**Міністерство освіти і науки України**

**Національний університет харчових технологій**

Кафедра інформаційних технологій, штучного

інтелекту та кібербезпеки

**Звіт**

із лабораторної роботи № **5**

з дисципліни «Прикладний аналіз даних мовами VBA та Python»

на тему: «Відновлення пропущених даних»

Виконав:

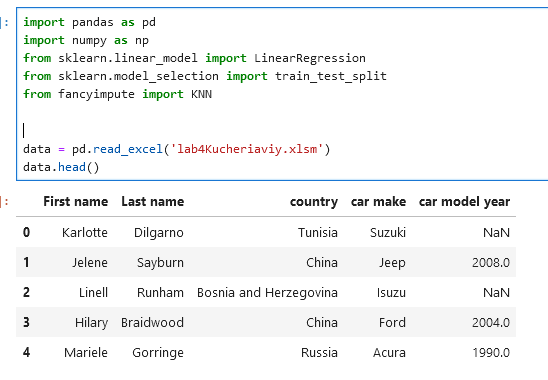
Студент групи КН-1-3М

Кучерявий М. В.

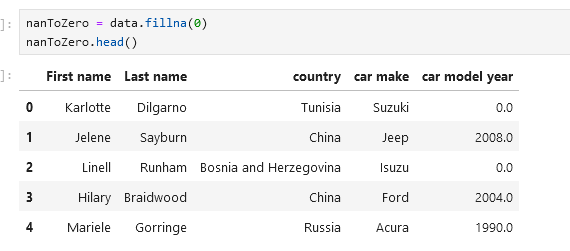
**Київ — 2024**

**Хід роботи:**

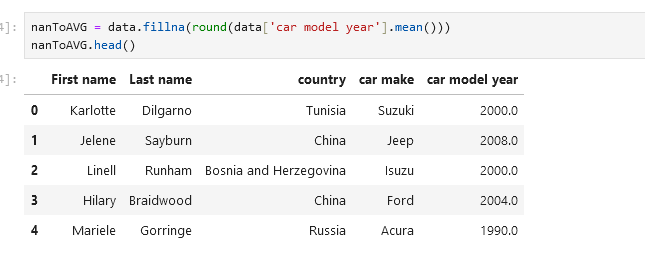
Оскільки дані в v2 json є цілісними, зробимо вручну певні видалення. Імпортуємо їх до як dataframe pandas:



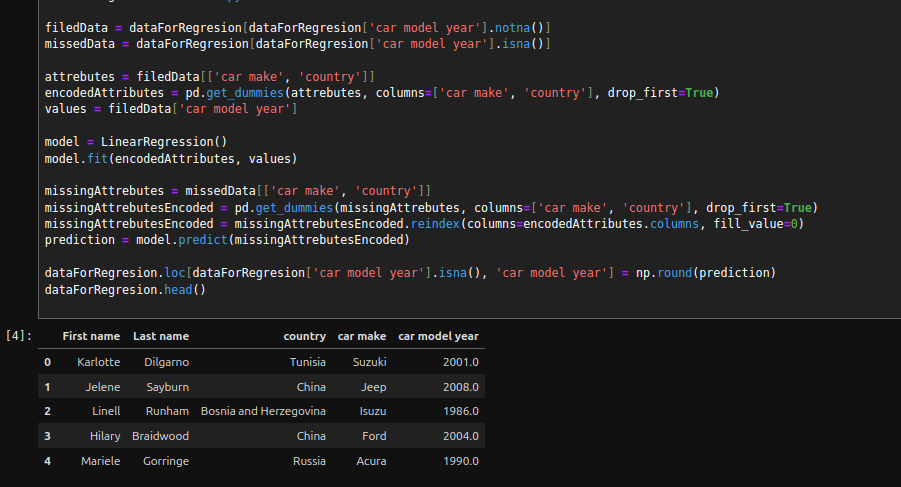
Виконаємо заповнення нулями:



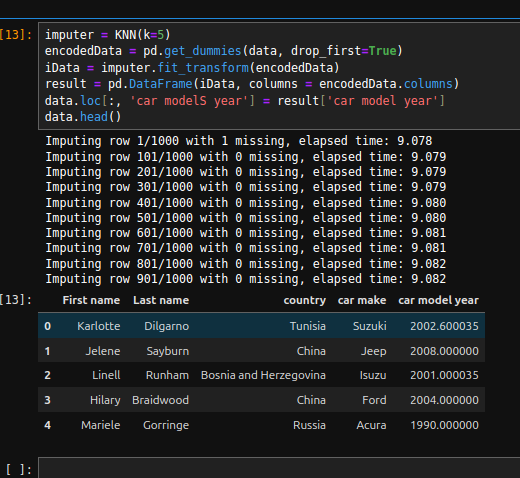
Виконаємо заповнення середніми значеннями:



Виконаємо регресійне заповнення. Для цього використаємо бібліотеку scikit-lear. Використаємо модель лінійної регресії, і навчимо її існуючими значеннями. Після приведення та перетворення далих створюємо передбачення і за допомогою нього заповнюємо пропуски.

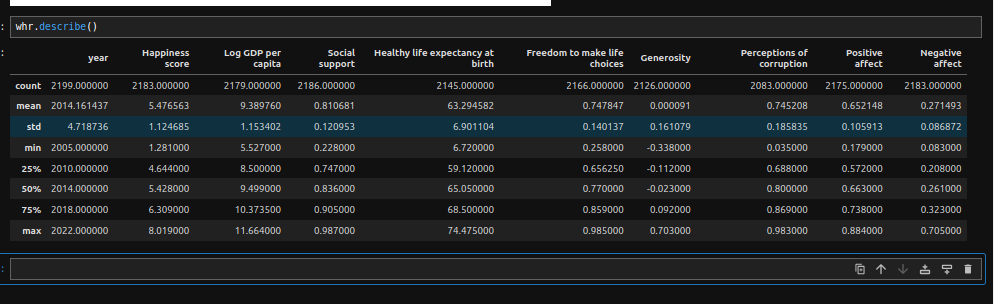
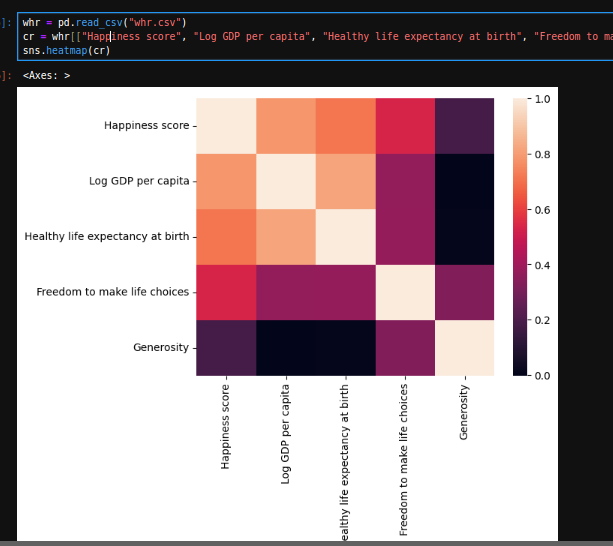


Виконаємо заповнення з використанням бібліотеки fancyimpute:



Для кореляції використаємо таблицю World Happines Report оскільки вона містить більше різноманітних числових даних.

Проведемо кореляцію таблиці:



Виділення метрик таблиці

