

Презентацію виконав
Студент другого курсу
Спеціальності "Прикладної математики"
Факультету інформатики
Савенко Нікіта
Викладач: Курочкін Андрій Володимирович



Final project report

Сенсовий ланцюг проробленої праці

Викачка даних,
особистих діалогів,
чатів з телеграму



Аналіз цих даних



Їх графічне
представлення



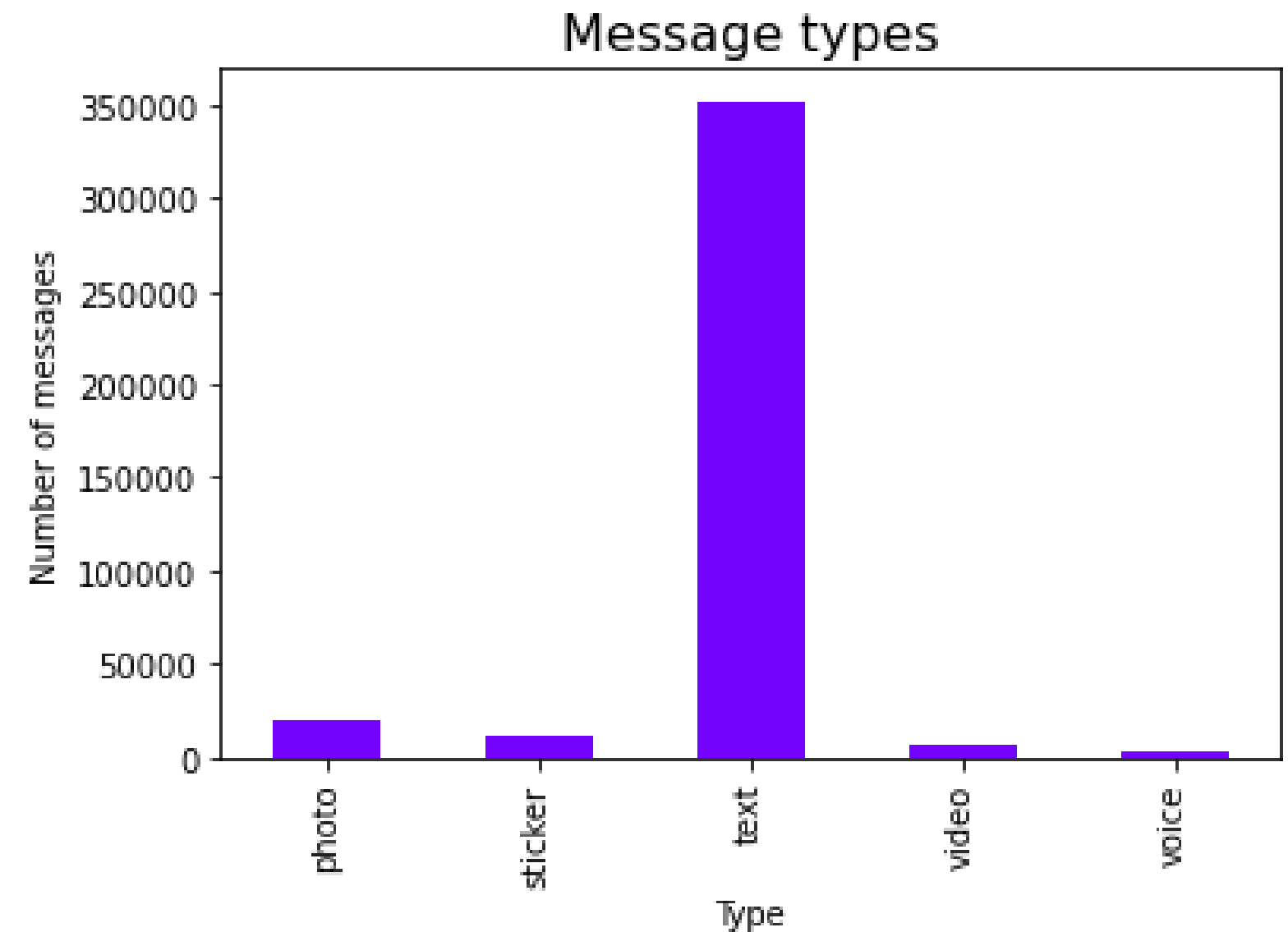
Певні висновки на основі побачених графіків, діаграм

Трохи статистики

- 1 На викачку всіх діалогів, чатів мені знадобилося близько **двох** годин
- 2 Сам етап завантаження пройшов нормально і ніяких проблем не виникало
- 3 Розмір датасету — **64,6** мб
- 4 Кількість усіх повідомлень — **421288**
- 5 Кількість усіх діалогів — **212**

"Типи повідомлень"

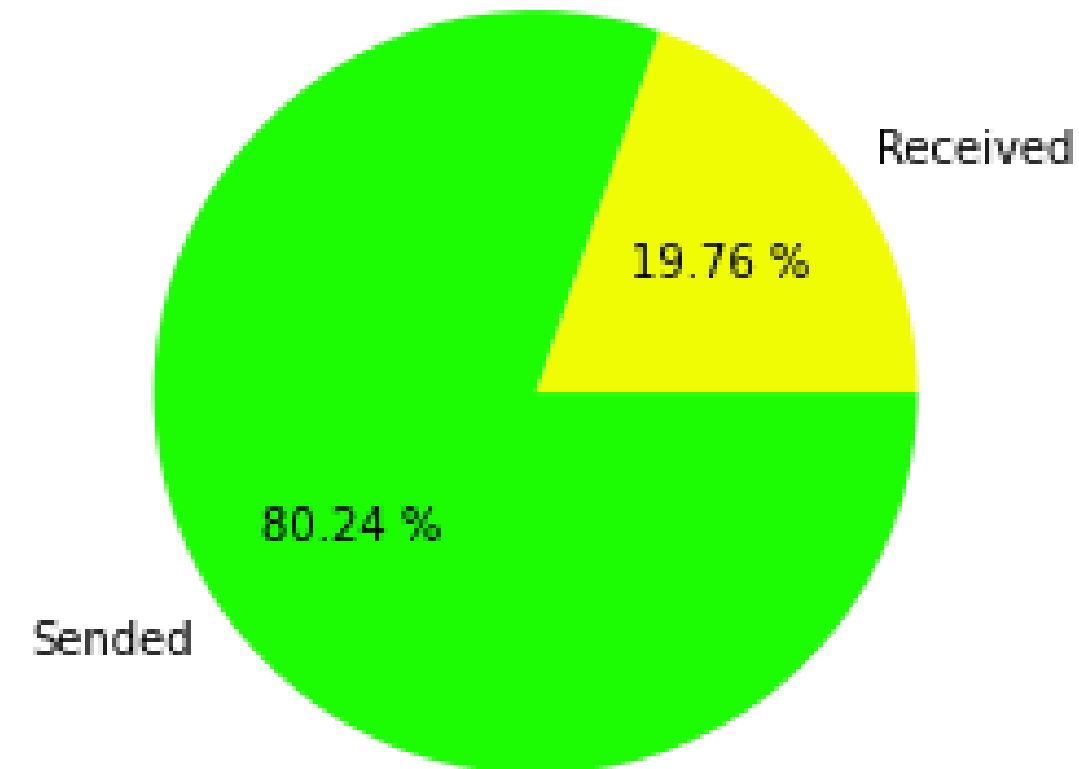
- Першопричиною створення даної гістограми була зацікавленість, наскільки більше людина вибирає текст як спосіб передачі інформації, порівнюючи з відео, аудіо та іншими.
- Можна дуже легко побачити, що найбільш популярним серед усіх типів повідомлень є текстові
- Далі за популярністю йдуть фото, стікери, відео, та аудіоповідомлення



"Хто більше пише, я чи мені?"

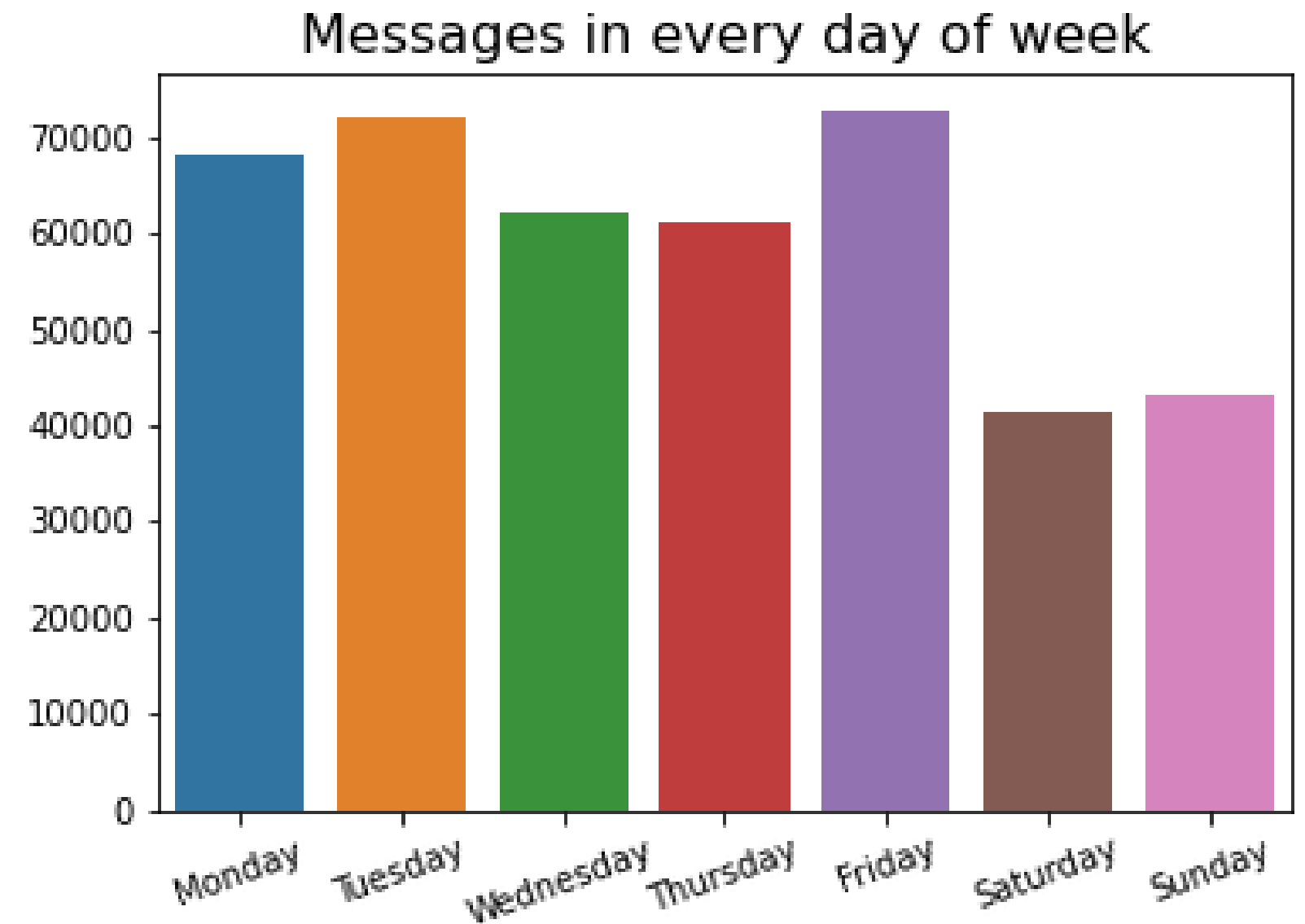
З цієї секторної діаграми помітно, що більшість повідомлень, а саме 80,2% були відправлені мною, лише 19,7% отримав я(

Who writes more?



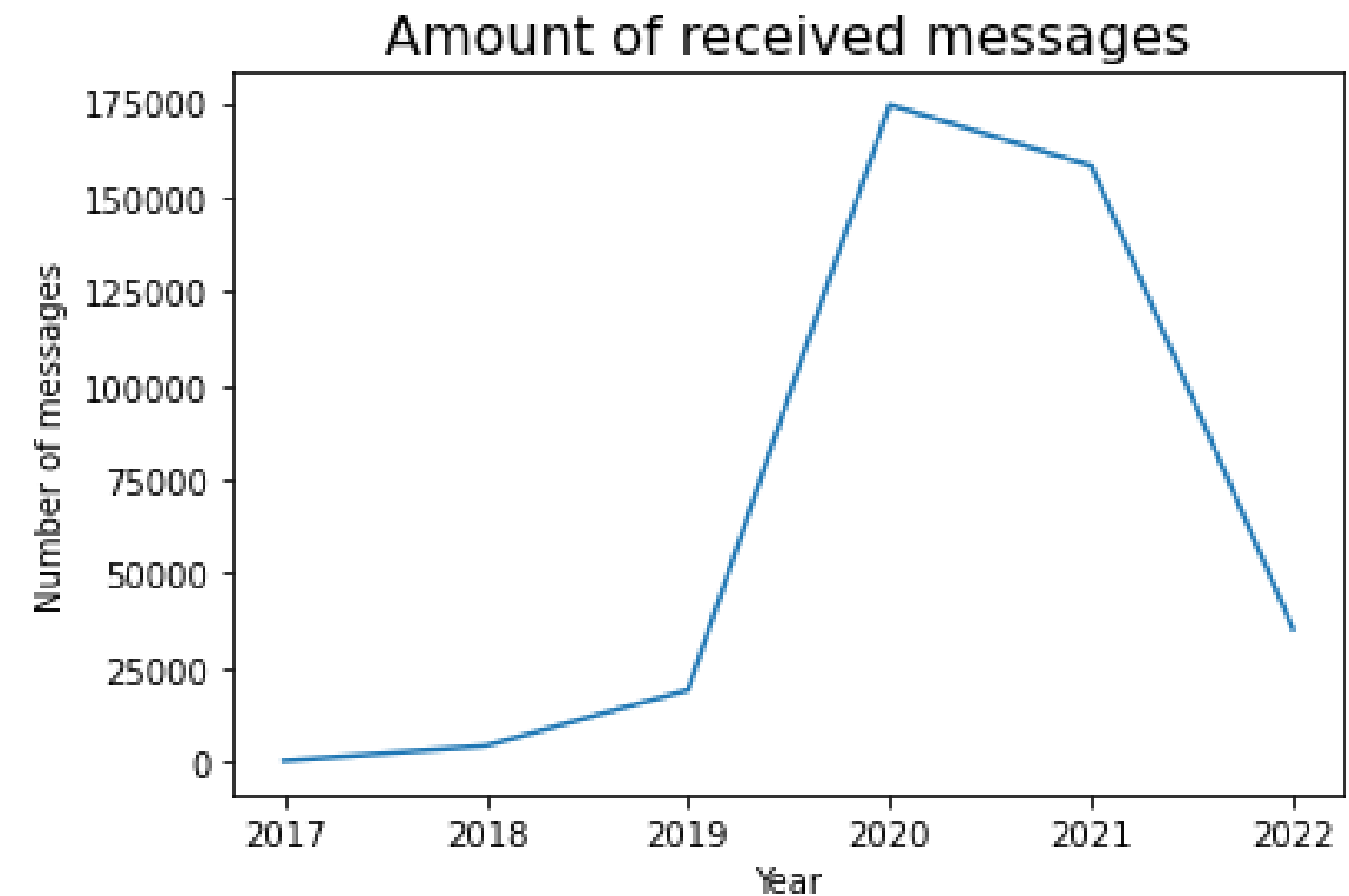
"Кількість повідомлень у різні дні тижня"

- Тут я вирішив дослідити, який серед днів тижня лідирує у кількості отриманих та відправлених повідомлень
- З невеликим відривом у вівторка перемагає п'ятниця
- Найменш "активним" днем, як виявилось, є субота



"Кількість отриманих повідомлень за 2017-2022 роки"

- Лінійна діаграма демонструє кількість отриманих мною повідомлень за 2017-2022 роки
- Бачимо, що найбільше писали мені у 2020-му році, скоріше за все це пов'язано із закінченням 11-го класу та вступу до вузу
- З 2021 по 2022 роки дуже помітний спад, який я пов'язую із коронавірусом та, так би мовити, певним відстороненням людей, а саме менше домовленостей про зустрічі, веселих бесід і так далі



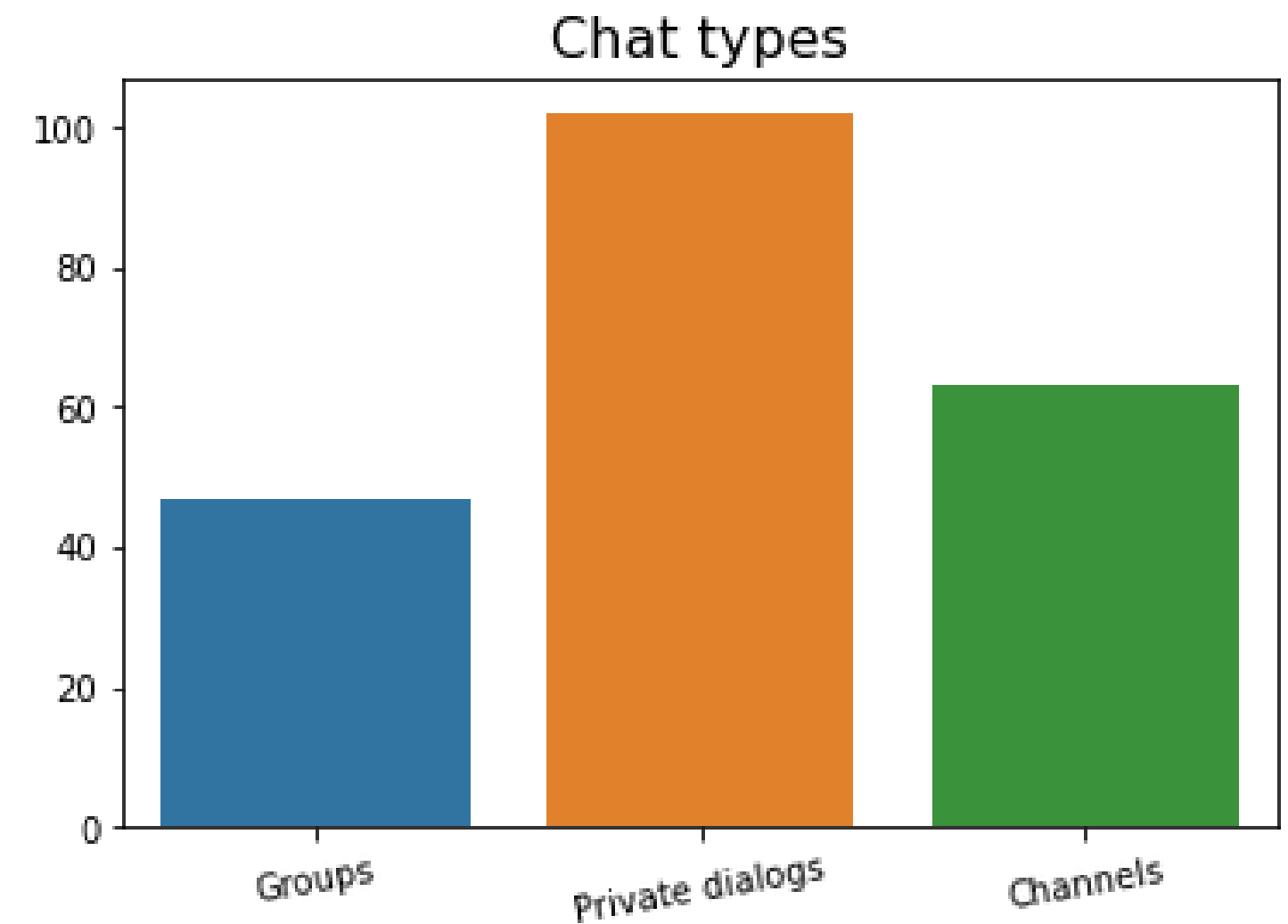
"Типи чатів"

Мотивом створення цієї діаграми стала моя особиста зацікавленість у тому, у скількох групах я є, кількість особистих діалогів та каналів блогерів, новин, розваг, на які я підписаний

Груп - **47**

Приватних діалогів - **102**

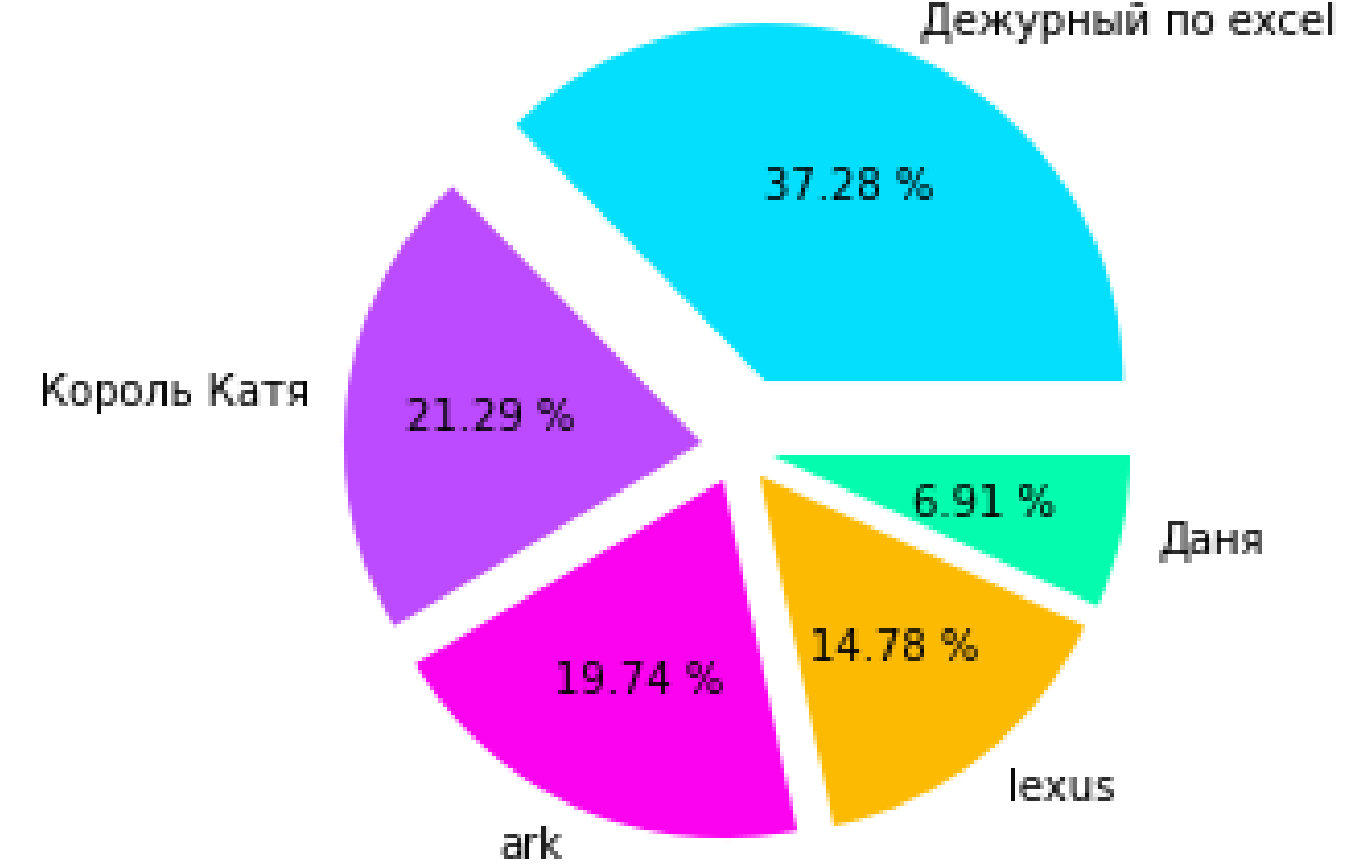
Каналів - **63**



"Топ 5 кому я писав найбільше"

Секторна діаграма, що відображає **топ 5** людей, яким я писав найбільше

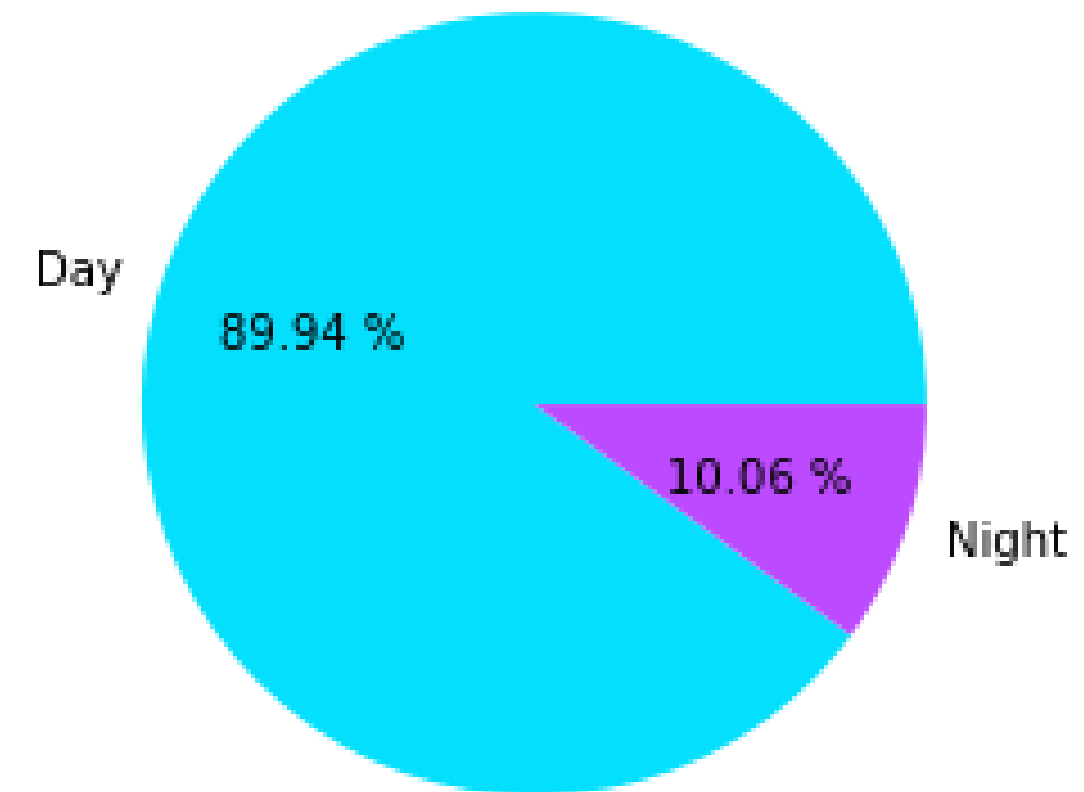
Top 5 persons to whom i wrote the most



"Денні та нічні повідомлення"

- Кількість повідомлень(отриманих і відправлених) в залежності від часу доби
- Як ніч, був взятий проміжок часу з 22:00 до 4:00
- А що тут скажеш, все правильно переписуватись потрібно вдень, а вночі спати)

Day/night messages



Наступні кроки

- Імпортувати csv файл зі статистикою коронавірусу, зробити графік росту хвороби та поєднати наприклад графік із сьомого слайду із створеним
- Відфільтрувати отримані та відправлені повідомлення за полом, проаналізувати кому чи хто пише більше мені, чоловіки чи жінки
- Зробити хмаринку слів у різних чатах, аби побачити які слова використовуються найчастіше

Дякую за увагу!

https://github.com/n1kitka/telegram_homework