

## День 2. Робота з кольорами

### Впровадження візуальних каналів на графіках

- 1) <sup>2 бали</sup> Програмно візуалізуйте категорійні дані з використанням найбільш рейтингових каналів візуалізації. У якості джерела даних обирайте [датасет](#) у формі csv-файлу, в якому дані щодо медичних страховок у США. Передбачено 7 стовпців:

- age: вік власника;
- sex: стать отримувача страховки (female або male);
- bmi: індекс маси тіла, який оцінює ступінь ожиріння організму (вимірюється в  $\text{кг}/\text{м}^2$ , норма в межах 18.5-24.9);
- children: кількість дітей (осіб), яких покриває медична страховка
- smoker: статус курця;
- region: регіон проживання власника страховки в США (northeast, southeast, southwest, northwest);
- charges: індивідуальні медичні витрати, оплачені за допомогою медичної страховки.

Інформацію щодо зчитування даних програмно засобами мови Python можете знайти [тут](#). Застосуйте метод head() до об'єкту зі зчитаними даними та виведіть на екран перші рядки таблиці. *Додайте відповідний код та скріншоти роботи до звіту.*

У випадку розташування на площині може бути доречним [графік розкиду](#) з бібліотеки matplotlib. Побудуйте залежність між індексом маси тіла та медичними витратами. *Додайте відповідний код та скріншоти до звіту та проаналізуйте отриману візуалізацію.*

Доповніть графік, враховуючи вплив паління. Для цього можемо застосувати наступний візуальний канал – колір. Разом з тим, тепер потрібно використовувати легенду, яка описуватиме значення відповідних кольорів (для статусу курця yes або no). Приклад побудови графіку з легендою та кольорами можете розглянути [тут](#). *Додайте оновлений код та скріншоти з графіком до звіту.*