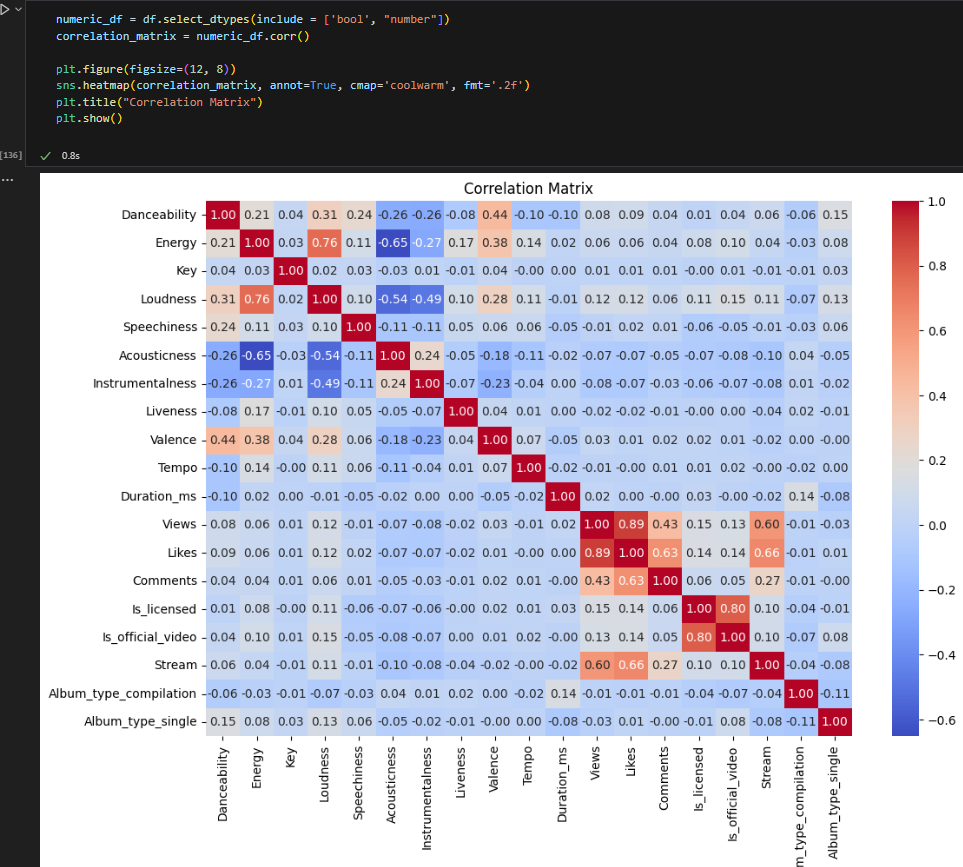
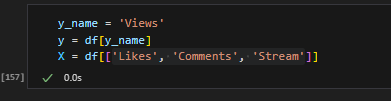
Загружаємо дані та форматуємо їх для роботи:

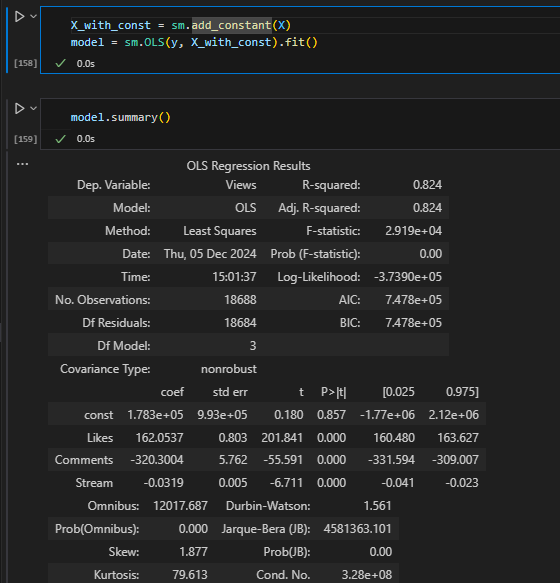
Будуємо матрицю кореляції щоб вибрати ознаки



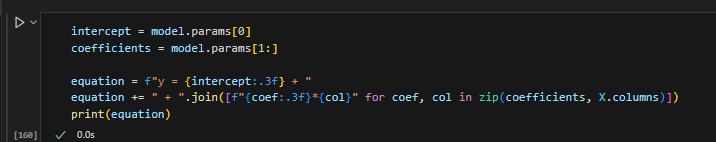
Будемо прогнозувати Views від Likes, Comments, Stream



Будуємо модель

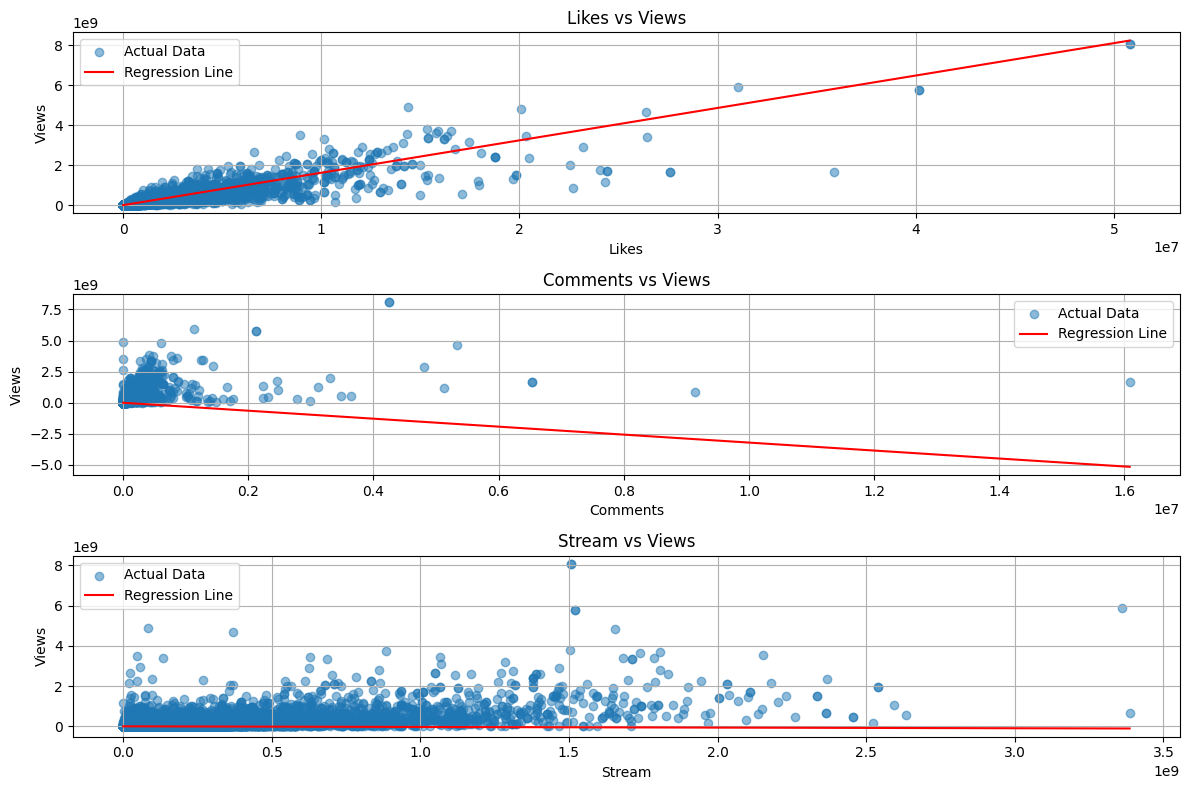
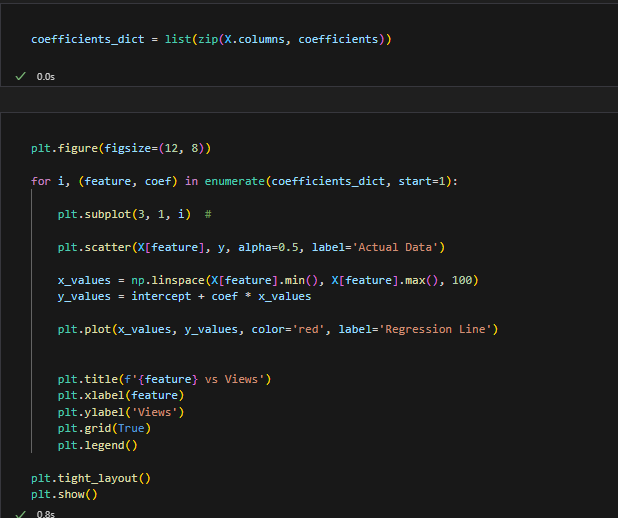


Отримуємо з моделі коефіцієнти

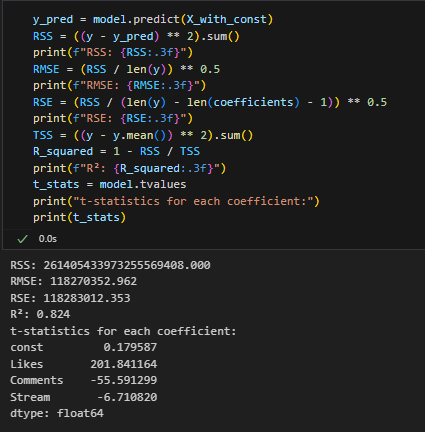
І маємо рівняння:

y = 178322.119 + 162.054\*Likes + -320.300\*Comments + -0.032\*Stream

Побудуємо графіки з лініями регресії

Маємо дивні лінії регресії для двух ознак. Напевно, оскільки модель будувалась для трьох змінних, то будувати регресії для окремих змінних є не гарною ідеєю.

Порахуємо показники: RSS, RMSE, RSE, R^2, t\_b.



Оскільки дані мають великі значення, то і RSS, RMSE, RSE набувають відповідно високих значень. R^2 показує гарний результат, модель пояснює 82.4% варіації в даних.

Кожне значення t-статистики показує, наскільки значущий коефіцієнт в моделі. Кількість лайків сильно пов'язана з Views. У Comments висока значущість, але негативний знак коефіцієнта вказує на зворотну залежність. Можливо, більше коментарів пов'язано з меншим числом переглядів (?). Stream виявився найменш значущим.

Побудуємо передбачення для три кроки вперед

