

## Дані:

Дано 3 файли- authors.csv, posts.csv та companies.csv

У першому файлі дані про блогерів:

Author ID- id автора

Name - назва автора

MEIBI - метрика ідентифікації впливу блогера

MEIBIX- метрика для визначення якості блогу щодо кількості та часу посилань на коментарі

Average Number Of Words In Posts - середня кількість слів у постах

Average Number Of Words In Posts (without stopwords) - середня кількість слів у постах без шумових слів

У другому файлі дані про пости:

Post ID - id поста

Title - заголовок

Blogger's Name - ім'я блогера

Blogger's ID - id блогера

Number of comments - кількість коментарів

Content - текст поста

Date - дата створення поста

Number of retrieved inlinks - Кількість отриманих посилань

Number of retrieved comments - Кількість отриманих коментарів

Дані можна з'єднати по Author ID (файл authors.csv) та Blogger's ID(файл posts.csv).

У третьому файлі - інформація про компанії:

Name - назва компанії

Year Founded - рік заснування

Industry - тематика компанії

Size Range - величина компанії

Country - країна реєстрації компанії

Current Employee Estimate - сумарна кількість співробітників

## Задача:

- виділити групи блогерів так, щоб ці групи якомога більше відрізнялися між собою.
- описати створені групи блогерів (чому групи розділились і як це можна використати)
- розробити рекомендаційний алгоритм для замовлення реклами у блогерів з метою просування продукту компанії