Ссылка на страницу, с которой была взята информация (я брала фотографии из главного альбома, т.к. именно публикаций на стене у меня мало):

**https://vk.com/kottmonavt**

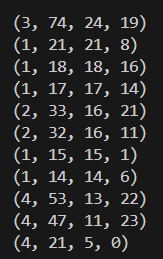
Вся работа была выполнена на языке Python при помощи библиотек sqlite3 и pandas для интеграции csv файла в таблицу sql.

1. Посмотрим, как влияет время публикации на количество лайков.

SQL запрос для данной проверки:

SELECT COUNT(id) AS count, SUM(likes) AS sum\_likes, SUM(likes)/COUNT(id) AS average, hour\_of\_publication FROM Data GROUP BY hour\_of\_publication ORDER BY average DESC

Результат:



(Первый столбец – количество постов в отдельный час, второй – сумма лайков по этим постам, третий столбец – среднее количество лайков в данный час, четвертый столбец – час публикации)

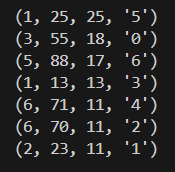
Больше всего постов было опубликовано после 18 часов вечера. Больше всего посты набирали лайков в промежутке от 8 утра до 21 вечера, т.е. +- время бодрствования обычного человека. Тогда как более поздние или более ранние посты набирали значительно меньше лайков.

1. Посмотрим, как влияет день недели публикации на количество лайков.

SQL запрос для данной проверки:

SELECT COUNT(id) AS count, SUM(likes) AS sum\_likes, SUM(likes)/COUNT(id) AS average, strftime("%w", date\_of\_publication) AS day from Data GROUP BY day ORDER BY average DESC

Результат:



(Первый столбец – количество постов в отдельный день, второй – сумма лайков по этим постам, третий столбец – среднее количество лайков в данный день, четвертый столбец – день недели, где 0 – воскресенье, 1 – понедельник, 2 – вторник, 3 – среда, 4 – четверг, 5 – пятница, 6 - суббота)

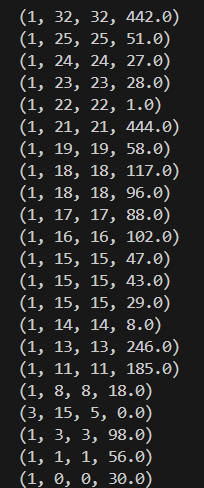
По результатам можем видеть, что больше всего лайков посты набирали именно в 3 последних дня недели, т.е. пятницу, субботу и воскресенье.

1. Посмотрим, как влияет количество дней между публикациями на количество лайков.

SQL запрос для данной проверки:

SELECT COUNT(id) AS count, SUM(likes) AS sum\_likes, SUM(likes)/COUNT(id) AS average, JULIANDAY(date\_of\_next\_publication) - JULIANDAY(date\_of\_publication) AS date\_difference from Data GROUP BY date\_difference ORDER BY average DESC

Результат:



(Первый столбец – количество постов, второй – сумма лайков по постам, третий столбец – среднее количество лайков, четвертый столбец – количество дней между текущей и прошлой публикацией)

Здесь нет какой-то определенной корреляции по лайкам… количество дней относительно прошлой публикации заметно не влияет на лайки.