**МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ**

**УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»**

**(РУТ (МИИТ)**

Институт/факультет «Академия “Высшая Инженерная Школа”, АВИШ»

Специальность/Направление подготовки Информатика и вычислительная техника

**КУРСОВАЯ РАБОТА**

**по дисциплине:** Технологии хранения больших данных

|  |  |
| --- | --- |
| **на тему:** | \_\_Аналитика игровых соревнований с помощью методов Big Data\_\_ |

|  |  |
| --- | --- |
| **Студент группы ШАД 212** | **(Гинда Д.А. )** |
| **Научный руководитель** | **(Зуев И.А. )** |

**Москва 2023 г.**

Оглавление

[Введение 3](#_Toc153827342)

[Исследование данных 4](#_Toc153827343)

[Заключение 13](#_Toc153827344)

[Список использованных источников и программ 14](#_Toc153827345)

[Приложение 15](#_Toc153827346)

## Введение

Актуальность данной темы заключается в том, что покер – одна из самых популярных карточных игр в мире, и онлайн-платформы предоставляют возможность любому желающему сыграть в эту игру, вне зависимости от его местоположения. Таким образом, онлайн турниры по покеру привлекают множество участников со всего мира.

Аналитика в онлайн покере может помочь игрокам стать более успешными игроками и получать больше прибыли. В основе покера лежит анализ ситуации, принятие рациональных решений и прогнозирование ходов соперников. Онлайн турниры предоставляют огромное количество статистических данных, которые можно использовать для анализа и принятия решений. В работе будет осуществлен анализ различных аспектов покерных турниров.

Результаты исследования позволят выявить ключевые факторы успеха данной отрасли, а также предложить рекомендации для улучшения результатов турниров. Полученные выводы могут быть полезными для игроков, инвесторов и тренеров в этой области, помогая им принимать обоснованные решения для достижения более высоких показателей.

В ходе работы будут использованы соответствующие теоретические и практические материалы, чтобы достичь поставленных целей.

## Исследование данных

Для проведения анализа был взят набор данных с результатами турниров одной из команд за 2022 год. Также в качестве аналитик было выбрано построение визуализации данных в виде графиков. Опишем эти графики:

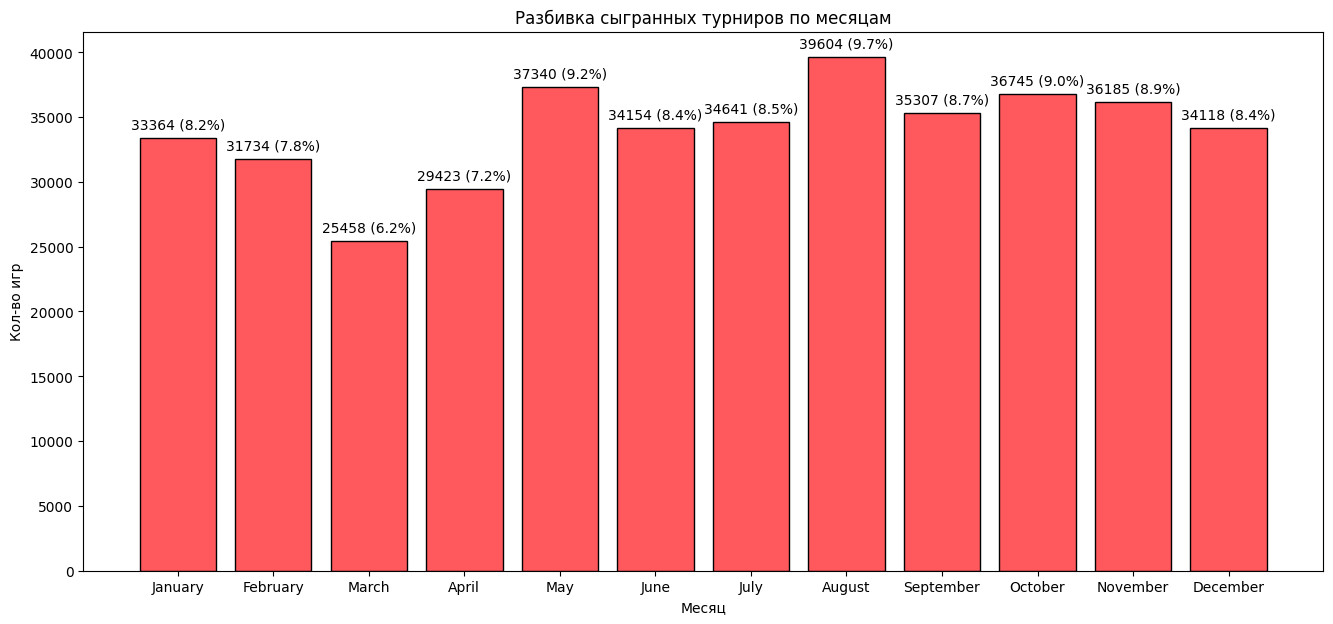


Рис. 1

В марте мы можем наблюдать спад активности, с последующим её увеличением. Из этого мы можем сделать вывод, что часть команды не могла участвовать в турнирах из-за каких-то обстоятельств, после чего, им нужны были дополнительные финансы, в связи с чем, в мае резко выросло число участий.

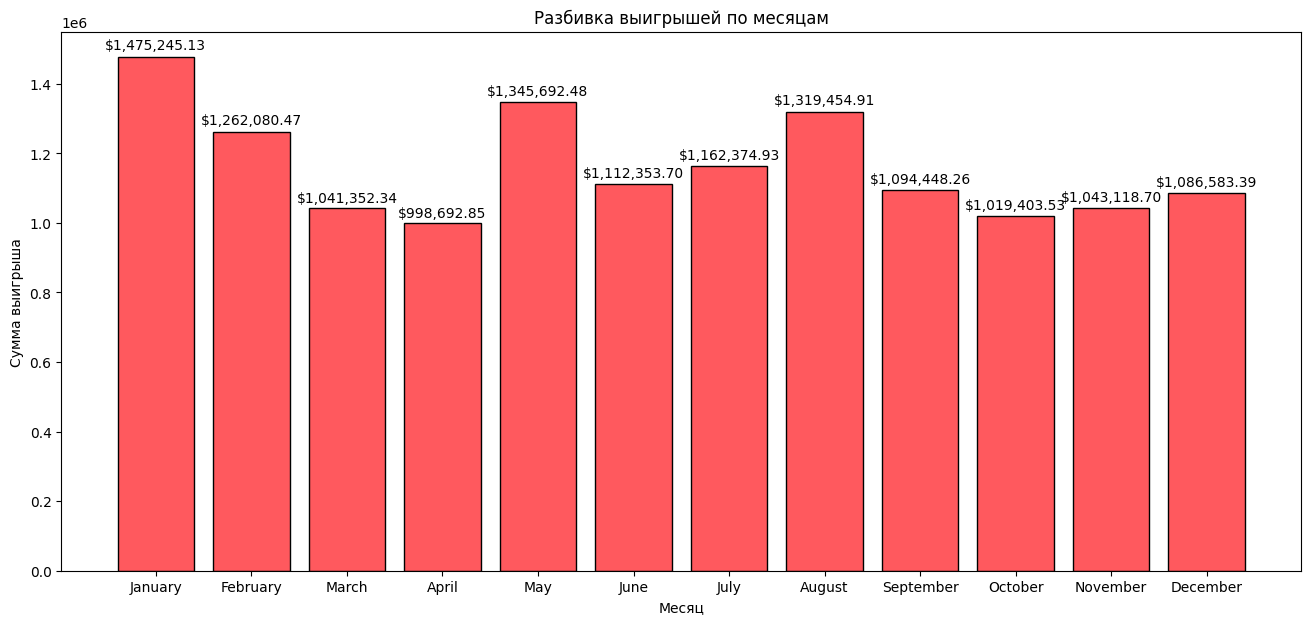


Рис. 2

И так, хоть в апреле было сыграно на 3965 турниров больше, чем в марте, однако общая сумма выигрышей в марте на $42659.49 больше чем в апреле. Некоторые могут сказать - "Вот, это да, какие цифры!", однако, это не показывает, сколько было получено прибыли, так как за большую часть турниров, нужно делать взнос, поэтому нужно ещё постараться, чтобы что-то заработать.

Чтобы узнать, на сколько прибыльной была инвестиция, используют ROI («Return On Investment», возврат инвестиций) — коэффициент рентабельности инвестиций, который помогает рассчитать окупаемость вложений в проект.

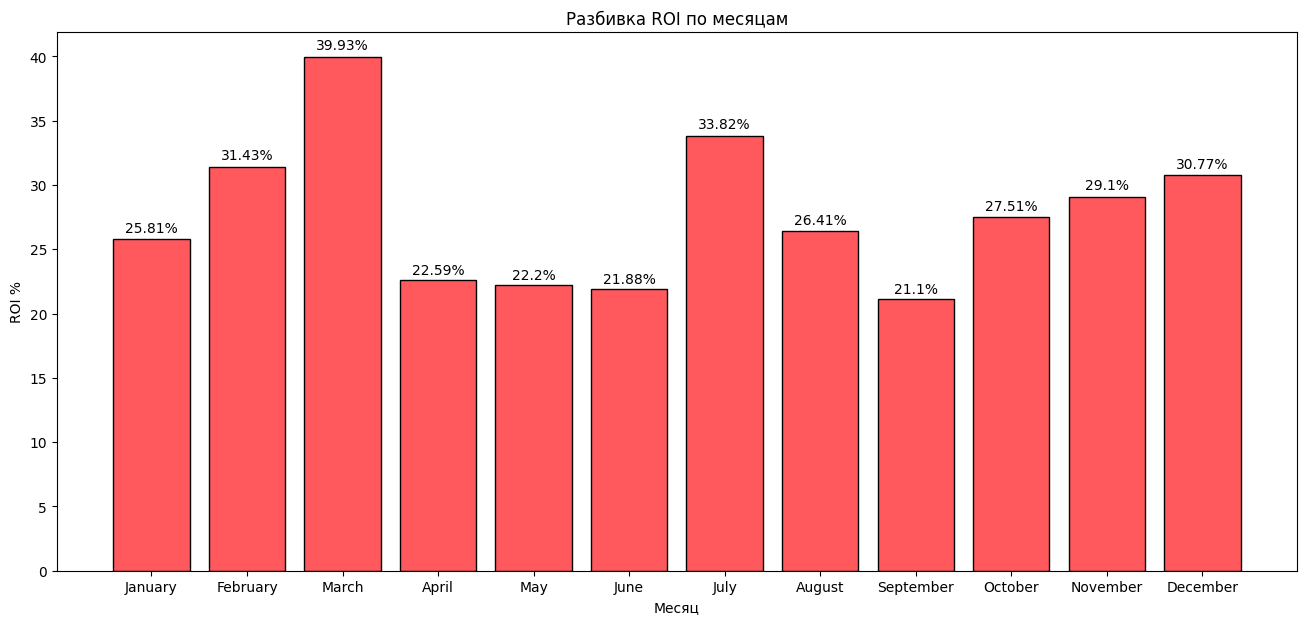


Рис. 3

Мы видим, что в марте прибыль составила почти 40%, эти данные вместе с предыдущими графиками, говорят нам о том, что был сделан упор не на количество, а на результат.

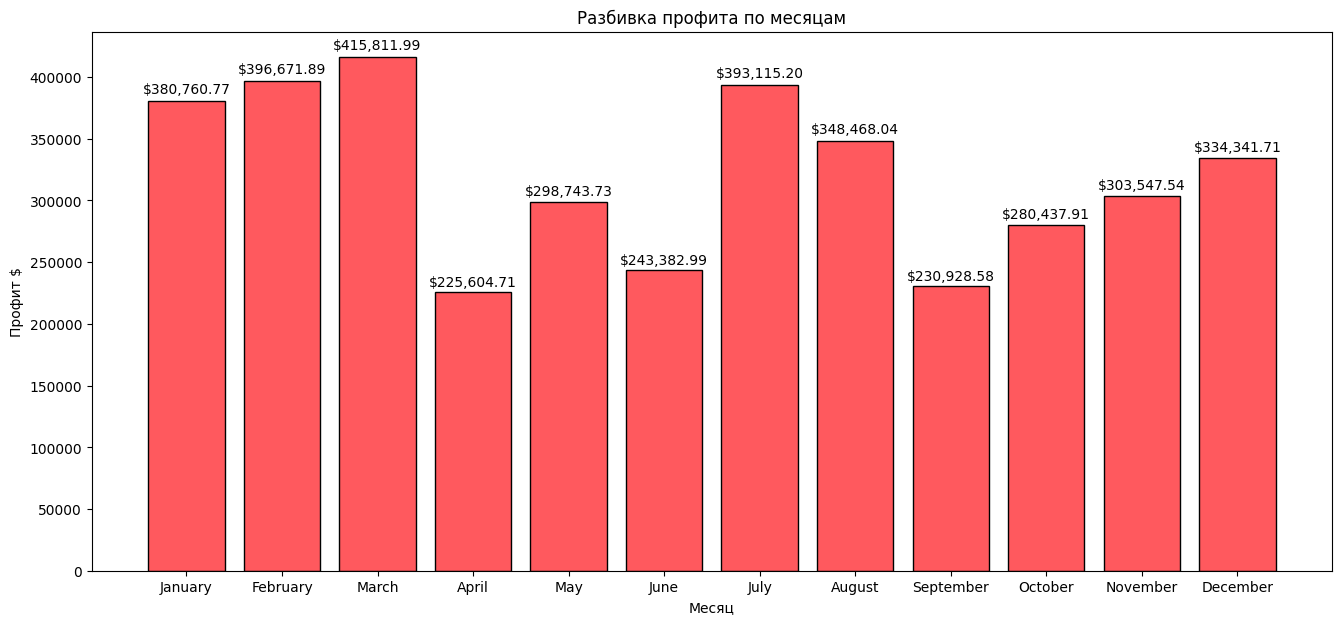


Рис. 4

Этот график нам показывает, что несмотря на то, что на Рис. 2 январь является самым продуктивным месяцем (суммарный выигрыш составил почти $1.5 миллиона), чистой прибыли он принёс всего на $380.760. Из этого мы делаем вывод: Для наиболее объективной оценки проделанной работы нужно проводить аналитику по различным параметрам.

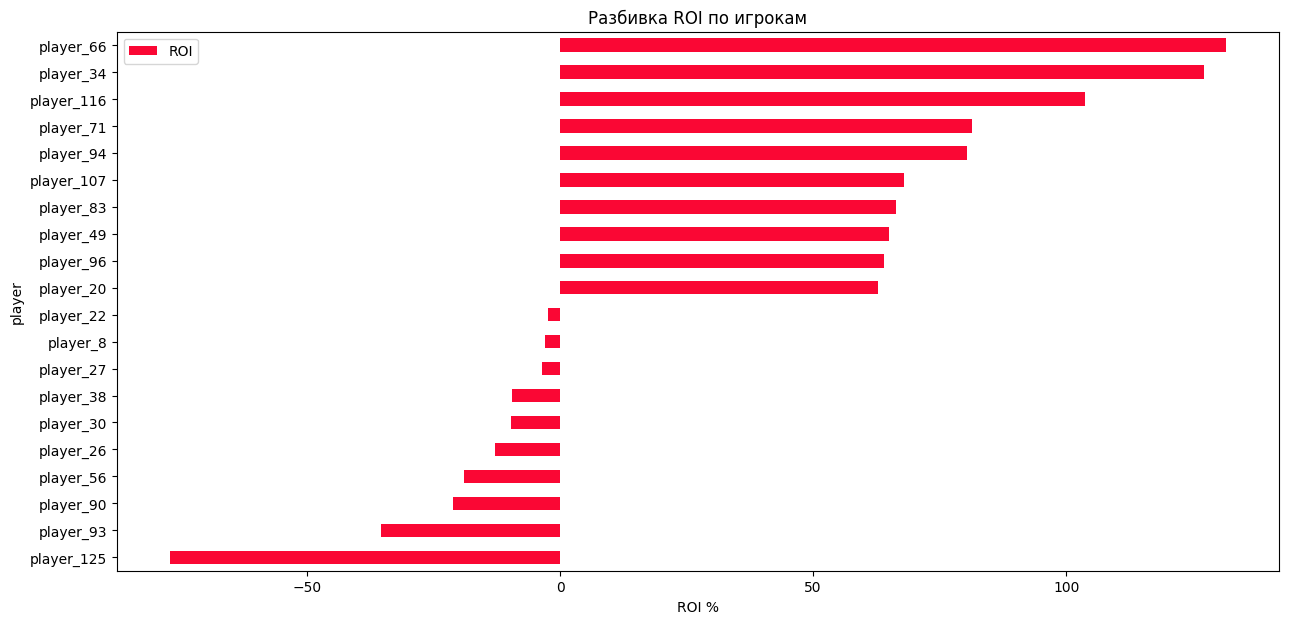


Рис. 5

На графике представлены топ 10 лучший и топ 10 худших игроков. Данный график свидетельствует нам о том, что не все игроки одинаково результативны – пока некоторые из них умножают вложенные инвесторами средства более чем в 2 раза, другие же эти средства теряют.

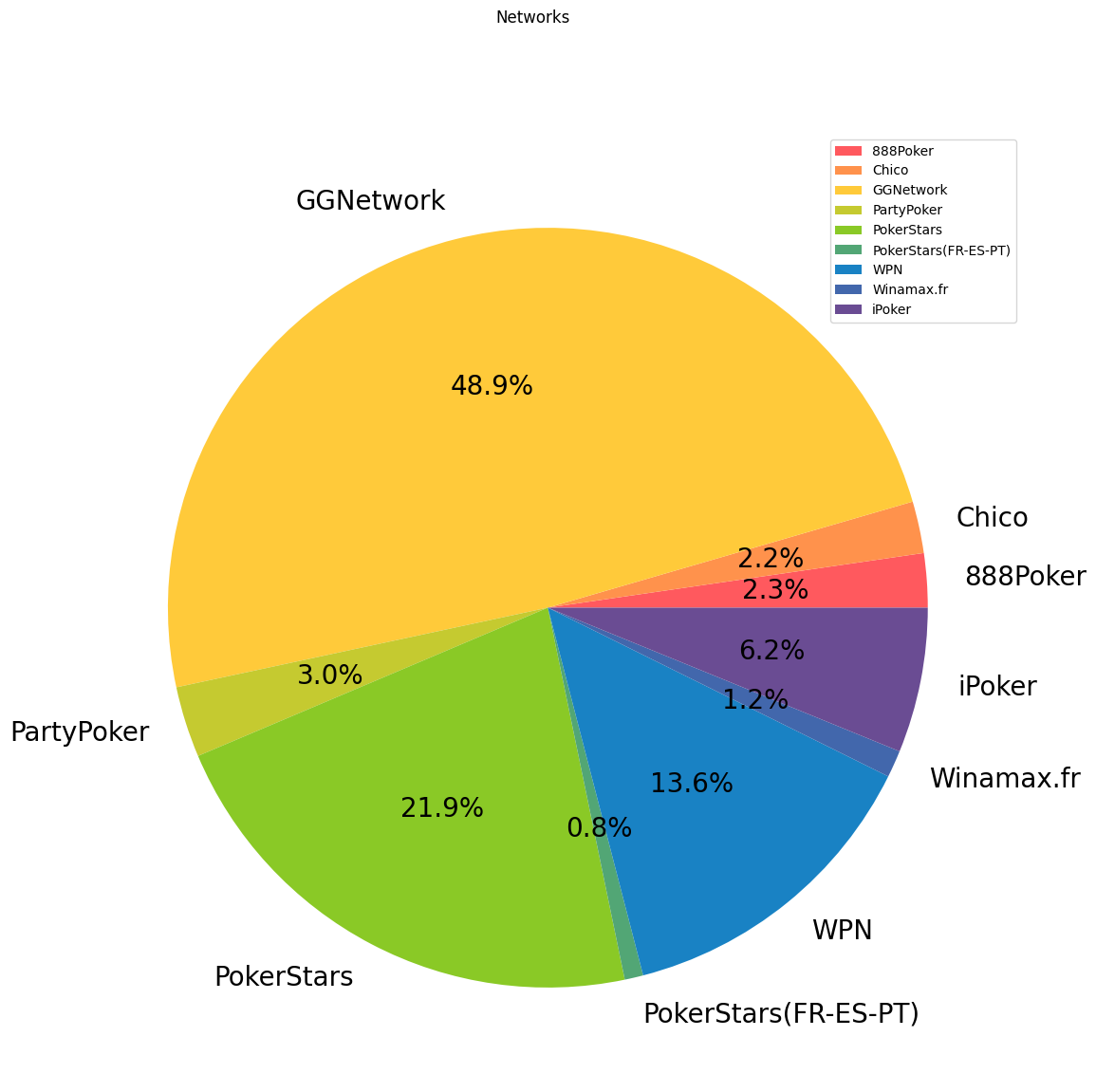


Рис. 6

На графике представлены данные о количестве турниров, проходящих в той или иной сети. Самая популярная из них GGNetwork (48.9%), на втором месте PokerStars (21.9%), за ним WPN (13.6%).

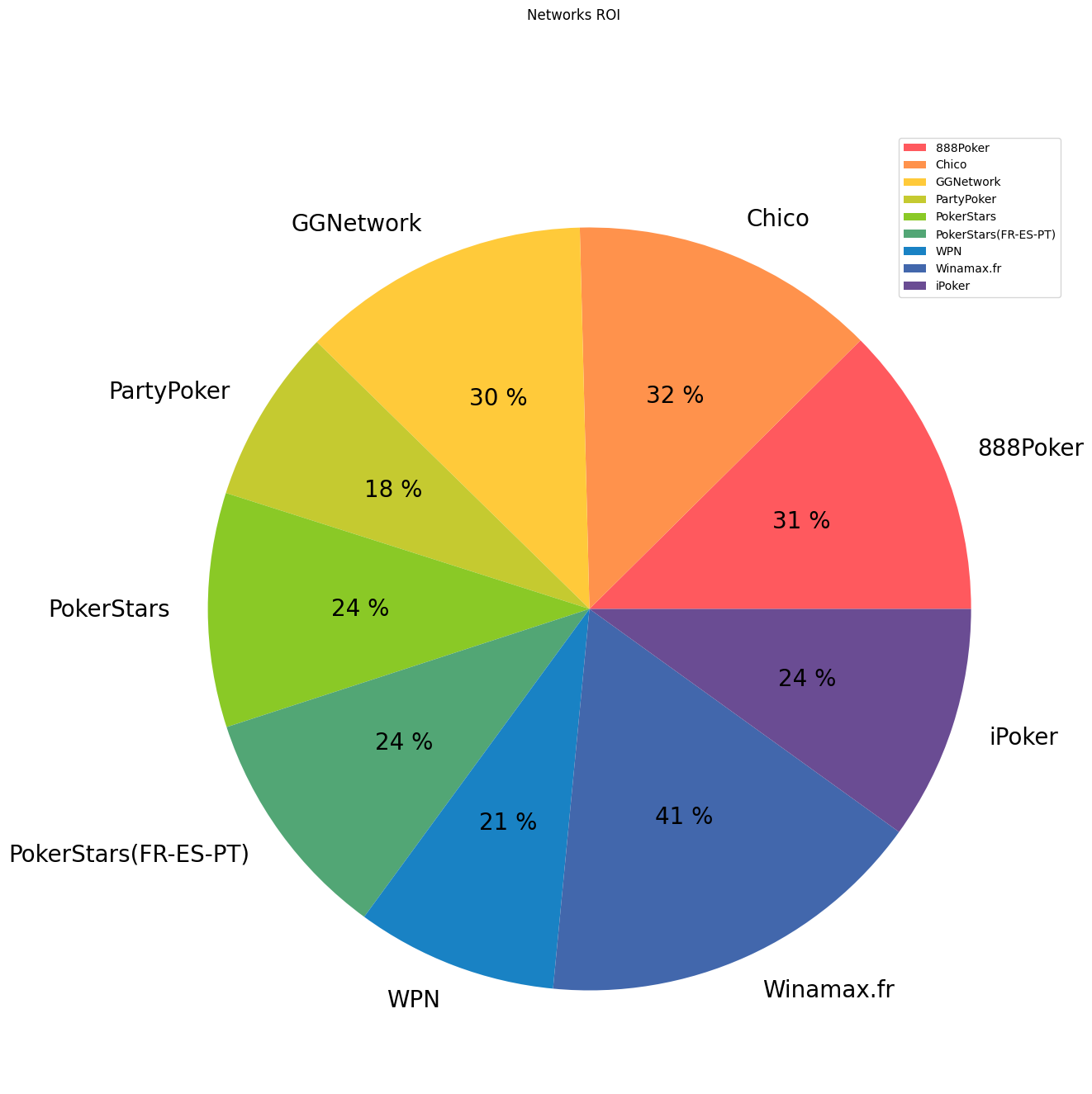


Рис. 7

В разбивке по сетям, турниры в сети Winamax.fr кажутся самыми прибыльными, однако, общее количество турниров в данной сети – 1.2%, об этом мы можем узнать из Рис. 6, так же обстоят дела с Chico и 888Poker – 2.2% и 2.3% от общего числа, что крайне мало. А вот в GGNetwork при обширном количестве турниров (48.9%) всё равно сохраняется высокий коэффициент возврата инвестиций – 30%, что крайне хорошо.

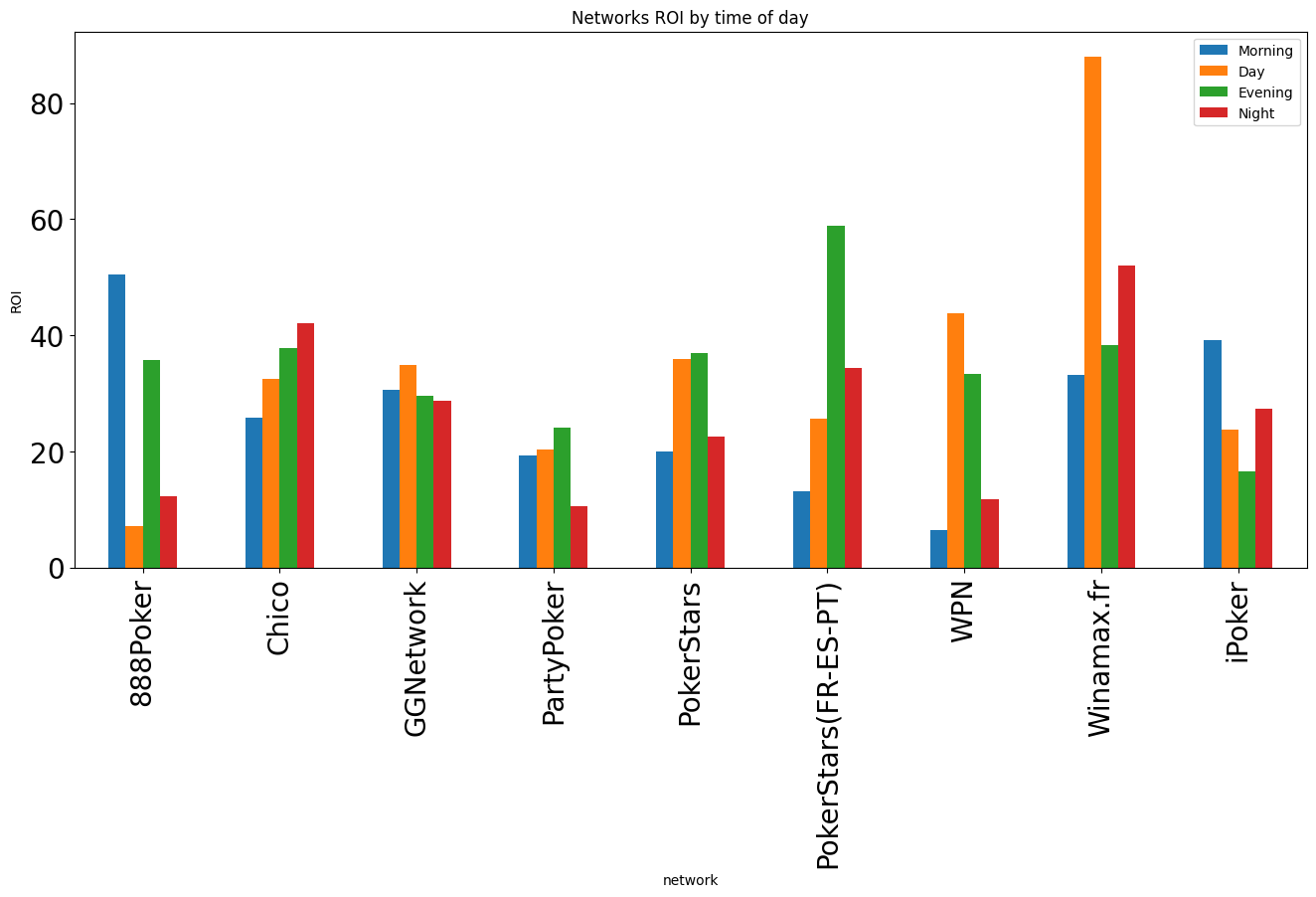


Рис. 8

Поскольку турниры проходят в разное время, игроки не всегда могут быть в хорошей форме, для того чтобы достичь хорошего результата. Мы может наблюдать, что в основном, наиболее хорошие результаты игроки показывают днём и вечером, и только в нескольких сетях утром и ночью.

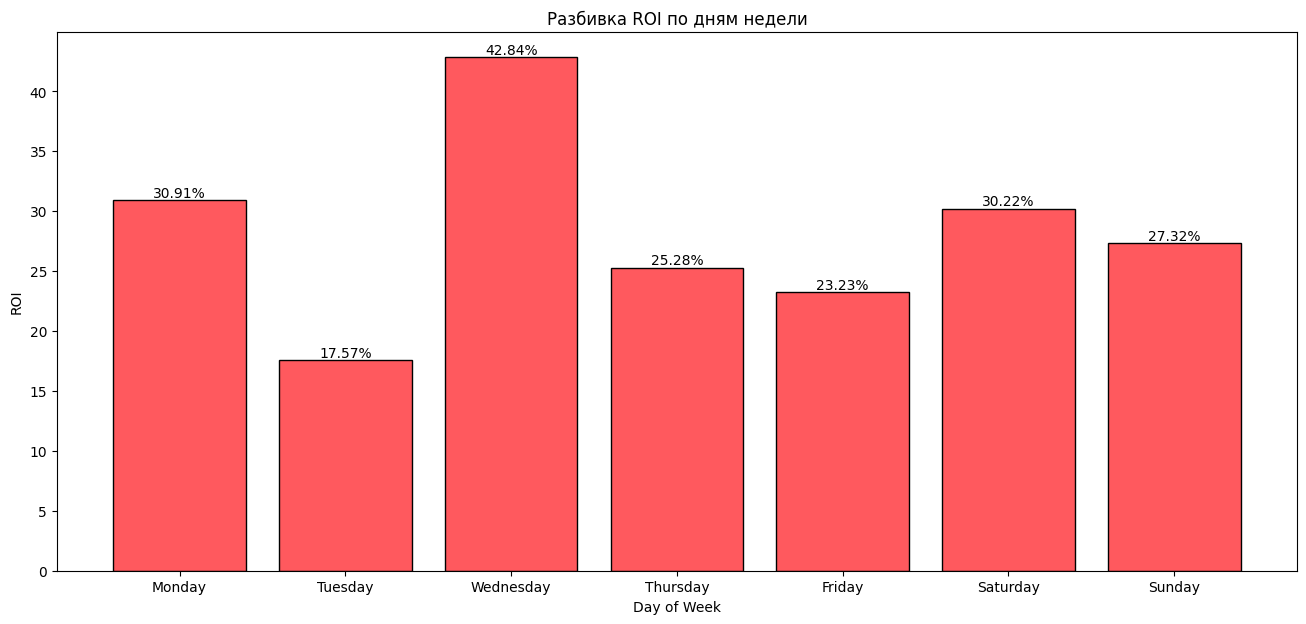


Рис. 9

В этом графике представлена разбивка по дням недели, и то как день недели влияет на ROI, мы можем увидеть, что по статистике, по средам прибыль выше.

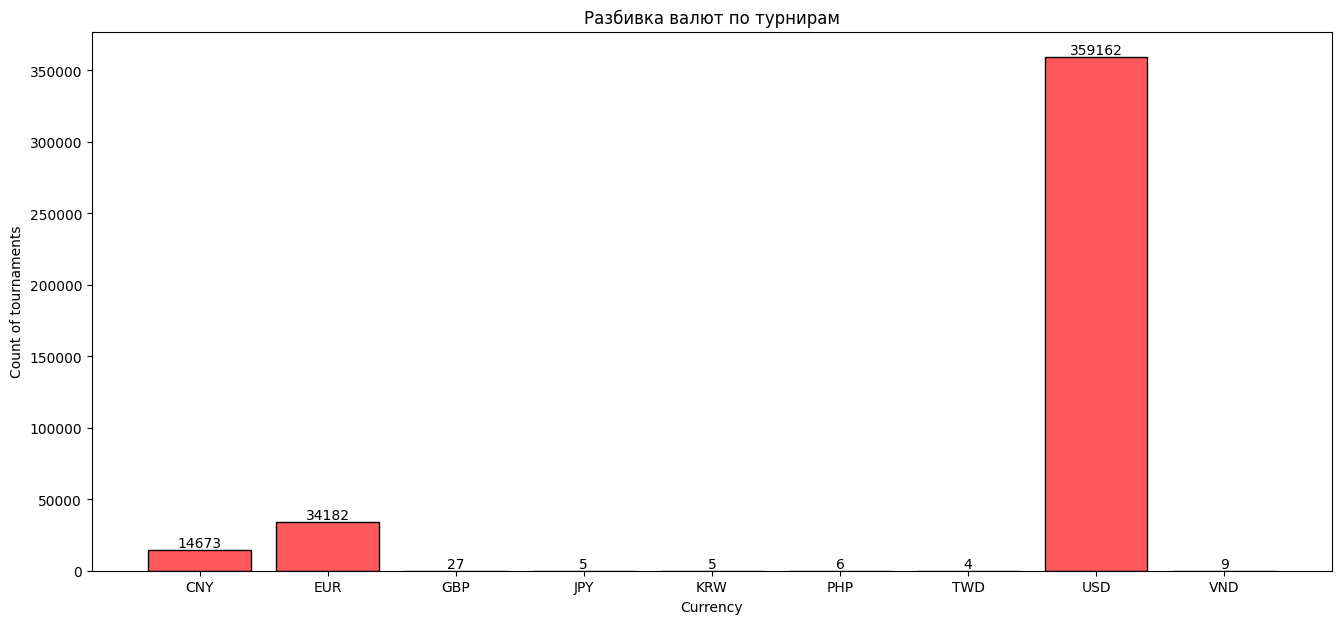


Рис. 10

Помимо разных сетей, в турнирах используют разные валюты, в данном наборе турниров, 9 уникальных валют, однако 6 из них так редко используются, что на графике они даже не видны. Основных валют 3 – Американский доллар, Евро и Китайский Юань.

## Заключение

1. Большое количество сыгранных турниров, не означает получение высокой прибыли.

2. Для наиболее объективной оценки проделанной работы нужно оценивать её по различным параметрам

3. Не все игроки в этой сфере получают прибыль – также как легко можно выиграть, так же легко всё можно проиграть.

4. Самая популярная сеть у представленной команды – GGNetwork, в ней проходит 48.9% всех турниров.

5. Для получения хорошего результата лучше выбирать турниры в GGNetwork в понедельник, среду или субботу, в утреннее время.

6. Самые популярные валюты: Американский доллар, Евро и Китайский Юань.

И подводя итог в результате анализа данных о покерных турнирах команды FireStorm Team были выполнены все поставленные задачи и выявлены ключевые факторы успеха данной отрасли, а также предложены рекомендации для улучшения результатов турниров.

В ходе работы были использованы соответствующие теоретические и практические материалы, которые помогли достичь поставленных целей.

## Список использованных источников и библиотек

1. Использовались обезличенные данные о турнирах игроков FireStorm Team за 2022 год, полученный с помощью сервиса <https://ru.sharkscope.com> и внутренних программ.

Доступно по ссылке: <https://disk.yandex.ru/d/OoN6jDyWz7Zzqg>

1. Библиотеки для проведения анализа и построения аналитик: numpy, pandas, matplotlib, ydata\_profiling.

## Приложение

Файл с кодом – <https://github.com/Ho1ms/bigdata-poker-stats-coursework>