МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

(ФГБОУ ВО «КубГУ»)

Кафедра прикладной математики

КУРСОВАЯ РАБОТА

АНАЛИЗ ПРЕДПОЧТЕНИЙ УЧАСТНИКОВ СООБЩЕСТВ С ПОМОЩЬЮ API "ВКОНТАКТЕ"

Работу выполнил \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Тихоненко В.В.

(подпись, дата) (инициалы, фамилия)

Факультет компьютерных технологий и прикладной математики курс 3

Направление 09.03.03 Прикладная информатика

Научный руководитель

профессор, к.ф. – м.н.,

доцент\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ В.Н. Кармазин

 (подпись, дата) (инициалы, фамилия)

Нормоконтролер,

к. ф.– м. н.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Г.В. Калайдина

(подпись, дата) (инициалы, фамилия)

Краснодар 2018

СОДЕРЖАНИЕ

[ВВЕДЕНИЕ 3](#_Toc516744155)

[1 API «ВКонтакте». Общие сведения. 4](#_Toc516744156)

[1.2 Мониторинг социальных сетей 5](#_Toc516744157)

[2 Разработка программы 7](#_Toc516744158)

[2.1.1 Методы и библиотеки 7](#_Toc516744159)

[2.1.2 Основной цикл программы 8](#_Toc516744160)

[2.2 Запись полученных данных в таблицу 11](#_Toc516744161)

[3 Анализ предпочтений студентов 13](#_Toc516744162)

[3.1 Сортировка по количеству просмотров 17](#_Toc516744163)

[3.2 Сортировка по количеству лайков 19](#_Toc516744164)

[3.3 Сортировка по количеству репостов 21](#_Toc516744165)

[ЗАКЛЮЧЕНИЕ 24](#_Toc516744166)

[ПРИЛОЖЕНИЕ 26](#_Toc516744167)

# ВВЕДЕНИЕ

Наиболее популярной социальной сетью в России является «ВКонтакте»

«ВКонтакте» – социальная сеть, российский аналог сервиса Facebook, самый посещаемый сайт Украины и Белоруссии, второй по популярности в России. Сайт изначально позиционировал себя в качестве социальной сети студентов и выпускников российских высших учебных заведений, позднее – как универсальный способ связи для всех социальных групп и возрастов.

«ВКонтакте» позволяет создавать свои сообщества. Создание и продвижение собственного сообщества является наиболее востребованным и часто применяемым инструментом продвижения во «ВКонтакте». Свое сообщество позволяет решить сразу несколько важных задач: во–первых, постоянно оставаться на связи с пользователями, сообщать им, к примеру, о новинках ассортимента или о скидочных акциях. Во–вторых, получать обратную связь от пользователей. В–третьих, информировать о компании новую аудиторию [1].

В ходе данной курсовой работы была написана программа на языке Python 3. Задача программы состоит в получении информации об интересах участников сообщества «ВКонтакте». Для анализа были выбраны группы «Подслушано в КубГУ» и «Подслушано в КубГМУ». Программа позволяет выводить данные в электронную таблицу в формате csv.

Из анализа полученных данных можно будет сделать вывод о том, какой контент больше нравится участникам сообщества.

# 1 API «ВКонтакте». Общие сведения.

API (от англ. application programming interface) – это интерфейс взаимодействия между сайтом и сторонними программами и серверами, набор готовых классов, процедур, функций, структур и констант, предоставляемых приложением (библиотекой, сервисом) или операционной системой для использования во внешних программных продуктах. Более просто – API это целый комплекс разнообразных готовых функций, готовых классов, констант и т.д., которые можно использовать в сторонних приложениях. API ВКонтакте постоянно дорабатывается и сегодня это – уже полнофункциональный набор дополнений, позволяющий создать любое приложение.

API ВКонтакте – позволяет получать информацию из базы данных vk.com с помощью http–запросов к специальному серверу. Вам не нужно знать в подробностях, как устроена база, из каких таблиц и полей каких типов она состоит – достаточно того, что API–запрос об этом «знает». Синтаксис запросов и тип возвращаемых ими данных строго определены на стороне самого сервиса [2].

Средства API ВКонтакте позволяют собирать и анализировать данные со страниц пользователей, страниц сообществ данной социальной сети, размещать разнообразные полезные виджеты на сайт.

Последнее глобальное преобразование этой среды разработки было проведено еще в апреле 2009 года. Конечно, и после этого сайт постоянно модифицировался.

Отметим основные возможности API ВКонтакте:

- На API был реализован доступ приложений к друзьям, их фото, видеозаписям. Приложения сейчас могут самостоятельно создавать альбомы, размещать в них информацию.

- Рассылка приложениями уведомлений.

- Приглашать друзей в приложения стало проще также благодаря интерфейсу прикладного программирования, на котором была реализована данная функция массового приглашения.

- Со страниц приложений исчезла реклама от ВК.

- Стало доступным API позволяющее использовать голоса ВКонтакте в качестве внутренних расчетов.

- Возможность запуска приложений в полноэкранном режиме.

- Порядок вывода приложений также реализован на API. Сегодня по умолчанию, те игры, которые запускаются чаще всего – находятся у вас на первых позициях. Которыми давно не пользовались – найдете в конце списка.

- Для самых популярных приложений стала возможна регистрация коротких поддоменов (например chat.vk.com и аналогичных) для простого запоминания и быстрого доступа.

- Появилась возможность вывода информации на личных страницах пользователей.

## 1.2 Мониторинг социальных сетей

Мониторинг социальных сетей – это выборка текстов пользователей социальных сетей, основанная на тех или иных критериях отбора данных текстов, а так же процесс сбора данной выборки и её интерпретации.

Мониторинг обязательно проводится перед стартом работ в социальных сетях, чтобы выяснить отношение аудитории к бренду, и регулярно повторяется, чтобы отследить динамику показателей.

Мониторинг социальных сетей дает ответы на важные вопросы:

* каково текущее восприятие бренда аудиторией;
* какие темы лучше обсуждать с аудиторией;
* как работают конкуренты в социальных сетях и каковы результаты

их PR–деятельности;

* на каких площадках целевая аудитория наиболее активна, и какие

из них больше всего влияют на репутацию бренда.

Мониторинг социальных сетей позволяет:

* быстро реагировать на потребности аудитории (часто на

негатив);

* принять решение о том, какая стратегия работы в социальных

сетях будет наиболее подходящей для бренда [3].

# 2 Разработка программы

### 2.1.1 Методы и библиотеки

Python – это высокоуровневый язык программирования с широким набором функций, наиболее универсальный из всех. Как правило, в первую очередь он направлен на высокую производительность и хорошую читаемость программного кода. Чтобы добиться этого синтаксис его ядра максимально упрощён, но вот в расширениях большое количество самых разных, в том числе и нестандартных функций. Его преимуществом является также поддержка сразу нескольких парадигм программирования, также он всегда позиционировался как объектно – ориентированный язык.

Данный язык программирования отличает огромное количество модулей, как входящих в стандартную поставку Python 3, так и сторонних.

В ходе выполнения были использованы следующие библиотеки:

* Requests – выполняет HTTP–запросы.
* csv – модуль для чтения и записи файла в формате SCV.
* Datetime – модуль предоставляет классы для обработки времени

и даты разными способами.

Во время разработки задачи необходимо было обратиться к серверу с запросом, для получения данных о каждой записи в сообществе. Это позволяет сделать метод API «wall.get». Он используется к методам для работы с записями на стене сообщества.

Данный метод возвращает список записей со стены пользователя или сообщества.

В таблице 1 приведены параметры данного метода:

Таблица 1 – Список параметров метода wall.get

|  |  |
| --- | --- |
| owner\_id | идентификатор пользователя или сообщества, со стены которого необходимо получить записи (по умолчанию – текущий пользователь) |
| domain | короткий адрес пользователя или сообщества |
| offset | смещение, необходимое для выборки определенного подмножества записей |
| count | количество записей, которое необходимо получить. |
| filter | определяет, какие типы записей на стене необходимо получить. Возможные значения:  suggests – предложенные записи на стене сообщества (доступно только при вызове с передачей access\_token);  postponed – отложенные записи (доступно только при вызове с передачей access\_token);  owner – записи владельца стены;  others – записи не от владельца стены;  all – все записи на стене (owner + others). |
| extended | 1 – в ответе будут возвращены дополнительные поля profiles и groups, содержащие информацию о пользователях и сообществах. По умолчанию: 0. |
|  | |

После успешного выполнения возвращает объект, содержащий число результатов в поле count и массив объектов записей на стене в поле items.

Если был задан параметр extended=1, возвращает число результатов в поле count, отдельно массив объектов записей на стене в поле items, пользователей в поле profiles и сообществ в поле groups [5].

### 2.1.2 Основной цикл программы

Перед тем как нам приступить к анализу данных, необходимо было получить записи со стены обоих сообществ. Для этого необходимо было в программе такие параметры, как owner\_id, offset и count. Owner\_id – идентификатор нужного нам сообщества («Подслушано в Кубгу»/«Подслушано в КубГМУ»). Параметром count было задано количество постов, которое будет выведено одним таким запрос. В данной программе было выбрано 50 постов. Параметр offset задавался динамически в основном цикле программы, чтобы можно было на каждой итерации цикла программы перемещаться по записям сообщества и получать записи, следующие после уже полученных.

В результате запрос к серверу и основной цикл программы приняли следующий вид:

group\_id = *'–* *59061778’ #’-* *55982026”*

offset = 0

date\_x = 1494374400

all\_posts = []

while True:

sleep(1)

r = requests.get(*'https://api.vk.com/method/wall.get'*,params={*'owner\_id'*:group\_id,*'count'*:50,*'offset'*:offset})

posts = r.json()[*'response'*]

all\_posts.extend(posts)

oldest\_post\_date = posts[–1][*'date'*]

offset += 50

print(offset)

if oldest\_post\_date < date\_x:

break

В переменной group\_id хранится идентификатор рассматриваемого нами сообщества. Для удобства он задается не в самом запросе, а в виде переменной, которая передается в запрос.

В переменной date\_x хранится дата в формате Unix epoch. UNIX–время или POSIX–время – способ кодирования времени, принятый в UNIX и других POSIX–совместимых операционных системах.

В данной программе будем рассматривать посты за прошедший год. Поэтому в переменной date\_x хранится дата, обозначающая тот же день год назад (10.05.2017). Она же является условием остановки цикла.

В цикле, с помощью запроса request.get, происходит обращение к базе данных vk.com для получения 50 постов анализируемой группы. Следующим шагом полученные посты записывается в массив posts, чтобы в дальнейшем их можно было записать в массив all\_posts, в котором содержатся посты полученные ранее. Затем происходит смещение (offset) на 50 постов вперед. Цикл программы выполняется пока значение переменной oldest\_post\_date не станет меньше значения даты 10.05.2017.

Действуют ограничения на число запросов к API. Для клиентского приложения лимит составляет 3 rps (раза в секунду). Для серверного приложения действует прогрессивная шкала, в зависимости от числа установок приложения (rps/число установок): 5/<10000, 8/<100000, 20/<1000000. 35/>1000000 [4].

Чтобы избежать ошибок в ходе выполнения программы, необходимо было установить ограничение на количество выполняемых запросов в секунду. Для этого была использована функция sleep в начале цикла. Она позволяет на каждой итерации приостановить цикл на одну секунду. Таким образом, выполнение запросов происходит не чаще одного раза в секунду.

Также в основной процедуре программы предусмотрен подсчет времени, за которое будут получены искомые данные. Для этой цели в начале процедуры в переменную start заносится текущее значение времени, а в конце процедуры, после работы цикла, в переменную end снова заносится текущее значение времени. Таким образом, отняв полученные значения, можно узнать, сколько времени прошло с момента запуска основной процедуры и до момента окончания ее работы.

В конце основной процедуры происходит вывод на экран значения в переменной total, в которой отражено время, затраченное на выполнение программы. Также на экран выводится длина массива all\_posts, чтобы можно было увидеть, сколько постов в итоге удалось получить.

## 2.2 Запись полученных данных в таблицу

Так как цель работы – получение данных о количестве просмотров, лайков и репостов для каждого поста за прошедший год в рассматриваемом сообществе, то для удобства следует записать эти данные в таблицу для удобства представления и последующей сортировки.

Для этого в программе используется цикл поочередного вызова двух процедур: get\_data и write\_csv:

post\_data = []

for post in all\_posts:

post\_data = get\_data(post)

write\_csv(post\_data)

Процедура get\_data получает на вход фрагмент данных из массива all\_posts, в котором хранятся все полученные посты за прошедший год, и выбирает из всех представленных полей записи только необходимые нам для анализа.

Далее процедура write\_csv производит непосредственную запись в файл формата csv. Если в указанной директории нет файла с таким именем, то он будет создан. Если есть – будет произведена запись в указанный файл. В каждую строку поочередно записываются просмотры, лайки, репосты и текст записи.

Обе процедуры были описаны в начале программы, до основной процедуры. Основная процедура просто осуществляла их вызов.

Следует также отметить, что разработанная программа является универсальной. Это значит, что при замене идентификатора группы можно будет получить и записать посты за прошедший год любого сообщества.

# 3 Анализ предпочтений студентов

После успешного выполнения программы было получено по 50 записей со стен групп «Подслушано в КубГУ» и «Подслушано в КубГМУ».

В марте 2017 «ВКонтакте» была введена функция подсчета просмотров. Просмотры записи – важный показатель статистики. Зачастую количество лайков не показывает насколько запись интересна. Реальное количество просмотров всегда в разы больше.

Фрагменты таблиц, полученных после работы программы:

Таблица 2 – Результат работы программы для сообщества «Подслушано в КубГУ»

|  |  |  |  |
| --- | --- | --- | --- |
| views | likes | reposts | Text |
| 5290 | 5 | 0 | Если ты боишься выступать, не знаешь как подать себя публике, хочешь научиться манипулировать вниманием слушателей или улучшить уже имеющиеся навыки публичных выступлений, тебе точно сюда! Специально для тебя, в рамках недели ОСО КубГУ, мы устраиваем ШОК! Ты всё ещё не понял? |
| 6760 | 12 | 0 | А вы бы выдержали эксперимент? https://m.lenta.ru/news/2018/03/23/student/ |
| 6189 | 3 | 0 | Товарищи студенты!!! Очень прошу вас помочь с опросом. Заранее всем огромная благодарность!!! https://docs.google.com/forms/d/e/1FAIpQLScLyOXDqwY9bJqbVLjPDALvImB746shKsHbvDmG60x31pED7Q/viewform |
| 5449 | 48 | 0 | Бесит, когда ты достаточно эрудирован, чтобы указать преподавателю, на момент, где ты с ним не согласен, но недостаточно эрудирован для того, чтобы вести после этого полноценную дискуссию. |
| 4145 | 13 | 0 | Постоянно на 5 этаже в корпусе интернет центра с подоконников убираю банки из под напитков и кофейные стаканы, если замечу. Не превращайте наш университет в помойку! |

Продолжение таблицы 2

|  |  |  |  |
| --- | --- | --- | --- |
| 4391 | 13 | 2 | На территории универа найдена собака, явно домашняя Если кто знает хозяина или может взять ее на передержку, онау заднего входа на эконом |
| 3566 | 0 | 0 | Сегодня, 19 марта, в аудитории 234, корпус с, химический факультет, в 7 вечера оставила флешку, черно-синяя без корпуса. На первом ряду от окна. |
| 3750 | 4 | 0 | Здравствуйте. Помогите, пожалуйста, в написании диплома нужно пройти вооот этот опрос. Очень буду признательна Нужно 380 человек;( https://docs.google.com/forms/d/1UCW92GooULw1DWfXrw |
| 3571 | 36 | 0 | Заходит как-то Дорошенко В.В. на пару: - Можно объявление? К нам приехала передвижная лаборатория, можно сходить, сделать тест на вич. В подарок дают ручку. Хорошо же? И блокнотик! Замечательно! И еще кое-что со вкусом ананаса. \*подмигивает и уходит\* |
| 3158 | 31 | 1 | Учитель в школе: я не буду навязывать вам свои политические убеждения, все сами потом решите. Преподаватель в универе: садитесь, а я сейчас объясню вам, какое у вас должно быть мнение по поводу гражданского общества и демократии в стране. |
| 11651 | 41 | 0 | "Вот вы узнали на паре что-то новое? Пополнили свой багаж знаний? О! День уже прошел не зря! Вас можно поздравить!" Гайденко В.В. |
| 10679 | 50 | 0 | Скоро сессия. Радуюсь, что могу теперь про нее шутить (шучу, не радуюсь) 1 курс. Юрфак. |
| 7230 | 1 | 0 | Уважаемые студенты, преподаватели и все-все-все:) Всем привет! Прошу принять участие в анонимном опросе, если Вас не затруднит! Нам нужна статистика для проекта по созданию web-квеста) Заранее большое всем спасибо!!! |
| 7367 | 58 | 0 | Кто самолет не рисовал, тот бух учета и не знал! |
| 5617 | 16 | 0 | У вас была/есть любимая пара в универе, вот чтобы прям интересно было и вы ни разу не пропустили? Какая?🤗 |

Таблица 3 – Результат работы программы для сообщества «Подслушано в КубГМУ»

|  |  |  |  |
| --- | --- | --- | --- |
| 792 | 33 | 0 | Я считаю, что 1 балла достаточно за волонтёрство! Ну реально, на половину пар не ходят, ничего не отрабатывают, да ещё и к ординатуре хорошую прибавку хотят, не много ли? Я бы ещё и забирала.Очень рада! |
| 1922 | 29 | 1 | Встречайте нового отечественного супергероя: ЧЕЛОВЕК-ВОЛОНТЕР Вы спросите, какими способностями он обладает? А я вам отвечу: +1 балл к ординатуре |
| 1140 | 30 | 0 | 1 балл за волонтерство! Как Вам такой подарочек от волонтерского центра? Ууу, закачаешься)) Че ты там делал? 5-6 лет как дибил пахал во благо вуза и государства? На держи свой 1 балл, мало? Скажи спасибо, что не числанули, половину занятий пропустил, ты в мед поступил или в кубик? Ты же волонтер, бескорыстный бесплатный работник, в этом вся суть, 1 балла тебе даже много. |
| 3593 | 93 | 1 | Отчисление - это как музыкальная карьера Бузовой. Видишь результат, а как это получилось - не понятно |
| 8435 | 208 | 0 | - Это плохая примета, если я тебя запомнил. (с) Байбаков С.Е. |
| 11692 | 137 | 0 | - Какие вы знаете медиаторы воспаления? - \*Начинает отвечать, но что-то пошло не так, и ответ не удался\* - Ребят, после таких ответов у меня работает все время один медиатор - БОЛИ... ©Кравченко |
| 10746 | 114 | 0 | На пересдаче: - Здравствуйте, можно итоговую ответить? - Нет, я уже ухожу, завтра приходите - Ну можно мы быстро? - Нуу быстро только кошки родят... (с) Похотько |
| 9481 | 119 | 0 | - Что можете рассказать про венозное полнокровие? - Я помню, что там происходит закупорка ствола! - Ствола какого? Яблони? Дуба? ©Кравченко |

Продолжение таблицы 3

|  |  |  |  |
| --- | --- | --- | --- |
| 10635 | 313 | 2 | На паре Жукова А.В.: -Почему "ядро одиночного пути" так называется? - Ну вот препарировали: на что было прохоже, так и назвали. Фибула - заколка для волос, Бука - музыкальный инструмент-труба, во время игры на которой сильно раздувались щеки. В ваше время были бы айфоноподобные образования или какой-нибудь "бугорок Оксимирона", "пучки Гнойного" - Как-то он даже репчик зачитывал по теме занятия Любимый преподаватель |
| 10188 | 473 | 3 | Шантыз, указывая на кость: -Возьми. -Кого? -Меня! |
| 19570 | 366 | 0 | Пара по пат.анатомии : Преподаватель: "как называется фермент , расщепляющий жиры ?" - Тишина... Преподаватель :" ну углеводы расщепляет амилаза , белки - протеиназы , а жиры ?" -студент : "Жиробаза ?" |
| 18321 | 207 | 1 | Концентрация ядов 2% - поздно пить боржоми (с) Литвинова Т.Н. |
| 16377 | 277 | 0 | Мало кто знает, что Славинский А.А. подрабатывает римским прокуратором. |
| 14530 | 363 | 2 | Началось Ежегодная рубрика "Сентябрьский Первокурсник" |
| 23063 | 268 | 2 | Пришло время сессии. А уровень твоих знаний по экзаменационным предметам на уровне консервированного горошка. |

В таблицах все данные расположены в той последовательности, в которой они были на стене сообщества. На данном этапе сложно сделать какие–либо выводы об актуальности тех или иных записей. Чтобы наши таблицы были информативными, необходимо выполнить сортировку и посмотреть какие записи набрали наибольшее количество лайков, репостов и просмотров.

## 3.1 Сортировка по количеству просмотров

Просмотры – это важный показатель, который гораздо лучше отражает какой контент пользователям больше интересен.

В результате сортировки получены следующие результаты, отраженные в таблице 4 и 5:

Таблица 4 – Записи, лидирующие по количеству просмотров «Подслушано в КубГУ»

|  |  |  |  |
| --- | --- | --- | --- |
| views | likes | reposts | Text |
| 11651 | 41 | 0 | "Вот вы узнали на паре что-то новое? Пополнили свой багаж знаний? О! День уже прошел не зря! Вас можно поздравить!" Гайденко В.В. |
| 10679 | 50 | 0 | Скоро сессия. Радуюсь, что могу теперь про нее шутить (шучу, не радуюсь) 1 курс. Юрфак. |
| 7367 | 58 | 0 | Кто самолет не рисовал, тот бух учета и не знал! |
| 7230 | 1 | 0 | Уважаемые студенты, преподаватели и все-все-все:) Всем привет! Прошу принять участие в анонимном опросе, если Вас не затруднит! Нам нужна статистика для проекта по созданию web-квеста) Заранее большое всем спасибо!!! |
| 6760 | 12 | 0 | А вы бы выдержали эксперимент? https://m.lenta.ru/news/2018/03/23/student/ |
| 11651 | 41 | 0 | "Вот вы узнали на паре что-то новое? Пополнили свой багаж знаний? О! День уже прошел не зря! Вас можно поздравить!" Гайденко В.В. |
| 6189 | 3 | 0 | Товарищи студенты!!! Очень прошу вас помочь с опросом. Заранее всем огромная благодарность!!! https://docs.google.com/forms/d/e/1FAIpQLScLyOXDqwY9bJqbVLjPDALvImB746shKsHbvDmG60x31pED7Q/viewform |

Количество просмотров говорит о том, сколько человек увидели данную запись. После сортировки мы имеем: лидерами по количеству просмотров являются записи с цитатами преподавателей и опросы, проводимые студентами.

Таблица 5 – Записи, лидирующие по количеству просмотров «Подслушано в КубГМУ»

|  |  |  |  |
| --- | --- | --- | --- |
| views | likes | reposts | Text |
| 23063 | 268 | 2 | Пришло время сессии. А уровень твоих знаний по экзаменационным предметам на уровне консервированного горошка. |
| 19570 | 366 | 0 | Пара по пат.анатомии : Преподаватель: "как называется фермент , расщепляющий жиры ?" - Тишина... Преподаватель :" ну углеводы расщепляет амилаза , белки - протеиназы , а жиры ?" -студент : "Жиробаза ?" |
| 18321 | 207 | 1 | Концентрация ядов 2% - поздно пить боржоми (с) Литвинова Т.Н. |
| 16377 | 277 | 0 | Мало кто знает, что Славинский А.А. подрабатывает римским прокуратором. |
| 14530 | 363 | 2 | Началось Ежегодная рубрика "Сентябрьский Первокурсник" |
| 11692 | 137 | 0 | - Какие вы знаете медиаторы воспаления? - \*Начинает отвечать, но что-то пошло не так, и ответ не удался\* - Ребят, после таких ответов у меня работает все время один медиатор - БОЛИ... ©Кравченко |
| 10746 | 114 | 0 | На пересдаче: - Здравствуйте, можно итоговую ответить? - Нет, я уже ухожу, завтра приходите - Ну можно мы быстро? - Нуу быстро только кошки родят... (с) Похотько |

После сортировки мы имеем: лидерами по количеству просмотров, как и в сообществе «Подслушано в КубГУ», являются записи с цитатами преподавателей. Однако студентам КубГМУ также интересны саркастические записи в сообществе.

## 3.2 Сортировка по количеству лайков

Произведем сортировку по количеству лайков (по убыванию). В результате в начале таблицы оказались следующие записи, приведенные в таблице 6 и 7:

Таблица 6 – Записи, лидирующие по количеству лайков «Подслушано КубГУ»

|  |  |  |  |
| --- | --- | --- | --- |
| views | likes | reposts | Text |
| 7367 | 58 | 0 | Кто самолет не рисовал, тот бух учета и не знал! |
| 10679 | 50 | 0 | Скоро сессия. Радуюсь, что могу теперь про нее шутить (шучу, не радуюсь) 1 курс. Юрфак. |
| 5449 | 48 | 0 | Бесит, когда ты достаточно эрудирован, чтобы указать преподавателю, на момент, где ты с ним не согласен, но недостаточно эрудирован для того, чтобы вести после этого полноценную дискуссию. |
| 11651 | 41 | 0 | "Вот вы узнали на паре что-то новое? Пополнили свой багаж знаний? О! День уже прошел не зря! Вас можно поздравить!" Гайденко В.В. |
| 3571 | 36 | 0 | Заходит как-то Дорошенко В.В. на пару: - Можно объявление? К нам приехала передвижная лаборатория, можно сходить, сделать тест на вич. В подарок дают ручку. Хорошо же? И блокнотик! Замечательно! И еще кое-что со вкусом ананаса. \*подмигивает и уходит\* |
| 3158 | 31 | 1 | Учитель в школе: я не буду навязывать вам свои политические убеждения, все сами потом решите. Преподаватель в универе: садитесь, а я сейчас объясню вам, какое у вас должно быть мнение по поводу гражданского общества и демократии в стране. |
| 5617 | 16 | 0 | У вас была/есть любимая пара в универе, вот чтобы прям интересно было и вы ни разу не пропустили? Какая? |

Исходя из полученных результатов, одну из лидирующих строчек занимают посты которые направлены на поднятие настроения. Это следовало ожидать, так как участники сообщества, как правило, лайкают то, что подняло им настроение.

Также оказались популярными посты, как и в предыдущем разделе, с темой «цитаты преподавателей».

Таблица 7 – Записи, лидирующие по количеству лайков «Подслушано КубГМУ»

|  |  |  |  |
| --- | --- | --- | --- |
| views | likes | reposts | Text |
| 10188 | 473 | 3 | Шантыз, указывая на кость: -Возьми. -Кого? -Меня! |
| 19570 | 366 | 0 | Пара по пат.анатомии : Преподаватель: "как называется фермент , расщепляющий жиры ?" - Тишина... Преподаватель :" ну углеводы расщепляет амилаза , белки - протеиназы , а жиры ?" -студент : "Жиробаза ?" |
| 14530 | 363 | 2 | Началось Ежегодная рубрика "Сентябрьский Первокурсник" |
| 10635 | 313 | 2 | На паре Жукова А.В.: -Почему "ядро одиночного пути" так называется? - Ну вот препарировали: на что было прохоже, так и назвали. Фибула - заколка для волос, Бука - музыкальный инструмент-труба, во время игры на которой сильно раздувались щеки. В ваше время были бы айфоноподобные образования или какой-нибудь "бугорок Оксимирона", "пучки Гнойного" - Как-то он даже репчик зачитывал по теме занятия Любимый преподаватель |
| 16377 | 277 | 0 | Мало кто знает, что Славинский А.А. подрабатывает римским прокуратором. |
| 23063 | 268 | 2 | Пришло время сессии. А уровень твоих знаний по экзаменационным предметам на уровне консервированного горошка. |
| 8435 | 208 | 0 | Это плохая примета, если я тебя запомнил. (с) Байбаков С.Е. |

После сортировки мы наблюдаем следующее: интересы студентов КубГУ и КубГМУ полностью совпали. Однако следует заметить, что КубГМУ являются более активными пользователями. Максимальное количество лайков в из группе составляет 473 когда а группе « Подслушано в КубГУ» всего 58, учитывая то, что количество просмотров примерно одинаково.

## 3.3 Сортировка по количеству репостов

Репост записи зачастую делается людьми, если информация кажется им достаточно интересной и полезной, чтобы ею можно было поделиться с друзьями, либо сохранить и прочитать еще раз в свободное время. Поэтому количеству репостов записей следует уделить достаточно внимания при анализе.

Произведем сортировку по количеству репостов. Результаты представлены в таблицах 8 и 9:

Таблица 8 – Записи, лидирующие по количеству репостов «Подслушано в КубГУ»

|  |  |  |  |
| --- | --- | --- | --- |
| views | likes | reposts | text |
| 4391 | 13 | 2 | На территории универа найдена собака, явно домашняя Если кто знает хозяина или может взять ее на передержку, онау заднего входа на эконом |
| 3158 | 31 | 1 | Учитель в школе: я не буду навязывать вам свои политические убеждения, все сами потом решите. Преподаватель в универе: садитесь, а я сейчас объясню вам, какое у вас должно быть мнение по поводу гражданского общества и демократии в стране. |
| 7367 | 58 | 0 | Кто самолет не рисовал, тот бух учета и не знал! |
| 10679 | 50 | 0 | Скоро сессия. Радуюсь, что могу теперь про нее шутить (шучу, не радуюсь) 1 курс. Юрфак. |

Таблица 9 – Записи, лидирующие по количеству репостов «Подслушано в КубГМУ»

|  |  |  |  |
| --- | --- | --- | --- |
| views | likes | reposts | text |
| 10188 | 473 | 3 | Шантыз, указывая на кость: -Возьми. -Кого? -Меня! |
| 14530 | 363 | 2 | Началось Ежегодная рубрика "Сентябрьский Первокурсник" |
| 10635 | 313 | 2 | На паре Жукова А.В.: -Почему "ядро одиночного пути" так называется? - Ну вот препарировали: на что было прохоже, так и назвали. Фибула - заколка для волос, Бука - музыкальный инструмент-труба, во время игры на которой сильно раздувались щеки. В ваше время были бы айфоноподобные образования или какой-нибудь "бугорок Оксимирона", "пучки Гнойного" - Как-то он даже репчик зачитывал по теме занятия Любимый преподаватель |
| 23063 | 268 | 2 | Пришло время сессии. А уровень твоих знаний по экзаменационным предметам на уровне консервированного горошка. |
| 18321 | 207 | 1 | Концентрация ядов 2% - поздно пить боржоми (с) Литвинова Т.Н. |
| 3593 | 93 | 1 | Отчисление - это как музыкальная карьера Бузовой. Видишь результат, а как это получилось - не понятно |
| 1922 | 29 | 1 | Встречайте нового отечественного супергероя: ЧЕЛОВЕК-ВОЛОНТЕР Вы спросите, какими способностями он обладает? А я вам отвечу: +1 балл к ординатуре |
| 19570 | 366 | 0 | Пара по пат.анатомии : Преподаватель: "как называется фермент , расщепляющий жиры ?" - Тишина... Преподаватель :" ну углеводы расщепляет амилаза , белки - протеиназы , а жиры ?" -студент : "Жиробаза ?" |

В данном случае выявить какую-то закономерность невозможно, так как количество репостов, по сравнению с лайками и просмотрами, ничтожно мало. Из этого следует вывод, что студенты неохотно делятся записями из сообществ своего университета.

# ЗАКЛЮЧЕНИЕ

В ходе данной курсовой работы была разработана программа, позволяющая получать информацию из базы данных vk.com, выбирать интересующие нас данные и записывать их в табличный вид.

Был произведен анализ записей сообществ «ВКонтакте» «Подслушано в КубГУ» и «Подслушано в КубГМУ» за прошедший год по трём основным показателям активности.

Можно сказать, что интересы участников сообществ совпадают на примерно на 70%, однако есть и существенные различия.

Проанализировав полученные результаты, можно сказать, что участникам сообщества, прежде всего, интересны веселые посты и статьи содержащую тематическую информацию сообщества.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Халилов Д. Маркетинг в социальных сетях: Манн, Иванов и Фербер; Москва; 2013 – 230 с.

2. Знакомство с API ВКонтакте URL: https://vk.com/dev/first\_guide (дата обращения: 10.03.2018)

3. Загребельный Г., Боровик М. «Интернет–маркетинг за 55 минут»; Москва; 2012 – 50 стр.

4. Д. А. Храмов «Сбор данных в Интернете на языке R»; Москва; 2017 – 450 стр.

5. https://vk.com/dev/wall.get (дата обращения 15.04.2018)

# ПРИЛОЖЕНИЕ

Парсер социальных сетей

import requests

import json

import csv

from datetime import datetime

from time import sleep

def write\_json(data):

with open('posts1.json','w',encoding="utf–8") as file:

json.dump(data, file, ensure\_ascii=False, indent=2)

def get\_data(post):

try:

post\_id = post['id']

except:

post\_id = 0

try:

views = post['views']['count']

except:

views = 0

try:

likes = post['likes']['count']

except:

likes = 'zero'

try:

reposts = post['reposts']['count']

except:

reposts = 'zero'

try:

text = post['text']

except:

text = '\*\*\*'

data = {

'id': post\_id,

'views': views,

'likes': likes,

'reposts': reposts,

'text': text

}

return data

def write\_csv(data):

with open('posts\_data.csv','a', encoding="utf–8") as file:

writer = csv.writer(file, delimiter=';')

writer.writerow((data['views'], data['likes'], data['reposts'], data['text']))

def main():

start = datetime.now()

group\_id = '– 59061778’ # '-55982026'

offset = 0

date\_x = 1494374400

all\_posts = []

while True:

sleep(1)

r = requests.get('https://api.vk.com/method/wall.get',params={'owner\_id':group\_id,'count':50,'offset':offset})

posts = r.json()['response']

all\_posts.extend(posts)

oldest\_post\_date = posts[–1]['date']

offset += 50

print(offset)

if oldest\_post\_date < date\_x:

break

post\_data = []

for post in all\_posts:

post\_data = get\_data(post)

write\_csv(post\_data)

end = datetime.now()

total = end – start

print(total)

print(len(all\_posts))

if \_\_name\_\_ == '\_\_main\_\_':

main()