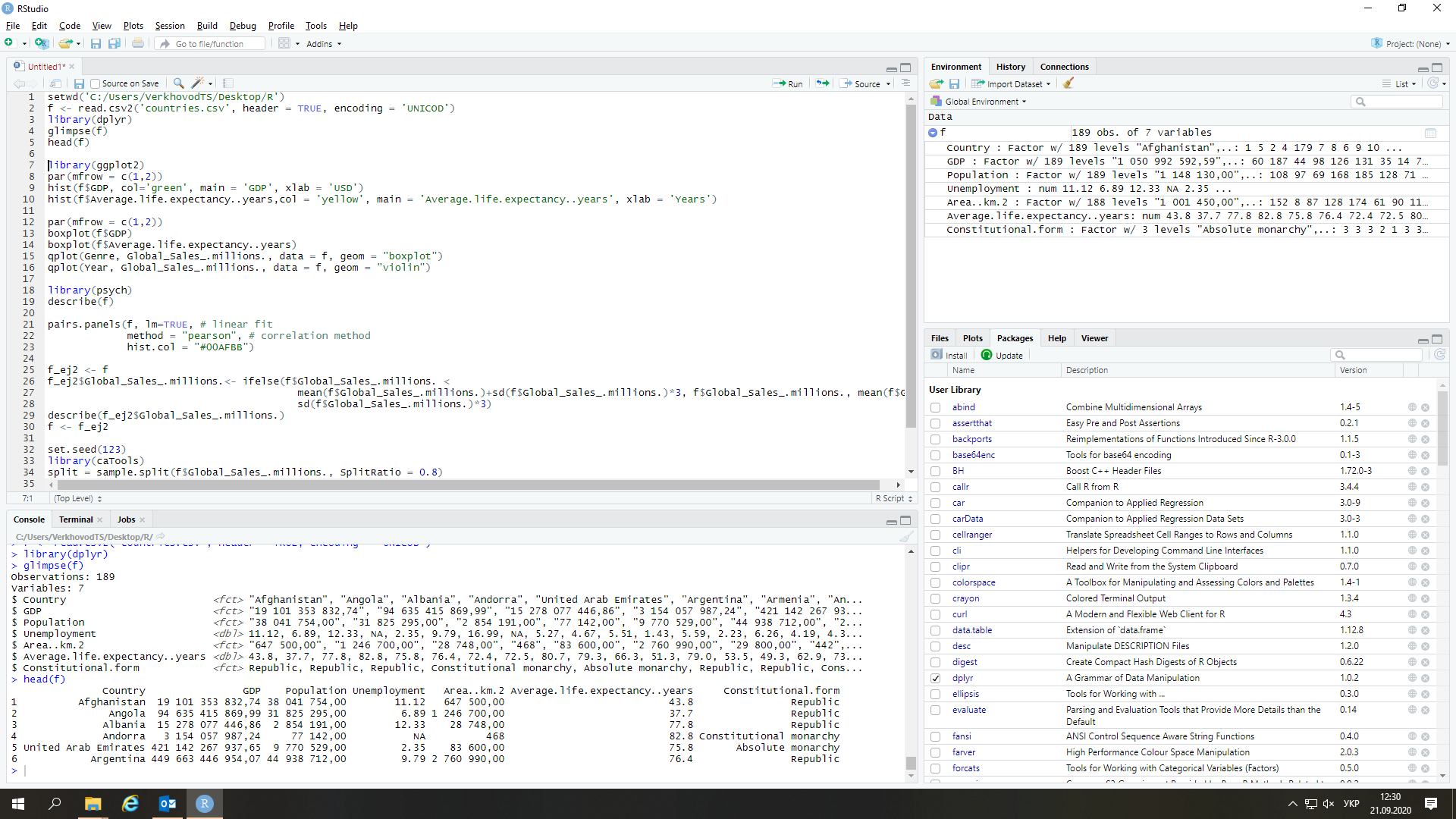
Звіт

за лабораторною роботою №2

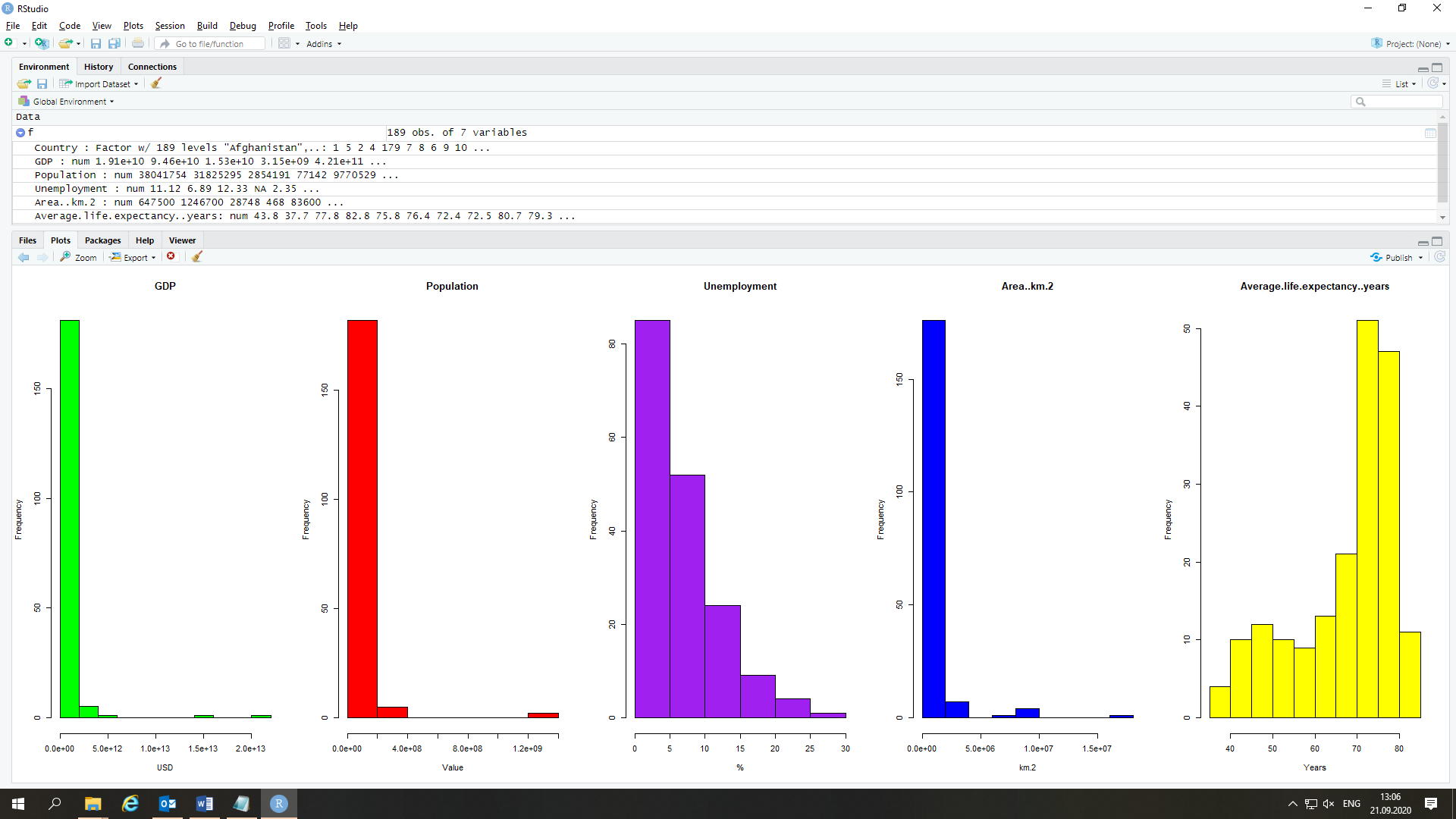
студентки групи ЕП-61

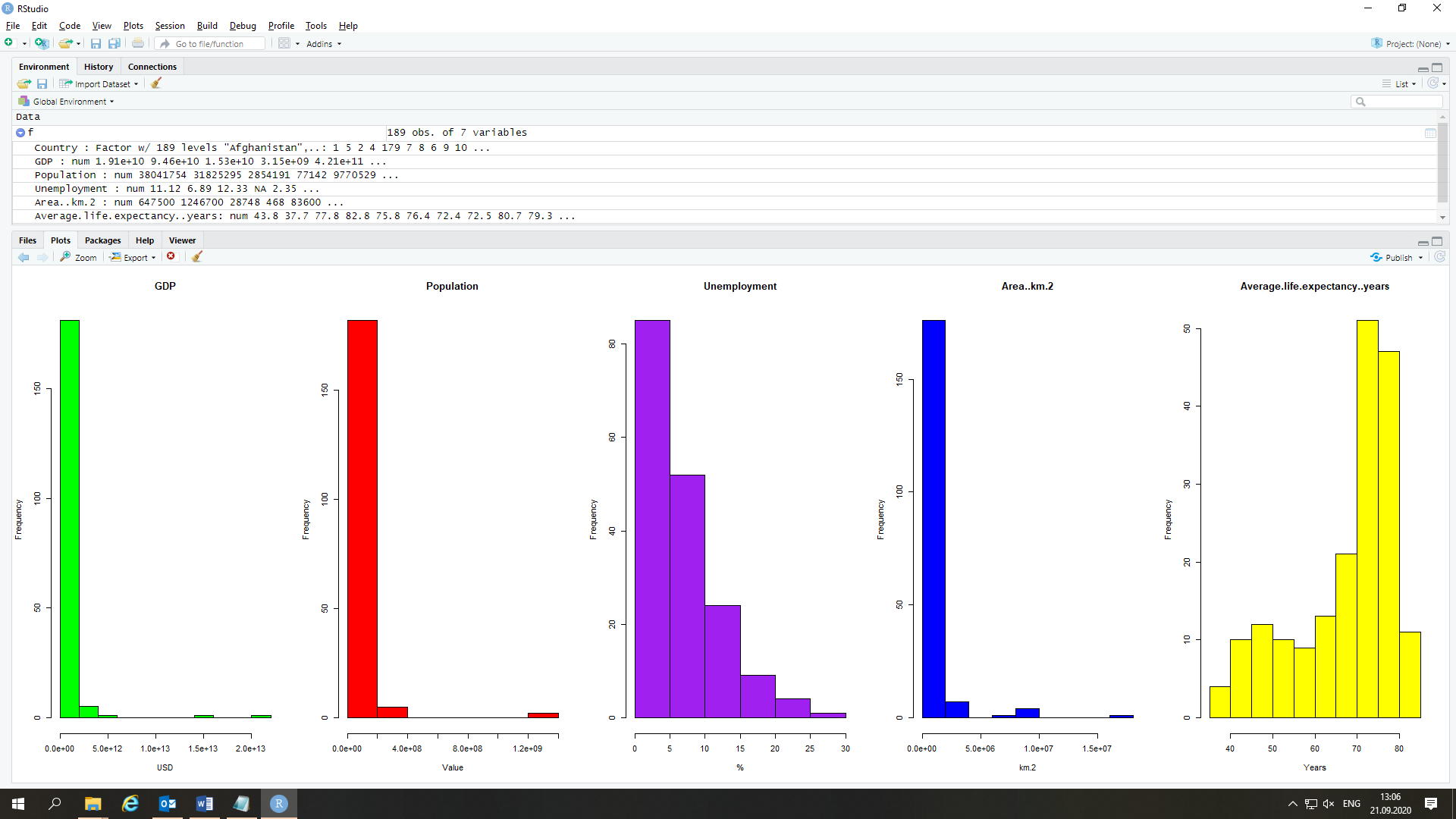
Верховод Тетяни

1. В масиві даних кількість спостережень – 189; кількість змінних – 7.

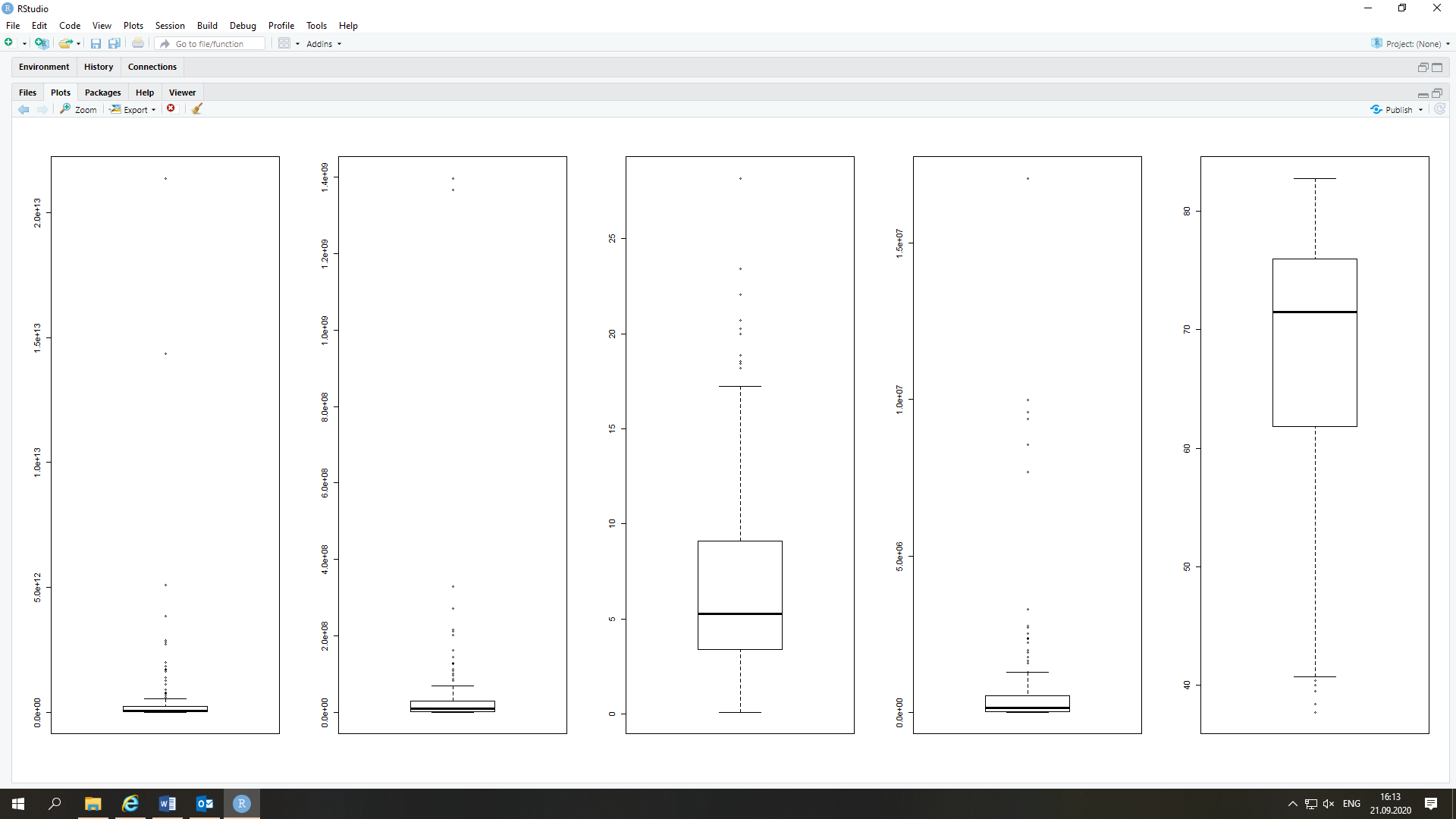


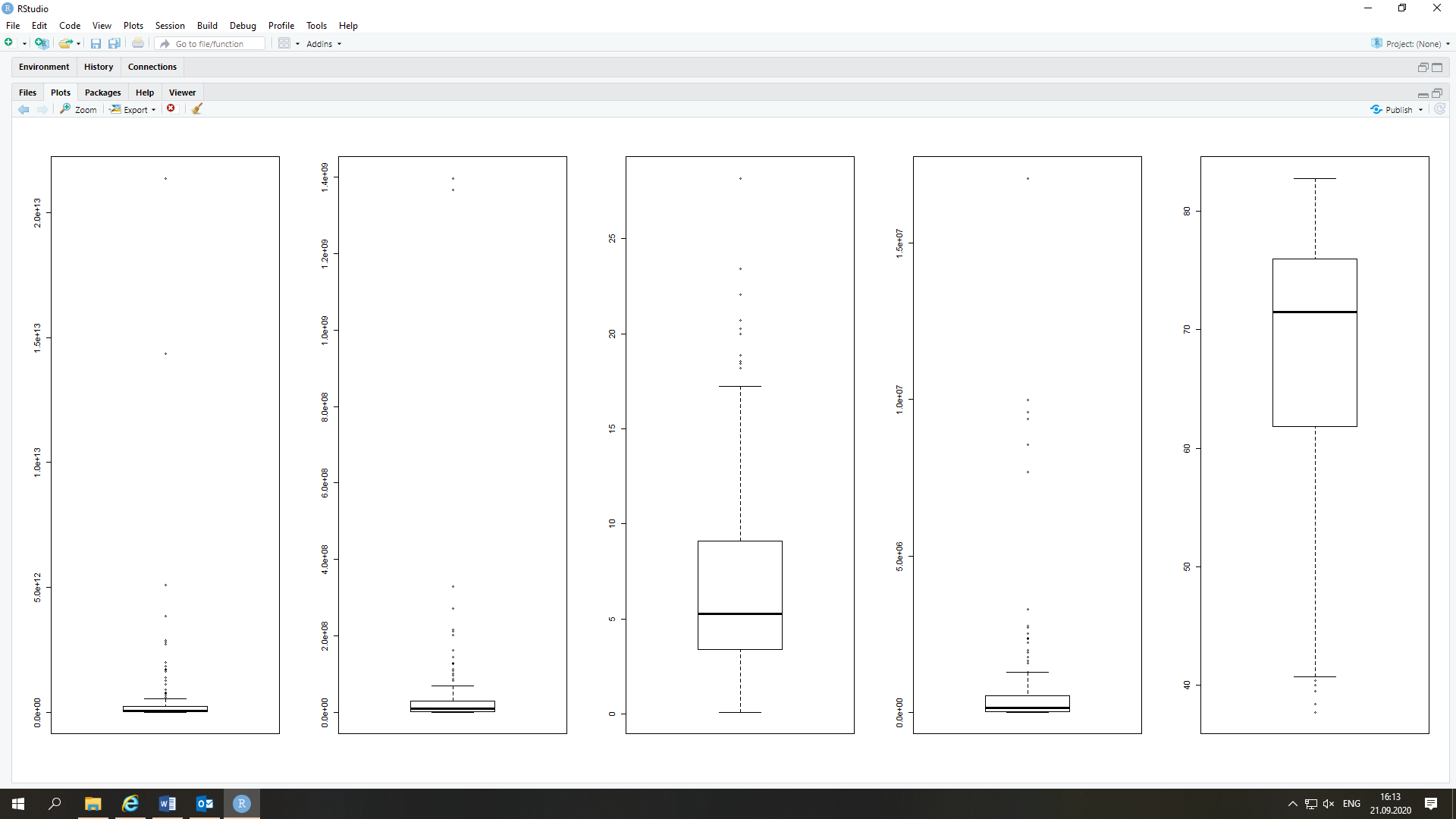
1. Розподіл змінних «GDP», «Population», «Unemployment», «Area..km.2» мають довгий хвіст.



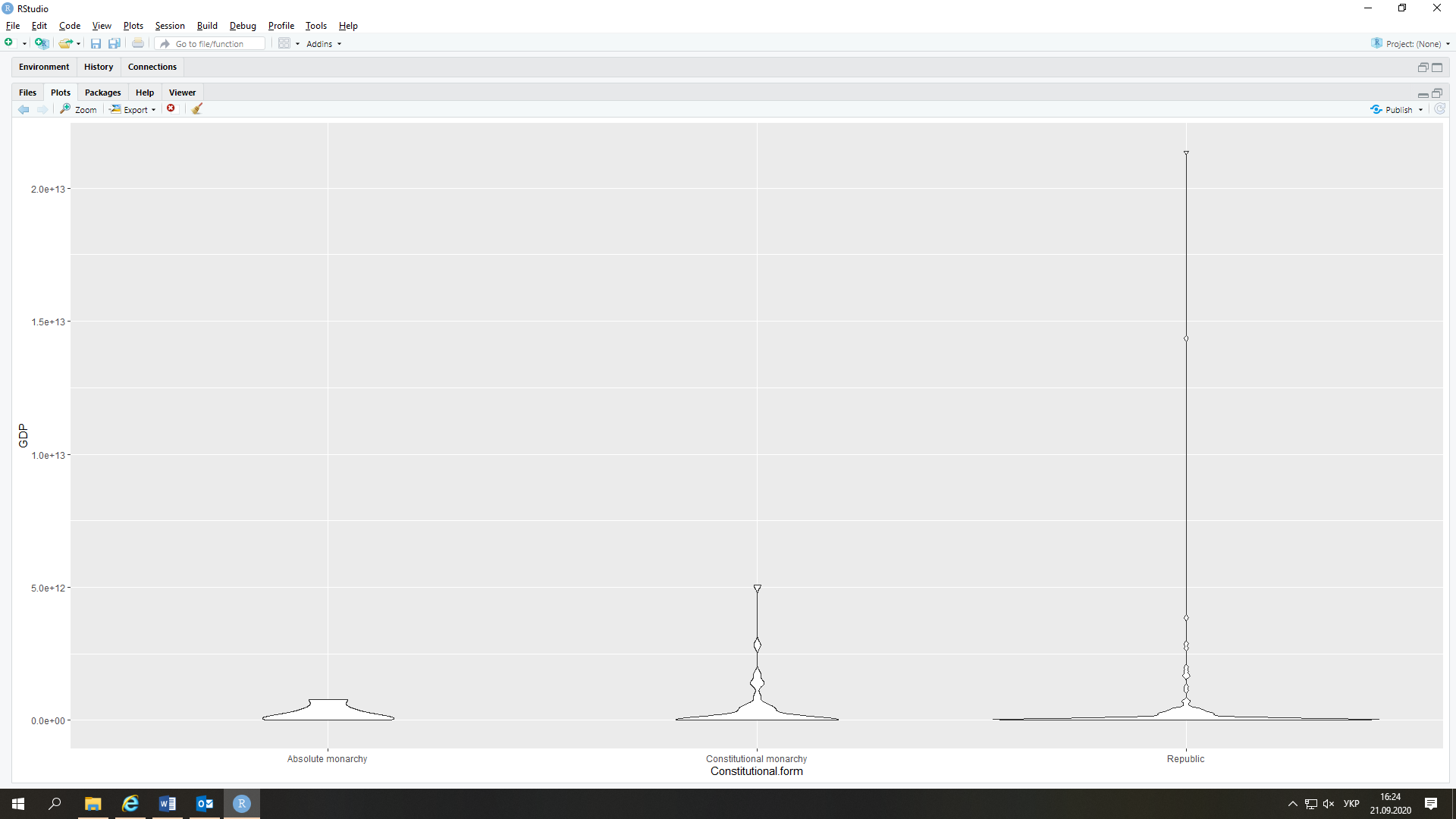


1. Змінні «GDP», «Population», «Unemployment», «Area..km.2» має викиди. Найбільші викиди спостерігаються в країнах з формою правління «Республіка».

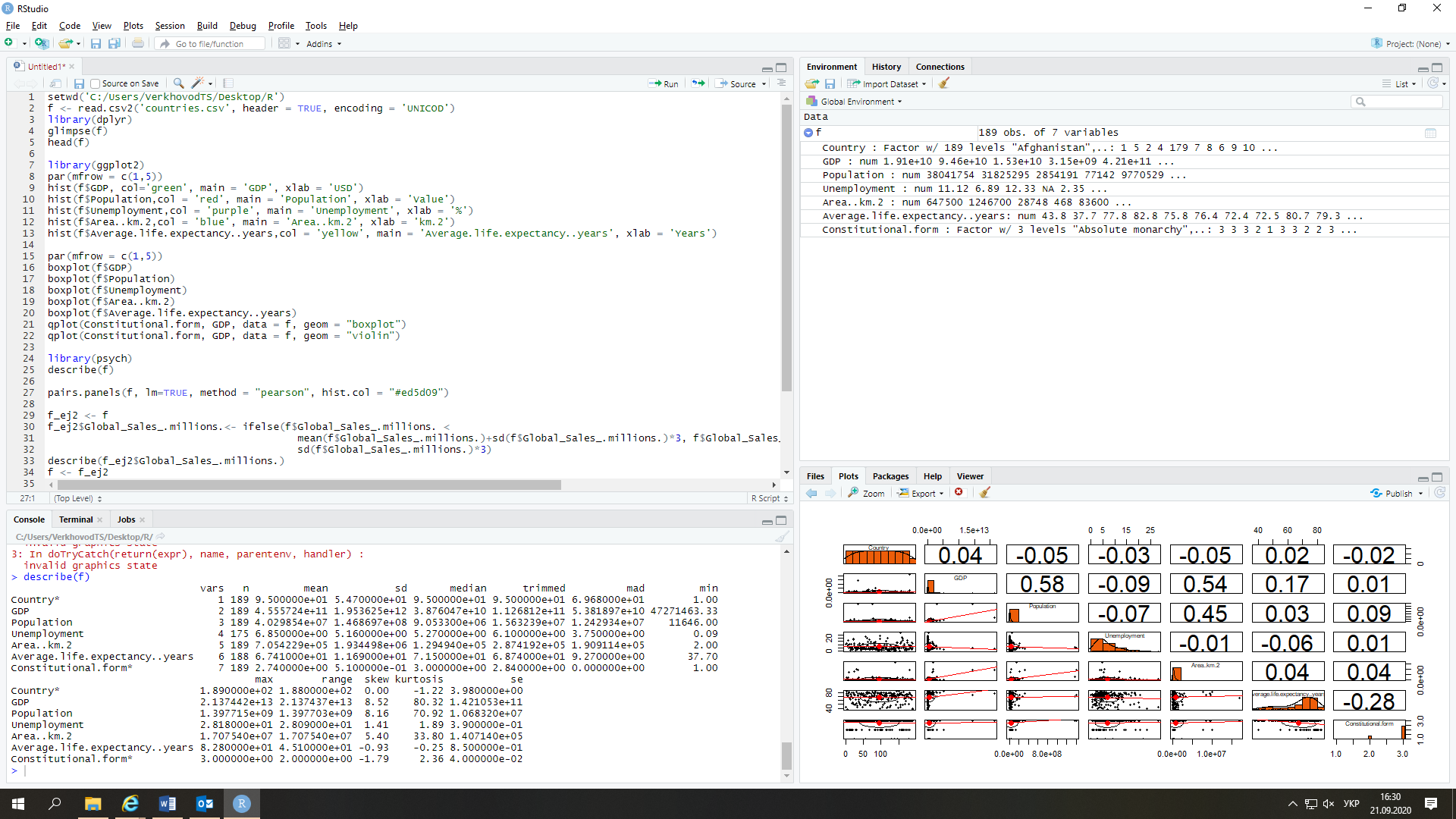




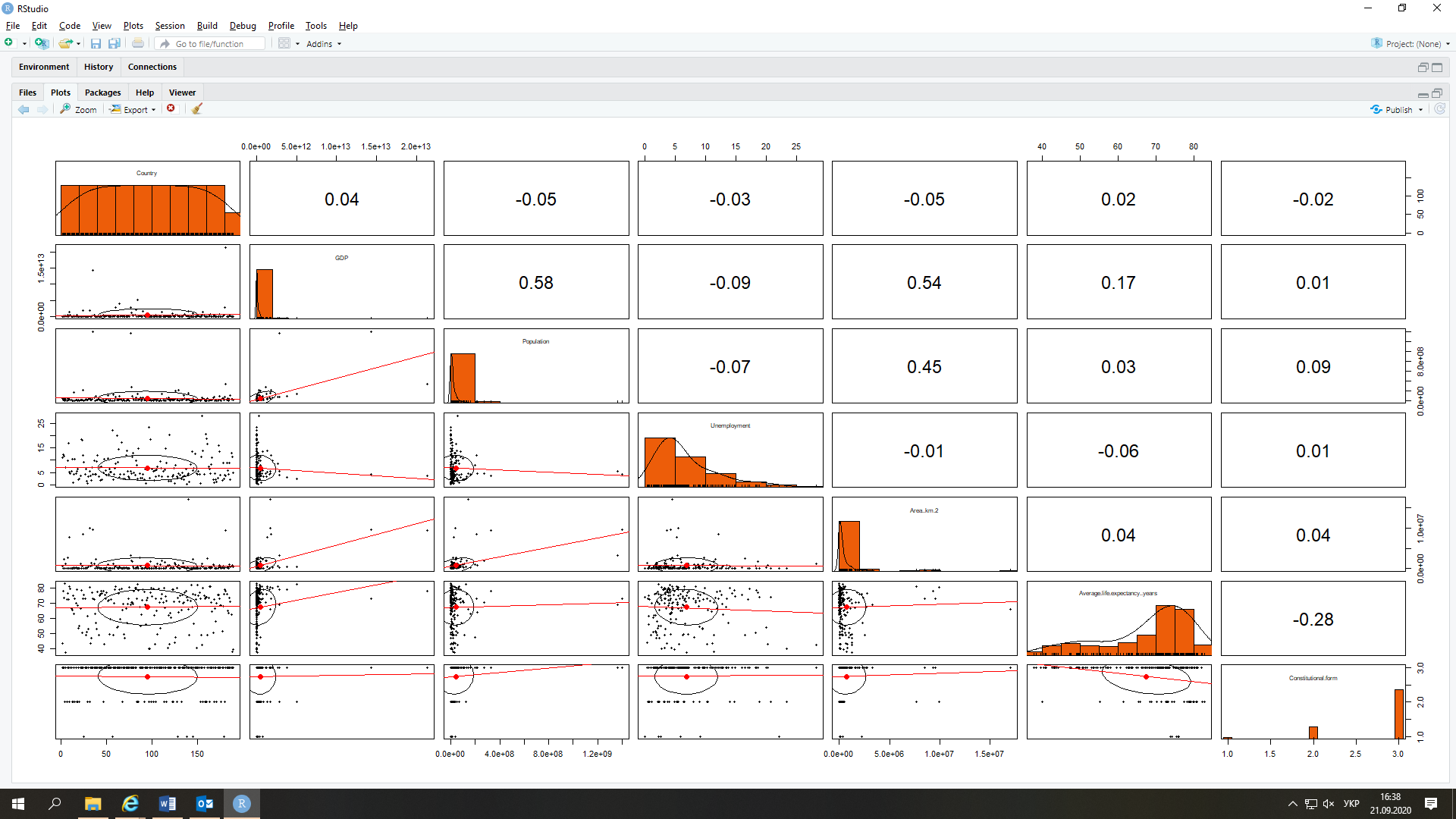
Violin



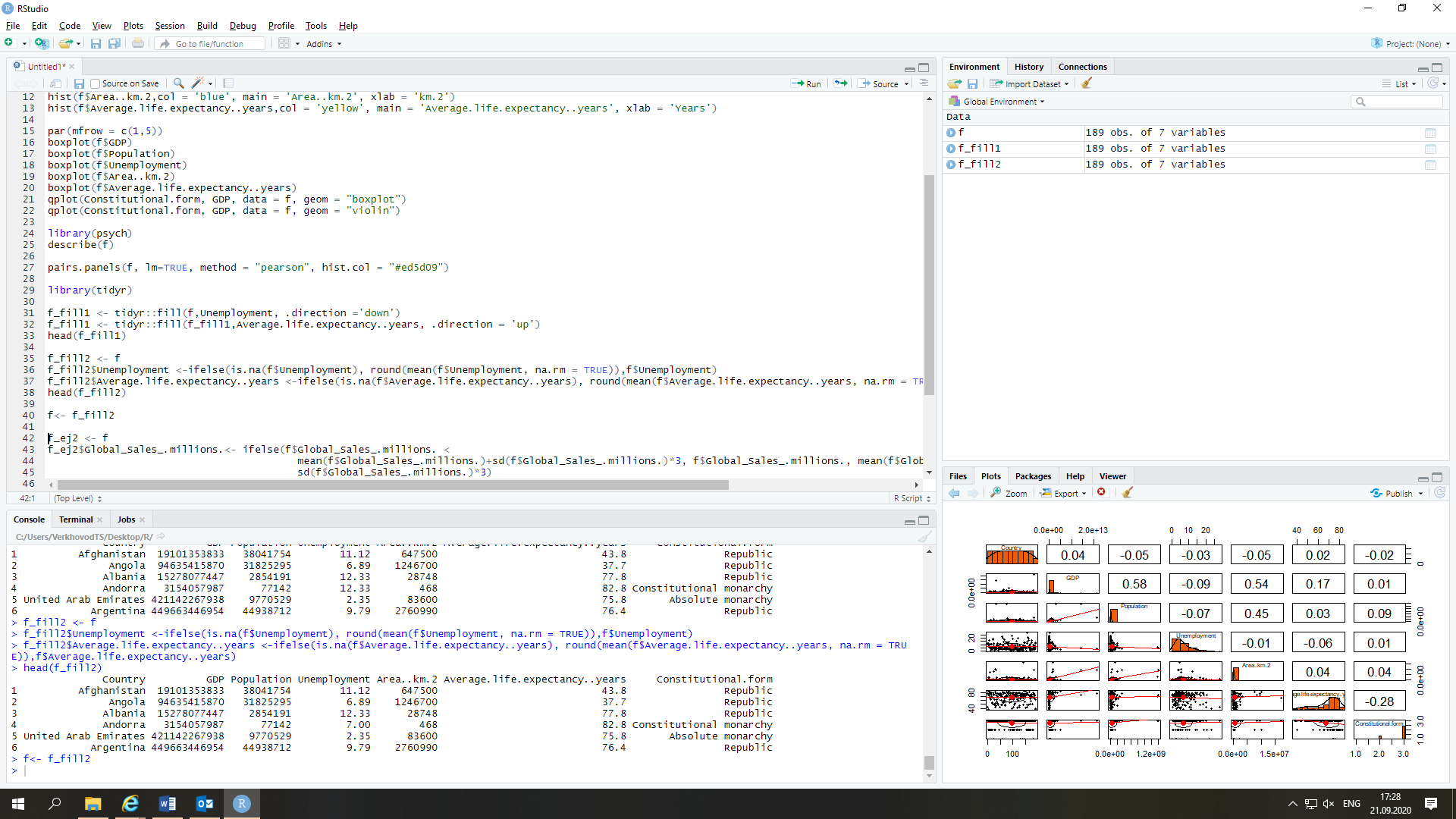
1. Аналіз основних показників описової статистики за кожною змінною показав, що пропущені значення є в змінних «Unemployment» та «Average.life.expectancy..years». Змінні «GDP», «Population», «Unemployment», «Area..km.2» мають викиди.



1. Найбільше на «GDP» мають вплив показники «Population» та «Area..km.2». Також, можемо простежити взаємозв’язок між «Population» та «Area..km.2», тобто кількістю населення та площею країни.



1. Пропуски були заповнені середніми значеннями.



1. Для корекції викидів в усіх випадках обраний варіант заповнення граничними значеннями.
2. Датасет розподілений на навчальну та тестову вибірки. Результати збережені в окремих файлах.