НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

Кафедра автоматизації проектування енергетичних процесів і систем

Лабораторна робота № 4

З дисципліни: «Big Data та аналіз даних»

**Виконав:**

студент 5 курсу

групи ТР-02мп, ТЕФ

Круглий Д.В.

**Перевірив:**

Матичин І.І.

Київ – 2021

**Постановка задачі**

1. Вчитати до змінної pisa файл з даними pisa2015.csv

2. Виділити підмножину всіх учнів жіночої статі з Німеччини (CNTRYID==”Germany”). (ST004D01T)

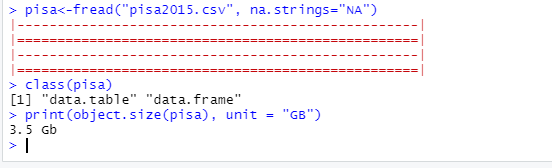
3. Підрахувати кількість учениць з Німеччини за допомогою функції .N

4. Підрахувати кількість учнів з Німеччини (Germany) та Уругваю (Uruguay), які мають вдома комп’ютер (змінна computer) та користуються навчальним програмним забезпеченням (змінна software).

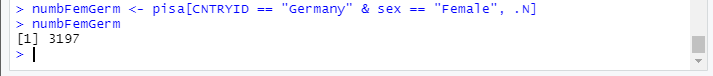
5. Підрахувати кількість учениць, які мають власну кімнату (own.room) або тихе місце для навчання (quiet.study).

**Хід роботи**

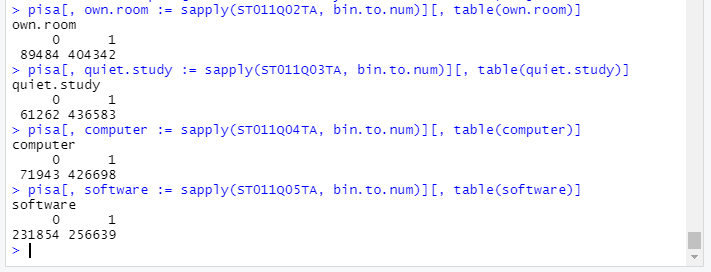
1. Вчитання до змінної pisa файлу з даними pisa2015.csv:



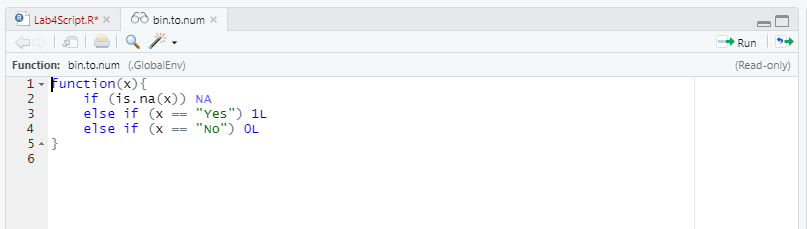
1. Виділення підмножини всіх учнів жіночої статі з Німеччини:
2. Кількість учениць з Німеччини у вибірці даних дорівнює ***3197***:



Перетворення бінарних змінних на цілий тип даних та перейменування колонок таблиці:



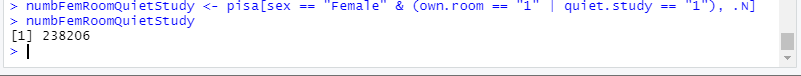
Код функції *bin.to.num*:



1. Кількість учнів з Німеччини та Уругваю, які мають вдома комп’ютер та користуються навчальним програмним забезпеченням дорівнює ***2320***:



1. Кількість учениць, які мають власну кімнату або тихе місце для навчання дорівнює ***238206***:



**Висновок:** виконавши дану лабораторну роботу, я засвоїв навички роботи з методом вчитування даних у середовище R за допомогою пакету data.table та методу fread(). Також було запрограмовано вибір підмножин та обчислення кількості записів у вибраних підмножинах за допомогою методу .N. Отримано навички перетворення бінарних змінних на цілий тип даних.