**Міністерство освіти і науки України**

**Запорізький національний університет**

**В. А. Ковпак**

**ОПРАЦЮВАННЯ ТА АНАЛІЗ СТАТИСТИЧНИХ ДАНИХ**

**Методичні рекомендації до самостійної роботи**

**для здобувачів ступеня вищої освіти бакалавра**

**спеціальності «Інформаційна, бібліотечна та архівна справа»**

**освітньо-професійної програми**

**«Інформаційно-комунікаційна справа»**

Затверджено

вченою радою ЗНУ

Протокол № 02 від 26.09.2023

Запоріжжя

2023

**УДК: 316.77:070]:311.17(076.5)**

**К 567**

Ковпак В. А. Опрацювання та аналіз статистичних даних : методичні рекомендації до організації самостійної роботи для здобувачів ступеня вищої освіти бакалавра спеціальності «Інформаційна, бібліотечна та архівна справа» освітньо-професійної програми «Інформаційно-комунікаційна справа». Запоріжжя : Запорізький національний університет, 2023. 68 с.

Методичні рекомендації містять пояснення та інструкції щодо виконання завдань із самостійної роботи з дисципліни «Опрацювання та аналіз статистичних даних», а також питання для самоконтролю, перелік рекомендованої літератури. Для формування необхідних компетентностей розроблено завдання для самостійної роботи різного рівня складності. Запропонований термінологічний мінімум із тлумаченням основних понять дисципліни, необхідний для формування базових знань.

Для здобувачів ступеня вищої освіти бакалавра спеціальності «Інформаційна, бібліотечна та архівна справа», які навчаються за освітньо-професійною програмою «Інформаційно-комунікаційна справа».

Рецензент

*Л. В. Чернявська,* доктор наук із соціальних комунікацій, доцент,завідувач кафедри журналістики

Відповідальний за випуск

*В. В.Березенко*, доктор наук із соціальних комунікацій, професор, завідувач кафедри соціальних комунікацій та інформаційної діяльності

**ЗМІСТ**

|  |
| --- |
| Вступ…………………………………………………………………… ………4 |
| Зміст самостійної роботи……………………………………………………….6 |
| Тема 1. Використання та поширення даних. Журналістика даних………….7 |
| Тема 2. Методологічні засади статистики. Статистичне спостереження …21 |
| Тема 3. Життєвий цикл даних. Техніки оброблення даних …….………….32 |
| Тема 4. Опрацювання та аналіз даних від збору до отримання цінної інформації (на прикладі реалізації публічних консультацій в інформаційній діяльності)……………………………………………………………………...47 |
| Індивідуальне практичне завдання………………..……….…………………57 |
| Тестові завдання для самоконтролю ………….………………….………….58 |
| Додатки …………………………..…………………………………………….60 |
| Список використаних джерел ………………………………………..………64 |

**ВСТУП**

Інформаційно-комунікаційна справа передбачає застосування сучасних інструментів і технологій пошуку, оброблення та аналізу інформації, зокрема, статистичних методів аналізу даних невеликого або великого обсягу, технологій роботи з відкритими даними, базами даних та інформаційними системами.

Тож укладені методичні рекомендації до самостійної роботи з курсу «Опрацювання та аналіз статистичних даних» є затребуваними, оскільки на цей вид роботи навчальним планом передбачено не менше 75 відсотків годин від їх загальної кількості. Відповідно для формування необхідних навичок запропоновано теоретичні кейси та практичні завдання, для контролю засвоєння теоретичних основ – тестові завдання та проходження онлайн-курсів, запропоновано тлумачення основних термінів і понять, правила, алгоритми, шаблони для виконання завдань.

*Метою* вивчення навчальної дисципліни «Опрацювання та аналіз статистичних даних» є формування знань майбутніх фахівців і набуття ними навичок користування техніками оброблення статистичних даних, тлумачення зібраних даних, використання результатів для формулювання висновків та застосування ефективних комунікаційних стратегій для поширення фінального звіту.

*Міждисциплінарні зв’язки.* Навчальна дисципліна «Опрацювання та аналіз статистичних даних» є складовою циклу професійної підготовки фахівців ступеня вищої освіти бакалавр, має синергетичні зв’язки з такими дисциплінами, як «Аналітико-синтетичне опрацювання документної інформації» (аналітичне опрацювання даних, документне оформлення інформації), «Інформаційно-пошукові системи» (системи даних), «Формування баз даних організації» (бази даних), «Візуалізація даних».

*Основними завданнями* освітньої компоненти «Опрацювання та аналіз статистичних даних» є такі позиції:

* опанувати поняття «дані», напрям «журналістики даних» як одного із найсучасніших та найбільш точних способів передачі інформації в засобах масової інформації (ЗМК), механізми використання журналістами та урядовими структурами відкритих даних під час воєнного стану;
* засвоїти методологічні засади статистики, форми, види, способи статистичного спостереження;
* проаналізувати життєвий цикл даних та техніки оброблення даних;
* дослідити механізми опрацювання та аналізу даних від збору до отримання цінної інформації (на прикладі реалізації публічних консультацій в інформаційній діяльності).

У результаті вивчення навчальної дисципліни студенти матимуть змогу набути таких результатів навчання (знання, уміння тощо) та компетентностей:

* загальні компетентності:
* здатність застосовувати знання у практичних ситуаціях;
* знання та розуміння предметної області та професійної діяльності;
* здатність реалізувати свої права та обов’язки як члена суспільства, усвідомлювати цінності громадянського (вільного демократичного) суспільства та необхідність його сталого розвитку, верховенства права, прав і свободи людини і громадянина в Україні;
* фахові компетентності:
* здатність здійснювати відбір, аналіз, оцінку, систематизацію, моніторинг, організацію, зберігання, розповсюдження та надання в користування інформації та знань у будь-яких форматах;
* здатність використовувати методи систематизації, пошуку, збереження, класифікації інформації для різних типів контенту та носіїв;
* здатність підтримувати ділову комунікацію з усіма суб’єктами інформаційного ринку, користувачами, партнерами, органами влади та управління, засобами масової інформації;
* програмні результати навчання:
* узагальнювати, аналізувати і синтезувати інформацію в діяльності, пов’язаній із її пошуком, накопиченням, зберіганням та використанням;
* використовувати знання технічних характеристик, конструкційних особливостей, призначення і правил експлуатації комп’ютерної техніки та офісного обладнання для вирішення технічних завдань спеціальності.

Виконання практичних завдань сприятиме не лише ознайомленню студентів із методологічними засадами статистичних спостережень, процесом аналізу даних, їх життєвого циклу та техніками їх оброблення, а й дасть можливість адаптувати запропоновані техніки до одного з видів інформаційно-комунікаційної діяльності, зокрема специфіки публічних консультацій в інформаційній сфері, та простежити зв’язок із цариною журналістики даних, використання інформаційниками та урядовими структурами відкритих даних у кризових комунікаціях, зокрема під час воєнного стану, що є особливо актуально в сьогоденні. Здобувачі набувають знання з використання комунікаційниками відкритих даних під час професійної діяльності, процесу аналізу даних, технік їх оброблення, стратегії збору даних, опрацювання цінної інформації для оприлюднення статистичних даних завдяки ефективним комунікаційним стратегіям, що стануть важливим складником циклу професійної підготовки фахівців освітнього ступеня «бакалавр» у контексті реалізації інформаційної діяльності в різних сферах суспільного життя. Саме цим і зумовлена необхідність вивчення дисципліни «Опрацювання та аналіз статистичних даних».

Знання, отримані в процесі ознайомлення та засвоєння матеріалу з дисципліни «Опрацювання та аналіз статистичних даних», сприятимуть набуттю студентами названих вище компетентностей, що і визначає актуальність видання методичних рекомендацій, що стануть базовою платформою в аналітиці даних інформаційно-комунікаційної сфери та своєрідним орієнтиром у вивченні закономірностей, тенденцій опрацювання та аналізу статистичних даних. Особливістю дисципліни «Опрацювання та аналіз статистичних даних» є набуття студентами навичок користування техніками оброблення даних, тлумачення зібраних даних, використання результатів для формулювання висновків у контексті конкретного виду інформаційно-комунікаційної діяльності, а саме на прикладі проведення публічних консультацій, застосування цих знань у розрізі журналістики даних та користування відкритими даними в межах роботи комунікаційних відділів установ, організацій тощо, зокрема з урахуванням обставин воєнного часу.

**ЗМІСТ САМОСТІЙНОЇ РОБОТИ**

***Тема* 1. Використання та поширення даних. Журналістика даних**

***Мета:*** схарактеризувати поняття даних та інформації, дефініціювати журналістику даних або інтерактивний сторітелінг як один із найсучасніших та найбільш точних способів передачі інформації в ЗМК, вивчити інструментарій використання журналістами та урядовими структурами відкритих даних під час воєнного стану.

*План*

1. Поняття даних та інформації.

2. Журналістика даних або інтерактивний сторітелінг як один із найсучасніших та найбільш точних способів передачі інформації в ЗМК.

3. Використання журналістами та урядовими структурами відкритих даних під час воєнного стану.

***Основні поняття:*** інформація,публічна інформація,дані,відкриті дані, журналістика даних, пошук даних, офіційні чат-боти

**📚 Методичні рекомендації та поради**

У процесі підготовки самостійної роботи студентам необхідно взяти до уваги таку інформацію.

Назва дисципліни «Опрацювання та аналіз статистичних даних» передбачає роботу з даними, зокрема з їх збором для аналізу в їхньому первісному вигляді. Коли дані обробляються та перетворюються таким чином, що вони стають *корисними* для користувачів, то, по суті, вони трансформуються в «інформацію».

«Хоча дані – це несистематичний факт чи деталь про щось, інформація – це систематизована та відфільтрована форма даних, яка є корисною» [3] (див. Табл. 1).

**Таблиця 1**

## Таблиця порівняння понять «дані» та «інформація» Джерело: [3].

|  |  |  |
| --- | --- | --- |
| **Основа для порівняння** | **Дані** | **Інформація** |
| Значення | Дані означають необроблені факти, зібрані про когось або щось, що є оголеними та випадковими. | Факти, що стосуються певної події чи теми, які уточнюються в процесі обробки, називаються інформацією. |
| Що це? | Це лише текст і цифри. | Це уточнені дані. |
| На основі | Записи та спостереження | Аналіз |
| Форма | Неорганізований | Організовано |
| Корисно | Може бути чи не бути корисним. | Завжди |
| Конкретні | Немає | Так |
| Залежність | Не залежить від інформації. | Без даних інформація не може бути оброблена. |

Отже, поняття **дані** (від лат. терміна «datum» – «щось дане») визначається як «сукупність фактів та деталей, зокрема тексту, цифр, малюнків, тверджень щодо спостережень, символів, опису речей, подій чи сутностей, зібраних із метою зробити висновки; це необроблені факти, які слід опрацювати (напр. фахівцям різних організацій, урядових відомств, установ, неурядових установ тощо) для отримання інформації. Відповідно, дані розрізняють первинні (якісні та кількісні) та вторинні (внутрішні та зовнішні)» [3].

**Інформація** (від лат. слова «informare» – «надати форму»), по суті, є тією формою даних, яка обробляється, організовується, є специфічною та структурованою, що представлена в певному дискурсі, позбавлена зайвих деталей, викристалізована в тому обсязі, що потрібно для дослідника. Так, наприклад, дані часто опрацьовують шляхом складання таблиць, діаграм та інших подібних операцій, що покращують їх пояснення та інтерпретацію.

Вимогою розвитку демократичного суспільства є формування простору відкритих даних. Проте органи влади досить часто ігнорують запити з проханням надати відкриті дані.

За словами експертки з питань відкритих даних, координаторки ініціативи OpenUp Ukraine Надії Бабинської (Вірної), вимога від владних органів відкритих даних є суспільно значущою, адже дає можливість контролювати, впливати і розвивати [4]. А для цього варто розуміти різницю між інформацією, даними, відкритими даними, публічною інформацією та публічною інформацією у форматі відкритих даних.

**Інформація** – будь-які відомості та/або дані, які можуть бути збережені на матеріальних носіях або відображені в електронному вигляді (відповідно до Закону України «Про інформацію», 1992 рік).

**Дані** – відомості, отримані шляхом вимірювання, спостереження, логічних або арифметичних операцій і представлені у формі, придатній для постійного зберігання, передачі і (автоматизованої) обробки.

**Відкриті дані** – дані, які можна вільно використовувати, повторно використовувати, поширювати будь-ким і будь-де, поєднувати.

**Публічна інформація** – відображена та задокументована будь-якими засобами та на будь-яких носіях інформація, що була отримана або створена в процесі виконання суб’єктами владних повноважень своїх обов’язків, передбачених чинним законодавством, або яка знаходиться у володінні суб’єктів владних повноважень, інших розпорядників публічної інформації, визначених цим Законом (ЗУ «Про доступ до публічної інформації», 2011).

**Публічна інформація у форматі відкритих даних** – це публічна інформація у форматі, що дозволяє її автоматизоване оброблення електронними засобами (машиночитний формат – xls, csv, json. xml тощо), вільний та безоплатний доступ до неї, а також її подальше використання (зміни до ЗУ «Про доступ до публічної інформації» від 09.04.2015). За правом доступу закон визначає два типи інформації: **відкрита інформація** та **інформація з обмеженим доступом** (конфіденційна, таємна, службова інформація). Таким чином, обмеження доступу можливе при дотриманні сукупності таких вимог:

* виключно в інтересах національної безпеки, територіальної цілісності або громадського порядку з метою запобігання заворушенням чи злочинам, для охорони здоров’я населення, для захисту репутації або прав інших людей, для запобігання розголошенню інформації, одержаної конфіденційно, або для підтримання авторитету і неупередженості правосуддя;
* розголошення інформації може завдати істотної шкоди цим інтересам;
* шкода від оприлюднення такої інформації переважає суспільний інтерес в її отриманні.

«Щоб розпорядник не надав інформацію, посилаючись на згадані причини, йому потрібно провести **трискладовий тест** і довести, що надання інформації може комусь чи чомусь зашкодити» [4] (див. Рис. 1).



Рис. 1. Трискладовий тест розпорядника щодо можливості надання інформації. Джерело: [5]

**Не може бути обмежено доступ до інформації про такі позиції** [5]:

* розпорядження бюджетними коштами;
* володіння, користування чи розпорядження державним,
* комунальним майном;
* не належать до інформації з обмеженим доступом відомості, зазначені у декларації особи, уповноваженої на виконання функцій держави або місцевого самоврядування, поданій відповідно до Закону України «Про запобігання корупції».

Крім декларацій, Закон України «Про запобігання корупції» також встановлює, що не може обмежуватись доступ до інформації про:

* розміри, види благодійної та іншої допомоги, що надається фізичним та юридичним особам чи одержується від них особами, уповноваженими на виконання функцій держави або місцевого самоврядування, або державними органами, органами місцевого самоврядування;
* розміри, види оплати праці, матеріальної допомоги та будь-яких інших виплат з бюджету особам, уповноваженими на виконання функцій держави або місцевого самоврядування, а також одержані цими особами за правочинами, які підлягають обов'язковій державній реєстрації;
* подарунки, які регулюються Законом України «Про запобігання корупції».

Антикорупційне законодавство також забороняє обмежувати в доступі інформацію про:

* передачу в управління належних особам підприємств та корпоративних прав, що здійснюється в порядку, передбаченому Законом, коли посадові особи обіймають посади, які передбачають вимоги несумісності з підприємницькою діяльністю. Це зокрема стосується народних депутатів, Президента, міністрів, суддів та ін.
* конфлікт інтересів та заходи з його врегулювання.

*До інформації з обмеженим доступом не можуть бути віднесені відомості*:

1) про стан довкілля, якість харчових продуктів і предметів побуту;

2) про аварії, катастрофи, небезпечні природні явища та інші надзвичайні ситуації, що сталися або можуть статися і загрожують безпеці людей;

3) про стан здоров’я населення, його життєвий рівень, включно з харчуванням, одягом, житлом, медичним обслуговуванням та соціальним забезпеченням, а також про соціально-демографічні показники, стан правопорядку, освіти і культури населення;

4) про факти порушення прав і свобод людини, включно з інформацією, що міститься в архівних документах колишніх радянських органів державної безпеки, пов’язаних з політичними репресіями, Голодомором 1932–1933 років в Україні та іншими злочинами, вчиненими представниками комуністичного та/або націонал-соціалістичного (нацистського) тоталітарних режимів;

5) про незаконні дії органів державної влади, органів місцевого самоврядування, їх посадових та службових осіб. Щодо діяльності державних та комунальних унітарних підприємств, господарських товариств, у статутному капіталі яких понад 50 відсотків акцій (часток) належать державі або територіальній громаді, а також господарських товариств, 50 і більше відсотків акцій (часток) яких належать господарському товариству, частка держави або територіальної громади в якому становить 100 відсотків, що підлягають обов’язковому оприлюдненню відповідно до закону;

6) інші відомості, доступ до яких не може бути обмежено відповідно до законів та міжнародних договорів України, згода на обов'язковість яких надана Верховною Радою України.

**Концепція відкритих даних** була запроваджена в українському інформаційному законодавстві у 2015 році й передбачає обов’язок розпорядників інформації надавати визначені переліки інформації у формі відкритих даних на запит, оприлюднювати й регулярно оновлювати таку інформацію на єдиному державному вебпорталі відкритих даних та на своїх вебсайтах.

Оприлюднення та надання на запит публічної інформації у форматі відкритих даних, які містять персональну інформацію можливе, якщо персональні дані знеособлені та захищені відповідно до Закону України «Про захист персональних даних» або наявна згода особи на таке поширення. Така інформація також може бути поширена, якщо її оприлюднення прямо передбачено законом чи заборонено відносити її до інформації з обмеженим доступом.

**Перелік наборів даних** (Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних, що затверджено постановою Кабінету Міністрів України від 21 жовтня 2015 р. № 835 [2]), що підлягають оприлюдненню у формі відкритих даних, визначаються Кабінетом Міністрів України.

Усі суб’єкти владних повноважень та деякі інші розпорядники мають оприлюднювати (нижче інформація наводиться з лекційного курсу Prometheus. «Доступ до публічної інформації: від А до Я»):

– Перелік підприємств, установ (закладів) та організацій розпорядника інформації та підпорядкованих йому організацій, у тому числі їх телефонів та адрес.

– Інформацію про організаційну структуру розпорядника інформації.

– Звіт про використання бюджетних коштів (для розпорядників інформації, що використовують бюджетні кошти), зокрема за окремими бюджетними програмами.

– Нормативи, що затверджуються розпорядником інформації та Переліки національних стандартів.

– Звіти, у тому числі щодо задоволення запитів на інформацію.

– Річні плани закупівель.

– Інформацію про систему обліку, види інформації, яка зберігається розпорядником.

– Переліки адміністративних послуг, інформаційні картки адміністративних послуг та бланки заяв, необхідних для звернення щодо надання адміністративної послуги.

– Фінансову звітність суб’єктів господарювання державного сектору економіки, що належать до сфери управління розпорядника інформації та ін.

Водночас, для окремих органів державної та місцевої влади обов’язковим є оприлюднення додаткових переліків інформації. Наприклад, МВС має поширювати у форматі відкритих даних інформацію про:

* осіб, які переховуються від органів влади;
* викрадені (вилучені) культурні цінності;
* транспортні засоби, що перебувають у розшуку у зв’язку з їх незаконним заволодінням;
* викрадену, втрачену зброю;
* стан аварійності на автомобільних дорогах та ін.

А органи місцевого самоврядування також повинні оприлюднювати відомості, наприклад, про:

* перелік перевізників, що надають послуги пасажирського автомобільного транспорту та маршрути перевезення;
* перелік розповсюджувачів реклами, що отримали дозвіл на розміщення зовнішньої реклами;
* перелік земельних ділянок, що пропонуються для здійснення забудови.

Вище згадана експертка Н. Бабинська радить завжди додавати в текст запиту на публічну інформацію прохання надавати відповідь у формі відкритих даних, відсилку на ст. 10-1 ЗУ «Про доступ до публічної інформації», а також згадувати Постанову Кабміну № 835 від 17.04.2019 року: «Можете вказати бажаний формат (але це як пощастить). Також вказуйте електронну скриньку. Будьте готові, що доведеться купувати диск чи флешку, щоб отримати дані…Якщо хтось оприлюднює на сайті міської ради відскановані документи і стверджує, що це відкриті дані, то це неправда. Бо комп’ютер не може сприймати цю інформацію, її треба заново перетворювати в машиночитний формат. Зокрема, відповідно структурувати або робити ”дерево”, щоб потім використовувати для машинної автоматизованої обробки» [4].

**Де оприлюднюють відкриті дані**?

*Єдиний державний портал відкритих даних*: https://data.gov.ua/

*Відкриті дані Верховної Ради України*: https://data.rada.gov.ua/open

*Єдиний вебпортал використання публічних коштів*: https://spending.gov.ua/new/

*Відкриті дані Prozorro*: https://prozorro.gov.ua/openprocurement

*Єдиний державний реєстр декларацій*: https://nazk.gov.ua/uk/

*Європейський портал відкритих даних*: https://data.europa.eu/en

Для поглибленого вивчення теми публічної інформації у форматі відкритих даних варто користуватися такими ресурсами, як:

- **Інфохаб** (про забезпечення інформаційного представлення і публічного доступу до відкритих державних даних громадянам, державним та іншим організаціям): https://data.gov.ua/pages/infohub;

- тематична підбірка від Наї Навак 2020 року «**Корисний карантин: книги і курси з open data, data science та data visualization**»: https://mind.ua/openmind/20209468-korisnij-karantin-knigi-i-kursi-z-open-data-data-science-ta-data-visualization

Навчальну інформацію про те, як записувати дані у таблицю, як її чистити, як зберегти таблицю у форматі відкритих даних, алгоритми створення форматів збереження відкритих даних: табличні та ієрархічні можна знайти у посібнику «Відкриті дані: формати і правила створення» від Texty.org.ua [1].

|  |
| --- |
| **Важлива думка**  *«Навіщо потрібні відкриті дані? Це сировина, завдяки якій можна створювати нові сервіси, нові продукти, нові робочі місця, економити час та бюджетні кошти і навіть зберігати життя», – каже експертка Н. Бабинська, наводячи приклади, як у США за допомогою відкритих даних розробляють сервіси для порятунку хворих на астму чи для того, щоб контролювати забруднення довкілля свинцем. [цит. за 4].* |

**Texty.org.ua** – незалежне видання, засноване Анатолієм Бондаренком та Романом Кульчинським у 2010 році, яке створює проєкти з журналістики даних та працює в традиційних журналістських жанрах: від розлогих репортажів до коротких повідомлень, має український погляд на світ, намагається максимально вивчити проблему, про яку пише, і показати, що відбувається насправді, а не просто публікує різні точки зору [7]. Серед останніх актуальних проєктів «Текстів» такі (див. Рис. 2):

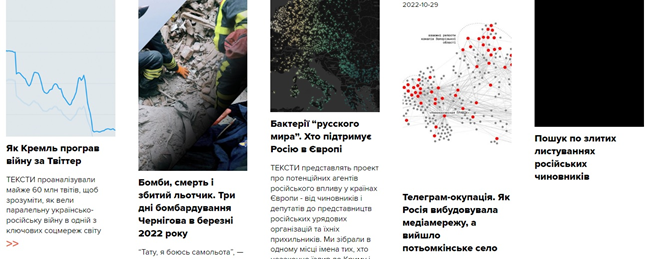


Рис. 2. Певна частина проєктів Texty.org. Джерело: [9]

Загалом, «**журналістика даних** або **інтерактивний сторітелінг** – це один із найсучасніших та найбільш точних способів передачі інформації в ЗМІ. Його активно використовують такі топові світові видання, як New York Times, Washington Post, The Guardian. Основною відмінністю візуалізації даних від традиційної форми подачі журналістського тексту є робота з великими масивами даних, та використання мов програмування. У квітні 2017 року світ побачив український переклад книги Альберто Каїро «The Functional Art. An introduction to informational graphics and visualization» (науковий редактор українського перекладу – аналітик texty.org Анатолій Бондаренко) [8].

За словами А. Бондренка, окрім texty.org «візуалізації на теми дня вже декілька років робить “Українська правда”, пробують себе у цьому напрямі “Bird in Flight”, “Liga.net” і “KyivPost”. Цікаво, що є сильні проекти журналістики даних у низці недержавних громадських організацій: Чесно, Опори, львівської CifraGroup, infolight.org.ua. Є приклади й у сусідніх країнах: особисто знаємо колег з Чехії, Польщі, Білорусі. Як не дивно, поки що Україна в цьому напрямі попереду» [8].

Наприклад, на основі алгоритму автоматичної класифікації супутникових знімків texty.org проаналізували руйнування 40 населених пунктів в Україні після повномасштабного вторгнення росії в Україну після 24 лютого 2022 року: «Ми не мали супутникових змінків достатньої якості, які покривають усіх населені пункти, які зазнали руйнувань. Тому картина не повна, зокрема на карті відсутній Маріуполь. Міста й села, які ми досліджували, розташовані в зоні активних бойових дій, а також на звільнених та окупованих територіях. Більшість із них – на сході й півдні країни, але для порівняння ми також дослідили Бородянку, Ірпінь і Бучу Київської області, Ягідне Чернігівської області та кілька населених пунктів Харківської області. Ми проаналізували супутникові знімки, отримані впродовж весни–літа 2022 року. У Сєвєродонецьку, Лисичанську, Ізюмі, Лимані й Авдіївці цілі квартали було зруйновано ще в середині літа 2022 року. За даними нашого дослідження, у Сєвєродонецькому районі повністю знищено смт Врубівка (близько 1000 мешканців), Ниркове, Тошківка (близько 4000 мешканців) й Новотошківське (понад 2000 мешканців). Через пів року руйнування в містах і селах поблизу лінії фронту, ймовірно, ще масштабніші. Зараз ми розуміємо, що також повністю зруйновано Бахмут, Соледар. У населених пунктах, зокрема окупованих, але розташованих поза зоною бойових дій (станом на лютий 2023 року), масштаби руйнувань відтоді здебільшого не змінилися. Карта нижче інтерактивна. Червоними квадратами позначені ділянки, у яких наш алгоритм знайшов руйнування. Наближайте карту до червоних квадратів, поки вони не стануть достатньо великими, тоді з’явиться супутниковий знімок» [9] (див. Рис. 3). Методологія алгоритму, що працює за принципом візуального розпізнавання, детально прописана.

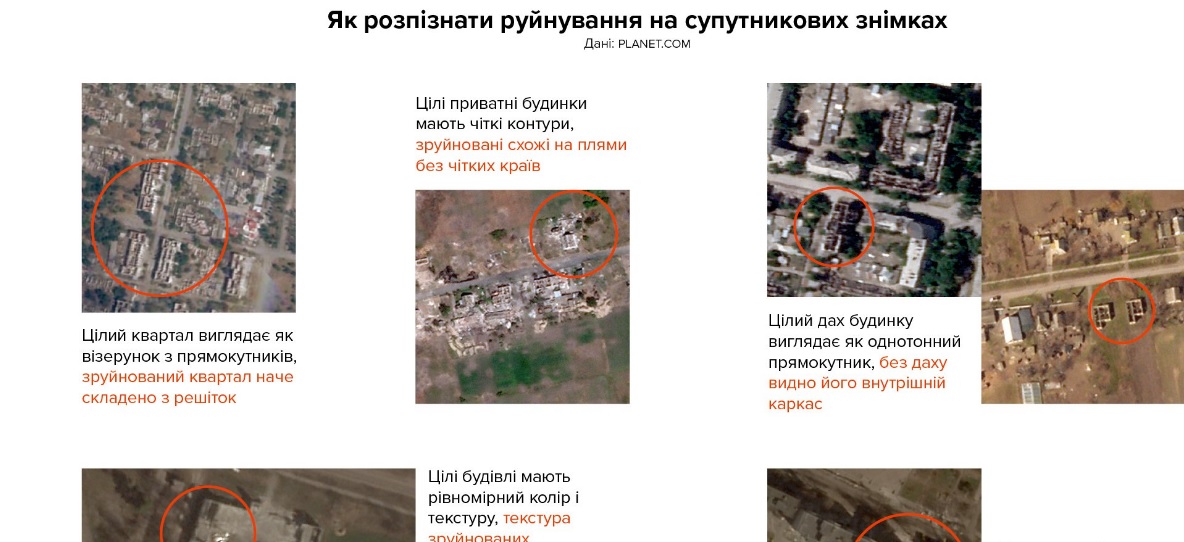


Рис. 3. Карта руйнувань. Джерело: [9]

На основі зібраних даних із відкритих джерел, таких як повідомлення Генштабу та звіти обласних військових адміністрацій, котрі оприлюднюються щодня, texty.org зробили карту бойових дій, яка постійно оновлюється, на ній не вказуються конкретні позиції – позначаються лише населені пункти, згадані в цих звітах, вона дає лише загальне уявлення, тож автори закликають не орієнтуватися на цю карту при ухваленні рішення про евакуацію. «Позиції ЗСУ позначені блакитним кольором, РФ – червоним (покинуті позиції РФ, або ж позиції, звідки українці вибили росіян тривалий час тому – жовтим). Бордовим – абстрактні "місця бойових дій". Сірим – місця колишніх бойових дій та обстрілів» [10] (див. Рис. 4).

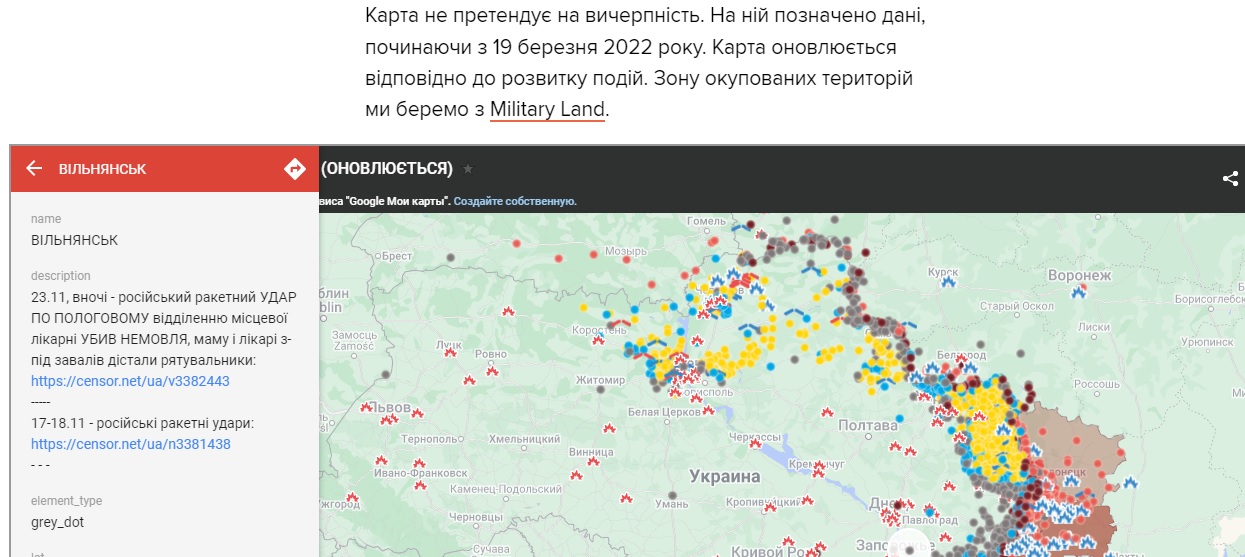


Рис. 4. Карта бойових дій. Активізовано опцію населеного пункту Вільнянськ станом на початок квітня 2023 року. Джерело: [10]

У 2019 році Texty.org. розробили інструмент для пошуку аптек – проєкт на основі відкритих даних НСЗУ, який складається з пошукової таблиці та інтерактивної мапи з локалізованими на ній аптеками та медичними закладами, покликаний допомагати кожному українському пацієнту шукати аптеки, що надають ліки за державною програмою «Доступні ліки»: У вікнах пошуку над таблицею пропонується ввести назву населеного пункту чи регіону, який цікавить. Інші фільтри для пошуку – назви аптек, а також вулиці, на яких вони розташовані. Таблиця автоматично оновлюється й підкаже найбільш відповідні варіанти. За необхідності детальнішої інформації про аптеки, заклад первинної медичної допомоги та амбулаторії можна перейти до інтерактивної карти під таблицею. «По кожній аптеці вказується назва, адреса, належність до аптечної мережі та сума коштів, компенсованих державою цій мережі за програмою “Доступні ліки”. По кожному закладу вказується назва, адреса та кількість пацієнтів, що підписали там декларації з лікарями» [11] (див. Рис. 5).

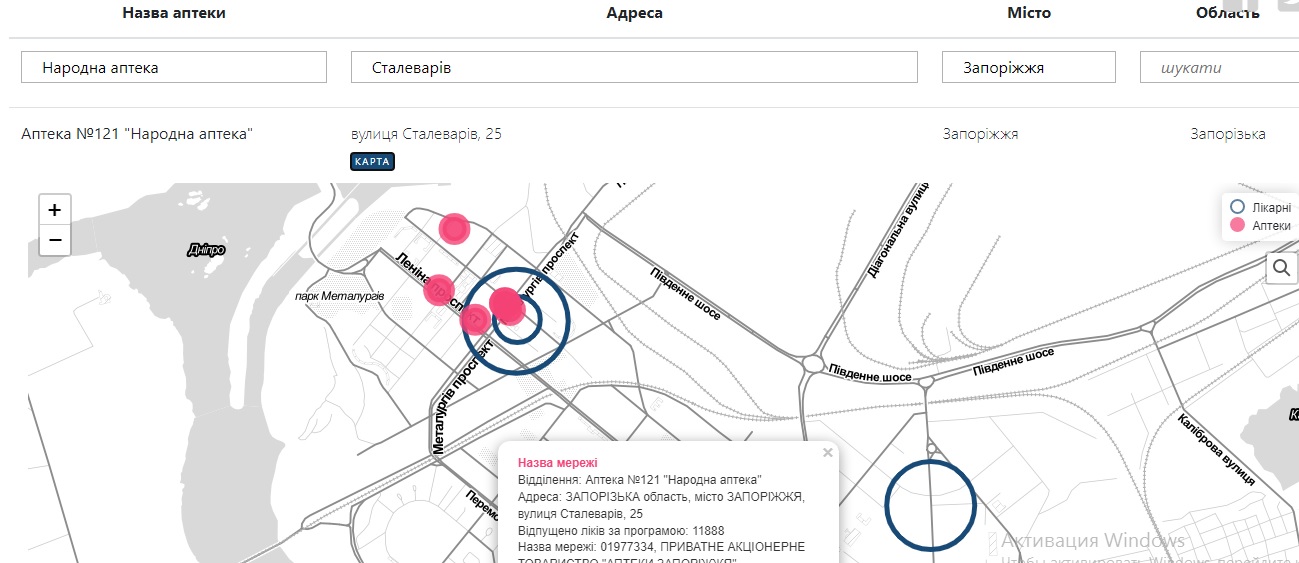


Рис. 5. Інструмент для пошуку аптек на основі відкритих даних. Джерело: [11]

Проте, як зазначає Яна Шекеряк у матеріалі «Карти, боти й пошук людей – як журналісти й урядові структури використовують відкриті дані під час воєнного стану», «із початком повномасштабного вторгнення російських військ в Україну медійна сфера, як і всі інші, сильно змінилася. Єдиний телемарафон, емоційно забарвлена лексика, цілодобова хроніка новин… Вплинуло на роботу медійників і закриття багатьох реєстрів. Та попри це журналісти продовжують використовувати відкриті дані для своєї роботи. Як урядові структури, так і волонтери й громадяни» [12].

Із початком введення воєнного стану припинилося функціонування багатьох реєстрів (Єдиний державний реєстр юридичних осіб, фізичних осіб-підприємців та громадських формувань, Державний реєстр речових прав на нерухоме майно, Автоматизованої системи виконавчого провадження, Єдиний реєстр боржників). Деякі з них уже відновили роботу (на тих територіях, де не ведуться бойові дії). Зокрема частково відновив роботу Єдиний державний реєстр юридичних осіб, фізичних осіб – підприємців та громадських формувань (ЄДР), Державний реєстр речових прав на нерухоме майно, Єдиний державний реєстр судових рішень. Станом на травень 2022 року не працював Єдиний державний вебпортал відкритих даних (https://data.gov.ua); до порталів spending.gov.ua та openbudget.gov.ua доступ був, проте дані за останні місяці там не оновлювалися; закритий доступ до порталів відкритих даних місцевих органів влади.

Дослідниця зауважує, що це виправдано, бо «атаки як на основні державні сервіси, так і на місцеві йдуть майже безперервно. Чи то з метою створення перебоїв у роботі, чи то з наміром опублікувати фейкові звернення про капітуляцію, чи то для викрадення даних. Наприклад, під час російської кібератаки 15 травня на сервіси Львівської міської ради викрали частину робочих файлів міста та опублікували її на ворожих телеграм-каналах. Та й згадані раніше реєстри, хоча й відновили свою роботу, доступні лише для певного кола людей» [12].

Серед способів отриманням даних активізувався **краудсорсинг** – тобто збирання даних серед аудиторії за допомогою, наприклад, гугл-форм, ботів, інших форм опитувань. «Зазвичай журналістам радять сприймати такі дані як підказку, а не основне джерело, адже не можливо перевірити точність цих даних, та й вибірка не є повною. Однак під час повномасштабної агресії росії ця форма збору даних набула нового значення у формі взаємодії між урядовими структурами та громадянами. Йдеться про різноманітні чат-боти. Перш за все, ті, які збирають дані у громадян» [12]:

* **@stop\_russian\_war\_bot** – офіційний бот, за допомогою якого можна повідомляти про переміщення диверсантів, військ і ворожої техніки. Ця інформація передається у відповідні штаби реагування.
* **@ukraine\_avanger\_bot** – бот Кіберполіції для повідомлень про виявлені мітки, які залишають окупанти та  пересування російської армії.
* **@dytyna\_ne\_sama\_bot** – бот для допомоги дітям. Сюди можна вносити дані про дітей, яких побачили без дорослих, або про загублених дітей.
* **@RuAssetsBot** –  бот для збирання даних про російські активи в Україні або за кордоном.
* **@rusrocketfly\_bot** –  бот для фіксації підозрілих об’єктів у повітрі.
* **@SaveEconomyBot** – бот для фіксації необґрунтованих спекуляцій на цінах та інших економічних злочинів
* **@Traitor\_Search\_bot** –  бот від СБУ для виявлення зрадників та мародерів (@stop\_marauder\_ua\_bot).
* **єВорог** – бот від Міністерства цифрової трансформації для фіксації воєнних злочинів.

Також є боти, які не лише збирають дані, а власне побудовані на основі відкритих даних, постійно оновлюють цю інформацію та доносять її до користувачів:

* **@cancel\_pushkin\_bot** – бот, який розповідає що не так з конкретним російським культурним діячем у контексті ставлення до України та її культури..
* **@BoycottRussiaBot** – бот для перевірки товарів на предмет взаємозв’язків із російським ринком.

У статті В. Ковпак та К. Таточенко «Офіційні чат-боти Telegram, Viber як засоби публічних кризових комунікацій» зауважується на ефективній інновації Telegram і Viber в активізації роботи чатботів – програмах, що імітують діалог із живою людиною (установлення Telegram → перехід за необхідним посиланням → дотримання вказівок програми), що дає можливість сортувати інформацію за темами та рубриками, доступна в гаджеті цілодобово для допомоги в боротьбі проти ворога, зокрема отриманні чи наданні корисної інформації: про ворожі об’єкти, волонтерські пропозиції, пошук загиблих чи загублених, прихистки для людей у різних регіонах тощо, що є ефективними інструментами кризових комунікацій публічного управління, адже вони створюють ефект згуртованості та залученості кожного громадянина, який отримує можливість стати частиною кібервійська та зробити свій внесок у перемогу [13].

В офіційному telegram-каналі Верховної Ради України час від часу, як зауважила Я. Шекеряк, публікують бази даних російських урядовців, медійників, пропагандистів, із закликами писати певні повідомлення для інформаційної атаки, поширюють певні шаблони повідомлень для твіттеру з позначенням іноземних діячів з метою просування інформації про події в Україні; МВС України запустило сайт **200rf.com**, де розміщує фото та документи загиблих та полонених окупантів, щоб їхні родичі могли їх знайти, а Головне управління розвідки публікує дані російських військових, які воювали на території України та здійснювали тут воєнні злочини (наприклад дані особового складу бригади армійської авіації з іменами військовослужбовців ВПС росії, які виконували бойові вильоти з аеродрому Ліда Республіки Білорусь для нанесення авіаударів по Україні; списки бригад, які чинили звірства у Бучі) [12].

Журналісти даних, журналісти-розслідувачі, громадські діячі також використовують технологію добування військової, політичної, економічної та іншої інформації з відкритих джерел за методологією, що узагальнюють терміном **OSINT (**Open source intelligence), тобто розвідкою за відкритими джерелами (Google-карти, фотографії, відео, пости які є в мережі, відкриті супутникові дані тощо), якою користується, наприклад, міжнародна група незалежних журналістів Bellingcat, зокрема для висвітлення війни в Україні. Вони збирають та картують випадки шкоди цивільним в Україні (руйнування інфраструктури, загибель людей), спростовують маніпуляції про так звану фальсифікацію злочинів російської армії, зокрема у Бучі та Краматорську, використовуючи фото і відео, виставлені у мережу, новинні повідомлення, супутникові знімки (ще одне велике джерело відкритих даних для журналістів).

У 2014 році **проєкт InformNapalm** досліджує російську збройну агресію проти України й інших країн, публікуючи бази даних із доказами цих злочинів А зараз аналізує сучасні воєнні злочини росіян на нашій території та поширює їх ідентифікаційні дані. Волонтерська ініціатива – платформа **#WorldForUkraine** на інтерактивній карті наносить ініціативи на підтримку України у всьому світі (мітинги, флешмоби, протести).

Детальніше про публічні консультації та краудсорсингові проєкти як актуальний спосіб комунікації та співпраці через опрацювання великих даних можна ознайомитися в тезах В. Ковпак «Публічні консультації та краудсорсингові проєкти як актуальний спосіб комунікації та співпраці через опрацювання великих даних» [14].

**❓***Питання для самоконтролю:*

1. Назвіть ресурси, на яких оприлюднюють відкриті дані.

2. До яких позицій не може бути обмежений доступ до інформації?

3. Схарактеризуйте найактуальніші проєкти видання Texty.org на основі відкритих даних.

4. Назвіть успішні краудсорсингові формати як способи отриманням даних під час воєнного стану.

***Рекомендації до виконання завдань***: насамперед варто переглянути визначену до плану заняття *основну літературу* та *додаткову*. Важливо у процесі підготовки користуватися *вказаними посиланнями*, а також звертатися до *категоріального апарату*.

**✍***Завдання для самостійної роботи студента:*

1. Опрацювати теоретичний матеріал (прочитати, зробити корисні скріни, нотатки).

2. У реєстрі http://spending.gov.ua спробувати реалізувати пошук даних (напр., за ЄДРПОУ).

3. Зробити список офіційних чат-ботів, які з’явилися під час російсько-української війни, та спробувати скористатися тими, принаймні, що працюють з відкритими даними, з’ясувавши кілька кейсів (тобто реально скористуватись пошуковим інструментарієм цих ботів): @cancel\_pushkin\_bot – бот, який розповідає, що не так з конкретним російським культурним діячем у контексті ставлення до України та її культури.

@BoycottRussiaBot – бот для перевірки товарів на предмет взаємозв’язків із російським ринком.

3. Проаналізувати один з актуальних проєктів Texty.org.ua: визначити мету, інструменти роботи з відкритими даними, протестувати роботу інструментарію – зробити скріни, інфографіку щодо такого тестування.

4. Пройти один онлайн-курс на Prometheus на вибір: «Смарт-громада: управління на основі даних» (про важливість відкритих даних та принципи їх публікації: https://courses.prometheus.org.ua/courses/course-v1:CID+SC101+2020\_T1/course/) або «Доступ до публічної інформації: від А до Я» (про важливість доступу до публічної інформації: https://courses.prometheus.org.ua/courses/course-v1:COE+PI101+2017\_T1/course/).

**🕮 *Література:***

***Основна:***

1. Відкриті дані: формати і правила створення : посібник. *Texty.org.ua*. Київ. URL: https://texty.org.ua/pdf/data2017.pdf (дата звернення: 08.04.2023).

2. Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних, що затверджено Постановою Кабінету Міністрів України від 21 жовтня 2015 р. № 835. URL: https://zakon.rada.gov.ua/laws/show/835-2015-%D0%BF/paran12#n12 (дата звернення: 08.04.2023).

***Додаткова:***

3. Різниця між даними та інформацією. URL: http://surl.li/fxfwz (дата звернення: 08.04.2023).

4. Кобиляцька В. Публічна інформація у форматі відкритих даних: чому це важливо і як її отримати? *Доступ до правди.* 27 квітня 2020. URL: https://dostup.pravda.com.ua/news/publications/publichna-informatsiia-u-formati-vidkrytykh-danykh-chomu-tse-vazhlyvo-i-iak-ii-otrymaty (дата звернення: 08.04.2023).

5. Доступ до публічної інформації: від А до Я. *Prometheus*. URL: https://courses.prometheus.org.ua/courses/course-v1:COE+PI101+2017\_T1/course (дата звернення: 08.04.2023).

6. Навак Н. Корисний карантин: книги і курси з open data, data science та data visualization. 2020. URL: https://mind.ua/openmind/20209468-korisnij-karantin-knigi-i-kursi-z-open-data-data-science-ta-data-visualization (дата звернення: 08.04.2023).

7. Texty.org.ua. Про нас. URL: https://texty.org.ua/p/about/ (дата звернення: 08.04.2023).

8. Василюк Д. Дані нам: як великі цифри змінюють журналістику. *Platfor.ma.* URL: https://platfor.ma/magazine/text-sq/media-innovations-lab/kayirobondarenko/ (дата звернення: 08.04.2023).

9. Губашов Д. Карта руйнувань. 21.03.2023. URL: https://texty.org.ua/projects/109019/karta-rujnuvan/ (дата звернення: 08.04.2023).

10. Тимошенко О., Кельм Н., Солодько П. Карта бойових дій. Оновлюється. 29.03.2022. *Texty.org*. URL: https://texty.org.ua/projects/106137/karta-bojovyh-dij-onovlyuyetsya/ (дата звернення: 08.04.2023).

11. Відкрита медреформа. Проєкт на основі відкритих даних НСЗУ. *Texty.org.* URL: https://texty.org.ua/d/2019/medical\_reform/?fbclid=IwAR1\_d\_3SdCMCZ7IIaoPuWFDE08Tkxz4xqoHMkPnogynFzsFJ75n19vlbGRo#/apteky (дата звернення: 08.04.2023).

12. Шекеряк Я. Карти, боти й пошук людей – як журналісти й урядові структури використовують відкриті дані під час воєнного стану. *Медіакритика.* URL: https://www.mediakrytyka.info/novi-tehnologii-media/karty-boty-y-poshuk-lyudey-yak-zhurnalisty-y-uryadovi-struktury-vykorystovuyut-vidkryti-dani-pid-chas-voyennoho-stanu.html (дата звернення: 08.04.2023).

13. Ковпак В., Таточенко К. Офіційні чат-боти Telegram, Viber як засоби публічних кризових комунікацій. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія Філологія. Соціальні комунікації.* 2022. Т. 33 (72). № 5. С. 211–217.

14. Ковпак В. Публічні консультації та краудсорсингові проєкти як актуальний спосіб комунікації та співпраці через опрацювання великих даних. *Суспільство і особистість у сучасному комунікаційному дискурсі :* матеріали ІV Всеукраїнської наук.-практичної конференції.2 листопада 2022 р., м. Запоріжжя / редкол. : В.Л. Погребна, В.В. Кузьмін, Н.В. Островська, Т.О. Бородулькіна та ін. Запоріжжя : НУ «Запорізька політехніка», 2022. С. 229–233. URL: https://zp.edu.ua/uploads/dept\_s&r/2022/conf/2.1/Materialy\_SiOuSKD-4.pdf (дата звернення: 08.04.2023).

***Тема* 2. Методологічні засади статистики. Статистичне спостереження**

***Мета*:** визначати типи даних, користуватися основними категоріями статистики, статистичного спостереження, формувати матрицю даних.

*План*

1. Процес аналізу даних. Типи даних. Матриця даних.

2. Види статистики. Основні категорії статистики.

3. Статистичне спостереження: форми, види, способи.

**Основні поняття:** статистика, статистичне спостереження, матриця даних, типи даних, середнє значення, мода, медіана

**📚 Методичні рекомендації та поради**

У процесі підготовки самостійної роботи студентам необхідно взяти до уваги таку інформацію.

**Аналіз даних** – це процес роботи з даними з метою їх правильного розміщення, пояснення, презентабельності та пошуку висновку із цих даних для винайдення корисної інформації для прийняття раціональних рішень.

Відповідно, **основною метою аналізу даних** є інтерпретація, оцінка, організація даних, а також їх ефективна та ефектна презентація.

**Процес аналізу даних** включає:

* Збір даних;
* Робота над якістю даних;
* Побудова моделі;
* Навчальна модель;
* Запуск моделі з повними даними.

Деякі поради щодо аналізу даних:

* Видаліть непотрібні дані перед аналізом.
* Не слід проводити аналіз на головній копії даних.

За даними спеціалізованого порталу myservername.com, варто розрізняти аналіз даних, видобуток даних та моделювання даних.

«**Аналіз даних** проводиться з метою пошуку відповідей на конкретні питання. Методи аналізу даних подібні до бізнес-аналітики та бізнес-аналітики.

**Видобуток даних** – це пошук різних моделей даних. Для цього до даних застосовуються різні математичні та обчислювальні алгоритми й генеруються нові дані.

**Моделювання даних** – це те, як компанії впорядковують дані або керують ними. Тут до даних застосовуються різні методології та методи. Для моделювання даних необхідний аналіз даних» [3].

Порівняння найкращих інструментів аналізу даних репрезентовано у Додатку А.

Як засвідчує А. Корнілова, для аналізу даних необхідно три основні компоненти:

* знання предметної області (це дозволяє розуміти, які проблеми потребують першочергового вирішення);
* знання математики та статистики (дозволяють формалізувати рішення, перевести його в алгоритм та оцінити, яка ймовірність отримати результат);
* вміння програмувати (дає можливість застосовувати величезні обчислювальні потужності) [1].



Рис. 6. Компоненти, необхідні для аналізу даних. Джерело: [1]

**Процес аналізу даних складається з трьох етапів.**

«Спочатку дані потрібно підготувати, тобто зібрати, очистити та відібрати ті, які потрібні для моделі. Цей процес займає близько 90% часу. Далі ми будуємо модель та валідуємо її результати. Останній етап – це презентація результатів. Тут ми демонструємо на яке питання ми шукали відповідь, які дані використовували та що отримали в результаті. Для того щоб це зробити максимально ефективно треба витрати ще 90% часу» [1].

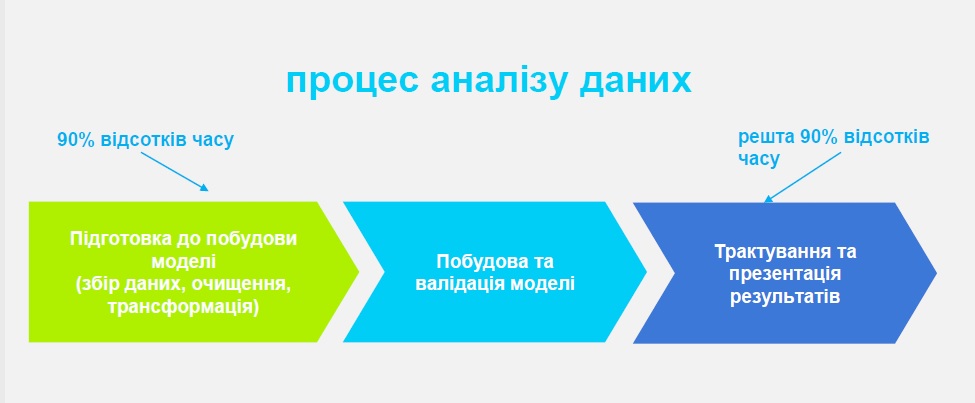


Рис. 7. Етапи процесу аналізу даних. Джерело: [1]

Існує **два методи аналізу даних**: якісний аналіз та кількісний аналіз.

* Якісний аналіз: якісний аналіз проводиться за допомогою інтерв’ю та спостережень.
* Кількісний аналіз: кількісний аналіз проводиться за допомогою опитувань та експериментів.

Відповідно, поняття «**статистика**» (від лат. status – стан речей) означає кількісний облік масових, насамперед соціально-економічних явищ і процесів.

Як окрема галузь науки статистика виникла з повсякденних практичних потреб людей, оскільки для державного управління потрібна інформація, наприклад, про наявність і склад земель, чисельність і склад населення, стан торгівлі, доходи і витрати населення, рівень і динаміку цін на товари та послуги, розвиток матеріального виробництва тощо; тобто як наука статистика об’єднує принципи та методи роботи з масовими числовими даними, тобто кількісними характеристиками зазначених явищ і процесів.

*Об’єкт її вивчення* – соціальні, економічні, політичні та культурні явища і процеси суспільного життя.

*Предмет* – розміри і кількісні співвідношення між масовими суспільними величинами, закономірність їх формування, розвитку і взаємозв’язку. Тобто, по-перше, статистика вивчає кількісний бік суспільних явищ, а по-друге, вона вивчає не поодинокі, а масові явища.

Статистика в сучасних умовах забезпечує державні органи всіх рівнів інформаційно аналітичними матеріалами, на основі яких розробляється податкова та цінова політика, приймаються засоби зі стимулювання або стримання розвитку ринку та окремих його сегментів, забезпечується соціальний захист населення і т. д. Комерційні підрозділи та спеціалізовані маркетингові фірми, поряд з обов’язковою звітністю, формують панелі споживачів, займаються опитуванням та анкетуванням, тобто всіма засобами вивчають цільову авдиторію та конкурентів за всіма параметрами.

Статистичні органи на комерційній основі надають фірмам, установам, організаціям статистичну інформацію – ряд відомостей, які утворюють, наприклад, маркетингове середовище (демографічні дані споживачів, індекси цін, індекси вартості життя тощо); формують прогнози, оцінки

**Основні категорії статистики**, за допомогою яких відображаються найбільш узагальнювальні та суттєві властивості, ознаки, зв’язки та відношення предметів та явищ (у базових наукових розвідках їх вивчали В.Л. Бек, А.Г. Гончарук, В.К. Горькавий):

1. **Статистична** [**сукупність**](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2440&displayformat=dictionary) – це певна множина елементів, поєднаних умовами існування і розвитку. Так, наприклад, статистичною є [сукупність](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2440&displayformat=dictionary) домогосподарств, [сукупність](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2440&displayformat=dictionary) родин, [сукупність](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2440&displayformat=dictionary) підприємств, фірм, об’єднань тощо. Сукупності можуть бути *однорідними* (якщо одна або декілька суттєвих ознак її об’єктів є узагальнювальними для всіх [одиниць](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2454&displayformat=dictionary)) та *різнорідними* (до сукупності входять явища різного типу).

* **Одиниця сукупності –** окремі об’єкти (елементи, явища тощо), які становлять статистичну [сукупність](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2440&displayformat=dictionary), характеризуються індивідуальними ознаками.

Так, наприклад, ознаками торгівельного підприємства є розміри основних та оборотних засобів, обсяг та структура реалізованої продукції тощо. Ознаками продукції є якість, термін зберігання, вартість і т. п.

Одні ознаки виражають числами, а інші – вербально. Їх називають відповідно якісними і атрибутивними (описовими).

Особливістю **статистичного дослідження** є те, що в ньому вивчаються тільки [варіативні ознаки](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=3383&displayformat=dictionary), тобто які мають в межах сукупності різні значення. Відмінність, коливання значення ознаки називається **варіацією.** Наприклад, ознаки підприємства: спеціалізація, форма власності, [рентабельність](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=3461&displayformat=dictionary) виробництва тощо.

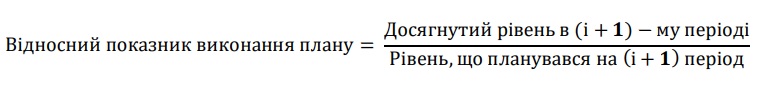
Якщо ж зміна явища відбувається в різні періоди часу, зберігаючи рівномірний характер, то говорять вже не про варіацію ознаки, а про її **динаміку**. **Статистичні показники** можуть бути *об’ємними* (чисельність населення, трудових ресурсів) та *розрахунковими* (середні величини), *плановими*, *звітними*, *прогностичними*.

За формою і змістом статистичні показники поділяються на *абсолютні* та *відносні*. Абсолютні показники характеризують розмір суспільних явищ, наприклад, розмір посівних площ, обсяг виробленої та реалізованої продукції, суму прибутку, вартість основних фондів, чисельність робітників, фонд заробітної плати підприємства тощо. Вони відповідають на запитання «скільки?» і завжди відображаються іменованими числами (га, ц, т, грн). Відносні показники відображають кількісні співвідношення між абсолютними показниками (напр. у відсотках, коефіцієнтах).

При розрахунку відносного показника абсолютний показник, що знаходиться в чисельнику одержуваного відносного, називається *поточним або порівнюваним.* Показник же, з яким проводиться порівняння і який знаходиться в знаменнику, називається *підставою або базою порівняння*. Таким чином, розраховувана відносна величина показує, у скільки разів порівнюваний абсолютний показник більше базисного, або яку становить від нього частку, або скільки одиниць першого припадає на 1, 100, 1000 і т. д. одиниць другої.

Так, собівартість виробництва 1 т продукції, грн (відносний показник) отримують внаслідок ділення собівартості виробництва всієї продукції, грн на кількість виробленої продукції, т.

Відносний показник виконання плану – це відношення фактично досягнутого рівня до планового завдання. У нижче наведеному прикладі показник відображає фактичний обсяг виробництва або реалізації у відсотках або коефіцієнтах порівняно з плановим рівнем.



**Відносні показники мають велике аналітичне значення**, їх обчислюють, щоб мати кількісну характеристику різноманітних сторін суспільного життя. За їхньою допомогою виражають ступінь виконання планів, ефективність та інтенсивність суспільного виробництва, продуктивність праці, ступінь задоволення матеріальних і культурних потреб людей, структуру і динаміку виробництва тощо. Завдяки ним можна порівнювати однойменні і різнойменні величини.

**Система статистичних показників** – це [сукупність](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2440&displayformat=dictionary) статистичних показників, що відображає взаємозв’язків, які об’єктивно існують між явищами.

Система статистичних показників охоплює всі напрями життя суспільства на різних рівнях: країни, регіонів – макрорівень; підприємств, фірм, об’єднань, родин, домогосподарств і т. п. – мікрорівень.

Системи статистичних показників мають такі особливості:

* вони мають історичний характер – змінюються умови життя населення, суспільства, змінюються і системи статистичних показників;
* методологія розрахунку статистичних показників безперервно удосконалюється.

Існують **три групи методів статистичного дослідження**:

[**Статистичне спостереження**](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=3469&displayformat=dictionary) – дає інформаційну базу для статистичних узагальнень та характеристики об’єктивних [закономірностей](https://cdn.snau.edu.ua/moodle/mod/glossary/showentry.php?eid=2447&displayformat=dictionary) ( тобто на цьому етапі здійснюється збирання первинного статистичного матеріалу). Це перший етап будь-якого статистичного дослідження.

**Метод групувань** – другий етап статистичного дослідження. На цій стадії зібрані факти класифікуються і систематизуються.

**Метод узагальнювальних показників** дає змогу надати характеристику досліджуваних явищ та процесів за допомогою статистичних величин – абсолютних, відносних, середніх. На цьому етапі статистичного дослідження виявляються взаємозв’язки та масштаби явищ, визначаються закономірності їх розвитку, даються оцінки на майбутнє.

**Закономірність** – це повторюваність, послідовність та порядок у розвитку соціальних явищ. Вона може проявлятися по-різному. Розрізняють два види закономірностей: *динамічну* та *статистичну.*

У природних явищах спостерігаються закономірності, які звуться динамічними, тобто певні ознаки того чи іншого явища спостерігаються у кожному окремому його прояві, зокрема, у фізичних, хімічних, математичних законах. Скажімо, площа кола змінюється на відповідну величину у зв’язку зі зміною його радіусу, закон Архімеда проявляється у кожному випадку занурення тіла у рідину тощо.

У суспільному житті в соціальних явищах такої закономірності не спостерігається, хоча вони в таких явищах теж існують. Але свій прояв закономірності соціальних явищ знаходять при масовому і послідовному спостереженні, і тому такі спостереження звуться *статистичними,* щоне виражають властивості кожного явища окремо, а тільки в масових процесах або в загальній системі.

**Статистична закономірність** – це така, яка виявляється лише в достатній кількості однорідних одиничних елементів, котрі й утворюють сукупність. Тобто кожний окремий елемент може не підтверджувати існування тієї чи іншої закономірності, тому що існування її в кожному елементі носить імовірний характер. Інакше кажучи, статистична закономірність властива лише сукупності одиниць, яка має назву статистичної сукупності, проявляється лише на підставі дослідження достатньо великої кількості одиниць спостереження. Наприклад, залежність злочинності від доходів родини, культурного рівня населення тощо.

Отже, як зауважує А. Корнілова, «статистика допомагає оцінити варіативність та зменшити невизначеність. Розрізняють описову та вивідну статистики.

* Описова – вивчає властивості спостережуваних даних.
* Вивідна статистика – виводимо припущення про властивості розподілу даних, з яких походять спостережувані дані.

За допомогою статистики можна дати відповідь на питання:

- чи є залежність між кількістю злочинів та фазою Місяця?

- яка ймовірність викликати Uber в Києві?

- побудувати довірчий інтервал часу, за який ви потрапляєте на роботу

- проводити опитування та трактувати їх результати» [1].



Рис. 7. Типи даних. Джерело: [1]

Матриця даних – «стартовий елемент для аналізу даних. Зазвичай йому передує етап збору, очищення та представлення у табличному вигляді. По рядках – респонденти, суб’єкти, учасники, спостереження По стовпцях – xарактеристики кожного запису (змінні). Також важливо звертати увагу на одиниці виміру, а також яким чином були зібрані ці дані. Ця таблиця включає в себе 6 рядків, однак зібрані дані мають майже 800 спостережень (спостереження зібрані з ресурсу і містять інформацію про квартири, які продаються). Для того, щоб описати вміст цієї таблиці в більш зрозумілій формі, використовують узагальнення та опис типових чи середніх значень. Для цього важливо знати тип даних» [1] (див. Рис. 8).



Рис. 8. Матриця даних. Джерело: [1]

Для узагальнення категоріальних даних використовують частотні таблиці. Наступна таблиця, презентована в рамках онлайн-курсу «Аналіз даних та статистичне виведення на мові R», містить кількість квартир, що продаються у кожному місті [1] (див. Рис. 9).



Рис. 9. Частотні таблиці. Джерело: [1]

**Центральна тенденція**. «Центральне або типове значення дозволяє зрозуміти основну характеристику даних.

**Середнє значення** підходить для узагальнення кількісних даних (як дискретних, так і неперервних). Формула обрахування проста: суму всіх чисел ділимо на їх кількість. Наприклад, якщо в нас є група з 5 учнів, оцінки яких 12, 3, 5, 10, 5. Сума їх оцінок дорівнює 35, а середнє значення 7. Однак із використанням середнього значення як опис центральної тенденції в даних є невелика проблема. Якщо є нетипово великі чи малі для даного набору значення – вони роблять великий внесок у значення середнього. Нехай у нас є певне невелике підприємство, яке має 5 працівників. Заробітні плати працівників в гривнях: 5000, 7000, 2000, 4000, 50 000. Середнє значення заробітної плати 13600 грн. Однак, якщо ми відкинемо екстремальне значення 50 000, то отримаємо, що середнє значення зменшилося до 4500» [1] (див. Рис. 10).

.

Рис. 10. Середнє значення. Джерело: [1]

**Медіана** – «це значення, яке ділить вибірку навпіл, тобто 50% є меншими за це значення, 50% більшими. Основна перевага використання медіани – менша чутливість до екстремальних значень. Для пошуку медіани дані треба розмістити в зростаючому порядку та поділити на дві частини. Якщо в нас парна кількість спостережень, то сусідні значення по краях сумуються та діляться на два. У випадку попереднього прикладу із заробіною платою: 2000, 4000, 5000, 7000, 50 000. Маємо, що по середині знаходиться значення 5000, то краще описує центральну тенденцію заробітної плати на підприємстві» [1] (див. Рис. 11).

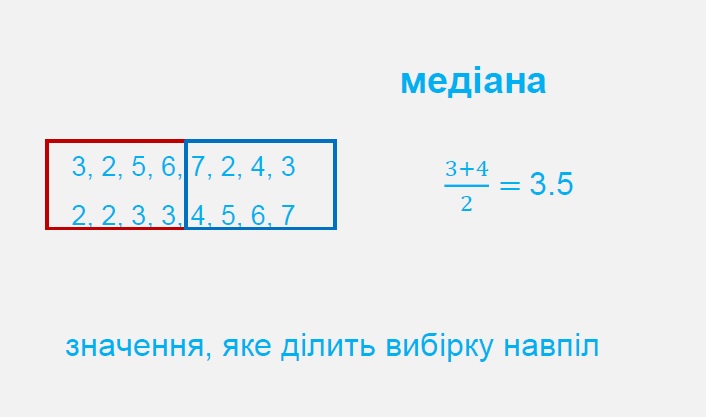


Рис. 11. Середнє значення. Джерело: [1]

**«Мода** використовується для визначення центральної тенденції категоріальних або кількісних дискретних даних. Мода – це значення, яке найчастіше трапляється. Наприклад, за інформацією Міністерства юстиції України у певний відрізок часу хлопчиків найчастіше називали Дмитром, Артемом, Максимом та Іваном, дівчаток – Анею, Анастасією, Софією та Дар’єю. (http://tyzhden.ua/News/137908). Ці імена були модою серед всіх імен» [1] (див. Рис. 12).



Рис. 12. Мода. Джерело: [1]

Власне, у статистичній практиці застосовують дві організаційні форми спостереження – **звітність** і **спеціально організовані статистичні спостереження** [2, с.19].

*Звітність* – це форма спостереження, при якій кожний суб’єкт діяльності подає свої дані в державні органи статистики та відомства у вигляді документів, звітів спеціально затвердженої форми.

*Спеціально організоване спостереження* – це форма спостереження, яка охоплює сфери життя та діяльності, що не відображуються звітністю. До числа таких спостережень належать: переписи, обліки, спеціальні обстеження, опитування.

*Перепис* – суцільне або вибіркове спостереження окремих масових явищ з метою вивчення їхнього розміру та складу на певну дату.

*Обліки* – суцільне спостереження окремих масових явищ згідно з певною тематикою, що виходить за межі звітності; можуть бути періодичними або одноразовими.

*Спеціальні обстеження* – несуцільні спостереження окремих масових явищ згідно з певною тематикою, що виходить за межі звітності; можуть бути періодичними та одноразовими.

*Опитування* – це, як правило, несуцільне спостереження за думками, мотивами, оцінками, що реєструються зі слів респондентів

**❓***Питання для самоконтролю:*

1. Назвіть етапи процесу аналізу даних.

2. Розкрийте основні категорії статистики та статистичного дослідження.

3. Які організаційні форми спостереження застосовують у статистичній практиці?

4. Які формули вирахування середнього значення, моди та медіани даних?

***Рекомендації до виконання завдань***: насамперед варто переглянути визначену до плану заняття *основну літературу* та *додаткову*. Важливо у процесі підготовки користуватисяматеріалами онлайн-курсу «*Аналіз даних та статистичне виведення на мові R» https://courses.prometheus.org.ua/courses/IRF/Stat101/2016\_T3/about*, а також звертатися до *категоріального апарату*.

**✍** *Завдання для самостійної роботи студента:*

1. Виконати завдання із визначення середнього значення, медіани, моди в заданій матриці даних (тема та дані обираються довільно).

2. Визначити актуальні тренди в конкретному регіоні в режимі реального часу. Скористайтеся інструментом «Трендові запити» для пошуку популярних тем та інфоприводів для інтернет-видань, тематичних контент-платформ та нішевих медіа: http://surl.li/gdawt.

3. Опрацювати статтю «Статистика відвідувань сайту»: https://wezom.com.ua/ua/blog/statistika-poseshchenij-sajta.

Скористатися одним із запропонованих інструментів.

Або проаналізувати статистику своєї сторінки в соціальній мережі (напр., за інструментарієм, описаним за посиланням: https://www.facebook.com/help/268680253165747?cms\_platform=www&helpref=platform\_switcher)

4. Апробувати ресурси з визначення індексів, рейтингів тощо, сформувати інформаційну довідку щодо знайдених статистичних показників, даних (у списку рекомендованої літератури [8–20].

**🕮 *Література:***

***Основна:***

1. Корнілова А. Prometheus. Аналіз даних та статистичне виведення на мові R : конспект лекцій. URL: http://surl.li/gdbhf (дата звернення: 08.04.2023).

2. Соціальна статистика : конспект лекцій для здобувачів вищої освіти першого (бакалаврського) рівня зі спеціальності 054 «Соціологія» / укладач А.В. Хмелюк. Кам’янське : ДДТУ, 2020. 155 с.

***Додаткова:***

3. 10 найкращих інструментів аналізу даних для ідеального управління даними [Список 2021]. URL: https://uk.myservername.com/10-best-data-analysis-tools (дата звернення: 08.04.2023).

4. Job Comparison – Data Scientist vs Data Engineer vs Statistician. Published On October 19, 2015 and Last Modified On June 24th, 2019. Analytics vidhya. URL: https://www.analyticsvidhya.com/blog/2015/10/job-comparison-data-scientist-data-engineer-statistician/ (дата звернення: 08.04.2023).

5. Nathan Yau. Flowing Data. URL: https://flowingdata.com/ (дата звернення: 08.04.2023).

6. 100 interesting data sets for statistics. URL: https://rs.io/100-interesting-data-sets-for-statistics/ (дата звернення: 08.04.2023).

7. R for Data Science. URL: https://r4ds.had.co.nz/index.html (дата звернення: 08.04.2023).

**Інформаційні ресурси:**

8. Державна служба статистики України. Статистика щодо економічної ситуації в Україні: ВВП, інфляція, безробіття, експорт / імпорт, реєстр підприємств, сільське господарство, будівництво, регіональна статистика та ін. URL: http://www.ukrstat.gov.ua/ (дата звернення: 08.04.2023).

9 Державний і місцеві бюджети: доходи, витрати, обслуговування , державного боргу, виконання бюджету. URL: http://www.treasury.gov.ua/ (дата звернення: 08.04.2023).

10. Євростат. Статистика з основних макропоказників в СС (ВВП, інфляція, безробіття та ін.). URL: http://ec.europa.eu/eurostat (дата звернення: 08.04.2023).

11. Єдиний веб-портал використання публічних коштів. URL: https://spending.gov.ua/new/ (дата звернення: 08.04.2023).

12. Індекс добробуту по країнах. URL: http://worldhappiness.report/ (дата звернення: 08.04.2023).

13. Індекс політичного розвитку в державах світу. URL: https://ec.europa.eu/ (дата звернення: 08.04.2023).

14. Індекс розвитку свободи слова. URL: https://rsf.org (дата звернення: 08.04.2023).

15. Індекс свободи ведення бізнесу, умови підприємництва в країнах. URL: http://www.heritage.org (дата звернення: 08.04.2023).

16. Рейтинг державних фондів. URL: https://www.fitchratings.com (дата звернення: 08.04.2023).

17. Статистика кількості та якості земельних ресурсів у світі. URL: http://www.iiasa.ac.at/ (дата звернення: 08.04.2023).

18. Статистика світових цін на газ, нафту. URL: https://www.eni.com/.

19. NGO: Energy Policy Group. Статистика з енергетичного ринку в СС, газринок. URL: http://energywatchgroup.org (дата звернення: 08.04.2023).

20. Тіньова економіка, офшори. URL: http://www.gfmtegrity.org/ (дата звернення: 08.04.2023).

***Тема* 3. Життєвий цикл даних. Техніки оброблення даних**

***Мета:*** сформувати навички визначення життєвого циклу даних як забезпечення результативного управління даними та їхнього збереження, вміти визначати мету, предмет, контекст, стейкхолдерів заходу, що передбачає оброблення даних, користуватися інструментами збору та аналізу даних у випадку офлайн- та онлайн-консультацій.

*План*

1. Життєвий цикл даних як забезпечення результативного управління даними та їхнього збереження (детальний аналіз версії, адаптованої до специфіки публічних консультацій в інформаційній діяльності).

2. Техніки оброблення даних: фільтрування; інтегрування; трансформування; узагальнення.

***Основні поняття:*** життєвий цикл даних, публічні консультації в інформаційній діяльності, техніки оброблення даних, фільтрування, інтегрування, трансформування, узагальнення

**📚 Методичні рекомендації та поради**

У процесі підготовки самостійної роботи студентам необхідно взяти до уваги таку інформацію.

Життєвий цикл даних як забезпечення результативного управління даними та їхнього збереження пропонуємо розглядати через детальний аналіз версії, адаптованої до специфіки публічних консультацій в інформаційній діяльності.

«В умовах інформаційної війни та повномасштабного вторгнення росії в Україну питання доступу до даних набуло статусу національної безпеки (зокрема, кібератаки на основні державні реєстри та портали для перешкоджання роботи, публікації фейкових звернень, викрадення даних тощо) та, разом із тим, механізму ефективної співпраці держави та громадян (наприклад, краудсорсинг у форматах роботи чат-ботів, опитувань через гуглформу тощо)» [5, с. 229].

Серед найефективніших платформ, що використовують відкриті дані, варто згадати «Канцелярську сотню», «яка розробляє успішні проєкти у сфері прозорості та протидії корупції, використовуючи відкриті дані і NLP» [7]. Наприклад, для відновлення документів В. Януковича та С. Курченка, що було знищено руками або пропущено через шредер та перемішано, було запущено процес відсканування шматочків документів волонтерами із застосуванням технології Computer Vision. Із відкриттям Міністерством юстиції реєстру нерухомості у січні 2015 року запрацював проєкт «Гарна хата», що «дало можливість відправляти запит із адресою і за невеликі гроші (17 грн) отримувати інформацію, хто має право власності на це житло…спочатку за допомогою краудсорсингу дослідили найбільш елітні адреси Києва – закриті вулички, новітні будівлі в історичній зоні, найдорожчі мікрорайони. Заплатили кошти, зібрали виписки з цих адрес, потім обробили ці дані, перевели їх у машиночитний формат… видалили особисті дані та виклали цей датасет у вільний доступ на сайт проєкту» [7]. Масове оцифрування декларацій українських чиновників, що розпочалося у зв’язку із необхідністю перевірки декларацій, яку мав робити ДФС, у процесі люстраційної перевірки (зокрема, у 2014 році Громадський люстраційний комітет оприлюднив паперові декларації), сформував у проєкті потужний датасет, адже надійшло більше 3 тис. декларацій у PDF з різних джерел, різних за параметрами, з некоректними назвами, а в результаті був створений сайт declarations.com.ua, де організували повноцінний пошук за даними з декларацій. «Ми замінили Google-форму для волонтерів на більш зручний інструмент – платформу для краудсорсингу під назвою “Вулик”, яку розробили самостійно. Вона включала в себе більш зручну форму з безліччю перевірок, підказок, корекцій, автозаповненням. Таким чином, за 2 роки ми розшифрували вручну більше 20 000 декларацій за допомогою 3 тис. волонтерів та 2 редакторів…Так було зібрано перший в Україні відкритий машиночитний датасет з інформацією про доходи та статки чиновників», – зазначає Д. Чаплинський, волонтер «Канцелярської сотні» [7]. Безперечно, ці дані стали незамінним інформаційним джерелом для журналістів-розслідувачів.

У контексті залучення громадськості та збору великих даних, високоякісних та актуальних, щодо концепції впровадженої політики, нового законопроєкту, упровадження нової послуги чи розробки проєкту значну роль відіграють **публічні консультації** як «формальний циклічний процес, ініційований урядом, у межах якого проінформовані громадяни беруть участь у змістовних дискусіях з метою впливу на політику уряду щодо питань, які їх (стейкхолдерів) стосуються» [1]. Зокрема, **складники консультацій можна розглядати як дані, а саме**: поставлені запитання; надані учасниками відповіді; характеристики консультацій, наприклад, дата, час та місце проведення заходу або кількість учасників офлайн та онлайн-консультацій; інформація про учасників, зокрема імʼя, вік, стать тощо.



Рис. 13. Матриця публічних консультацій. Джерело: [1]

«Використання підходу, орієнтованого на дані, допоможе отримати великі обсяги даних (у різних форматах і з різних джерел) та систематично проаналізувати їх, щоб одержати корисні знання. Він упроваджений у цикл публічних консультацій таким чином, що дозволяє перейти від простого збору необроблених даних від громадян до отримання цінної інформації з їхніх відповідей та генерування дієвих знань, корисних для розробників політики» [1] (див. Рис. 14). Серед **ключових показників для оцінювання успішності таких консультацій** автори посібника «Публічні консультації: крок за кроком» пропонують такі: 1) *Представленість учасників консультацій* (залучення усіх груп стейкхолдерів, з різними інтересами й поглядами, що забезпечить мобілізацію та вислуховування всіх, що особливо важливо для онлайн-консультацій, де потрібно використати додаткові ресурси для залучення людей з обмеженим доступом до інтернету й електронних пристроїв чи тих, хто менше ними користується (наприклад, людей старшого віку). Такі ознаки як місце проживання, вік, стать, культурна належність, наявність професійних або експертних знань допоможуть оцінити, наскільки учасники консультацій представляють цільову аудиторію. Хоч варто пам’ятати, що неможливо досягти повної репрезентативності населення. 2) *Якість наданої учасникам інформації*. (Попереднє інформування стейкхолдерів забезпечує їхню змістовну участь. Чітку, актуальну та повну інформацію про політику та про те, як долучитися до консультацій, потрібно надавати в різних форматах зрозумілою учасникам мовою. 3) *Актуальність внеску учасників* (варто спонукати учасників до конкретики: запитання, поставлені під час консультацій, є настільки ж важливими, як і попередня інформація, надана учасникам для підготовки. Запитання, які передбачають чіткі пропозиції та рекомендації, дають більше користі при ухваленні рішень, ніж запитання, які ведуть до розмитих відповідей). 4) *Якість отриманих даних та їхній аналіз* (здатність чітко зафіксувати, звести дані та проаналізувати обговорення визначатиме користь публічних консультацій для формування політики. Без даних неможливо зрозуміти, що непокоїть громадян, а також зробити висновки, потрібні для ухвалення рішень) [1]. А уряд Канади користується підходом так званого *«відкритого діалогу»*, що передбачає в контексті якісної взаємодії з громадськістю дотримання прозорості, актуальності, інклюзивності, підзвітності та адаптивності.



Рис. 14. Публічні консультації, орієнтовані на дані. Джерело: [1]

У матеріалі «Публічні консультації та краудсорсингові проєкти як актуальний спосіб комунікації та співпраці через опрацювання великих даних» [5] ми розглядаємо приклади ефективних публічних консультацій. Зокрема, у грудні 2021 року Національна рада України з питань телебачення та радіомовлення ухвалила рішення провести публічні консультації з громадськими та галузевими організаціями для обговорення проєкту Порядку проведення перевірки ділової репутації покупця (потенційного покупця) засобу масової інформації (Порядок), що розробляється на виконання Закону України «Про запобігання загрозам національній безпеці, пов’язаним із надмірним впливом осіб, які мають значну економічну та політичну вагу в суспільному житті (олігархів)» [6]. У 2017 році громадська організація «Центр UA» (співзасновники «ЧЕСНО», Реанімаційного пакету реформ) провела 160 публічних консультацій у 63 містах України [2]. Зокрема, у Полтавській області, за свідченнями А. Гончаренко, координаторки ГО «Центр UA» в Полтавській області, перша в місті публічна консультація була покликана унеможливити непрозорий збір батьківських внесків на потреби полтавських шкіл та дитячих садочків, дійти згоди щодо формату розвитку Прирічкового парку, розробити механізми захисту об’єктів історичної спадщини міста від рук недоброчесних орендарів чи власників.

У вересні 2022 року у Дрогобичі тривали публічні консультації та збір громадських думок щодо питання демонтажу «Вічного вогню» як символу тоталітарного комуністичного режиму та подальшого благоустрою даної території: фізичні особи, представники громадських та політичних об’єднань надсилали електронні листи, письмові пропозиції або ж коментарі на сервісі електронних консультацій, що лягли в основу обговорення робочою групою [4].

У продовження ще більш актуалізованого в умовах війни обговорення концепту декомунізації та дерусифікації в громадах, варто згадати, наприклад, як після консультацій у червні 2022 року, наданих О. Пустовгаром, регіональним представником Українського інституту національної пам’яті, заступником директора Путивльського державного історико-культурного заповідника в Сумській області, де є «Парк радянського періоду», та спеціалістами з охорони культурної спадщини, гадяцька громада дійшла спільної думки по закінченню війни створити на базі Гадяцького історико-краєзнавчого музею експозицію під відкритим небом «Територія терору», де будуть зібрані зразки монументальної пропаганди радянського тоталітарного режиму [3]. Міські ради також активно залучають містян до участі в публічних консультаціях щодо перейменування вулиць.

Отже, «опрацювання та аналіз великих даних у різних форматах, отриманих у результаті співпраці держави та громадян – публічних консультацій, проєктних форм для одержання корисних знань – контекстуалізованої інформації на основі категоризованих даних має стати в умовах війни механізмом постійної комунікації, що сприяє залученості стейкхолдерів до просування та реалізації нової політики» [5, с. 232].

Тож для прототипування опрацювання та аналізу даних у рамках певної публічної консультації варто слідувати алгоритму **7 етапів циклу публічних консультацій** (за посібником «Публічні консультації: крок за кроком»):

1. **Визначити цілі** з урахуванням *предмета* й *контексту* проведення консультацій, стейкхолдерів та показників успішності. Для цього варто розробити план заходу із зазначенням ключових етапів, із визначенням ролі та обов’язків кожного члена команди у конкретних часових та бюджетних межах.

Обов’язково слід осмислити контекст проведення консультацій. Почато варто із ознайомлення з текстом концепції політики заходу, аналізу поточного стану речей та прогностичного впливу нової концепції на цільову авдиторію, обставини (на цьому етапі варто отримати цінні відомості від експертного середовища).

І тоді можна говорити про формулювання мети проведення консультацій з перспективою обрання вибору їх формату, залучених стейкхолдерів. «Наприклад, консультації можуть проводити, щоб отримати експертні думки про технічні зміни до політики, або вони можуть мати набагато ширшу мету — наприклад, залучення загалу для збору ідей удосконалення державної установи.

Консультації можуть відбуватися на початкових етапах процесу формування політики, як у цьому сценарії, або пізніше – щоб зрозуміти занепокоєння громадян щодо нової політики» [1]. Загалом, у рисунку 13 окреслено простір цільових прицілів консультацій – *інформування* громадськості (напр., День відкритих дверей, комунікації через соціальні мережі тощо); *консультування* (напр., робота фокус-груп, опитування тощо); *залучення* (залучити громадськість до обговорення у рамках, наприклад, семінарів для вразування всіх їх вимог і скарг); *співпраця* (напр., у форматі онлайн-форумів, дорадчих громадських рад); *надання повноважень* (голосування, форуми тощо).

Очікувані наслідки консультацій, наприклад, для державної політики: поширення результатів консультацій, використання отриманої інформації для визначення політики або прийняти нормативно-правового акту чи документа державної політики (див. Рис. 15).



Рис. 15. Контекст консультацій. Джерело: [1]

**Увага!** За посиланням у посібнику «Публічні консультації: крок за кроком» можна знайти **шаблон для формулювання контексту консультацій** за обраною вами темою для заповнення: https://drive.google.com/file/d/1ACDLQNNq5re-8DEQUJ0ORaG0hyNsOKeN/view

**2. Розробити консультацію,** зокремавсі складники проведення консультаційної події, визначити **стейкхолдерів** (всі люди та організації, що зазнають прямого або опосередкованого впливу політики чи рішення, ті, хто мають можливості впливати на рішення (як позитивно, так і негативно), а також ті, хто просто зацікавлені у відповідному питанні; «консультації з охорони морського середовища не зацікавлять тих самих людей, що й консультації з реформування початкової освіти. Деякі теми стосуватимуться суспільства загалом, тоді як інші можуть стосуватися лише окремих груп» [1]), обрати формати консультацій, зібрати інформацію для надання учасникам (їх кількість може бути абсолютно різною), підготувати стратегію збору даних та розробити детальний план дій, у якому визначені ролі та обов'язки кожного члена вашої команди.

**«Стейкхолдерів (**хто може мати вплив на рішення? хто має експертні знання з цієї теми? хто має зацікавленість у цій темі? хто може не погоджуватися з цим рішенням? І т. д.) **розділяємо на 4 основні типи**:

**експерти** – спеціалісти, які мають багато досвіду та знань із теми консультацій;

**залучені та поінформовані** – громадяни, на яких безпосередньо впливає предмет консультацій, які добре обізнані з контекстом і зробили персональний внесок у результати консультацій;

**незалучені, але зазнають впливу** – громадяни, на яких вплинуть результати консультацій, але які не обізнані з питанням та зазвичай не долучаються до консультацій;

**впливові особи та лідери думок** – громадяни, які можуть впливати на суспільну думку щодо процесу консультацій та його результати (преса, спілки, політична опозиція, громадські активісти) (див. Рис. 16).



Рис. 16. Визначення стейкхолдерів відповідно до предмета. Джерело: [1]

**Увага!** За посиланням у посібнику «Публічні консультації: крок за кроком» можна знайти шаблон для визначення стейкхолдерів, їх цілей, мотивації, впливовості відповідно до предмета консультацій за обраною вами темою для заповнення: https://drive.google.com/file/d/14TCvyM2XhJ9Jq3OsQwV8rt-9r5JJWbko/view

«Публічні консультації можна проводити як онлайн, так і офлайн. Обидва типи пропонують різні формати консультацій. Наприклад, можна організувати невеликий офлайн-воркшоп або зібрати сотні громадян на зустріч громади, можна провести онлайн-форум або розробити просте опитування. Це різноманіття форматів дає можливість адаптувати публічну консультацію до типу взаємодії, якої організатори сподіваються досягти, врахувавши особливості та потреби стейкхолдерів. Кожен тип консультацій охоплює іншу аудиторію, передбачає певний спосіб зібрання даних та має різні виклики й переваги. Успішні публічні консультації часто поєднують офлайн- і онлайн-заходи, щоб охопити більшу аудиторію» [1].

**Поширені формати онлайн-консультацій** [1] :форми для коментарів (призначені для зауважень та рекомендацій стейкхолдерів), опитування (використовують для збору детальної інформації з одного або кількох питань, переважно в них поєднані відкриті та закриті запитання,), дискусійні форуми (зазвичай відкриті платформи, доступ до яких надають на певний період часу, де учасники можуть викласти та обговорити свої думки, ідеї та історії щодо різних тем), системи голосування (ефективний метод для ранжування та визначення пріоритетності ідей чи коментарів), вебінари (презентація або воркшоп, що транслюється через інтернет із використанням відео-конференції).

**Поширені формати офлайн-консультацій** [1]:особисті інтерв’ю, круглі столи, фокус-групи, семінари, групові обговорення («світове кафе»: аудиторія ділиться на невеликі групи, кожна група обговорює одне конкретне питання протягом 15–20 хвилин, потім переходить до наступного столу, за яким обговорюють інше питання; за кожним столом «модератор столу» збирає та групує думки від кожної групи, а потім повідомляє їх усім учасникам), таунхол-зустрічі (зустрічі громади з певних географічних місцевостей для обговорення питань, актуальних саме для неї), громадські слухання (громадяни, організації та експерти представляють свої думки й рекомендації стосовно предмета консультацій у письмовому або усному вигляді)

Наступний крок: постановка питань. Якість та чіткість кожного запитання буде визначати, які дані отримає в результаті організатор. Існує багато різних **типів запитань**: запитання, що передбачають відповідь «так/ні», відкриті запитання, запитання, де потрібно обрати одну із запропонованих відповідей.

Для проведення онлайн-консультацій створені **численні інструменти автоматичного збору та аналізу даних**. У випадку офлайн-консультацій, коли дані збирають вручну, для використання цих засобів потрібно спочатку **ввести дані в машиночитаний формат, наприклад, електронні таблиці. Одним із способів для цього є використання шаблонів,** які створюють стандартизовану систему для збору даних та забезпечують учасникам або тим, хто веде записи, можливість вводити дані самостійно у відповідні поля (зокрема ім’я, стать, відповіді на запитання). Такі шаблони можна використовувати безпосередньо у комп'ютерному та паперовому форматі за умови, що дані, зібрані на папері, потім транскрибуєте у такі самі стандартизовані електронні таблиці.

План дій, часові межі, ресурси, відповідальні особи – основні елементи публічних консультацій, що потребують детального опрацювання (див. Рис. 17).



Рис. 17. Основні елементи консультацій. Джерело: [1]

**Увага!** За посиланням у посібнику «Публічні консультації: крок за кроком» можна знайти шаблон для підготовки плану дій консультацій за обраною вами темою для заповнення: https://drive.google.com/file/d/1\_zV2Bn\_CrMwIyglcDpRBKgbK1pGL7zZP/view

https://drive.google.com/file/d/1058mf2\_kx2JVg1bnuUntsMqRe2IPp-iC/view

**3. Розробити комунікаційний план: прокомунікувати і залучити** (рис. 18) з урахуванням цільової авдиторії, яка реагуватиме на різні комунікаційні інструменти та стратегії по-різному, хоч має певні спільні ознаки: демографічні (наприклад, жінки віком 50 років та старші); географічні (наприклад, мешканці Запоріжжя); професійні (наприклад, лікарі, медійники, юристи, народні депутати); рівень експертних знань (наприклад, громадські організації, експерти, широкий загал) та багато інших. За посиланням приклад: https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/shabloni-ta-instrumenti/pidgotuite-komunikaciinii-plan



Рис. 18. Комунікаційний план консультацій. Джерело: [1]

**Комунікаційні інструменти з урахуванням протоколів брендингу (уніфікація використання логотипів, кольорів тощо)**[1]:

* вебсайт міністерства, інформаційний бюлетень або розсилка;
* співпраця зі ЗМІ (пресрелізи, пресконференції, інтерв'ю, медіа-тури тощо);
* публікації в соцмережах, месенджерах та іншому (фейсбук, твітер, інстаґрам, ютуб, телеґрам, воцап тощо);
* друковані матеріали (листівки, брошури, плакати, банери тощо);
* реклама в газетах, журналах, розсилках, на телебаченні, радіо, в інтернеті тощо;
* автоматичні дзвінки чи SMS;
* рекламні ролики, інфографіка або публічні виступи.

Актуальну ключову інформацію про тему консультацій треба звести в **короткий офіційний консультаційний документ**. Наприклад, Консультаційний документ до проєкту Стратегії розвитку системи охорони здоров’я до 2030 року: https://inrespublica.org.ua/aktyvna-hromada/konsultatsijnyj-dokument-do-proyektu-strategiyi-rozvytku-systemy-ohorony-zdorov-ya-do-2030-roku.html

**Увага!** За посиланням у посібнику «Публічні консультації: крок за кроком» можна знайти шаблон для підготовки комунікаційного плану консультацій за обраною вами темою для заповнення:

https://drive.google.com/file/d/1GAl3JEi2zalVQimZ08mlTpX1znYzwyri/view

**4. Провести консультації та зібрати відповіді**

Підготовка приміщення, технічного забезпечення (мікрофони, динаміки, проєктор, екран, wi-fi), підтримка ІТ-спеціаліста або звукотехніка, підготовка матеріалів, зокрема: список для реєстрації стейкхолдерів; анкети для учасників;

шаблони для тих, хто буде вести записи; програми події; будь-які інформаційні чи рекламні матеріали, які підготовлені на попередніх етапах, або інші матеріали, які слід мати у друкованому вигляді.

Організація онлайн-реєстрації, роботи з реєстраційним листком, забезпечення присутності кваліфікованого модератора

Приклад програми заходу у Додатку Б.

Зібрані відповіді можуть поставати у вигляді **якісних даних**, як-от неструктурований текст або аудіо (детальніше – у таблиці 1). «Хоча вони є корисними для розуміння уподобань, думок чи цінностей, тлумачити їх досить складно. Робота з якісними даними може стати ще складнішою у випадку офлайн-консультацій, на яких багато учасників можуть говорити одночасно. Запис та структурування таких форматів потребують значного часу та досвіду. Чим послідовнішими будуть шаблони для запису інформації, тим простіше буде отримати структуровані та якісні дані для аналізу» [1].

Для заняття ефективним буде використання шаблонів:

**Список учасників:** https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/shabloni-ta-instrumenti/skladit-spisok-uchasnikiv-kozhnogo-zakhodu

(шаблон визначає, яку інформацію треба зафіксувати про кожного присутнього на заході; поля (ім'я, вік, стать, електронна адреса) можна адаптувати та розширити відповідно до потреб та контексту консультацій);

**Нотатки обговорення:** https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/shabloni-ta-instrumenti/vikoristaite-shablon-dlya-vedennya-zapisiv-pid-chas-grupovikh-diskusii (можна записувати окремі думки або формування групового консенсусу, або навіть розбіжності чи важливі цитати; іноді не оминають увагою й фізичні дії, як-от «вдарив рукою об стіл», щоб передати емоції під час розмови);

**Опитувальники:** https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/shabloni-ta-instrumenti/pidgotuite-prostii-opituvalnik-dlya-uchasnikiv-vashogo-zakhodu (для збору відповідей учасників).

**5. Проаналізувати отримані дані**

**Життєвий цикл даних** забезпечує результативне управління даними та їхнє збереження. Існує кілька варіантів цього процесу. Зокрема, розглянемо версію, адаптовану до специфіки публічних консультацій, репрезентовану у посібнику «Публічні консультації: крок за кроком».

**Основні фази**: планування, збір, оброблення, аналіз та поширення даних (Рис. 19).



Рис. 19. Життєвий цикл даних. Джерело: [1]

**Види даних у контексті публічних консультацій**

|  |  |
| --- | --- |
| **Визначення** | **Приклад** |
| **➡️ Кількісні дані** відображають величину, обсяг або діапазон у числовому вимірі. | Кількість учасників консультацій, їхній вік, кількість зібраних вами коментарів тощо. |
| **➡️ Якісні дані** відображають властивості об'єктів. Вони часто є неструктурованими та представленими у вигляді тексту, їх неможливо подати в числовому вимірі. | Текст зібраних вами коментарів, нотатки, зроблені вашими аналітиками для узагальнення обговорень, імена учасників тощо. |
| **➡️ Метадані** — це дані, що надають інформацію про інші дані. Вони можуть бути як якісні, так і кількісні. | Дата та час заходу, місто, де його провели, джерело деяких даних (опитування або офлайн-захід) тощо. |

Табл. 2. Види даних у контексті публічних консультацій. Джерело: [1]

«Першим кроком в аналізі публічних консультацій є **транскрибування здобутих даних** (написані від руки нотатки, аудіо-файли, паперові опитувальники та інші дані, зібрані під час консультацій, та сортують у машиночитані дані, які потім можна аналізувати). Аналітики можуть робити це вручну, друкуючи нотатки та вносячи дані до електронних таблиць, або вони можуть використовувати різні транскрибувальні програми.

Транскрибування має відбуватись у вигляді стандартизованого процесу. Це означає, що кожен, хто займається такою діяльністю, має дотримуватись однакових правил, щоб забезпечити точність даних та послідовність розшифрування всіма аналітиками.

Під час офлайн-заходів ви можете отримати різні види даних, наприклад, аудіо-стенограми, інформація про заходи (метадані) та письмові нотатки. Кожне з цих джерел спочатку слід транскрибувати в окремий Excel-аркуш з використанням шаблону. Це забезпечує чітке розмежування при зборі даних і допоможе в їхньому обробленні, якщо хтось виявить помилки під час подальшого аналізу» [1].

Наступний крок: очищення даних, створення резервної копії оригінальних файлів з даними у захищеному місці, перш ніж переходити до подальшого оброблення. **Опрацювання даних** означає здійснення маніпуляцій та змін, а це пов’язано з ризиком втрати якості. Варто закодувати всі якісні дані, наприклад «тематичним кодуванням» – для виявлення та групування основних тем, зафіксованих під час консультації. Цей метод дає змогу проаналізувати якісні дані, послуговуючись простими засобами кількісного аналізу.

Аналіз відповідей на запитання з множинним вибором зазвичай простий: підрахувати, скільки людей вибрали варіанти А чи Б. Однак у випадку неструктурованих відповідей на відкриті запитання потрібно робити **систематичний контент-аналіз**, перш ніж ви зможете відповісти на зазначені запитання (Рис. 20).

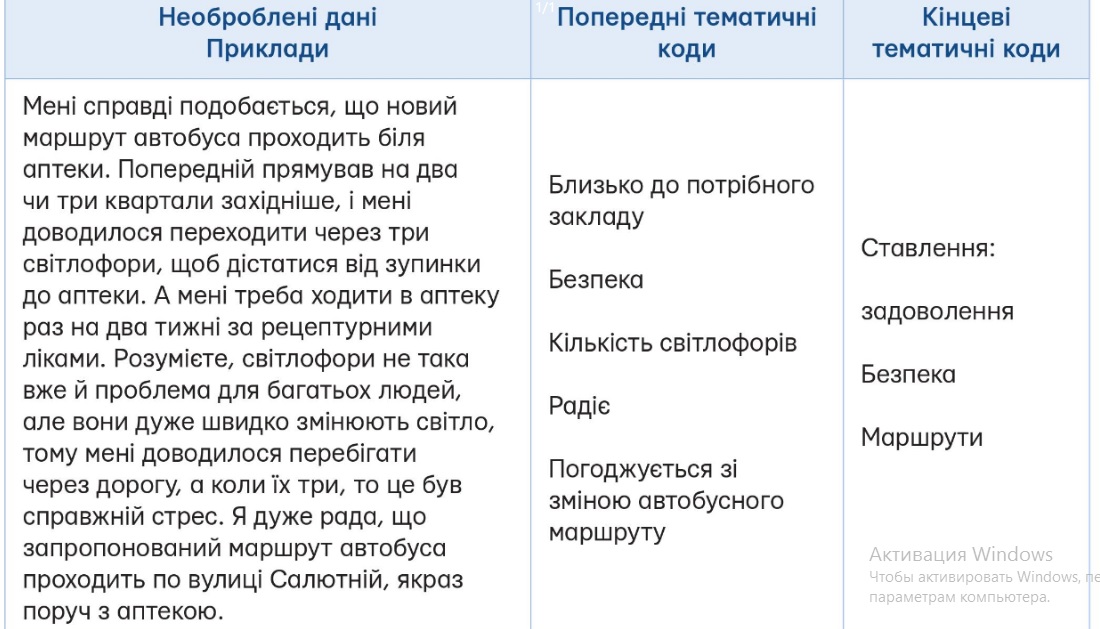


Рис. 20. Приклад аналізу необроблених даних під час проведення консультацій щодо зміни маршруту місцевого автобуса. Джерело: [1]

**Техніки оброблення даних [1]:**

**фільтрування** – перегляд даних на основі різних критеріїв для вилучення найменш релевантних;

**інтегрування** – поєднання різних наборів даних;

**трансформування** – конвертування, реорганізація або переформатування даних без зміни їхнього значення (наприклад, для використання іншого програмного забезпечення);

**узагальнення** – групування даних та підготовка статистики.

**Процес аналізу** передбачає всі дії та методи, які допоможуть описати факти, виявити закономірності, знайти пояснення, визначити подібності та відмінності, зрозуміти зв’язки між змінними та перевірити гіпотези. Для цього можна застосувати статистичний аналіз, візуалізувати зібрані дані або просто розтлумачити отримані результати.

Наприклад, отримані 200 коментарів про зручність нового транспортного маршруту можна було б узагальнити та порівняти різні змінні (Рис. 21).



Рис. 21. Опрацьовані дані коментарів про зручність нового транспортного маршруту. Джерело: [1]

**Увага!** За посиланням у посібнику «Публічні консультації: крок за кроком» можна знайти шаблон для опрацювання даних консультацій за обраною вами темою для заповнення:

https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/shabloni-ta-instrumenti/zakoduite-otrimani-yakisni-dani

**❓***Питання для самоконтролю:*

1. Які основні фази життєвого циклу даних?

2. Назвіть етапи циклу публічних консультацій як механізму постійної комунікації з громадськістю, що базується на отриманні контекстуалізованої інформації на основі категоризованих даних.

3. Назвіть інструменти автоматичного збору та аналізу даних у випадку офлайн- та онлайн-консультацій.

4. Які комунікаційні інструменти з урахуванням протоколів брендингу варто використовувати під час пблічних консультацій?

5. Які види даних у контексті публічних консультацій?

6. Розтлумачте поняття «транскрибування здобутих даних».

7. Назвіть приклад аналізу необроблених даних під час проведення консультацій.

8. Які є техніки оброблення даних?

***Рекомендації до виконання завдань***: насамперед варто переглянути визначену до плану заняття *основну літературу* та *додаткову*. Корисно у процесі підготовки користуватися *онлайн-посібником «Публічні консультації: крок за кроком» за посиланням: https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/*

**✍** *Завдання для самостійної роботи студента:*

1. Опрацювати онлайн-посібник «Публічні консультації: крок за кроком» https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/: розділи «Роль публічних комунікацій», «Нормативно-правова база, що регулює проведення публічних консультацій в Україні», «Публічні консультації: ключові концепції та керівні принципи», «Проведення публічних консультацій. Крок за кроком: Етап 1, 2, 3, 4, 5», заповнивши шаблони, які там пропонуються.

2. Запропонувати ідею публічних консультацій (це не лише урядові ініціативи, а й заходи на кшталт Дня відкритих дверей, перфомансних комунікацій та ін. – це теж публічні консультації), продумати їх формат, збір та опрацювання даних (спробувати реалізувати бодай на 10 респондентах, опитування у соцмережах).

**🕮 *Література:***

***Основна:***

1. Публічні консультації: крок за кроком : посібник. *Edera.* URL: https://edera.gitbook.io/publ-chnkonsultacz-krok-za-krokom/provedennya-publichnikh-konsultacii.-krok-za-krokom/etap-5.-proanalizuite-otrimani-dani (дата звернення: 08.04.2023).

***Додаткова:***

2. Гончаренко А. Впливати на місцеву політику: що таке публічна консультація та як це працює. Зміст. *Зміни створюєш ти*. 21.03.2018. URL: https://zmist.pl.ua/blogs/vplivati-namiscevu-politiku-shcho-take-publichna-konsultaciya-ta-yak-ce-pracyuje (дата звернення: 08.04.2023).

3. Декілька слів про декомунізацію і демонтаж пам’ятників. *Гадяцька міська рада*. 14.06.2022. URL: https://hadiach-rada.gov.ua/news/1655194738/ (дата звернення: 08.04.2023).

4. Завершились громадські обговорення питання демонтажу «Вічного вогню». *Дрогобицька міська рада*. URL: http://surl.li/dmfly (дата звернення: 08.04.2023).

5. Ковпак В. Публічні консультації та краудсорсингові проєкти як актуальний спосіб комунікації та співпраці через опрацювання великих даних. *Суспільство і особистість у сучасному комунікаційному дискурсі :* матеріали ІV Всеукраїнської науково-практичної конференції (м. Запоріжжя, 2 листопада 2022 р.) / редкол.: В.Л. Погребна, В.В. Кузьмін, Н.В. Островська, Т.О. Бородулькіна та ін. Запоріжжя: НУ «Запорізька політехніка», 2022. C. 229–233. URL: https://zp.edu.ua/uploads/dept\_s&r/2022/conf/2.1/Materialy\_SiOuSKD-4.pdf (дата звернення: 08.04.2023).

6. Регулятор проведе публічні консультації з громадськістю щодо Порядку проведення перевірки ділової репутації покупця (потенційного покупця) ЗМІ. *Національна рада України з питань телебачення і радіомовлення*. URL: https://www.nrada.gov.ua/regulyator-provede-publichni-konsultatsiyi-z-gromadskistyu-shhodo-poryadku-provedennya-perevirky-dilovoyi-reputatsiyi-pokuptsya-potentsijnogo-pokuptsya-zmi/ (дата звернення: 08.04.2023).

7. Шимкович В. Дмитро Чаплинський – про те, як IT-засоби допомагають боротися з корупцією. *DOU.* 17 травня 2017. URL: https://dou.ua/lenta/interviews/white-collar-hundred/ (дата звернення: 08.04.2023).

***Тема* 4. Опрацювання та аналіз даних від збору до отримання цінної інформації (на прикладі реалізації публічних консультацій в інформаційній діяльності)**

***Мета:*** набути навички використання типів шкал та змінних, одновимірного аналізу та «крос-табуляції», вміти користуватися перевагами середнього значення, міри та медіани як мір центральної тенденції, структурувати результати оброблення даних у фінальному звіті та ефективно їх оприлюднювати.

*План*

1. Стратегія збору даних.

2. Опрацювання цінної інформації.

3. Структура для фінального звіту. Оприлюднення статистичних даних завдяки ефективним комунікаційним стратегіям.

***Основні поняття:*** стратегія збору даних, опрацювання інформації, оприлюднення статистичних даних, комунікаційні стратегії фінального звітування

**📚 Методичні рекомендації та поради**

У процесі підготовки самостійної роботи студентам необхідно взяти до уваги таку інформацію.

У посібнику від «Українського центру суспільних даних» [1] зазначається, що статистичні методи можуть бути використані для отримання відповідей на кшталт:

1. Якого типу дані і в якій кількості потрібно зібрати?
2. Як варто організувати та узагальнити дані?
3. Як ми можемо проаналізувати дані та дійти висновків?
4. Як ми можемо оцінити силу висновків та оцінити їхню непевність?

Стратегії збору даних ми вже частково розглядали в попередніх розділах, другому та третьому, тож зупинимося переважно на узагальненні та аналізі даних. Наступні тези щодо видів змінних, шкал для статистичного аналізу репрезентовано на основі матеріалів **«Відкритого посібника з відкритих даних для громадських організацій, журналістів, і не тільки…», упорядкованого Д. Кубаєм, А. Горбалем** на платформі «Українського центру суспільних даних» за посиланням: https://socialdata.org.ua/manual/ [1].

**Види шкал та змінних.**

Для цілей статистичного аналізу варто дефініціювати основні типи змінних (та типів шкал), аби зрозуміти, які математичні операції можна робити з різними змінними для статистичного аналізу.

**Змінна** – будь-яка характеристика об’єкта, що вимірюється чи досліджується, може набувати різних значень для різних об’єктів. Виділяють такі типи шкал:

**Номінальна шкала –** категорійна змінна, значення якої не можуть бути впорядковані – наприклад ім’я людини («Якби вибори Президента України відбулися у найближчу неділю, за кого б Ви проголосували?») або стать.Єдина операція, яку можна застосовувати до змінних з номінальною шкалою – це перелік, підрахунок частоти появи того чи іншого значення в масиві.

**Порядкова шкала** – змінна, значення якої за означенням впорядковані (наприклад «слабо», «помірно», «сильно»). Прикладом порядкової змінної в соціологічному опитування буде питання типу: «Наскільки ви довіряєте самому собі?» – з варіантами відповіді «Зовсім не довіряю», «Скоріше не довіряю, ніж довіряю», «Важко сказати довіряю чи ні», «Скоріше довіряю ніж не довіряю», «Цілком довіряю». Або ж значення в полі «Освіта» (неповна середня, середня, середня спеціальна, незакінчена вища, вища, вчений ступінь).

Значення такої змінної можна порівнювати на тотожність, на «більше-менше», але при цьому не можна сказати, **наскільки** більше чи менше.

**Метричні шкали:**

**Інтервальна – з**начення змінної впорядковані як в порядковій шкалі, але при цьому відмінності між значеннями є змістовними – наприклад, температура за Цельсієм, часові інтервали, тощо.

Допустимі операції: додавання та віднімання, але не множення чи ділення.

**Шкали відношень**

**Шкала вимірювань кількісної властивості** – змінні мають всі властивості, що й інтервальні, але також мають «природню» нульову точку «відліку» — наприклад вік, вага, температура за Кельвіном.

Додавання, віднімання, множення, ділення є *змістовними операціями*.

Також змінні розділяють на дискретні та неперервні. Так, кількість дітей у якійсь родині, кількість автомобільних аварій на певній території, або кількість людей, що прочитали цей текст за останній місяць – це результат простого підрахунку випадків настання певного явища, і тому є **значеннями дискретних змінних**. Дискретною є змінна, значеннями якої можуть бути лише цілі числа.

А **неперервні змінні** — такі як довжина, вага, температура, вік – можуть бути виміряні достатньо точно і не мають мінімальної неподільної одиниці виміру. Вага може бути виміряна з точністю до грама, а може й до мікрограма, а вік – з точністю до днів, годин, хвилин, секунд.

У невеликих масивах даних можна знайти мінімальне та максимальне значення, побачити, наскільки повторюються значення або наскільки багато значень є близькими один до одного. Для цього досить відсортувати масив даних в електронній таблиці за певною колонкою (змінною), тобто впорядкувати масив за певною ознакою. Але коли масиви даних надто великі, потрібно застосовувати інші методи для стиснення інформації, отримання розуміння про характеристики масиву.

**Описові статистики** єнайпростішими способами статистичного аналізу, які на найпростішому рівні кількісно описують або підсумовують масив даних.

**Одновимірний розподіл** – дозволяє підсумувати частоту, з якою різні значення певної змінної спостерігаються в наборі даних. Наприклад, ми маємо набір даних, у якому представлено 26 дітей віком від 1 до 6 років. Тоді **частотний розподіл змінної «вік» може матити в таблиці такий вигляд**:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Вік | 1 | 2 | 3 | 4 | 5 | 6 |
| Частота | 5 | 3 | 7 | 5 | 4 | 2 |

Звісно, коли змінні інтервальні, і значень їх дуже багато, доцільно розбити на певні інтервали, об’єднати значення змінних за певними категоріями:

|  |  |  |  |
| --- | --- | --- | --- |
| Вікова група | 1–2 | 3–4 | 5–6 |
| Частота | 5 | 3 | 7 |

Можна також побачити частоту того чи іншого значення у вигляді частки чи відсотка від усіх спостережень. У таблиці нижче представлено **частотний розподіл та відносний частотний розподіл.**

|  |  |  |
| --- | --- | --- |
| Оцінка студентів | Частота | Відносна частота |
| 0 – 25 | 2 | 0,13 |
| 26 – 50 | 5 | 0.31 |
| 51 – 75 | 4 | 0.25 |
| 76 – 100 | 5 | 0.31 |
| Разом | 16 | 1,00 |

Як видно з таблиці, підсумок відносної частоти в кожній із категорій дорівнює 1 – або 100% – оскільки включає в себе всі спостереження.

Одновимірний розподіл може бути реалізований щодо якісних та кількісних змінних, дискретних та неперервних, для відкритих та закритих категорій (так, **«відкритою» категорією** буде «люди старші 60 років», або «дохід нижче 3000 грн»). Ми можемо класифікувати дані **за кількісними характеристиками** – такими як вікові групи, зарплата, рівень доходу тощо; можемо також класифікувати **за якісними характеристиками** – такими як стать, професія, національність, тощо.

Приклад: **частотний розподіл кількісних дискретних даних з відкритою категорією**

| **Рівень доходу, грн** | **Частота** | **Відносна частота** |
| --- | --- | --- |
| 0 – 5000 | 15 | 0,30 |
| 5000 – 10000 | 25 | 0,50 |
| 10000 – | 10 | 0,20 |
| Разом | 50 | 1,00 |

Приклад: **частотний розподіл даних номінальної шкали виміру**

| **Стать** | **Частота** | **Відносна частота** |
| --- | --- | --- |
| Чоловіки | 20 | 0,40 |
| Жінки | 30 | 0,60 |
| Разом | 50 | 1,00 |

**Двовимірний розподіл або крос-табуляція –** ще один спосіб узагальнення даних в таблиці. Це таблиця, що представляє дані, зіставляючи дві окремі характеристики (змінні) – по колонках і рядках. Так, в колонках – різні значення (чи групи значень, категорії) однієї змінної, а в рядках – значення іншої змінної.

Приклад: таблиця з даними Держстату

| **Область** | **2012** | **2013** | **2014** |
| --- | --- | --- | --- |
| Вінницька | 35441 | 37323 | 39184 |
| Волинська | 19546 | 20609 | 21971 |
| Дніпропетровська | 95349 | 99995 | 109545 |
| Донецька | 128767 | 135362 | 114135 |

Важливо розуміти, що з точки зору грамотного і охайного структурування даних – до початку аналізу – у наборі даних (у таблиці) всі змінні мають бути в колонках, всі спостереження в рядках, а на їх перетині – значення змінних для спостережень.

Але для аналізу даних можна порівнювати одні змінні з іншими, **робити крос-табуляцію**, **будувати двовимірні розподіли** – наприклад, для того, щоб на найпростішому рівні виявити залежності між змінними.

Приклад: два одновимірних розподіли наведених вище – за доходом і статтю.

| **Стать** | **Дохід** | **Разом у рядках** |
| --- | --- | --- |
|  | 0 – 5000 | 5000 – 10000 |
| Чоловіки | 7 | 9 |
| Жінки | 8 | 16 |
| Разом у колонках | 15 | 25 |

**Будуючи таблицю розподілу для набору даних, що містить велику кількість спостережень, варто дотримуватися таких основних правил** (за Д. Кубаєм, А. Горбалем):

* знайти найменше та найбільше значення змінної;
* визначити «ширину» категорій, за якими будуть згруповані значення змінної;
* включити всі можливі значення змінної.

Визначаючись із шириною інтервалу, потрібно знайти компроміс: щоб інтервали були достатньо вузькими (щоб всі або більшість спостережень не потрапили в один інтервал), і при цьому достатньо широкими (щоб не вийшло так, що буде по одному спостереженню в інтервалі). Також важливо, щоб інтервали були взаємно виключними.

Міри центральної тенденції – **середнє**, **моду** та **медіану** – ми вже розглядали у змістовому модулі 2, а саме їхню механіку, а ось правила, переваги та недоліки варто розглянути на основі тез «Відкритого посібника з відкритих даних для громадських організацій, журналістів, і не тільки…» [1].

Ці показники демонструють загальні характеристики розподілу даних за певною змінною, дозволяють виявити одне значення, що описує весь розподіл; це окремі значення, що представляють весь набір даних, типові для всіх значень у групі.

**Міри центральної тенденції потрібні з таких міркувань:**

1. Щоб отримати загальну картину розподілу. Ми не можемо запам’ятати кожен факт, що стосується сфери дослідження.
2. Щоб отримати чітку картину щодо досліджуваної сфери для розуміння та отримання потрібних висновків.
3. Щоб отримати чіткий опис групи в цілому та мати змогу порівнювати дві або більше груп у термінах типової «поведінки».

Найвідомішою мірою центральної тенденції — і найбільш вживаною в повсякденному побуті – є **середнє**, або ж просте середнє, або ж арифметичне середнє (arithmetic mean) – просто середнє значення ряду даних.

Для його обчислення досить скласти разом всі значення в розподілі і поділити на кількість спостережень. В Екселі чи Google Spreadsheets для цього є функція MEAN. Є різні математичні способи підрахунки середнього, але в усіх сучасних електронних таблицях та спеціальних програмних пакетах для роботи з даними і статистикою є ця функція.

**Є певні загальні правила для використання середнього, зокрема:**

1. Середнє – це «центр тяжіння» розподілу, і кожне значення дає внесок у визначення середнього значення, коли поширення значень є симетричними довкола центральної точки.
2. Середнє значення більш стабільне, ніж медіана чи мода. Тому, коли потрібно знайти найбільш стабільну міру центральної тенденції, використовують середнє.

**Переваги середнього:**

1. Середнє визначене дуже жорстко, тому не виникає питань чи нерозуміння щодо його значення та суті.
2. Це найбільш поширена міра центральної тенденції, оскільки її легко зрозуміти.
3. Середнє легко підрахувати.
4. Враховує всі значення розподілу.

**Обмеження чи недоліки середнього:**

1. На значення середнього впливають екстремальні значення (відомий іронічний жарт про «середню температуру по лікарні»).
2. Часом середнім є значення, що не присутнє в розподілі.
3. Часом результатом можуть бути абсурдні значення. Наприклад, маємо 41, 44, та 42 учнів у 5а, 5б та 5в класах певної школи. Виходить, що середня кількість учнів у 5 класах школи – 42,3 (3). А так не буває.

**Медіану** можна визначити як точку на ряді розподілу (впорядкований набір значень змінної для різних спостережень – наприклад від найменшого до найбільшого значення) – до цієї точки розташовано половина всіх значень, і після цієї точки теж половина значень. Тобто медіана – це значення, що ділить впорядкований ряд навпіл. Якщо кількість значень непарна, то береться одне зі значень – те, що стоїть у розподілі рівно по центру.

Коли значень парна кількість, то беруть два центральні значення і знаходять їхнє середнє.

**Для чого використовують медіану?**

1. Коли потрібно знайти точну середню точку, точку на «півдорозі» від найменшого значення до найбільшого.
2. Коли екстремальні значення впливають на середнє, то медіана є найкращою мірою центральної тенденції.
3. Медіану використовують, коли потрібно, щоб певні значення впливали на центральну тенденцію, але все, що про них відомо, що вони «нижче» або «вище» медіани.

**Переваги медіани:**

1. Легко вирахувати та зрозуміти.
2. Для підрахунку медіани не потрібні всі значення в розподілі.
3. Екстремальні значення розподілу не впливають на медіану.
4. Її можна визначити і для «відкритих» категорій / класів інтервалів.

**Обмеження медіани:**

1. Вона не так жорстко визначена як середнє, оскільки її значення не так вираховується, як знаходиться (серед значень в розподілі).
2. Не враховує всі спостереження (значення для всіх спостережень).
3. З медіаною потім не можна робити алгебраїчні перетворення так, як із середнім.
4. Потребує впорядкування значень або класів інтервалів у висхідному чи спадному порядку.
5. Часом медіаною може бути значення, не присутнє у самому розподілі.

Третя міра центральної тенденції – це **мода** – значення, що найчастіше зустрічається в розподілі. Як правило, вона представляє найбільш типове значення. На моду ніколи не впливають екстремальні значення в розподілі, а впливають – екстремальні частоти значень, наскільки часто те чи інше значення змінної зустрічається в розподілі.

**Мода використовується:**

1. Коли нам треба швидка і приблизна міра центральної тенденції.
2. Коли потрібна міра центральної тенденції, що має бути типовим значенням.

**Переваги моди:**

1. Мода показує найбільш поширене значення в розподілі.
2. На моду не впливають екстремальні значення – так як на середнє.
3. Моду можна визначити для відкритих інтервалів / категорій.
4. Допомагає аналізувати якісні дані.
5. Моду можна виявити просто побудувавши графік розподілу чи стовпчасту діаграму.

**Обмеження:**

1. Не включає до визначення / розрахунку всі спостереження розподілу, а лише концентрацію частот.
2. Подальші алгебраїчні перетворення неможливі – на відміну від середнього.
3. Буває важко визначити моду у випадку багатомодального чи бімодального розподілу

Розподіл може мати більше двох популярних значень, але якщо має більше ніж трьох мод, опис такого розподілу в термінах найбільш частих значень може втрачати будь-який сенс.

Д. Кубай, А. Горбаль як упорядники посібника наголошують на активному використанні графічного відображення, візуалізації даних, без якої статистичний аналіз є неможливим фактично.

Таким чином, графічне відображення стає важливим інструментом у підготовці детального звіту за підсумками отриманої інформації та презентації висновків.

Наприклад, після завершення публічних консультацій, які ми обрали як «робочий дискурс», у рамках якого ми проводимо аналіз даних, потрібно підготувати детальний звіт, який міститиме результати, що доцільно надати тим, хто ухвалює рішення, а також поширити на широкий загал. Такий документ-звіт є дуже корисним, оскільки:

* «він чітко описує те, як внесок громадян використають для розробки політики;
* він додає прозорості та легітимності публічним консультаціям, що може переконати інших людей долучатися до такої діяльності в майбутньому;
* він наголошує на тому, яку кропітку роботу виконали під час проведення консультацій;
* у ньому документують процес публічних консультацій, пропонують можливості для поступового вдосконалення практик, методів та засобів;
* і найважливіше — він слугує інструментом розробки політики» [2].

**Орієнтовна структура для фінального звіту** за [2]:

1) *вступ* (коротке пояснення, про що йдеться у звіті та чому уряд, установа взялися проводити консультації. Ця частина дає відповідь на запитання: «Яку проблему хотів вирішити уряд/організація/установа?»;

2) *контекст* (підсумовують довідкову інформацію, яку надавали громадянам, щоб забезпечити їхнє розуміння теми консультацій та сприяти їхній змістовній участі в обговореннях);

3) *процес проведення консультацій* (перераховують заходи, проведені протягом усього процесу консультацій, описують консультаційні події, комунікаційні стратегії та перераховують залучених партнерів, пояснюють методологію, яку використали для збору й аналізу даних);

4) *учасники* (детальна інформація про те, хто брав участь у консультаціях і наскільки добре вони представляли вашу цільову авдиторію);

5) «*що ми почули*» (найважливіший у звіті й ключовий при формуванні політики фрагмент, де узагальнюють коментарі та дискусії учасників, виокремлюють важливі висновки та групують провідні теми консультацій, ґрунтується на отриманих даних);

6) *рекомендації щодо розробки політики та засвоєні уроки* (різні рекомендації за результатами публічних консультацій; інформування щодо наступних кроків процесу розробки політики, якщо вони вже відомі; зазнаються ті уроки, що засвоїли під час публічних консультацій для удосконалення процесу в майбутньому);

7) *додатки* (всі документи, що мають відношення до проведених консультацій, зокрема офіційні документи, копії онлайн-опитувальників, приклади оголошень, розроблених для контактування з цільовою авдиторією тощо) [2].

«Щоб визначити, хто має бути присутній на цій зустрічі, варто проаналізувати, хто має повноваження затверджувати рішення та коли, а також звернути увагу на ключові моменти ухвалення рішень, зокрема зустрічі з обговорення політики, процеси планування бюджету, експертів, що проводили збір даних, експертів тематичної проблематики, представників комунікаційного відділу міністерства, організації тощотощо» [2].

**Звіт** варто опублікувати за підсумками проведених консультацій **на урядовому вебсайті, на офіційному сайті установи, організації**, де будуть розміщені звіт та всі відповідні консультаційні документи. Таку сторінку можна в подальшому оновлювати на кожному етапі процесу розробки політики, як-от подання законопроєкту до парламенту, його затвердження або імплементація. Пошук цієї інформації має бути легким для громадськості та всіх, хто брав участь у консультаціях,

Утім, буде недостатнім просто оприлюднити звіт за результатами консультацій у мережі. Треба застосувати ефективні комунікаційні стратегії й переконатися, що стейкхолдери та широкий загал дізнаються про результати консультацій. У посібнику «Публічні консультації: крок за кроком» [2] пропонують такі варіанти:

* організувати **пресконференцію** для презентації підсумків консультацій;

зняти **відеоролик** та/або розробити **брошуру,** в яких узагальнені ключові моменти проведених консультацій;

* опублікувати **звіт** у соціальних мережах та на будь-яких каналах, що виявилися успішними для залучення стейкхолдерів на попередніх етапах **консультацій**;
* роздрукувати **примірники звіту** та розіслати їх відповідним групам і обʼєднанням.

**Інформаційно-комунікаційна робота** не закінчується оприлюдненням та поширенням фінального звіту, оскільки у стейкхолдерів можуть виникнути подальші запитання або зауваження. Публічні консультації спрямовані на підтримку діалогу, тому стейкхолдерам іноді пропонують залишатися залученими за допомогою різних механізмів комунікації, як-от:

* інформаційні бюлетені;
* соціальні медіа та застосунки для обміну повідомленнями;
* групи для обговорень.

Постійна комунікація сприяє участі в майбутніх консультаціях, а залучені стейкхолдери можуть стати союзниками у просуванні та реалізації нової політики.

Приклади системи збору та аналізу даних для прийняття стратегічних рішень у різних царинах презентовано в додаткових джерелах використаної літератури до теми.

**❓***Питання для самоконтролю:*

1. Які типи шкал виділяють у статистичній практиці?

2. Розтлумачте поняття «одновимірний аналіз» та «крос-табуляція».

3. Яких основних правил варто дотримуватися при побудові таблиці розподілу для набору даних, що містить велику кількість спостережень?

4. Назвіть переваги, обмеження та недоліки середнього значення, міри та медіани як мір центральної тенденції.

5. Окресліть складники орієнтовної структури для фінального звіту публічних консультацій.

***Рекомендації до виконання завдань***: насамперед варто переглянути визначену до плану заняття *основну літературу* та *додаткову*.

**✍** *Завдання для самостійної роботи студента:*

1. Опрацювати онлайн-посібник «Публічні консультації: крок за кроком» https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/: розділ «Проведення публічних консультацій. Крок за кроком: Етап 6, 7», заповнивши шаблони, які там пропонуються. Сформувати звіт за проаналізованими даними. Презентувати проєкт.

2. Пройти онлайн-курс «Публічні консультації» на платформі Edera за посиланням: https://courses.ed-era.com/courses/course-v1:NACS\_NDI\_EDERA+PC101+2019/about.

**🕮 *Література:***

***Основна:***

1. Кубай Д., Горбаль А. Відкритий посібник з відкритих даних для громадських організацій, журналістів, і не тільки… Київ : Український центр суспільних даних, 2016. URL: https://socialdata.org.ua/manual/ (дата звернення: 08.04.2023).

2. Публічні консультації: крок за кроком : посібник. *Edera.* URL: https://edera.gitbook.io/publ-chnkonsultacz-krok-za-krokom/provedennya-publichnikh-konsultacii.-krok-za-krokom/etap-5.-proanalizuite-otrimani-dani (дата звернення: 08.04.2023).

***Додаткова:***

3. Боброва А., Вербицький І. Дані та статистика про міста і громади: основні тези дискусії. URL: https://mistosite.org.ua/articles/dani-ta-statystyka-pro-mista-i-hromady-osnovni-tezy-dyskusii (дата звернення: 08.04.2023).

4. Вдосконалення розвитку статистики в Україні. *Методи і стандарти розробки запитальників і анкет.* 13 липня 2021 року. URL: https://www.oecd.org/eurasia/countries/ukraine/Draft\_Proceedings\_Webinar2\_StatisticsUKR.pdf (дата звернення: 08.04.2023).

5. Семінар щодо вдосконалення системи збору та аналізу судових даних для прийняття стратегічних рішень. URL: https://www.coe.int/uk/web/kyiv/-/workshop-on-improving-the-collection-and-analysis-of-judicial-data-for-policy-making-purposes (дата звернення: 08.04.2023).

6. Система збору даних про злочини на ґрунті ненависті, мову ворожнечі та дискримінацію в Україні: рекомендації щодо вдосконалення та застосування спільного підходу та ситуаційний аналіз. вересень 2020 р. *Проєкт «Посилення доступу до правосуддя через несудові механізми правового захисту для жертв дискримінації, злочинів на ґрунті ненависті та мови ворожнечі в країнах Східного партнерства»* URL: https://rm.coe.int/final-data-collection-report-ukraine-ua/16809fac71 (дата звернення: 08.04.2023).

**ІНДИВІДУАЛЬНЕ ПРАКТИЧНЕ ЗАВДАННЯ З КУРСУ «ОПРАЦЮВАННЯ ТА АНАЛІЗ СТАТИСТИЧНИХ ДАНИХ»**

Індивідуальне практичне завдання передбачає виконання таких позицій:

1. Зареєструватися на платформі Prometheus. Пройти онлайн-курс «Аналіз даних та статистичне виведення на мові R» за посиланням: https://courses.prometheus.org.ua/courses/IRF/Stat101/2016\_T3/about. Це є базовий курс для майбутніх комунікаційників. Завдяки отриманим знанням можна проводити первинний статистичний аналіз даних, висувати гіпотези та оцінювати їх на основі зібраних даних. Також мовою програмування обрана мова R, яка є найбільш вживаною мовою для статистиків на сьогодні. Курс складається з онлайн-лекторію, тестових завдань до кожної теми та фінального тесту. Для того, щоб отримати сертифікат за результатами цього курсу, необхідно набрати не менше 60 зі 100 можливих балів.

2. Зареєструватися на платформі Еd-era. Пройти онлайн-курс «Публічні консультації» за посиланням: https://courses.ed-era.com/courses/course-v1:NACS\_NDI\_EDERA+PC101+2019/info. Цей курс містить ключову інформацію про процес публічних консультацій та надає практичні поради щодо всіх етапів їхнього проведення – від визначення цілей консультацій, збору даних, їх обробки до звітування та оприлюднення результатів. Курс складається із чотирьох модулів, де наприкінці кожної теми, що розкривається у відеолекції, можна перевірити отримані знання за допомогою проміжного тестування та фінального тесту. Для генерування сертифіката за результатами курсу необхідно набрати не менше 60 зі 100 можливих балів.

Студент може отримати максимально 20 балів за результативне проходження онлайн-курсів.

**ТЕСТОВІ ЗАВДАННЯ ДЛЯ САМОКОНТРОЛЮ**

**1.** **Який із запропонованих варіантів найкраще відображає зміст поняття «відкриті дані»?**

а)  публічна інформація, що наявна в онлайн-сервісах органів державної влади та місцевого самоврядування;

б)  публічна інформація надана на запит фізичних та юридичних осіб через електронні засоби;

в) це публічна інформація, що оприлюднена у форматі для автоматизованої обробки.

**2**. **Чи потрібно публікувати у формі відкритих даних інформацію про дозволи на зовнішню рекламу міста?**

а)  так;

б)  ні;

в) за рішенням органу місцевого самоврядування.

**3**. **Знайдіть середнє значення вибірки: 7, 5, 3, 7, 6, 7, 4, 6, 8, 5:**

а)  7;

б)  5,8;

в)  6.

**4**. **Знайдіть медіану вибірки: 7, 5, 3, 7, 6, 7, 4, 6, 8, 5:**

а)  7;

б)  6;

в)  5,8.

**5**. **Знайдіть моду вибірки: 7, 5, 3, 7, 6, 7, 4, 6, 8, 5:**

а)  5;

б)  7;

в)  6.

**6**. **Фільтрування як техніка обробки даних це:**

а)  конвертування, реорганізація або переформатування даних без зміни їхнього значення;

б)  поєднання різних наборів даних;

в)  перегляд даних на основі різних критеріїв для вилучення найменш релевантних.

**7**. **До інформації з обмеженим доступом не можуть бути віднесені відомості:**

а)  дані про її національність, освіту, сімейний стан, релігійні переконання, стан здоров'я, а також адреса, дата і місце народження обов’язково;

б)  про факти порушення прав і свобод людини, включно з інформацією, що міститься в архівних документах колишніх радянських органів державної безпеки, пов’язаних з політичними репресіями, Голодомором 1932–1933 років в Україні та іншими злочинами, вчиненими представниками комуністичного та/або націонал-соціалістичного (нацистського) тоталітарних режимів;

в) про аварії, катастрофи, небезпечні природні явища та інші надзвичайні ситуації, що сталися або можуть статися і загрожують безпеці людей.

**8**. **За формою і змістом статистичні показники поділяються на такі:**

а) об’ємні та розрахункові;

б)  абсолютні та відносні.;

в) однорідні та різнорідні.

**9**. **За типом даних змінна «бакалавр» і «магістр» є даними:**

а)  категоріальними впорядковними;

б)  категоріальними невпорядкованими;

в)  категоріальними бінарними.

**10**. **Який тип даних має змінна «вага»?**

а)  категоріальна;

б)  кількісна дискретна;

в)  кількісна неперервна.

**ДОДАТОК А**

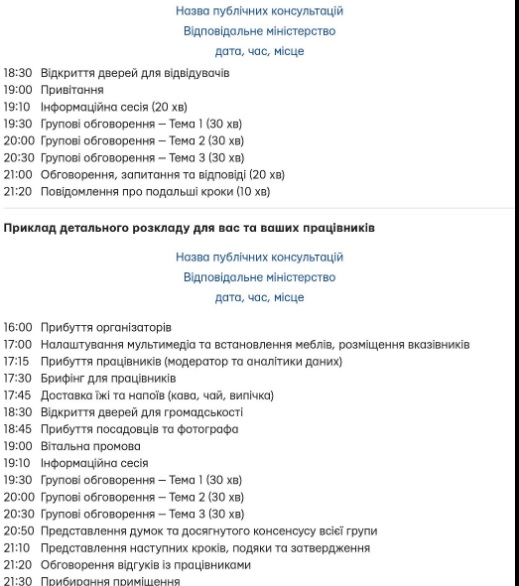
**Порівняння найкращих інструментів аналізу даних**

**(Джерело:** 10 найкращих інструментів аналізу даних для ідеального управління даними [Список 2021]. URL: https://uk.myservername.com/10-best-data-analysis-tools**)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Інструмент аналізу даних** | **Платформа** | **Рейтинги** | **Вирок** | **Ціна** |
| [**Швидкий Майнер**](https://rapidminer.com/)  [RapidMiner_Logo](https://rapidminer.com/) | Крос-платформа | 5 зірок | Система проста у використанні. Потужний графічний інтерфейс. П’ять товарів на вибір. | Безкоштовно: 10000 рядків даних. Невеликий: 2500 доларів на користувача на рік. Середній: 5000 доларів на користувача на рік. Великий: 10000 доларів на користувача / рік. |
| [**Microsoft Power BI**](https://www.scnsoft.com/services/business-intelligence/microsoft/power-bi)  Логотип Power BI | Window, Mac, Linux, Android, iOS. | 5 зірок | Power BI включає набір продуктів для випадків використання, починаючи від візуалізації даних та вдосконаленого аналізу великих даних. | Power BI Desktop, Power BI Mobile - Безкоштовно Послуга Power BI - від $ 9,99 / користувач / місяць |
| [**Xplenty**](https://try.xplenty.com/software-testing-help/)  Логотип Xplenty | Windows і Mac | 5 зірок | Xplenty - це повний набір інструментів для побудови конвеєрів даних. | Отримати цитату |
| [**Zoho Analytics**](https://www.zoho.com/analytics/)  https://myservername.com/img/other/87/10-best-data-analysis-tools-2.jpg | Хмара, Windows, Linux, Мак, Android, ios | 5 зірок | Зручний інструмент візуалізації даних. Співвідношення ціни та якості. | Безкоштовний план. Хмара: починається з $ 22 / місяць (базовий); На місці: починається від $ 150 / місяць. |
| [**Посипати**](https://www.sprinkledata.com/)  [Логотип - посипати](https://www.sprinkledata.com/) | На основі хмари | 5 зірок | Спринкл здатність легко створювати звіти, сегменти та інформаційні панелі для нетехнічних користувачів виділяється серед інших інструментів бізнес-аналітики на ринку. | [**Посипати**](https://www.sprinkledata.com/)Отримайте ціну на цінові плани, Essentials, Professional та Enterprise.  https://www.sprinkledata.com/ |
| [**HubSpot**](https://mbsy.co/wTzmH)  HubSpot_Logo | Windows, Mac, Android, iOS, Windows Phone, Інтернет | 5 зірок | Платформа надасть вам усі дані, необхідні для прийняття розумніших рішень, керованих даними. | Він починається з 40 доларів на місяць. |
| [**Громадський стіл**](https://public.tableau.com/en-us/s/)  [TablePublic_Logo](https://public.tableau.com/en-u) | Windows, Мак, Інтернет, Android, ios | 5 зірок | Хороший інструмент доступний безкоштовно з хорошими функціями та функціональними можливостями. | Громадський стіл: Безкоштовно Creator Tableau: 70 доларів на користувача на місяць. |
| [**КНІМ**](https://www.knime.com/)  [knime логотип](https://www.knime.com/) | Windows, Мак, Linux. | 4 зірки | Працює з Microsoft Azure та AWS. Просте в освоєнні програмне забезпечення. | Платформа KNIME Analytics: Безкоштовно. Сервер KNIME: починається з $ 8500 |
| [**Помаранчевий**](https://orange.biolab.si/)  [Orange_Logo](https://orange.biolab.si/) | Windows, Мак, Linux. | 4 зірки | Зручний графічний інтерфейс | Безкоштовно  https://public.tableau.com/app/discover |
| [**OpenRefine**](http://openrefine.org/)  [OpenRefine_Logo](http://openrefine.org/) | Windows, Мак, Linux. | 4 зірки | Настільна програма Вибір декількох рядків із фільтрами. | Безкоштовно  https://openrefine.org/ |

**ДОДАТОК Б**

**Приклад програми публічного заходу. Джерело: https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/shabloni-ta-instrumenti/pidgotuite-programu-zakhodu**

****

**СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ**

1. Боброва А., Вербицький І. Дані та статистика про міста і громади: основні тези дискусії. URL: https://mistosite.org.ua/articles/dani-ta-statystyka-pro-mista-i-hromady-osnovni-tezy-dyskusii (дата звернення: 08.04.2023).

2. Василюк Д. Дані нам: як великі цифри змінюють журналістику. *Platfor.ma.* URL: https://platfor.ma/magazine/text-sq/media-innovations-lab/kayirobondarenko/ (дата звернення: 08.04.2023).

3. Вдосконалення розвитку статистики в Україні. Методи і стандарти розробки запитальників і анкет. 13 липня 2021 року. URL: https://www.oecd.org/eurasia/countries/ukraine/Draft\_Proceedings\_Webinar2\_StatisticsUKR.pdf (дата звернення: 08.04.2023).

4. Відкриті дані: формати і правила створення : посібник. *Texty.org.ua*. URL: https://texty.org.ua/pdf/data2017.pdf. (дата звернення: 08.04.2023).

5. Відкрита медреформа. Проєкт на основі відкритих даних НСЗУ. *Texty.org.* URL: https://texty.org.ua/d/2019/medical\_reform/?fbclid=IwAR1\_d\_3SdCMCZ7IIaoPuWFDE08Tkxz4xqoHMkPnogynFzsFJ75n19vlbGRo#/apteky (дата звернення: 08.04.2023).

6. Гончаренко А. Впливати на місцеву політику: що таке публічна консультація та як це працює. Зміст. *Зміни створюєш ти*. 21.03.2018. URL: https://zmist.pl.ua/blogs/vplivati-namiscevu-politiku-shcho-take-publichna-konsultaciya-ta-yak-ce-pracyuje (дата звернення: 08.04.2023).

7. Губашов Д. Карта руйнувань. 21.03.2023. URL: https://texty.org.ua/projects/109019/karta-rujnuvan/ (дата звернення: 08.04.2023).

8. Декілька слів про декомунізацію і демонтаж пам’ятників. *Гадяцька міська рада*. 14.06.2022. URL: https://hadiach-rada.gov.ua/news/1655194738/ (дата звернення: 08.04.2023).

9. Завершились громадські обговорення питання демонтажу «Вічного вогню». *Дрогобицька міська рада*. URL: http://surl.li/dmfly (дата звернення: 08.04.2023).

10. Доступ до публічної інформації: від А до Я. *Prometheus*. URL: https://courses.prometheus.org.ua/courses/course-v1:COE+PI101+2017\_T1/course/

11. Кобиляцька В. Публічна інформація у форматі відкритих даних: чому це важливо і як її отримати? *Доступ до правди.* 27 квітня 2020. URL: https://dostup.pravda.com.ua/news/publications/publichna-informatsiia-u-formati-vidkrytykh-danykh-chomu-tse-vazhlyvo-i-iak-ii-otrymaty (дата звернення: 08.04.2023).

12. Ковпак В. Публічні консультації та краудсорсингові проєкти як актуальний спосіб комунікації та співпраці через опрацювання великих даних. *Суспільство і особистість у сучасному комунікаційному дискурсі :* матеріали ІV Всеукраїнської науково-практичної конференції (м. Запоріжжя, 2 листопада 2022 р.) / редкол. : В.Л. Погребна, В.В. Кузьмін, Н.В. Островська, Т.О. Бородулькіна та ін. Запоріжжя : НУ «Запорізька політехніка», 2022. С. 229–233.

13. Ковпак В., Таточенко К. Офіційні чат-боти Telegram, Viber як засоби публічних кризових комунікацій. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія Філологія. Соціальні комунікації.* 2022. Т. 33 (72). № 5. С. 211–217.

14. Корнілова А. Prometheus. Аналіз даних та статистичне виведення на мові R : конспект лекцій. URL: http://surl.li/gdbhf (дата звернення: 08.04.2023).

15. Кубай Д., Горбаль А. Відкритий посібник з відкритих даних для громадських організацій, журналістів, і не тільки… Київ : Український центр суспільних даних. URL: https://socialdata.org.ua/manual/ (дата звернення: 08.04.2023).

16. Навак Н. Корисний карантин: книги і курси з open data, data science та data visualization. 2020. URL: https://mind.ua/openmind/20209468-korisnij-karantin-knigi-i-kursi-z-open-data-data-science-ta-data-visualization (дата звернення: 08.04.2023).

17. Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних, що затверджено Постановою Кабінету Міністрів України від 21 жовтня 2015 р. № 835. URL: https://zakon.rada.gov.ua/laws/show/835-2015-%D0%BF/paran12#n12 (дата звернення: 08.04.2023).

18. Про друковані засоби масової інформації (пресу) в Україні : Закон України від 16.11.92 [№ 2783-XII. Дата втрати чинності: 13.12.2022.](https://zakon.rada.gov.ua/laws/show/2783-12) URL: https://zakon.rada.gov.ua/laws/show/2782-12#Text (дата звернення: 08.04.2023).

19. Про інформацію : Закон України від 02.10.92 р. [№ 2658-XII](https://zakon.rada.gov.ua/laws/show/2658-12" \t "_blank)*[.](https://zakon.rada.gov.ua/laws/show/2658-12" \t "_blank)* Дата оновлення: 13.12.2022. URL: https://zakon.rada.gov.ua/laws/show/2657-12#Text (дата звернення: 08.04.2023).

20. Про нас. *Texty.org.ua.* URL: https://texty.org.ua/p/about/ (дата звернення: 08.04.2023).

21. Про офіційну статистику : Закон України від 16.08.2022 р. № 2524-IX. URL: https://zakon.rada.gov.ua/laws/show/2524-20#Text (дата звернення: 08.04.2023).

22. Про телебачення і радіомовлення : Закон України від 21.12.93 р. № 3760-XII. Дата втрати чинності: 13.12.22. URL: https://zakon.rada.gov.ua/laws/show/3759-12 (дата звернення: 08.04.2023).

23. Публічні консультації: крок за кроком : навчальний посібник OpenNorth для Національного демократичного інституту (НДІ) у межах програми «Залучення громадян у процес прийняття рішень шляхом проведення публічних консультацій» за підтримки Фонду розвитку ефективного врядування Сполученого Королівства (GGF). URL: https://edera.gitbook.io/publ-chn-konsultacz-krok-za-krokom/ (дата звернення: 08.04.2023).

24. Регулятор проведе публічні консультації з громадськістю щодо Порядку проведення перевірки ділової репутації покупця (потенційного покупця) ЗМІ. *Національна рада України з питань телебачення і радіомовлення*. URL: https://www.nrada.gov.ua/regulyator-provede-publichni-konsultatsiyi-z-gromadskistyu-shhodo-poryadku-provedennya-perevirky-dilovoyi-reputatsiyi-pokuptsya-potentsijnogo-pokuptsya-zmi/ (дата звернення: 08.04.2023).

25. Різниця між даними та інформацією. URL: http://surl.li/fxfwz (дата звернення: 08.04.2023).

26. Семінар щодо вдосконалення системи збору та аналізу судових даних для прийняття стратегічних рішень. URL: https://www.coe.int/uk/web/kyiv/-/workshop-on-improving-the-collection-and-analysis-of-judicial-data-for-policy-making-purposes (дата звернення: 08.04.2023).

27. Система збору даних про злочини на ґрунті ненависті, мову ворожнечі та дискримінацію в Україні: рекомендації щодо вдосконалення та застосування спільного підходу та ситуаційний аналіз. вересень 2020 р. *Проєкт «Посилення доступу до правосуддя через несудові механізми правового захисту для жертв дискримінації, злочинів на ґрунті ненависті та мови ворожнечі в країнах Східного партнерства».* URL: https://rm.coe.int/final-data-collection-report-ukraine-ua/16809fac71 (дата звернення: 08.04.2023).

28. Соціальна статистка : конспект лекцій для здобувачів вищої освіти першого (бакалаврського) рівня зі спеціальності 054 «Соціологія» / укладач А.В. Хмелюк. Кам’янське : ДДТУ, 2020. 155 с.

29. Тимошенко О., Кельм Н., Солодько П. Карта бойових дій. Оновлюється. 29.03.2022. *Texty.org*. URL: https://texty.org.ua/projects/106137/karta-bojovyh-dij-onovlyuyetsya/ (дата звернення: 08.04.2023).

30. Шекеряк Я. Карти, боти й пошук людей – як журналісти й урядові структури використовують відкриті дані під час воєнного стану. *Медіакритика.* URL: https://www.mediakrytyka.info/novi-tehnologii-media/karty-boty-y-poshuk-lyudey-yak-zhurnalisty-y-uryadovi-struktury-vykorystovuyut-vidkryti-dani-pid-chas-voyennoho-stanu.html (дата звернення: 08.04.2023).

31. Шимкович В. Дмитро Чаплинський – про те, як IT-засоби допомагають боротися з корупцією. 17 травня 2017. DOU:https://dou.ua/lenta/interviews/white-collar-hundred/ (дата звернення: 08.04.2023).

32. 10 найкращих інструментів аналізу даних для ідеального управління даними [Список 2021]. URL: https://uk.myservername.com/10-best-data-analysis-tools (дата звернення: 08.04.2023).

33. 100 interesting data sets for statistics. URL: https://rs.io/100-interesting-data-sets-for-statistics/ (дата звернення: 08.04.2023).

34. Job Comparison – Data Scientist vs Data Engineer vs Statistician. Published On October 19, 2015 and Last Modified On June 24th, 2019. *Analytics vidhya.* URL: https://www.analyticsvidhya.com/blog/2015/10/job-comparison-data-scientist-data-engineer-statistician/ (дата звернення: 08.04.2023).

35. Nathan Yau. Flowing Data. URL: https://flowingdata.com/ (дата звернення: 08.04.2023).

36. R for Data Science. URL: https://r4ds.had.co.nz/index.html (дата звернення: 08.04.2023).

**Інформаційні ресурси:**

37. Державна служба статистики України. Статистика щодо економічної ситуації в Україні: ВВП, інфляція, безробіття, експорт / імпорт, реєстр підприємств, сільське господарство, будівництво, регіональна статистика та ін. URL: http://www.ukrstat.gov.ua/ (дата звернення: 08.04.2023).

38 Державний і місцеві бюджети: доходи, витрати, обслуговування , державного боргу, виконання бюджету. URL: http://www.treasury.gov.ua/ (дата звернення: 08.04.2023).

39. Євростат. Статистика з основних макропоказників в СС (ВВП, інфляція, безробіття та ін.). URL: http://ec.europa.eu/eurostat (дата звернення: 08.04.2023).

40. Єдиний веб-портал використання публічних коштів. URL: https://spending.gov.ua/new/ (дата звернення: 08.04.2023).

41. Індекс добробуту по країнах. URL: http://worldhappiness.report/.

42. Індекс політичного розвитку в державах світу. URL: https://ec.europa.eu/ (дата звернення: 08.04.2023).

43. Індекс розвитку свободи слова. URL: https://rsf.org.

44. Індекс свободи ведення бізнесу, умови підприємництва в країнах. URL: http://www.heritage.org (дата звернення: 08.04.2023).

45. Рейтинг державних фондів. URL: https://www.fitchratings.com (дата звернення: 08.04.2023).

46. Статистика кількості та якості земельних ресурсів у світі. URL: http://www.iiasa.ac.at/ (дата звернення: 08.04.2023).

47. Статистика світових цін на газ, нафту. URL: https://www.eni.com/ (дата звернення: 08.04.2023).

48. NGO: Energy Policy Group. Статистика з енергетичного ринку в СС, газринок. URL: http://energywatchgroup.org (дата звернення: 08.04.2023).

49. Тіньова економіка, офшори. URL: http://www.gfmtegrity.org/ (дата звернення: 08.04.2023).

Навчально-методичне видання

(українською мовою)

Ковпак Вікторія Анатоліївна

Методичні рекомендації до самостійної роботи для здобувачів ступеня вищої освіти бакалавра спеціальності «Інформаційна, бібліотечна та архівна справа»

освітньо-професійної програми «Інформаційно-комунікаційна справа»

Рецензент *Л.В.Чернявська*

Відповідальний за випуск *В.В. Березенко*

Коректор *В.А. Ковпак*