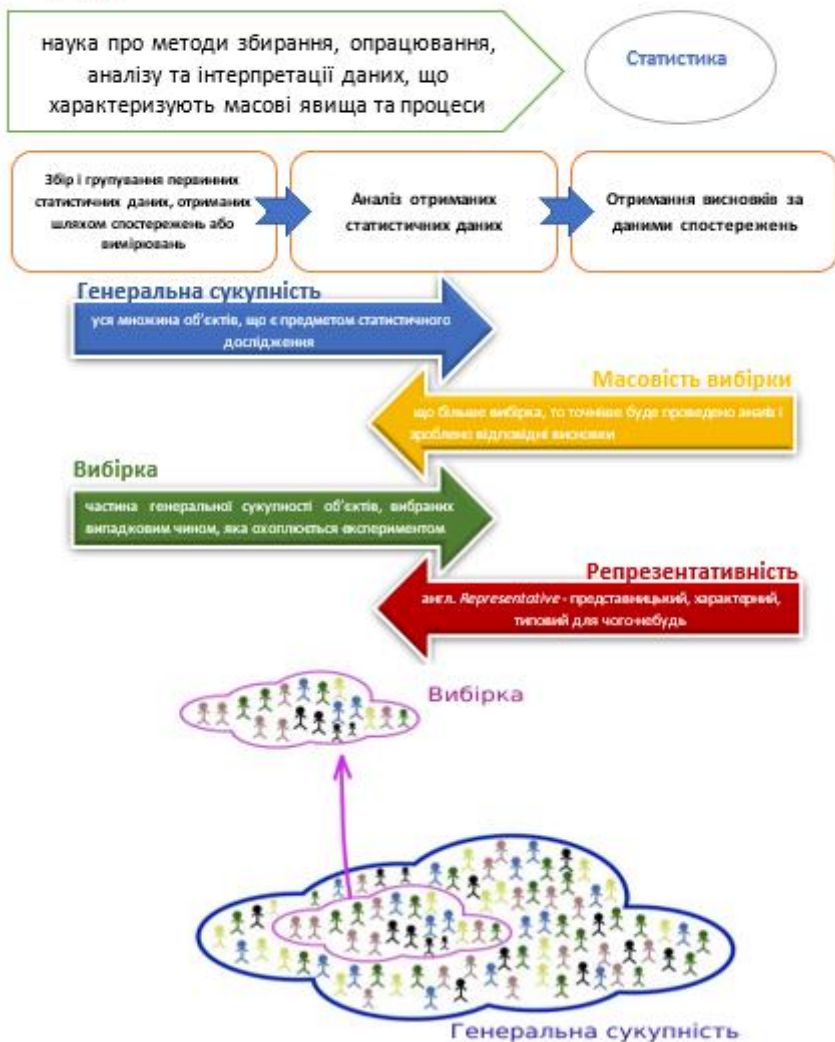
ОСНОВНІ СТАТИСТИЧНІ ФУНКЦІЇ:

Назва функції (англ.)	Призначення
AVERAGE(аргумент1;{аргумент2};...)	
MIN(аргумент1;{аргумент2};...)	
MAX(аргумент1;{аргумент2};...)	
MEDIAN(аргумент1;{аргумент2};...)	
MODE.SNGL(аргумент1;{аргумент2};...)	
STDEV.S(аргумент1;{аргумент2};...)	
RANK (аргумент; посилання; порядок)	



ОПОРНИЙ КОНСПЕКТ

Тема уроку: Обчислення основних статистичних характеристик



Дискретні варіаційні - це такі ряди розподілу, в яких варіанта як величина кількісної ознаки може набувати тільки певного значення

варіанти є цілими числами

Інтервальні варіаційні - такі ряди розподілу, в яких значення варіанти дано у вигляді інтервалів

варіанти мають вигляд інтервалів

ОСНОВНІ СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ ВИБІРКИ

ОБСЯГ ВИБІРКИ	кількість випадків, включених у вибірку сукупність
РОЗМАХ ВИБІРКИ	різниця між максимальним і мінімальним значеннями елементів вибірки
СЕРЕДНЄ ЗНАЧЕННЯ	це частка від ділення суми цих чисел на їх кількість (обсяг вибірки)
СТАНДАРТНЕ ВІДХИЛЕННЯ	характеризує, наскільки широко розташовані значення ряду даних відносно їх середнього арифметичного
МОДА ВИМІРЮВАННЯ	це значення, яке найчастіше трапляється
МЕДІАНА ВИБІРКИ	середнє число із групи чисел, тобто значення половини чисел більші за медіану, а значення іншої половини – менші

ОСНОВНІ СТАТИСТИЧНІ ФУНКЦІЇ:

Назва функції (англ.)	Призначення
AVERAGE(аргумент1;{аргумент2};...)	середнє (арифметичне) аргументів
MIN(аргумент1;{аргумент2};...)	найменше число в наборі значень
MAX(аргумент1;{аргумент2};...)	найбільше значення з набору значень
MEDIAN(аргумент1;{аргумент2};...)	медіана вказаних чисел
MODE.SNGL(аргумент1;{аргумент2};...)	мода вибірки
STDEV.S(аргумент1;{аргумент2};...)	стандартне відхилення вибірки