**Український католицький університет  
факультет суспільних наук  
кафедра медіакомунікацій**

|  |  |
| --- | --- |
| **Назва курсу** | **"Візуалізація даних"**  **(із використанням мови програмування R)** |
| **Викладачка** | Надія Романенко |
| **Профайл викладачів** |  |
| **Контактний тел.** | (+38-032) 240-99-40 (внутрішній - 2073) |
| **E-mail:** | nadiia.romanenko@gmail.com |
| **Сторінка курсу в CMS** |  |
| **Консультації** |  |

1. **Анотація до курсу**

Концептуально етапи курсу “Візуалізація даних” співпадають з базовими етапами будь якого проєкту з візуалізації даних: отримання даних, обробка, аналіз, візуалізація. Впродовж курсу студенти отримують теоретичні знання з візуалізації даних, вчаться знаходити та скрейпити дані, чистити дані для аналізу, виконувати основні операції агрегації даних, на практиці вивчають мінімальні основи математичної статистики та готують візуалізації. Впродовж курсу використовуємо “забігання” — дуже швидко пробуємо щось нове/складне і отримуємо результат. Потім пізніше повертаємося до цієї теми і детально робираємо її. Курс націлений на отримання практичних навичок: ми пробуємо, помиляємося і виправляємо помилки, і знову пробуємо.

1. **Мета та цілі курсу**

**Мета курсу** «Візуалізація даних» — студенти розвивають інфографічний смак, орієнтуються у роботі з даними, можуть використати дані у свїх проєктах, за потреби скориставшись програмуванням.

1. **Формат курсу**

Курс складається з лекцій, практичних занять і самостійної роботи. Знання та навички, отримані під час аудиторних занять, застосовуються для виконання завдань в рамках підсумкового проєкту. У кінці курсу студенти готують проєкт, журналістський матеріал з історією, розказаною за допомогою даних.

**4. Результати навчання**

У результаті вивчення навчальної дисципліни студенти повинні

**знати:**

1. поняття візуалізації даних;
2. основні принципи та терміни візуалізації даних;
3. основні посібники, в яких викладені правила візуалізації даних;
4. основні принципи інфодизайну;
5. основні типи даних та правила їхньої візуалізації;
6. основні види візуалізацій;
7. засади HTML: теги, дерево елементів DOM, css, css-selectors;

**вміти:**

1. працювати з табличними (csv, xls), ієрархічними(json) та слабко структурованими (html) форматами даних;
2. перетворювати дані у форму, потрібну для аналізу та візуалізації — обробляти дані;
3. виокремлювати дані, які відповідають на журналістське або дослідницьке питання;
4. створювати візуалізації, оптимальні для задачі та даних;
5. оформлювати візуалізації даних згідно з правилами інфодизайну;
6. програмувати у R;
7. працювати в HTML;

**5.   Обсяг курсу**

|  |  |  |  |
| --- | --- | --- | --- |
| **Вид заняття** | лекції | практичні заняття | самостійна робота |
| **К-сть годин** | 12 | 18 | 60 |

**6. Ознаки курсу**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Рік викладання** | **семестр** | **спеціальність** | **Курс**, **(рік навчання)** | **Нормативний\вибірковий** |
| **2019** | 3 | **Візуалізація даних** | **2** | **нормативний** |

**7. Технічне й програмне забезпечення /обладнання**

Студенти отримують теоретичні і практичні навички для візуалізації даних, працюючи з фаховою літературою та комп’ютерними програмами.

**8. Політики курсу**

Загальні політики УКУ. Крім того, студетів заохочують до використання якнайширшого кола ресурсів, у тому числі фрагментів готового коду або рішень для візуалізації чи обробки даних, за умови дотримання цитування та інших правил академічонї доброчесності.

Роботи в межах курсу можна поєднувати з навчальними завданнями з інших курсів за умови погодження з викладачами цих курсів (наприклад, використати візуалізацію даних у буклеті, статті, або ессеї).

**9. Схема курсу**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Тиж. /**  дата /  **год.** | **Тема, план заняття, короткі тези** | **Форма діяльності (заняття) / Формат** | **Матеріали** | **Література**  **Ресурси в інтернеті** | **Завдання для самостійної роботи,**  **год** | **Вага оцінки** | **Термін виконання** |
| **Тиж. 1.**  **2 акад. год.** | **Модуль І.**  **Тема 1. Вступ до візуалізації даних.**  Світові тенденції у візуалізації. Як слідкувати за найкращим? Функції візуалізації? Приклади візуалізацій. Поняття «дані», «обробка», «аналіз»,-«візуалізація».Г оловна ідея візуалізації: наратив, історія. Особливості візуального сприйняття: Око і мозок, когнітивні особливості – звідки беруться принципи ефективної подачі - інформації. Як правильно кодувати інформацію у візуальну форму. Архітектура візуалізації | Лекція | Презентація,  відеоматеріали |  | Передивитись презентацію,  Опрацювати задану літературу.  *2 год.* |  |  |
| **Тиж.2**  **2 акад. год.** | **Тема 1. Вступ до візуалізації даних.**  Світові тенденції у візуалізації. Як слідкувати за найкращим? Функції візуалізації? Приклади візуалізацій. Поняття «дані», «обробка», «аналіз»,-«візуалізація».Г оловна ідея візуалізації: наратив, історія. Особливості візуального сприйняття: Око і мозок, когнітивні особливості – звідки беруться принципи ефективної подачі - інформації. Як правильно кодувати інформацію у візуальну форму. Архітектура візуалізації | Практичне заняття |  |  | 1. Знайти і передивитися всі відео від Darkhorse: мапа, таблиця, стовпчики, млинець  ДЗ: інсталяція R та RStudio  **Опрацювати на вибір**:  1,4,5,6,8  *4 год.* | 10 |  |
| **Тиж.3**  **2 акад. год.** | **Тема 2. Візуалізації без програмування. RAWgraphs та Google Spreadsheets як інструменти візуалізації даних**  Табличні дані. Типи і формати даних (csv, json - API, shp). Широкий і довгий формат, який і для чого використовується? Джерела даних: Інформаційні запити, Відкриті API, створюємо дані власноруч (приклади) | Лекція-практичне. |  | 2,3 | Заводимо персональний github, робимо перший сайт на github.io  *4 год.* |  |  |
| **Тиж 4**  **2 акад. год.** | **Тема 2. Візуалізації без програмування. RAWgraphs та Google Spreadsheets як інструменти візуалізації даних**   * Веб-Скрейпінг: Практичний воркшоп Як влаштована сторінка HTML (редагуємо сайт через інструменти розробника) * Що таке теги, дерево елементів DOM, css, css-selectors * Знайомство з інструментами розробника * Перша програма на R (rvest) | Практичне |  | 10,8 | Завдання:   * зробити запит через API declarations.com.ua; * дослідницький аналіз даних (електронні таблиці); * сортування, фільтрація, умовне форматування; * робота з строковими функціями; * обчислення нових змінних (трансформація значень у колонках); * групування та агрегація (pivot table).   4 год. | 10 |  |
| Тиж.5  **2 акад. год.** | **Тема 3.** **Основні концепції програмування на прикладі Excel**   * Трансформація даних в інші дані, за допомогою функцій (числові, строкові тощо). Що таке функція (гра у функції) * Програмування у R - основні поняття. Синтаксис: змінна, вектор, список, датафрейм, функція * векторизація | Лекція-практичне | Презентація, відеоматеірали |  | Переглянути презентацію, прочитати.  Опрацювати задану літературу.  *4 год.* |  |  |
| **Тиж.6**  **2 акад. год.** | **Тема 3. Основні концепції програмування на прикладі Excel**   * Трансформація даних в інші дані, за допомогою функцій (числові, строкові тощо). Що таке функція (гра у функції) * Програмування у R - основні поняття. Синтаксис: змінна, вектор, список, датафрейм, функція * Векторизація | Практичне |  |  | Готуємо сирі дані, отримані скрейпінгом, до аналізу та візуалізації.  *4 год.* | 10 |  |
| **Тиж.7**  **2 акад. год.** | **Тема 1, Тема 2, Тема 3**  **Презентація проєктів** | Практичний модуль |  |  | *Підготовка практичних проєктів*  *7 год* | 17 |  |
| **Тиж.8**  **2 акад. год.** | **Тема 4. Граматика даних.**   * Поняття граматики даних. * [Добре структуровані дані](https://tidyr.tidyverse.org/articles/tidy-data.html) tidyverse. * Основні дієслова для роботи з даними. * Pivot table за допомогою dplyr. | Лекція *F2F* |  | 7 | Переглянути презентацію, опрацювати задану літературу.  Почати роботу з документацією ([dplyr](https://dplyr.tidyverse.org/))  *4 год.* |  |  |
| **Тиж.8**  **2 акад. год.** | **Тема 4. Поняття граматики даних**.   * Поняття граматики даних. * Добре структуровані дані tidyverse. * Основні дієслова для роботи з даними. * Pivot table за допомогою dplyr. | Практичне |  |  | Створюємо таблиці даних як результат.  *4 год.* | 10 |  |
| **Тиж.9**  **2 акад. год.** | **Тема 5. Граматика графіки**   * Анатомія ggplot та побудова графіків. * Функції візуалізації даних. * Типологія графіки. * Ефективні види графіків, які нечасто використовуються: графік розсіяння, slope chart, dot chart, small multiples, гістограма | практичний workshop | Презентація,  відеоматеріали |  | Переглянути презентацію, опрацювати задану літературу.  Продивитись типи графіків у [ggplot2](https://ggplot2.tidyverse.org/reference/index.html) (geom\_…)  *4 год.* |  |  |
| **Тиж.10**  **2 акад. год.** | **Тема 5. Граматика графіки.**  Інструментів для роботи з даними та візуалізацій. Як будувати роботу і вибирати тему. | Практичне | Презентація,  відеоматеріали |  | **Прочитати:**  R for Data Science: Import, Tidy, Transform, Visualize, and Model Data  Опрацювати  інструменти для роботи з даними та візуалізацій  - MOAL: http://bit.ly/12uokPm  - Raw (https://rawgraphs.io/)  - Datawrapper (https://www.datawrapper.de/)  - Vega / Vega Lite (https://vega.github.io/vega-lite/)  - Idyll (http://idyll-lang.org/)  - Floorish (https://flourish.studio/)  - Observable (https://beta.observablehq.com/)  - R/Python notebooks  - Tableau (https://www.tableau.com/)  *4 год.* | 10 |  |
| **Тиж.11**  **2 акад. год.** | **Тема 6. Картографія.**   * Картографія. Статичні та інтерактивні карти на R. * Що таке GIS. * Статична карта. * Інтерактивна карта. * Гео-кодування. Онлайн-сервіси: Carto, Mapbox. * Відтворювана журналістика. | Лекція  *F2F* | Презентація,  відеоматеріали | 3,4 | Переглянути презентацію, опрацювати задану літературу.  *2 год* |  |  |
| **Тиж.13**  **2 акад. год.** | **Тема 6. Картографія.**  Консультаційна сесія: підготовка до фінального проєкту | **Модуль** |  |  | Говоримо про:  Візуалізації в R  - Текст + Дані + Код  - "Золоті"" принципи організації роботи з даними  *4 год* | 10 |  |
| **Тиж.14**  **2 акад. год.** | Тема 4, Тема 5, Тема 6 | Практичне заняття |  |  | Закріплення матеріалу  *7 год* |  |  |
| **Тиж. 15**  **2 акад год.** | Презентація підсумкового проєкту. | *Підсумкове заняття* |  |  | Підготувати проєкт з візуалізації даних.  *6 год.* | 23 |  |
| **Тиж. 16**  **2 акад год.** | **Презентація підсумкового проєкту.** | *Підсумкове заняття* |  |  | Підготувати проєкт з візуалізації даних.  *6 год.* |  |  |

**10. Система оцінювання та вимоги**

Розрахунок суми балів здійснюється за наступною формулою:

* максимальна оцінка за домашні завдання та роботу в авдиторії — 60 балів.
* максимальна оцінка власного проєкту в межах Модуль 1 — 17 балів.
* максимальна оцінка за виконання фінального проєкту — 23 бали.

**Критерії оцінювання практичних завдань (6 балів)**

дотримання дедлайнів — 1 бал

виконання всіх пунктів завдання — 4 бали

оформлення завдання — 1 бал

**Критерії оцінювання роботи в авдиторії (4 бали)**

готовність до заняття — 2 бали

активність в груповій роботі — 2 бали

**11. Рекомендована література**

**Базова**

1. William Cleveland, Visualizing data. Hobart Press; 1 edition (March 1, 1993)
2. Hadley Wickham, Garrett Grolemund. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data
3. Few Stephen. Information Dashboard Design: Displaying Data for At-a-glance Monitoring
4. Noah Iliinsky. Beautiful Visualization, Looking at Data Through the Eyes of Experts by Julie Steele,.
5. Eileen and Stephen McDaniel. The Accidental Analyst: Show Your Data Who’s Boss
6. Альберто Каїро. Функціональне мистецтво: вступ до інфографіки та візуалізації / переклад з англ. Л. Белея за ред. Р. Скакуна. Львів: Видавництво Українського католицького університету 2017. xviii + 350 с.
7. Edward R.Tufte. The Visual Display of Quantitative Information – Graphics Press; 2nd edition edition (May 1, 2001, 200 p/
8. Daniel Rosenberg, Anthony Grafton. Cartographies of Time: A History of the Timeline, Princeton Architectural Press (February 1, 2012), 272 p.
9. Sandra Rendgen, Julius Wiedemann. (2012). Information Graphics. Taschen Publishing
10. Jason Lankow, Josh Ritchie, Ross Crooks (2012). Infographics: The Power of Visual Storytelling
11. Edward R. Tufte (2006). Beautiful Evidence. Cheshire. CT: Graphics Press.
12. John Wilder Tukey (1977). Exploratory Data Analysis. Addison-Wesley
13. Doug Newsom and Jim Haynes (2004). Public Relations Writing: Form and Style. p.236.
14. Mark Smiciklas (2012). The Power of Infographics: Using Pictures to Communicate and Connect with Your Audience.

### Додаткова

1. Heer, J., Bostock, M., & Ogievetskey, V. (2010). A tour through the visualization zoo. Communications of the ACM, 53(6), 59-67.
2. Card, Scott (2009). Information visualization. In A. Sears & J. A. Jacko (Eds.), Human-Computer Interaction: Design Issues, Solutions, and Applications (pp. 510-543). Boca Raton, FL: CRC Press.
3. Zaman, Jazib (2019-05-31). Infographics; a form of visual content
4. Murray, Scott (2017). Interactive Data Visualization for the Web, 2nd Ed. O'Reilly Media, 457p.
5. Ware., C. (2008). Visual Thinking for Design, Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann, 256 p.
6. Knaflic, C. N. (2015). Storytelling With Data: A Data Visualization Guide for Business Professionals. Wiley, 288p.
7. McCandless, D. (October 21, 2014). Knowledge is Beautiful. Harper Design, 256p.
8. Yau, Nathan (July 20, 2011). Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics. Wiley; 1 edition, 384p.
9. Wexler, S., Shaffer, J., and Cotgreave, J. (April 24, 2017). The Big Book of Dashboards, Wiley; 1 edition, 448p.