

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Центр непрерывного образования

Факультета компьютерных наук

ИТОГОВЫЙ ПРОЕКТ

Дашборд с аналитикой результатов гонок формулы 1

в зависимости от типов трасс

Выполнила:

Шаталова Анна Николаевна

Руководитель:

Бир Анастасия Павловна

Москва 2024

Оглавление

I. Введение

II. Обзор литературы

III. Методы

IV. Эксперименты

V. Заключение

Приложение

Список литературы:

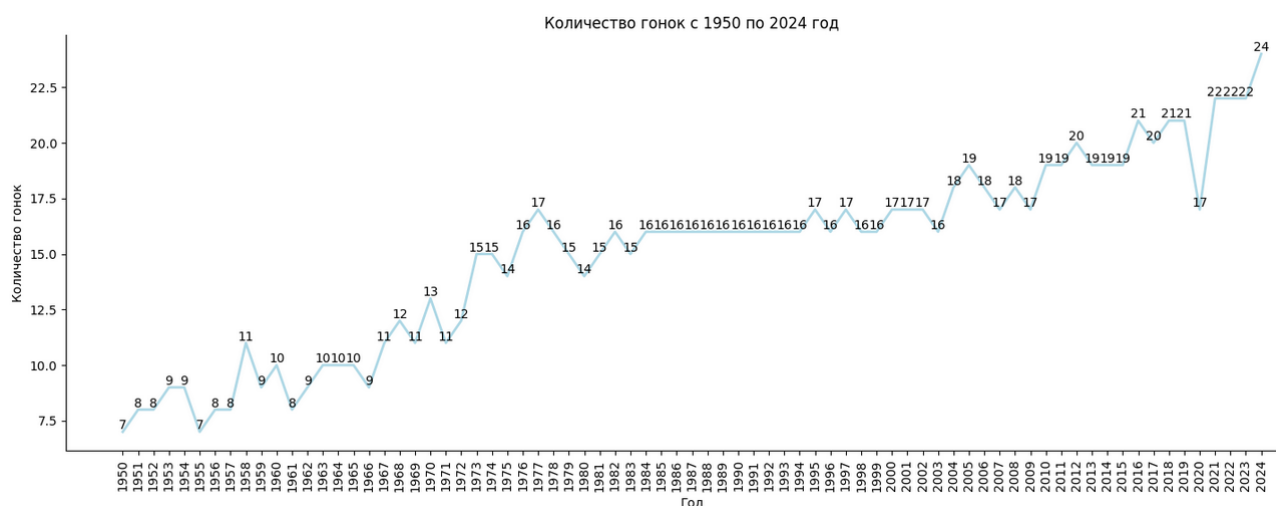
I ВВЕДЕНИЕ

«Мир Формулы 1 — это мир, движимый инновациями и данными»

© Джессика «Джесс» Томкинс

Формула 1 является одним из наиболее престижных и высокотехнологичных видов автоспорта, где каждый элемент гонки играет решающую роль в достижении победы. На протяжении многих десятилетий Формула 1 была не только состязанием лучших пилотов мира, но и ареной, где передовые инженерные разработки и стратегии команд проходят проверку в экстремальных условиях. В современном мире, где технологии и данные играют всё более важную роль в спорте, важнейшим аспектом является анализ влияния разнообразных факторов на результаты гонок. Одним из них является тип и конфигурация гоночных трасс, выбранные в качестве основного аспекта изучения для данной работы.

Популярность гонок на протяжении всей их истории продолжает неуклонно расти. Как видно из графика ниже, количество гонок в сезоне с годами увеличивалось. Сейчас в год проводится в 3 раза больше заездов.



- сезон 1950-го: 7 гонок
- сезон 2024-го: 24 гонки

«Провал» 2020 года связан с эпидемией COVID-19, когда из запланированных 22-х гран-при было проведено только 17.

В рамках дипломной работы будут рассмотрены данные по гонкам, проводившимся за период с 2011 по 2023 год. Это ограничение обусловлено проведенными исследованиями по изменениям технического регламента. Его требования постоянно претерпевают изменения в связи с внесением новых правил и ограничений.

Рассмотрим факторы, которые менялись (или не менялись) на протяжении рассматриваемых 13 лет:

- с 2010 года – запрещена дозаправка топлива
- с 2010 года – введена единая 25-бальная система победных очков
- с 2011 года - внедрена система DRS (*изменения затронули как трассы, так и болиды*)
- с 2011 года - единый производитель шин (*монополия PIRELLI*)
- с 2014 год - появление LCD-экранов на рулевом колесе
- с 2018 года – обязательная система защиты гонщиков halo
- 2022 год - возвращение граунд-эффекта¹
- вес болидов увеличился почти на 160кг (*с 640 до 798*)
- с 2022 год - стандартизация коробки передач
- постоянный диаметр шин (*изменения были внесены в 2022² году*)

Таким образом, наиболее схожим являются года с 2011 по 2021. Но для исследования мы захватим еще 2 года нового регламента, чтобы посмотреть будут ли там значительные отклонения в наблюдениях.

Актуальность этого проекта обусловлена тем, что команды Формулы 1 оперируют огромными объёмами данных, которые необходимо

¹ граунд-эффект в автоспорте — это появление дополнительной прижимной силы у автомобиля за счет создания зоны низкого давления и разреженного воздуха между днищем и трассой

² до 2022 года диаметр шин составлял 13 дюймов. FIA внесла новые изменения в технический регламент и теперь диаметр колес составляет 18 дюймов, что повлекло за собой полное изменение конструктивной концепции болида. Помимо этого, увеличилась ширина передних шин с 305 до 270 мм. Ширина задних шин осталась прежней — 405 мм

анализировать для принятия оперативных решений. Эти данные включают не только результаты гонок, но и детализированную информацию о характеристиках каждой трассы и о поведении автомобилей в различных условиях. Однако, несмотря на доступность этих данных, их сложность и объём могут затруднить быструю интерпретацию и применение для практических целей.

Дашборд, разрабатываемый в рамках этого проекта, позволит пользователям анализировать ряд характеристик, и влияние их параметров на конечные результаты гонок

- визуализировать результаты гонок в зависимости от типа трассы
- выявить скрытые зависимости между различными параметрами трасс и результатами гонок.

Для этого потребуется собрать и обработать данные по гонкам, провести их анализ, а также выбрать наиболее релевантные метрики для визуализации.

Особое внимание будет уделено таким параметрам трасс, как:

- высота над уровнем моря;
- количество поворотов;
- длина главной прямой;
- происшествия на трассе;
- скорость болидов

Логичным и важным элементом анализа является исследование влияния погодных условий, таких как дождь или высокая температура, на производительность болидов на разных типах трасс. Например, на городских трассах с узкими поворотами дождь может существенно увеличить количество аварий, что влияет на стратегию команд и продолжительность гонок. Использование данных о погоде в контексте анализа трасс может дать более полную картину. Но эти исследования потребуют более масштабной работы, которую нет возможности

реализовать в текущие сроки. Поэтому основной фокус внимания этой работы сосредоточен на менее подверженных изменениям параметрах.

В качестве исходных данных использованы 12 датасетов по гонкам Формулы 1 с 1950 года по первую половину сезона 2024 года (до так называемой летней паузы), размещенных на платформе *Kaggle.com*. Помимо этого, из-за специфики темы часть необходимой информации отсутствовала – по характеристикам трасс – эти данные были собраны самостоятельно и оформлены в несколько небольших датасетов, которые интегрировались в сводный рабочий датасет.

Размер датасетов указан ниже:

трассы : (77, 9)
результаты команд : (12495, 5)
позиции команд : (13261, 7)
команды : (212, 5)
гонщики : (859, 9)
результаты гонщиков : (34574, 7)
круги гонок : (575029, 6)
пит-стопы : (10945, 7)
квалификации : (10234, 9)
гран-при : (1125, 18)
результаты гонок : (26499, 18)
сезоны : (75, 2)
результаты спринтов : (300, 16)
статусы : (139, 2)
аварии красные флаги : (32, 6)
аварии желтые флаги : (265, 7)
красные флаги : (61, 4)
изображения машин : (20, 1)
данные по трассам : (35, 16)
двигатели : (159, 3)

Выявленные зависимости между типами трасс, характеристиками автомобилей и стратегиями команд могут предоставить командам ценные аналитические инсайты, которые можно использовать для улучшения своих результатов и повышения конкурентоспособности.

Помимо этого, результаты будут понятны и простым обывателям, мало погруженным в эту тему.

II ОБЗОР ЛИТЕРАТУРЫ

Основным источником информации по выбранной теме являлись интернет-статьи и публикации в специализированных порталах. Формат книжных изданий менее эффективен, поскольку не учитывает динамику постоянных изменений в Формуле 1 и быстро теряет актуальность.

Несмотря на то, что гоночные трассы не меняют своего расположения (хотя могут открываться новые), но может изменяться их конфигурация, уточняться длина треков, повороты и их количество, устраиваться новые зоны DRS. Наиболее достоверные данные чаще размещены непосредственно в новостях по текущему этапу.

1. "Формула-1: Знания изнутри" (The Science of Formula 1 Design)

Автор: Дэвид Тротт,

2. Data Driven at 200 MPH: How Analytics Transforms Formula One Racing

Журнал Forbes, автор: Joel Shapiro, 2023

3. F1 Technology & Data Analysis: How Data is Transforming Race Performance

Источник: CATAPULT.COM, 2024

4. The Influence of F1 Circuit Design on Race Outcomes The design of Formula 1 circuits plays a vital role in shaping race outcomes

Источник: scuderiafans.com, 2024

5. "Формула-1: Полная история" (Formula One: The Complete Story)

Автор: Марк Хьюз,

Всесторонний обзор истории Формулы 1, включая информацию о различных трассах и их влиянии на результаты гонок

6. "Как построить гоночный автомобиль" (How to Build a Car)

Автор: Эдриан Ньюи (технический директор Red Bull Racing)

7. "Track Characteristics and Their Impact on Formula 1 Car Performance"

Авторы: Michael Brown, Sophie Thompson, 2019

8. "Formula 1 Strategies: How Track Layout Affects Race Outcomes"
Автор: David Garcia, 2021
9. "Dashboards in Sports Analytics: Visualizing Data for Strategic Decisions"
Авторы: Linda White, Robert Adams, 2022
10. "Telematics and Track Performance in Motorsport: A Comparative Study"
Автор: Sarah Johnson, 2020
11. "Race Track Design and its Effects on Vehicle Dynamics"
Авторы: Mark Williams, Jessica Lee, 2018
12. "Advanced Visual Analytics in Motorsports: A Case Study on Formula 1"
Автор: Peter Nguyen, 2021
13. "Pit Stop Strategy and Track Layouts in Formula 1 Racing"
Авторы: Kevin Martin, Laura Green, 2022

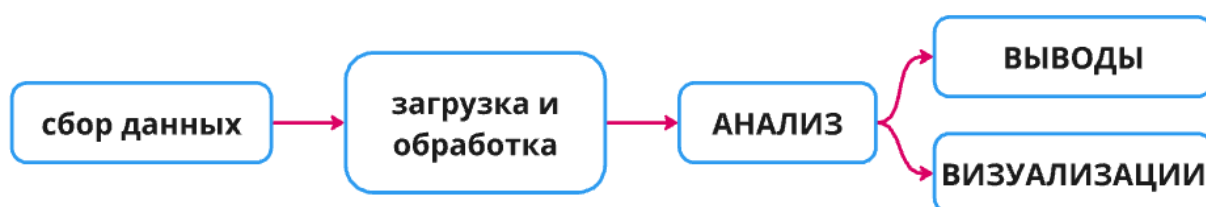
III МЕТОДЫ

«Нельзя забывать, что наука о данных и искусственный интеллект на самом деле очень сложны, и путь от получения данных до решения проблемы очень длинный»
© Harvard Business Review

Работа с данными проводилась в несколько условных этапов, которые переходят из одного в другой, а иногда разрабатываются параллельно (например, делались предварительные визуализации в процессе формулирования вопросов и знакомства с данными).

1. Сбор данных

Основные датасеты взяты на платформе *Kaggle.com*. Это данные по командам, гонщикам, трассам, результатам гонок, пит-стопам, кругам и



событиям на трассе.

Недостающие данные, такие как:

- тип трассы (*уличная или гоночная, движение по или против часовой стрелки*)
- прижимная сила (*максимальная, высокая, средняя, средне-низкая, средневысокая, низкая*)
- количество зон DRS (*не менее 1*)
- длина главной прямой (*не более 2 км*)
- количество поворотов
- типы происшествий на трассе (*технические проблемы, навыки пилотов, случайности*)
- красные и желтые флаги во время гонки
- потери времени на пит-стопах (*минимальные-максимальные*)

были собраны из интернет-источников. Наиболее полезными оказались сайты: www.statsf1.com, f1.tfeed.net, www.racingcircuits.info, www.reddit.com, f1.fandom.com.

После формирования новых данных в таблицах *Google Docs* они сохранялись в формате **.csv** для возможности дальнейшей загрузки и интеграции в работе.

Несмотря на отсутствие на этом этапе необходимости непосредственно работать с данными по времени он оказался ощутимо продолжительным.

2. Загрузка и обработка

Для загрузки данных использовался специально созданный под проект репозиторий на *GitHub.com*.

GitHub — это платформа для хостинга и совместной работы над проектами с использованием системы контроля версий Git. На *GitHub* удобно хранить датасеты, так как платформа обеспечивает версионирование, позволяет отслеживать изменения и улучшать данные с помощью pull request'ов и комментариев, а



также легко делиться ими. Это делает процесс хранения и обновления данных прозрачным и управляемым.

Помимо этого, часть данных была загружена в созданную базу данных, чтобы иметь возможность делать SQL-запросы. Использовалась реляционная база *PostgreSQL*.

Она используется для хранения, управления и анализа структурированных данных и обеспечивает высокую надёжность, безопасность и масштабируемость.

Основная часть работы проводилась в среде *Google Colab* с использованием языка программирования *Python*.



Применялись следующие библиотеки для построения графиков и взаимодействия с данными:

```
import pandas as pd
import numpy as np
import random
from numpy.random import rand, randn
from scipy.stats import norm
from matplotlib import pyplot as plt
import seaborn as sns
```

Загруженные датасеты проверялись на наличие пропусков, дубликатов, выбросов, типы данных. В результате часть столбцов были переименованы (*rename*), часть удалены (*drop*), часть данных преобразовано в нужные



форматы. Удалены были данные, которые не использовались в исследовании не влияли на его результаты, например, URL со ссылками.

Для удобства работы разрозненные таблицы были объединены при помощи метода *pd.merge*. В большинстве случаев использовался метод *left join*, позволяющий добавлять новые значения из второго датасета по имеющимся с первым совпадениям.

Пример работы с датасетом по гонщикам:

```
# создание нового параметра имя гонщика из двух столбцов
df_drv['driver_name'] = df_drv['forename'] + ' ' + df_drv['surname']

# удаление не нужных более столбцов
df_drv.drop(['surname', 'forename'], axis = 1, inplace = True)

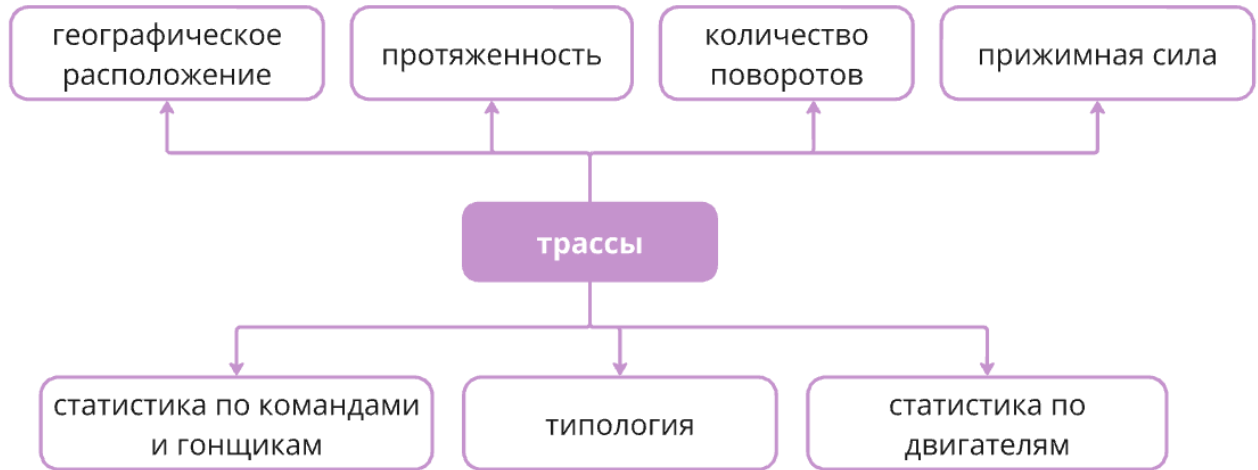
# переименование для дальнейшего удобства работы
df_drvst.rename(columns={'points': 'driver_points', 'position' :
'driver_position_in_race', 'positionText' : 'DriverPositionText'},
inplace=True)

# объединение датасетов
drivers = pd.merge(df_drvst, df_drv, how = "left", on = ["driverId"])
```

После завершения работ этого этапа переходим к анализу сформированных данных.

3. Анализ данных

Самый «затягивающий» этап работы. Когда одно открытие влечет за собой все новые и новые вопросы. Поэтому, чтобы сконцентрироваться на



теме определим основные параметры, по которым будем рассматривать гоночные трассы и искать зависимости между их характеристиками.

Предлагается сосредоточить внимание и проанализировать вышеуказанные показатели, а также ответить на следующие вопросы:

- какие есть типы трасс и какая у них специфика?
- одинаково ли количество городских и гоночных треков?
- какое направление движение преобладает для трасс с высокой или низкой прижимной силой?
- где происходит больше всего поломок двигателей?
- какие двигатели самые надежные?
- как выглядит смена позиций на трассе?
- все ли трассы вписываются в регламент по таким требованиям как длина трасс и протяженность гонки?
- правда ли, что с увеличением количества поворотом максимальная скорость не снижается, а растет?
- часто ли пилоты выигрывают гонки стартуя с первой позиции?
- какие команды уверенно чувствуют себя на любом треке?

IV ЭКСПЕРИМЕНТЫ

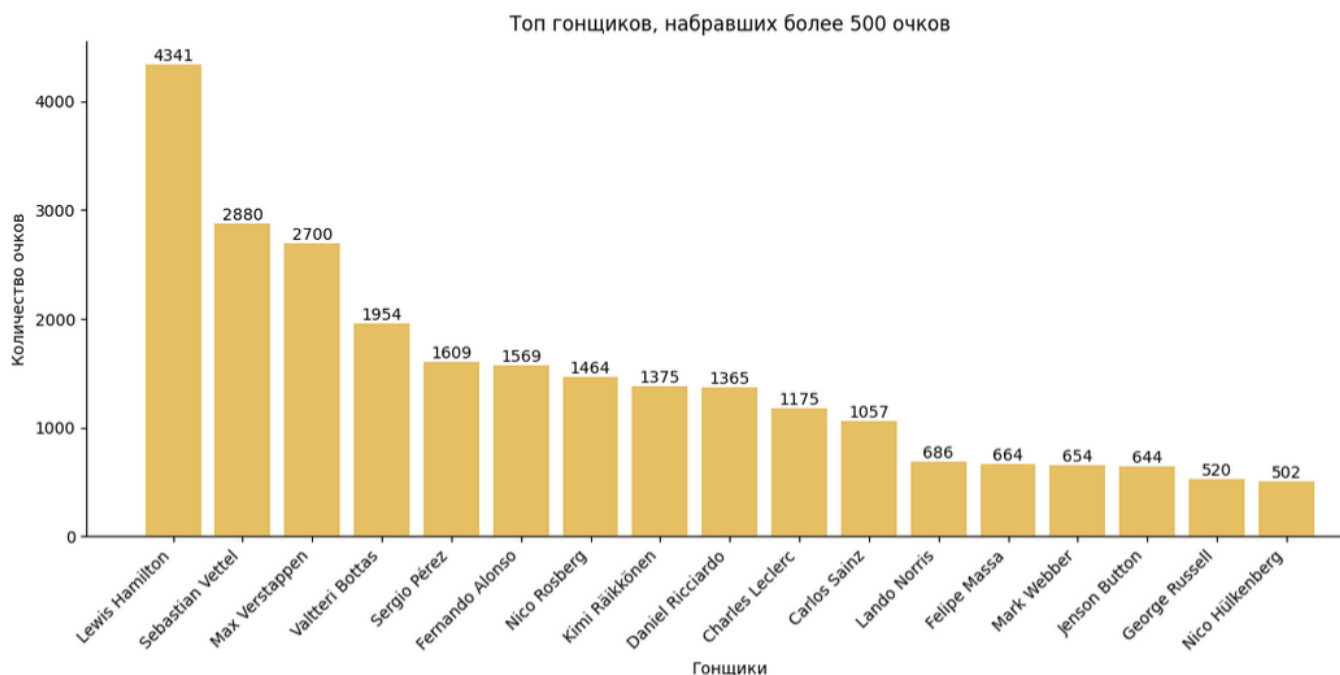
*«Никаким количеством экспериментов нельзя доказать теорию,
но достаточно одного эксперимента, чтобы ее опровергнуть»*
©Альберт Эйнштейн

Перед началом ознакомимся с имеющимися данными и с общими фактами. Посмотрим 30 гонщиков за все время набравших наибольшее количество победных очков в гонках:

Lewis Hamilton: 4341.55	George Russell: 520.80
Sebastian Vettel: 2880.20	Nico Hülkenberg: 502.90
Max Verstappen: 2700.15	Esteban Ocon: 436.40
Valtteri Bottas: 1954.10	Pierre Gasly: 428.10
Sergio Pérez: 1609.60	Romain Grosjean: 408.80
Fernando Alonso: 1569.70	Lance Stroll: 285.40
Nico Rosberg: 1464.60	Kevin Magnussen: 250.90
Kimi Räikkönen: 1375.70	Daniil Kvyat: 205.20
Daniel Ricciardo: 1365.00	Alexander Albon: 188.20
Charles Leclerc: 1175.00	Paul di Resta: 134.40
Carlos Sainz: 1057.85	Michael Schumacher: 116.70
Lando Norris: 686.50	Kamui Kobayashi: 102.80
Felipe Massa: 664.90	Oscar Piastri: 80.70
Mark Webber: 654.10	Pastor Maldonado: 77.60
Jenson Button: 644.90	Adrian Sutil: 69.50

Как видно только 17 из них набрали более 500 очков. При этом всего в гонках участвовало (некоторые продолжают участвовать) - 852 гонщика.

Посмотрим, как выглядит распределение победных очков между пилотами:



Аналогично рассмотрим 30 команд и их результаты:

Топ команд с наибольшим количеством победных очков:

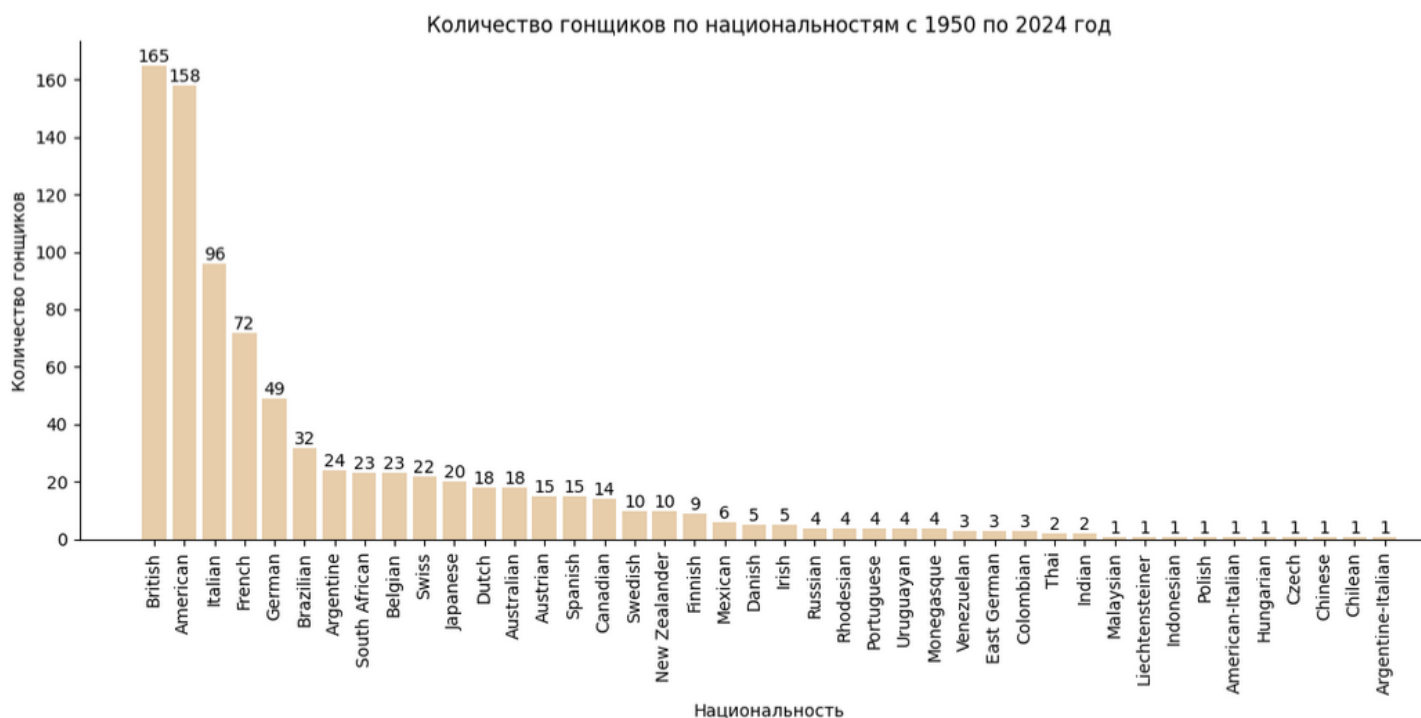
team_name			
Ferrari	9990.000	Alpine F1 Team	457.000
Red Bull	7603.000	BRM	439.000
Mercedes	7482.500	Ligier	388.000
McLaren	6658.500	AlphaTauri	309.000
Williams	3628.000	BMW Sauber	308.000
Renault	1777.000	Jordan	291.000
Force India	1098.000	Racing Point	283.000
Team Lotus	918.000	Toyota	278.500
Benetton	861.500	Haas F1 Team	268.000
Lotus F1	706.000	Lotus-Climax	253.000
Tyrrell	621.000	Cooper-Climax	229.000
Brabham	594.000	BAR	227.000
Sauber	557.000	Alfa Romeo	199.000
Toro Rosso	500.000	Lotus-Ford	174.000
Aston Martin	470.000	Brawn	172.000

Всего в гонках зафиксировано участие 212 различных команд. Разница по очкам ощутимая даже в первой десятке. Лидерами является гоночный коллектив Ferrari, которые смогли набрать почти 10 тыс. очков.

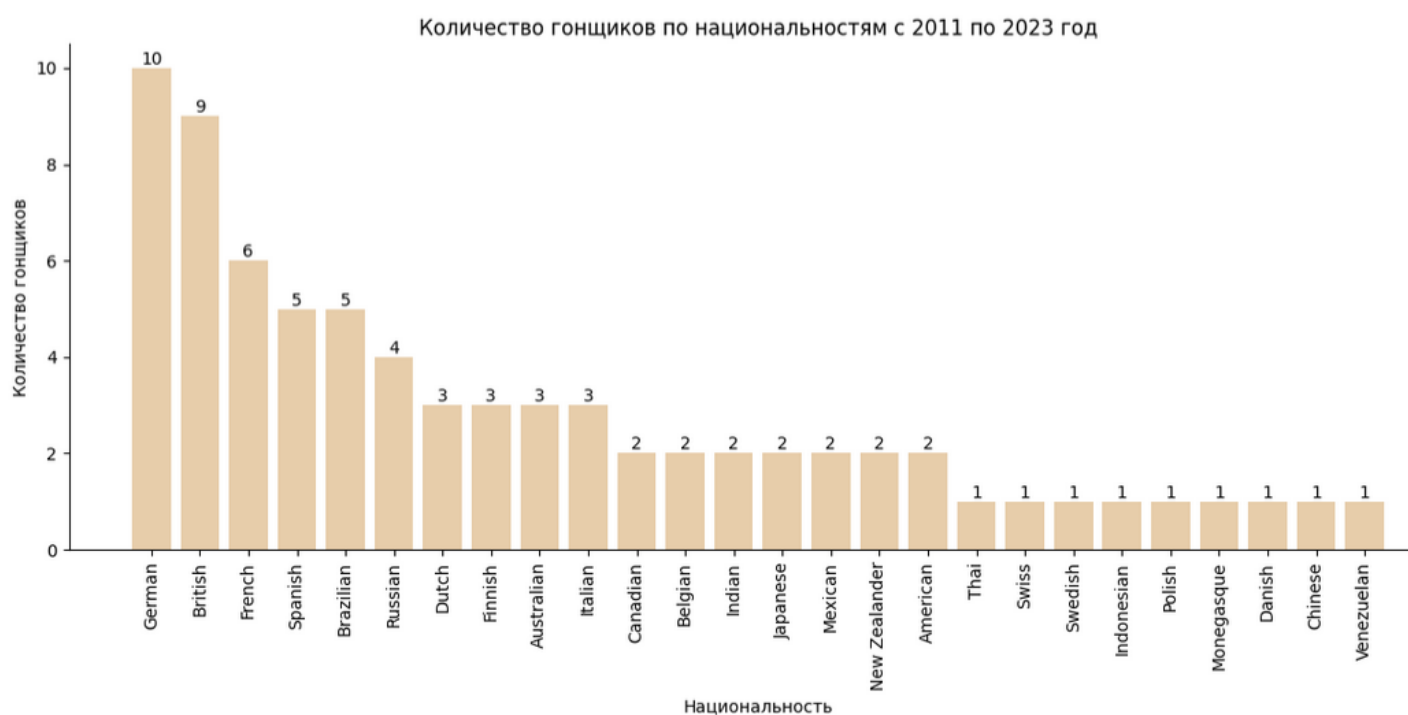
Посмотрим команды какой национальности являются рекорсменами в этом виде спорта. Всего представлено 23 национальности и только 15 из них набрали хотя бы 1 победное очко. Лидеры – британцы и итальянцы.



Если посмотреть на всю историю гонок Формулы 1, то наибольшее количество гонщиков были из Великобритании и Америки. Следом Италия, Франция, Германия.



С 2011 по 2023 годы картина выглядит иначе. Преобладают гонщики из Германии, Британии, Франции, Испании и Бразилии. Всего 2 американца.



Посмотрим сколько всего гран-при было проведено:

grand_prix			
British Grand Prix	75	Turkish Grand Prix	9
Italian Grand Prix	75	Russian Grand Prix	8
Monaco Grand Prix	70	United States Grand Prix West	8
Belgian Grand Prix	69	Detroit Grand Prix	7
German Grand Prix	64	Azerbaijan Grand Prix	7
French Grand Prix	62	Swiss Grand Prix	6
Spanish Grand Prix	54	Swedish Grand Prix	6
Canadian Grand Prix	53	Mexico City Grand Prix	4
Brazilian Grand Prix	47	Korean Grand Prix	4
United States Grand Prix	45	Emilia Romagna Grand Prix	4
Hungarian Grand Prix	39	São Paulo Grand Prix	4
Japanese Grand Prix	38	Saudi Arabian Grand Prix	4
Australian Grand Prix	38	Miami Grand Prix	3
Austrian Grand Prix	37	Indian Grand Prix	3
Dutch Grand Prix	34	Qatar Grand Prix	3
San Marino Grand Prix	26	Styrian Grand Prix	2
European Grand Prix	23	Las Vegas Grand Prix	2
South African Grand Prix	23	Pacific Grand Prix	2
Bahrain Grand Prix	20	Luxembourg Grand Prix	2
Argentine Grand Prix	20	Caesars Palace Grand Prix	2
Mexican Grand Prix	20	Sakhir Grand Prix	1
Malaysian Grand Prix	19	Pescara Grand Prix	1
Portuguese Grand Prix	18	Moroccan Grand Prix	1
Chinese Grand Prix	17	Eifel Grand Prix	1
Abu Dhabi Grand Prix	16	Dallas Grand Prix	1
Singapore Grand Prix	15	Tuscan Grand Prix	1
Indianapolis 500	11	70th Anniversary Grand Prix	1

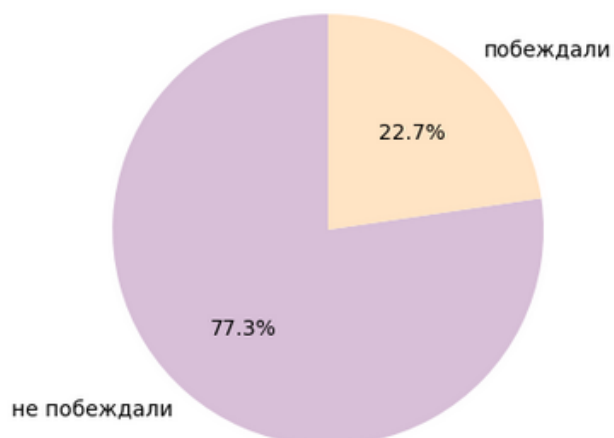
Как видно, некоторые гран-при проходят чаще или регулярно. Какие-то трассы больше не используются. Так, с 2022-го года прекращен гран-при России. Трасса в Сочи больше не принимает у себя гонки Формулы 1.

Немного статистики с 2011 года: в гонках участвует 75 гонщиков. Из них 58 никогда не были на подиуме на 1 месте.

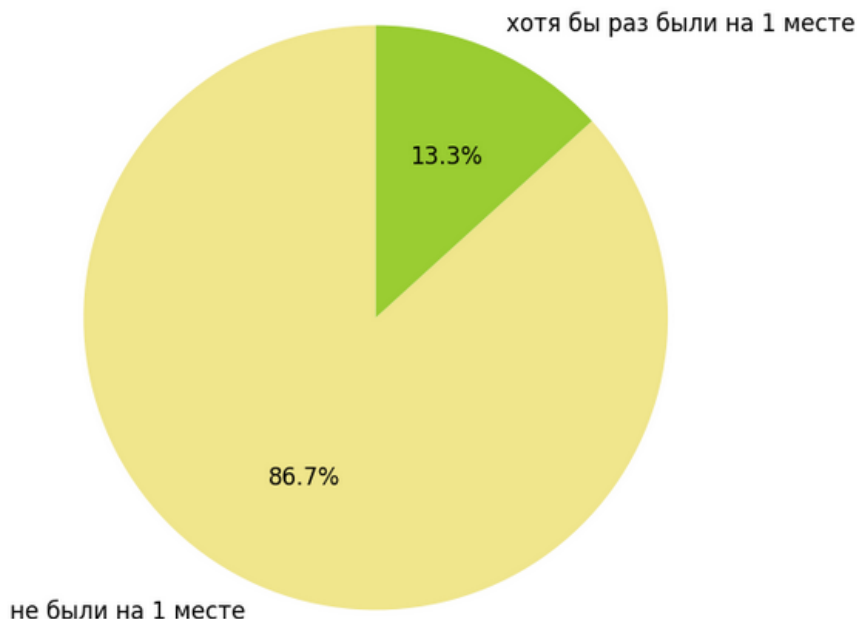
Таким образом, только 22.7% гонщиков сумели стать победителями гонки.

Сравним с аналогичным показателем, но за все время – с 1950-го года.

Процент победителей в гонках Формулы 1



Процент гонщиков, никогда не финишировавших на 1 месте (с 1950 по 2024 год)



Из 852 гонщиков 746 пилотов никогда не становились победителями. За последнее десятилетие у большего количества участников есть возможность побывать на подиуме на 1-ом месте.

Перейдем непосредственно к трассам. Пройдемся по различным аспектам, относящимся к их характеристикам. Рассмотрим каждый из них по отдельности.

1. Повороты

Количество поворотов на различных треках варьируется от 10 поворотов на трассе Red Bull Ring до 27 поворотов на Jeddah Corniche Circuit. 16 – среднее количество поворотов для трасс Формулы 1.

3 трассы с максимальным количеством поворотов:

	track_name	turns
4643	Jeddah Corniche Circuit	27
176	Valencia Street Circuit	25
4095	Autodromo Enzo e Dino Ferrari	21

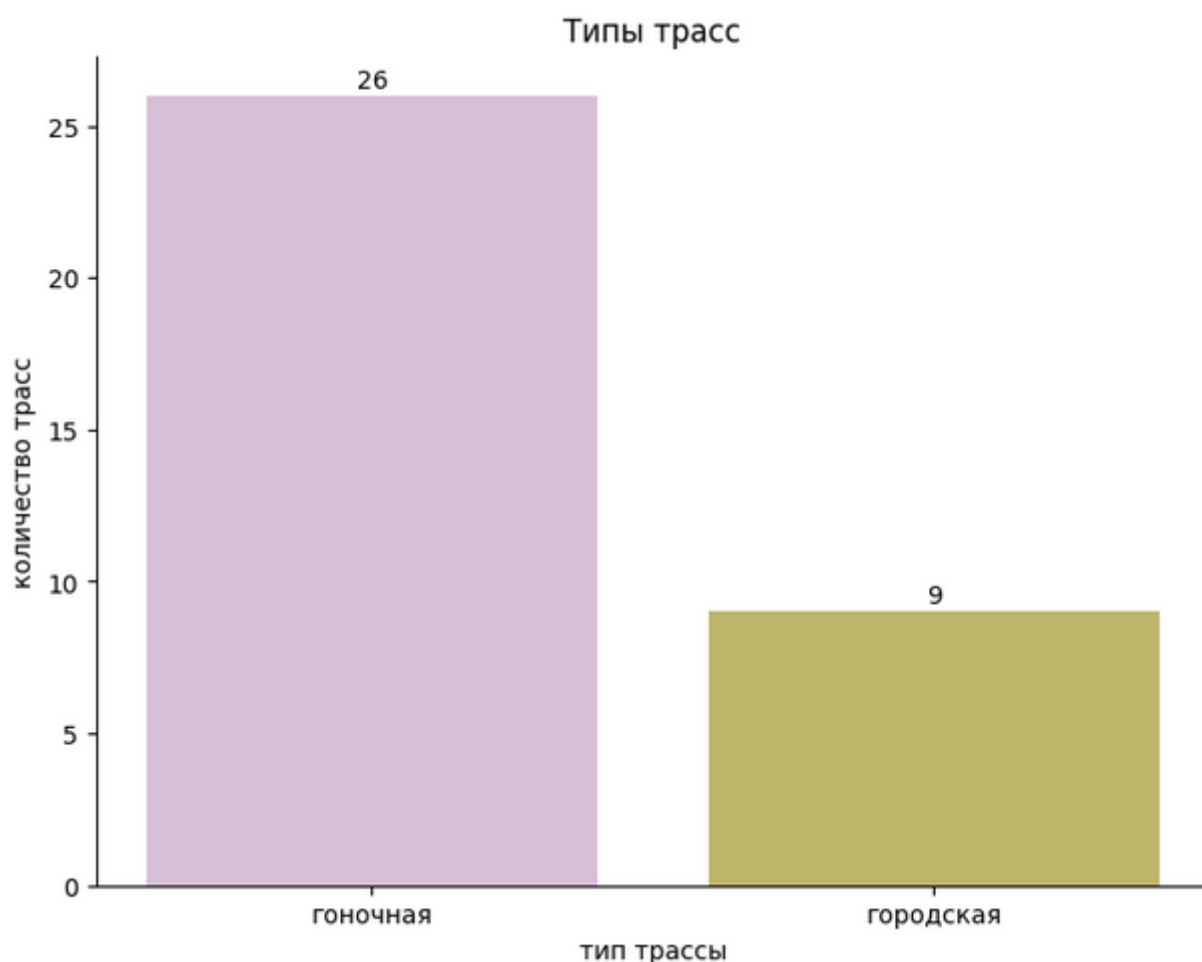
3 трассы с минимальным количеством поворотов:

	track_name	turns
4395	Red Bull Ring	10
4481	Autodromo Nazionale di Monza	11
952	Albert Park Grand Prix Circuit	14

Средняя максимальная скорость круга для трасс с более чем 16-ю поворотами: 198.95км/ч, а с менее 16-ти поворотов: 205.93км/ч.

2. Типы трасс

Несмотря на то, что так называемых городских трасс в календаре не так много, однако они пользуются популярностью и их количество растет.



Гоночные трассы (автодромы):

- Специально построены для автоспорта с широкими участками, позволяющими безопасно обгонять.
- Имеют разнообразие типов поворотов, длинные прямые, зоны безопасности и покрытия, оптимизированные для высоких скоростей и минимизации аварий.
- Обладают большими зонами безопасности, покрытыми гравием или асфальтом, которые помогают замедлить машину в случае вылета.
- Покрытие специально подготовлено для гонок, обеспечивает оптимальное сцепление, стабильность и равномерность.
- Обычно более длинные, с плавными поворотами и длинными прямыми, позволяющими развивать высокие скорости и использовать различные гоночные стратегии.

- Часто располагаются в местах с предсказуемыми погодными условиями, что позволяет планировать гонку заранее.

Городские трассы:

- Прокладываются по улицам города и часто включают узкие участки, что усложняет обгоны.
- Ширина трассы ограничена архитектурой города, что увеличивает риск аварий и требует от гонщиков большей концентрации.
- Зоны безопасности минимальны или отсутствуют, ограничены барьерами и стенами, что делает каждую ошибку более опасной.
- Покрытие менее однородное, зачастую скользкое, может включать брусчатку, люки, разметку, что снижает сцепление и создаёт дополнительные трудности для гонщиков.
- Часто имеют короткие прямые и много острых поворотов, что ограничивает максимальную скорость и делает гонку более напряжённой и требовательной к навыкам пилотов.
- Уязвимы к неожиданным изменениям погоды, особенно к дождю, который может значительно повлиять на сцепление и безопасность.



С учетом описаний выше можно предположить, что количество гоночных происшествий на городских трассах больше. Проверим это.

На 9 городских трассах произошло 284 гоночных инцидента:

	track_name	incident_count
0	Albert Park Grand Prix Circuit	66
3	Circuit de Monaco	59
6	Marina Bay Street Circuit	52
2	Circuit Gilles Villeneuve	46
1	Baku City Circuit	33
4	Jeddah Corniche Circuit	14
8	Valencia Street Circuit	6
7	Miami International Autodrome	5
5	Las Vegas Strip Street Circuit	3

На гоночных трассах произошло 696 гоночных инцидентов.

Таким образом, на городских трассах произошло 54,1% всех инцидентов, в то время как на гоночных трассах — 45,9%.

Хотя на городских трассах проводится значительно меньше гонок (9 против 26), процент происшествий на них значительно выше. Это подтверждает, что городские трассы более опасны и сложны для пилотов

3. Протяженность

В 1989 году было установлено требование, о том, что оптимальная протяженность гоночной дистанции должна составлять не менее 305 км (190 миль), а общая длина трасс не менее 3,5км и не более 7км.

Проверим так ли это. Посчитаем общую дистанцию трасс:

3 трассы с максимальной протяженностью:

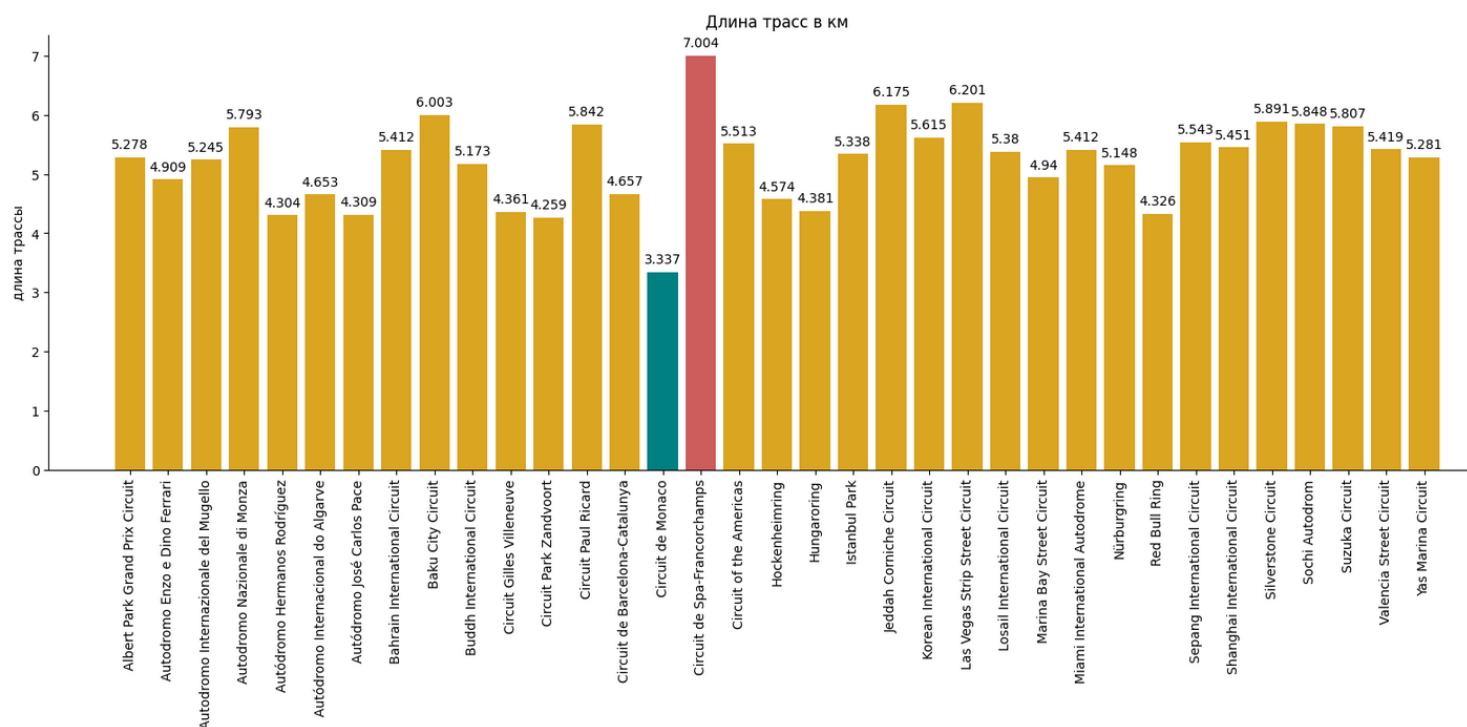
	track_name	race_length_km
28	Sepang International Circuit	310.408
9	Buddh International Circuit	310.380
22	Las Vegas Strip Street Circuit	310.050

3 трассы с минимальной протяженностью:

	track_name	race_length_km
29	Shanghai International Circuit	305.256
34	Yas Marina Circuit	290.455
14	Circuit de Monaco	260.286

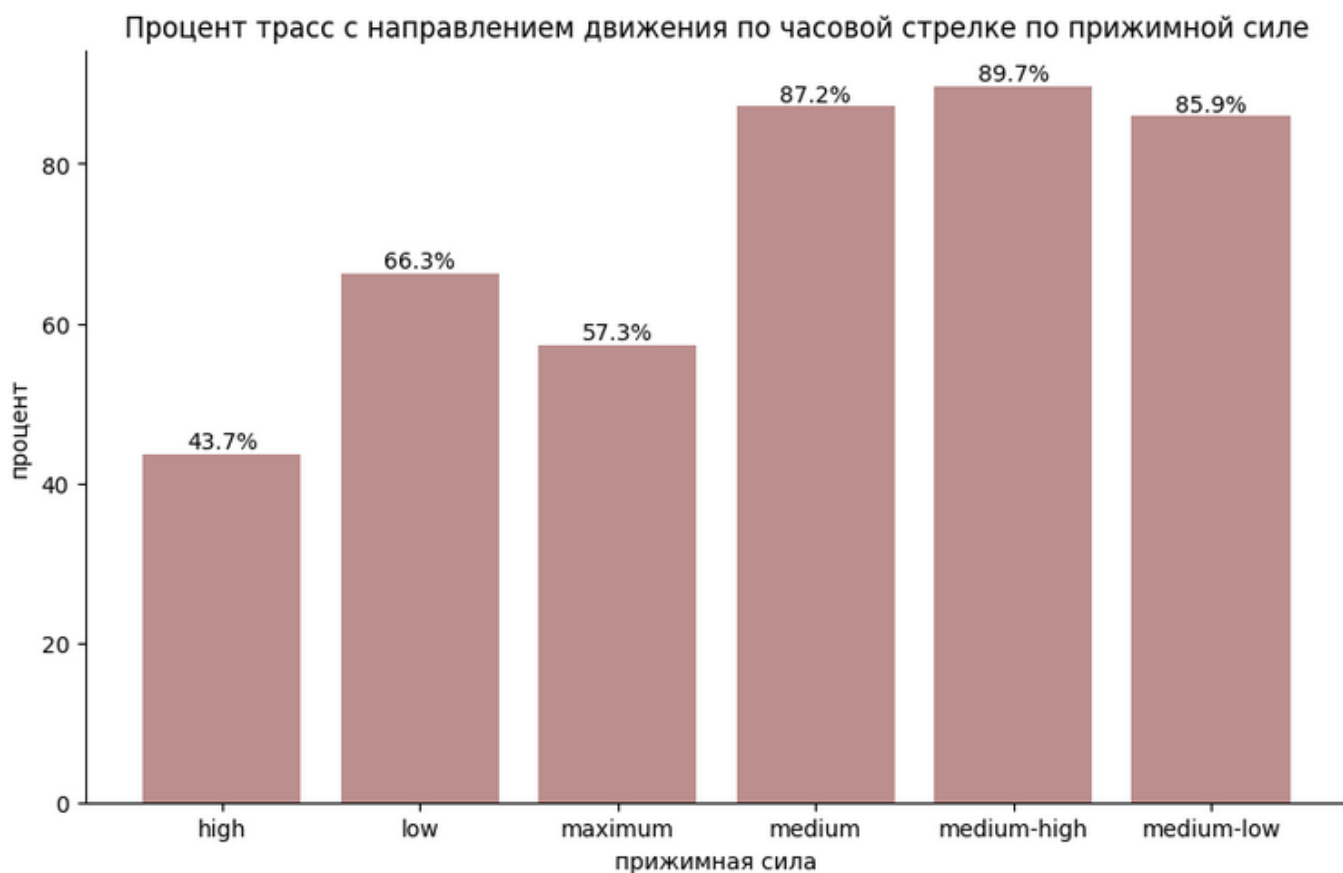
Только две трассы не дотягивают до этого требования: Yas Marina Circuit (290.455км) и Circuit de Monaco (260.286). Превышение длины дистанции некоторых трасс не более 5.5км.

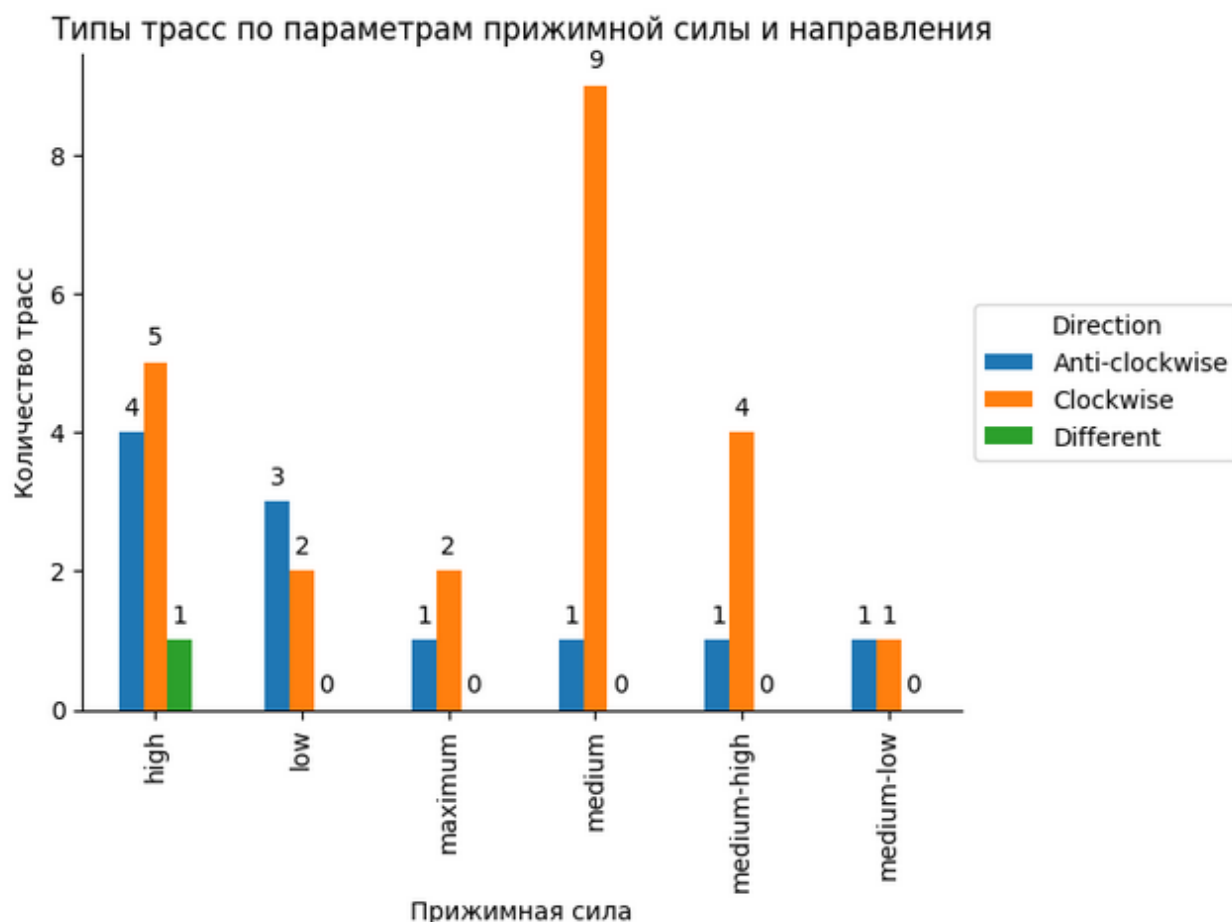
Выведем график длины трасс. Снова две выбиваются из требований FIA — это Circuit de Spa-Francorchamps (чуть больше 7км) и Circuit de Monaco (менее 3.5км).



4. Прижимная сила и направление движения

Трассы с высокой и максимальной прижимной силой в процентном соотношении чаще имеют направление движения против часовой стрелки.





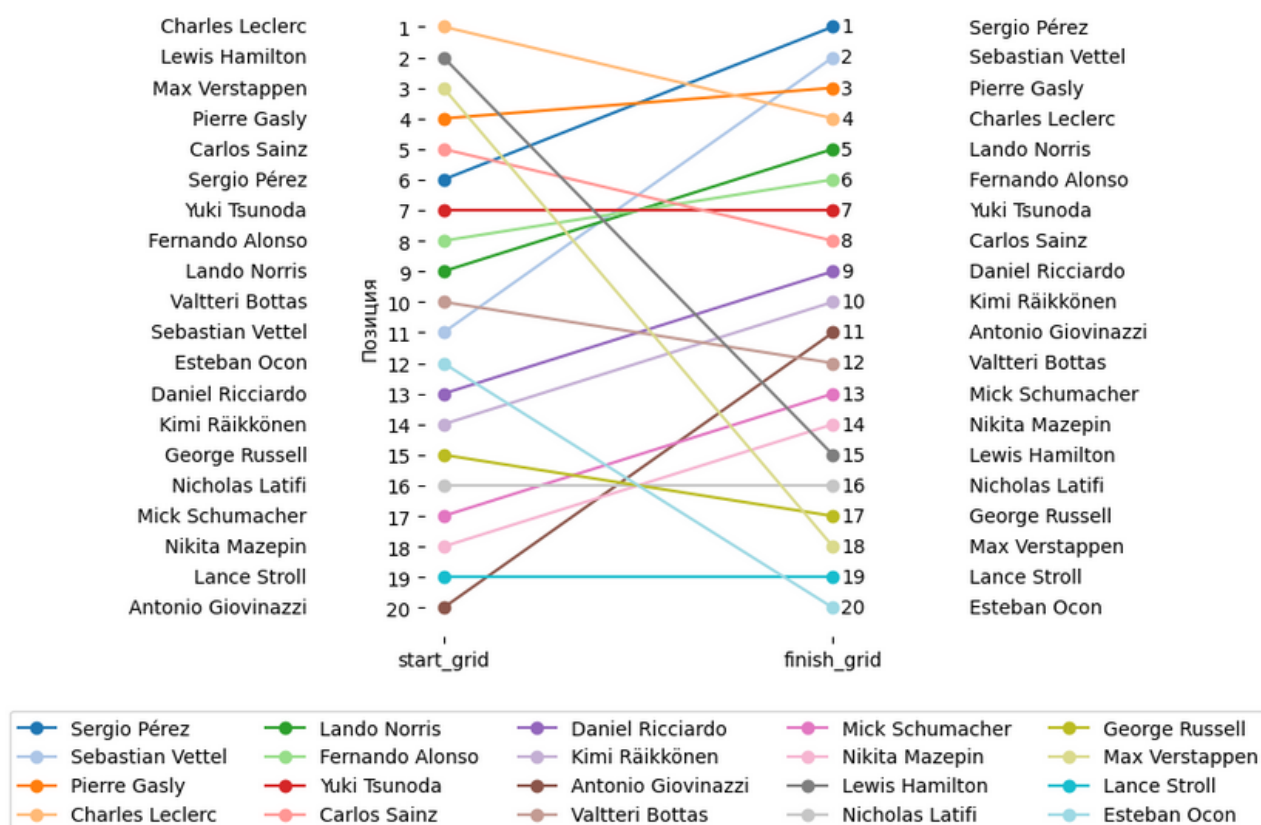
5. Смены позиций

Одной из особенностей трасс, которая привлекает пристальное внимание зрителей, является ее обгонность. В данном случае, будем рассматривать немного другой показатель – «смену позиций». Принципиальная разница в том, что на протяжении гонки может происходить любое количество размен позиций каждым из гонщиков и показатель «обгон» будет увеличиваться, а параметр смены позиций фиксирует фактическое положение пилотов на финише и сравнивает со стартовой позицией. Это может ввести в заблуждение, т.к. отсутствие смены позиций не говорит об отсутствии гоночной борьбы и обгонов.

Используем специальный Slope график, который в контексте аналитики гонок Формулы 1 можно интерпретировать как изменение позиций гонщиков от старта до финиша (насколько гонщик улучшил или ухудшил свою стартовую позицию к моменту финиша).

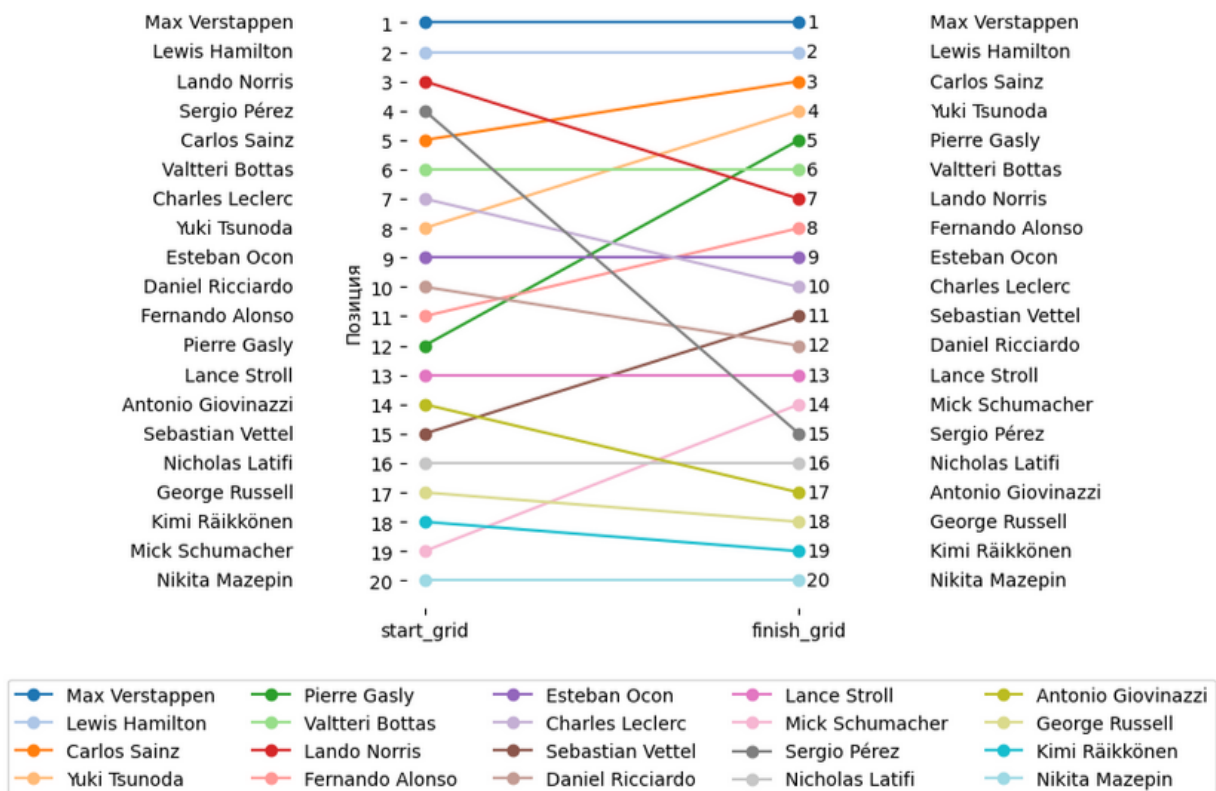
Городская трасса в Баку. Динамичная смена позиций

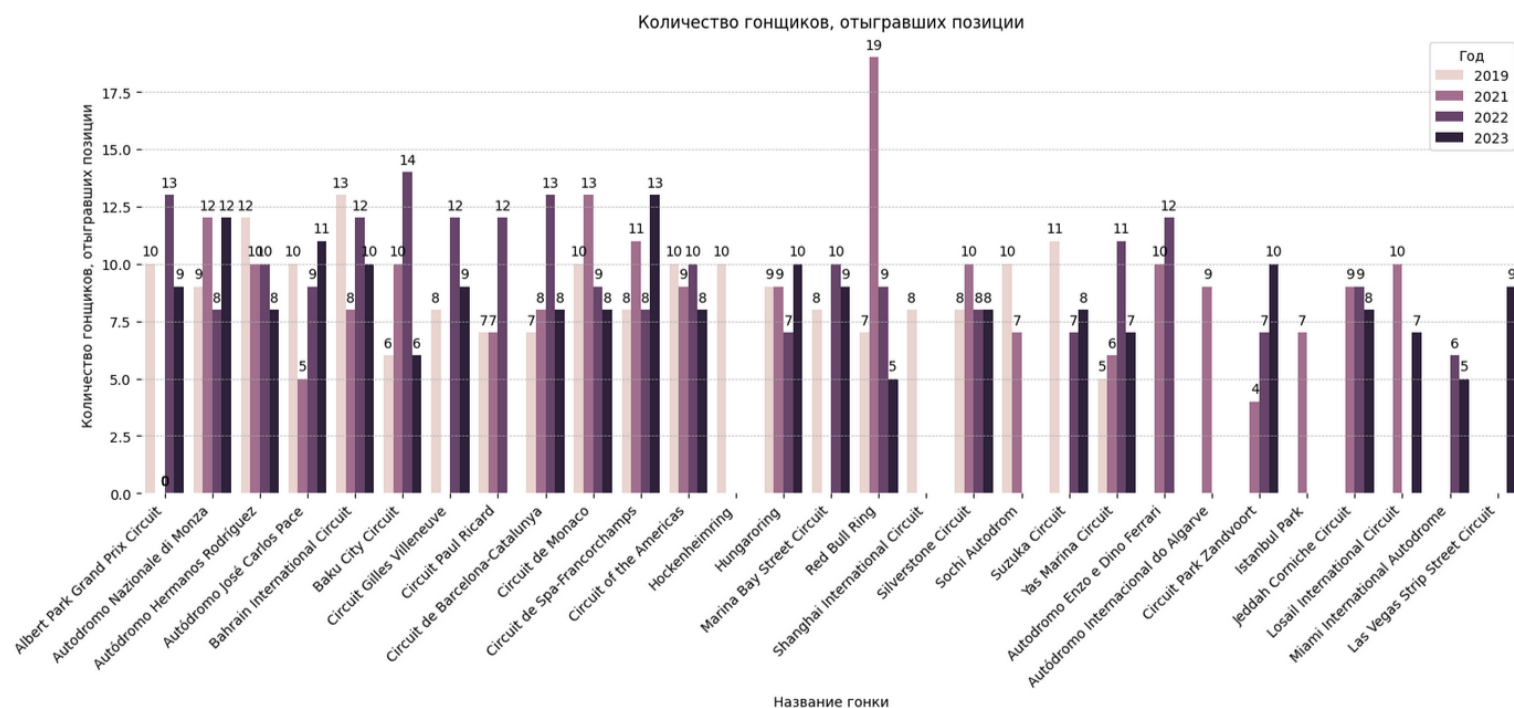
График смены позиций для Azerbaijan Grand Prix (2021)



Гоночная трасса Yas Marina. Более спокойная смена позиций.

График смены позиций для Abu Dhabi Grand Prix (2021)





Больше всего позиций – 21 – было отыграно в 2012 году гонщиком команды Red Bull Себастианом Феттелем на гран-при Абу-Даби.

Гонка с максимальным количеством отыгранных позиций:

Год: 2012

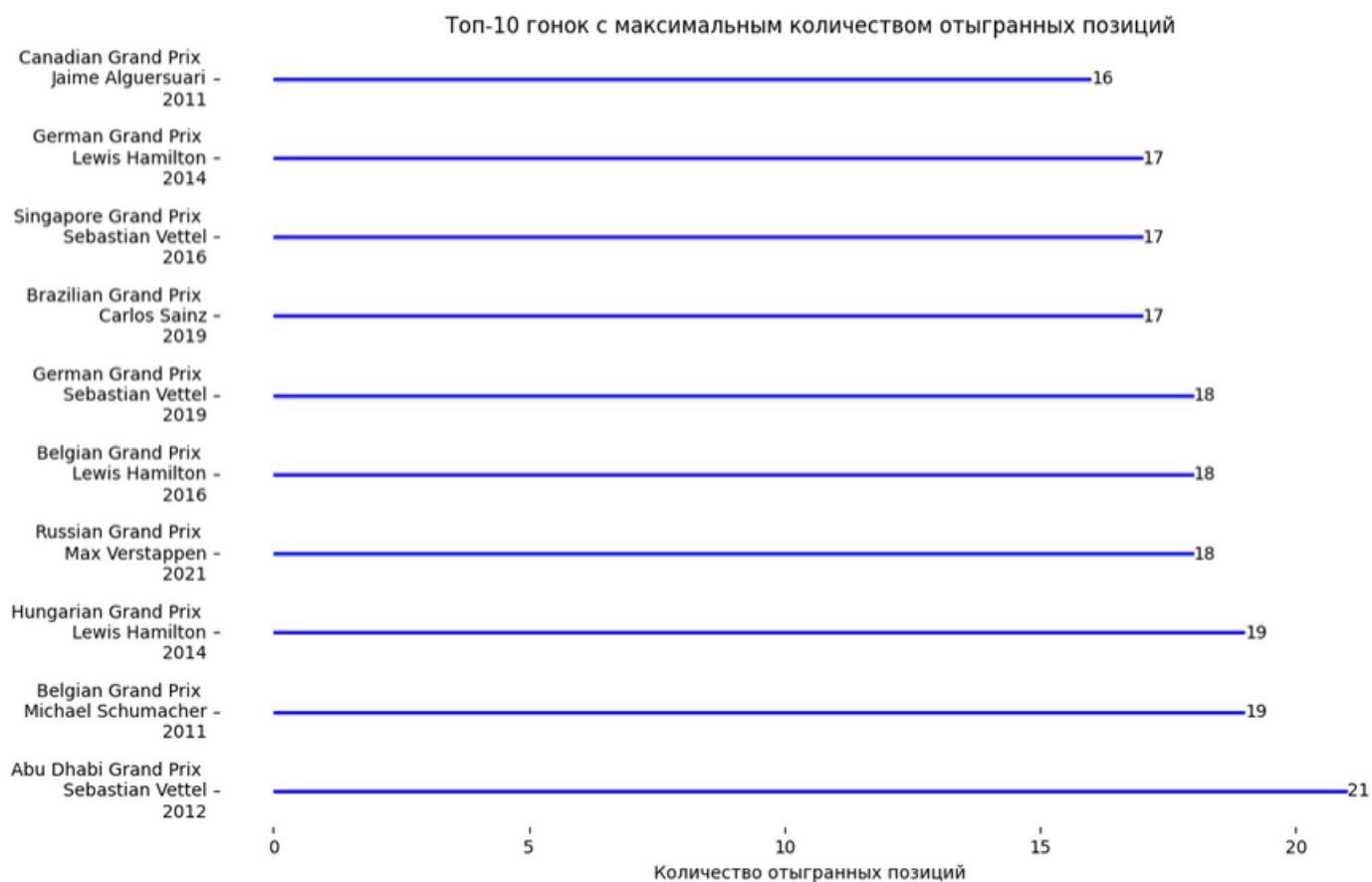
Название гран-при: Abu Dhabi Grand Prix

Трек: Yas Marina Circuit

Гонщик: Sebastian Vettel

Количество отыгранных позиций: 21.0





Большое количество отыгрываемых позиций на трассе может говорить о следующем:

Высокая динамика гонки: возможности для обгонов и смены позиций, наличие длинных прямых, эффективных зон DRS и участков для позднего торможения.

Стратегические возможности: эффективная стратегия пит-стопов и выбора шин.

Скорость и мастерство пилота: высокое мастерство и уверенность гонщика при обгонах, позволили ему минимизировать ошибки и использовать возможности для атаки.

Проблемы у конкурентов: всегда есть шанс, что лидеры столкнулись с техническими проблемами, инцидентами или неэффективной стратегией

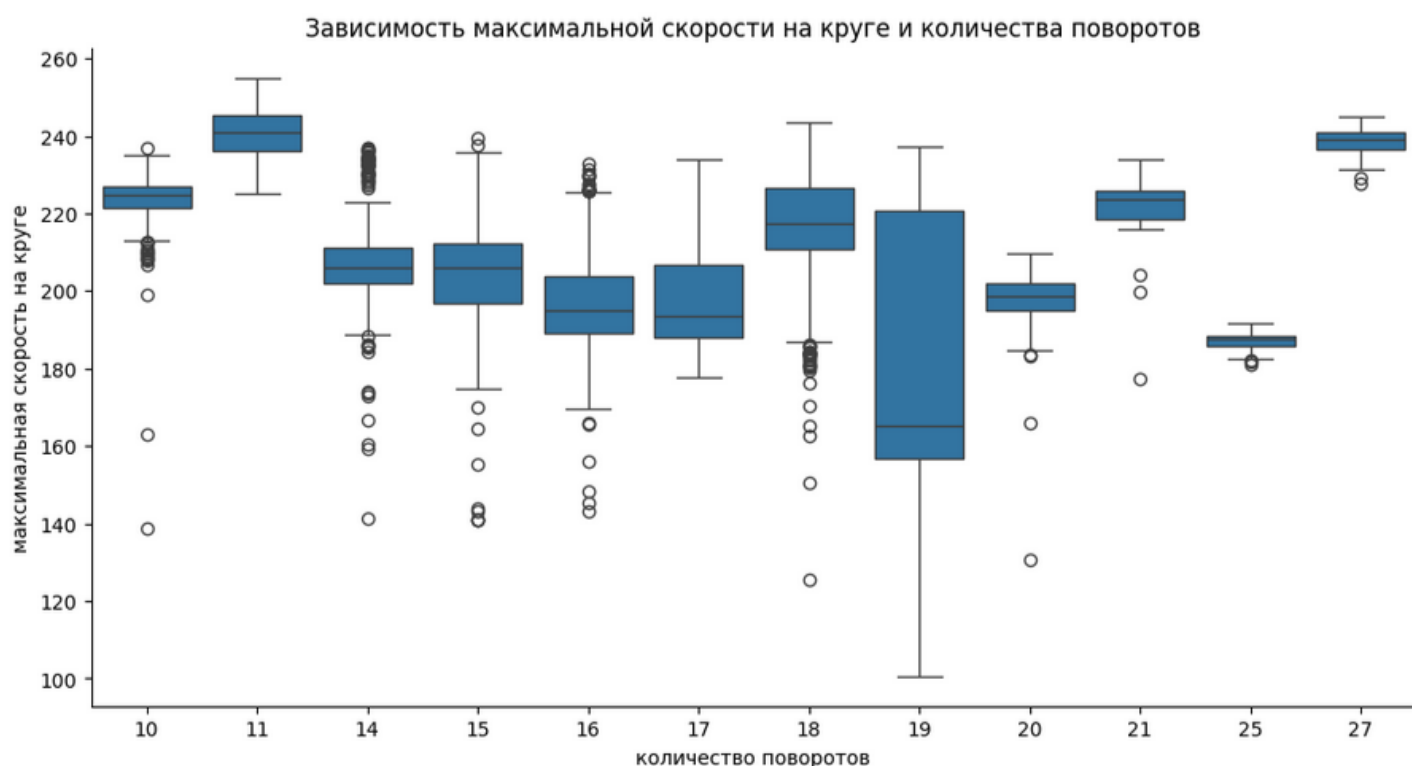
Важность: трасса интересна для стратегического планирования и анализа поведения автомобилей.

6. Скорость на круге и повороты

Попробуем найти как между собой связаны такие показатели как максимальная скорость на круге и количество поворотов. Используем построение боксплотов («ящичков с усами»).

В целом, при увеличении количества поворотов на трассе наблюдается снижение медианной (средней) максимальной скорости. Это соответствует ожиданиям, так как большее количество поворотов требует частого торможения, что снижает максимальные скорости.

Для трасс с 10-11 поворотами медианные максимальные скорости значительно выше, достигая около 240 км/ч. Это логично, так как на трассах с меньшим количеством поворотов гонщики могут поддерживать высокие скорости.



При увеличении числа поворотов до 14-19 медианные максимальные скорости снижаются до диапазона 180-200 км/ч. Диапазоны значений (размеры коробок) для 14, 15 и 18 поворотов относительно узкие, что указывает на более стабильные максимальные скорости. Трассы с 19 поворотами демонстрируют наиболее широкий разброс значений, что

говорит о большом разнообразии в скорости между разными гонками или автомобилями.

При дальнейшем увеличении числа поворотов (20, 21, 25, 27) наблюдается стабилизация и даже небольшое повышение медианных скоростей (около 200 км/ч для некоторых трасс). Возможно, это связано с тем, что трассы с большим количеством поворотов могут иметь и длинные прямые участки, где гонщики компенсируют потерю скорости.

На каждом графике присутствуют точки выбросов — данные, значительно отклоняющиеся от основной массы. Это может быть связано с авариями, техническими неполадками, различными условиями гонок (например, дождь) или особенностями стратегии гонщиков.

Выводы: количество поворотов сильно влияет на максимальные скорости. На трассах с меньшим количеством поворотов гонщики могут развивать более высокие максимальные скорости. На трассах с 14-19 поворотами максимальные скорости значительно падают, а затем слегка повышаются на трассах с большим числом поворотов.

Стабильность скоростей на разных трассах варьируется. Для некоторых значений количества поворотов (например, 15 и 18) скорость демонстрирует небольшое отклонение, в то время как трассы с 19 поворотами показывают значительно большее разнообразие скоростей.

Аномальные значения (выбросы) следует дополнительно исследовать, так как они могут содержать информацию о технических неисправностях, авариях или других критических событиях.

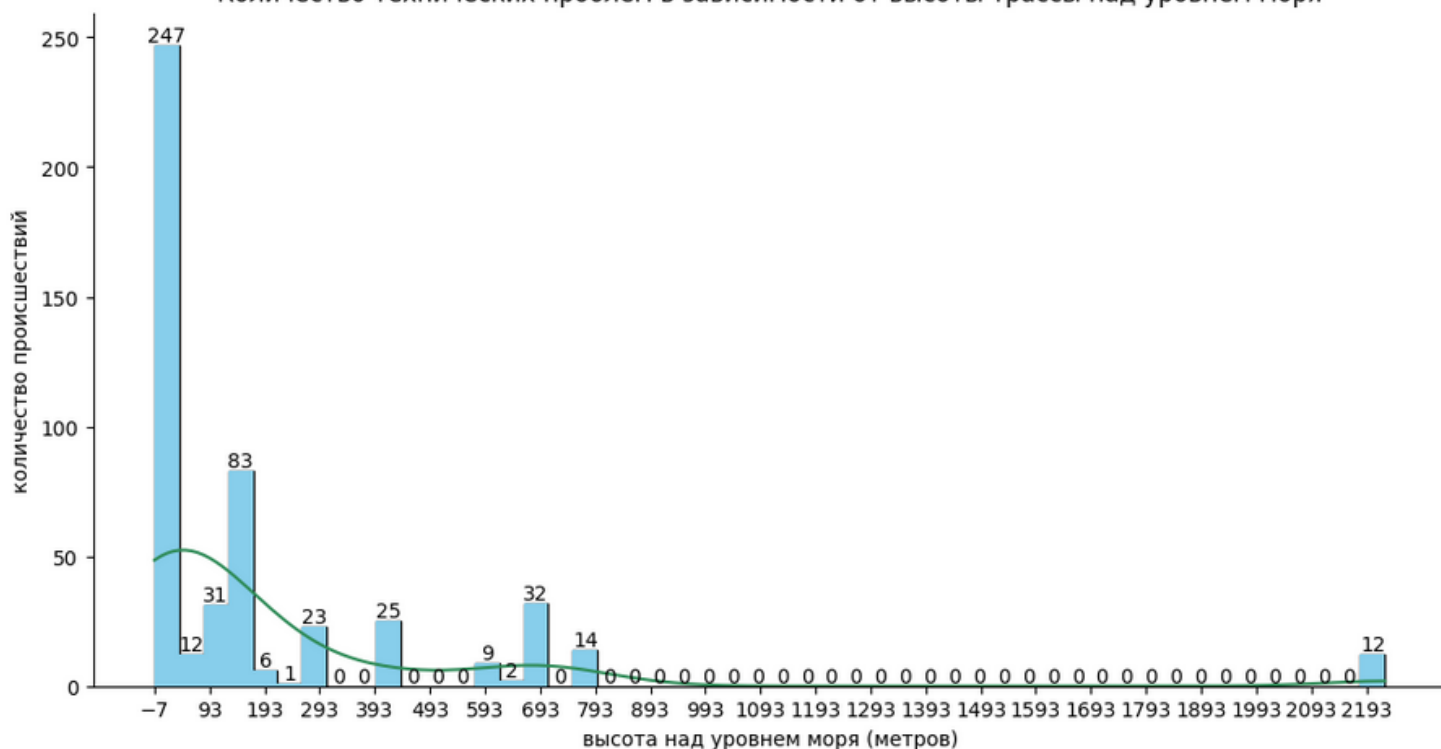
7. Высота расположения трасс

Географическое расположение трассы (высота над уровнем моря), может значительно влиять на несколько аспектов гонок Формулы 1:

Мощность двигателя: на больших высотах воздух разреженнее, что снижает количество кислорода, поступающего в двигатель. Это приводит к уменьшению мощности, так как двигатели внутреннего сгорания зависят от

кислорода для горения топлива. Турбированные двигатели, как в современных болидах F1, менее чувствительны к этому эффекту, так как турбина помогает компенсировать разреженность воздуха. Тем не менее, на больших высотах мощность всё равно может быть слегка снижена.

Аэродинамика: Разреженный воздух уменьшает сопротивление и количество технических проблем в зависимости от высоты трассы над уровнем моря



прижимную силу. Это значит, что машины на трассах, расположенных высоко над уровнем моря, могут иметь меньшую аэродинамическую устойчивость и хуже удерживать сцепление с трассой, особенно в поворотах. Инженеры могут пытаться компенсировать это изменениями в настройках машины, увеличивая углы атаки аэродинамических элементов.

Температурные условия: Высота влияет на температуру воздуха. Как правило, чем выше над уровнем моря, тем прохладнее становится. Это может влиять на охлаждение двигателя и шин, а также на их производительность. Прохладный воздух может улучшать охлаждение машины, но в то же время усложняет прогрев шин до оптимальной рабочей температуры.

Наименьшее количество инцидентов происходит на средних и высоких высотах (от 600 до 900 метров), что может указывать на то, что автомобили лучше справляются с такими условиями.

Максимальное количество инцидентов наблюдается на низких высотах, особенно ниже 100 метров над уровнем моря, что может свидетельствовать о большей плотности трасс в этих регионах или о специфике условий на этих трассах.

Проверим:

Список трасс с высотой ниже 100 метров:

	track_name	alt			
4764	Baku City Circuit	-7	4122	Bahrain International Circuit	7
364	Korean International Circuit	0	5120	Albert Park Grand Prix Circuit	10
5148	Miami International Autodrome	0	5387	Losail International Circuit	12
2205	Sochi Autodrom	2	4793	Circuit Gilles Villeneuve	13
1288	Yas Marina Circuit	3	4609	Jeddah Corniche Circuit	15
644	Valencia Street Circuit	4	1882	Sepang International Circuit	18
2622	Shanghai International Circuit	5	2000	Marina Bay Street Circuit	18
4476	Circuit Park Zandvoort	6	4699	Autodromo Enzo e Dino Ferrari	37
2254	Circuit de Monaco	7	2037	Suzuka Circuit	45

Количество трасс с высотой ниже 100 метров: 18

Действительно, на этих высотах сконцентрировано больше половины всех рассматриваемых трасс, задействованных в гонках с 2011 по 2023 год.

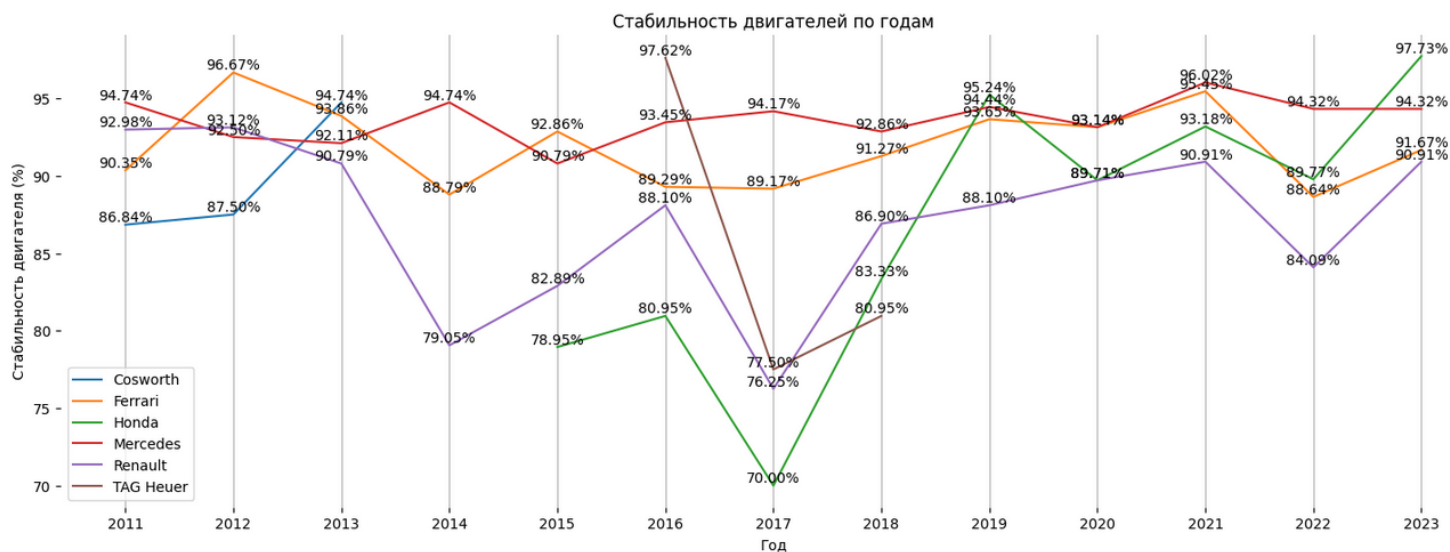
На экстремально высоких трассах (более 2000 метров), как в Мехико, также присутствуют технические проблемы, вероятно, связанные с разреженностью воздуха, что увеличивает нагрузку на двигатели и системы охлаждения.

8. Надежность двигателей

Mercedes и Ferrari демонстрируют наивысшие показатели стабильности, что объясняет их успехи и доминирование в исследуемые годы.

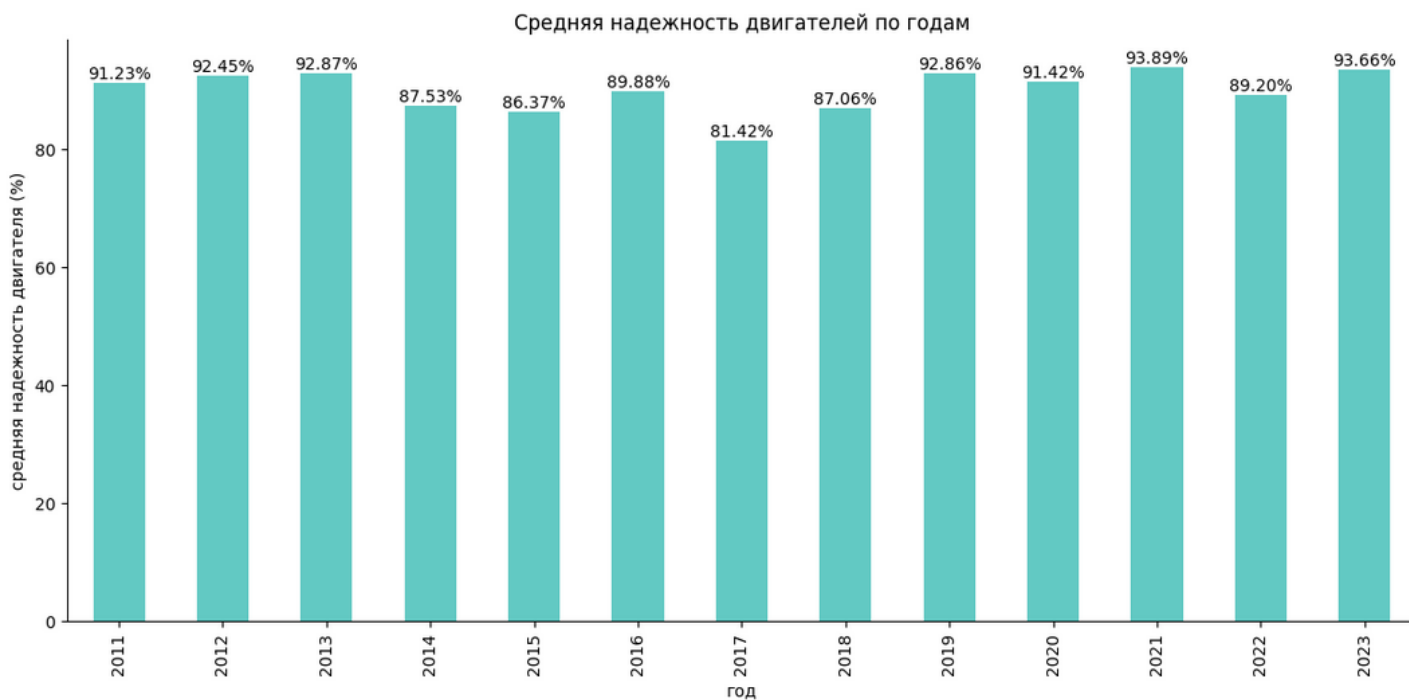
Honda переживала серьезные трудности в 2017 году, но затем смогла значительно улучшить показатели надежности, что сыграло важную роль в успехах команд с их двигателями.

Renault показывает непостоянные результаты, с падениями и восстановлением стабильности, что указывает на нестабильность их двигателей в разные годы.



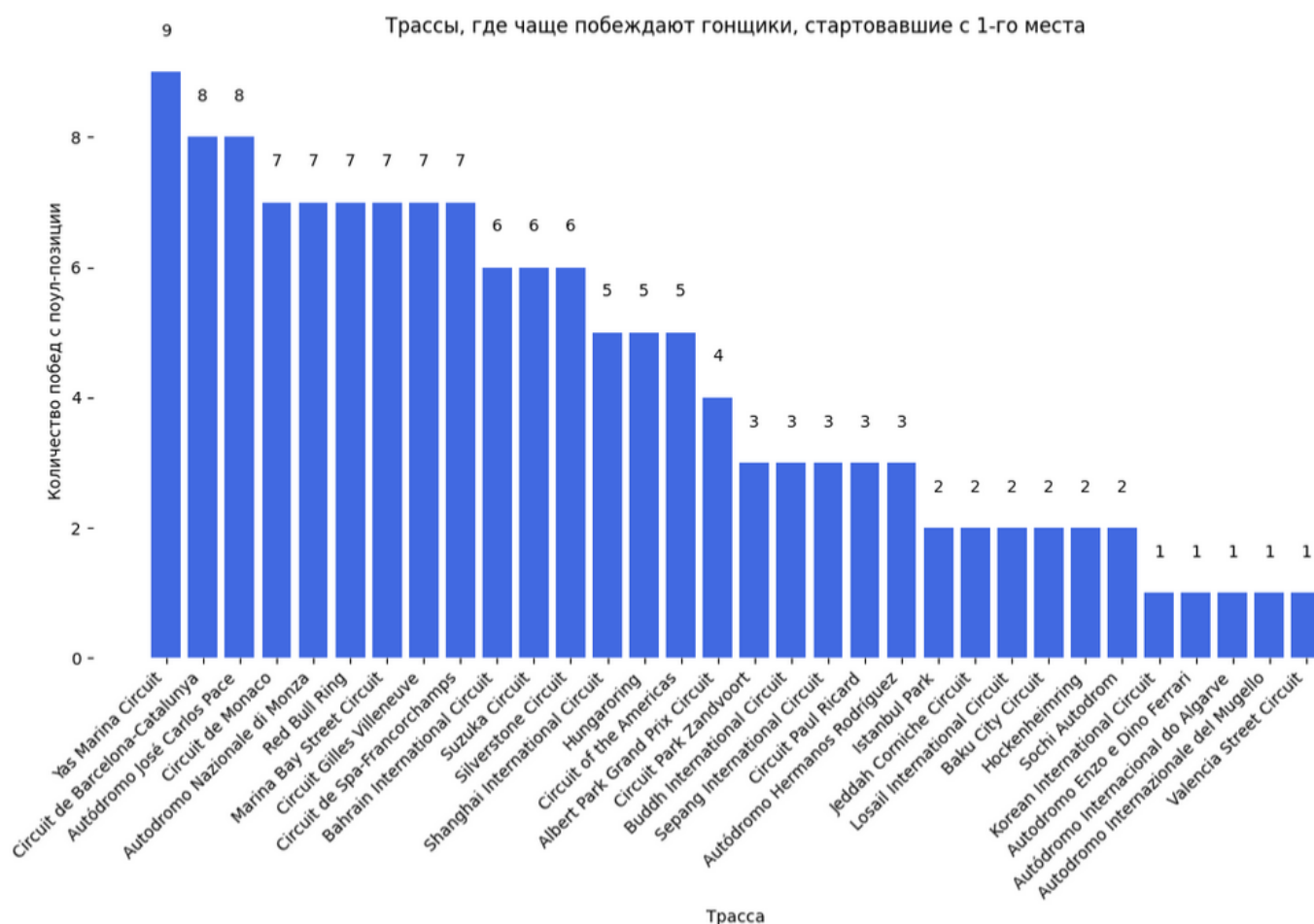
Наблюдается общая тенденция к повышению стабильности: большинство производителей двигателей стремятся повысить надёжность и уменьшить количество технических проблем, что отражается на повышении показателей в последние годы, что хорошо отражается на графике ниже.

Mercedes как производитель и поставщик двигателей для других команд чувствует себя весьма уверенно.



9. Поул-позиция³

На трассах с более широкими и длинными прямыми (например, Монца) значение поул-позиции уменьшается, так как высокие скорости и зоны DRS позволяют конкурентам осуществлять обгоны и менять позиции на первых кругах.



На городских трассах, таких как Монако и Сингапур, где обгонять крайне сложно из-за узких улиц и большого количества поворотов, пилот на поул-позиции имеет значительное преимущество. Здесь сохранение поул-позиции до финиша вероятно, если не случатся технические проблемы или инциденты. И действительно, если посмотреть на график многие городские трассы имеют высокий процент побед при старте с первого места.

³ Поул-позиция (pole position) в Формуле 1 — это результат, показанный гонщиком в квалификации, который позволяет ему стартовать с первой позиции на решётке в гонке. Достижение поул-позиции часто является показателем высокой скорости и мастерства пилота, а также эффективности болида на конкретной трассе.

10. Происшествия на трассах

Всего в изучаемых датасетах представлено 139 различных происшествий, которые происходили и могут произойти во время гонки. Для удобства работы и большей конкретики в запросах был добавлен новый столбец с обозначением 5 типов событий: технические неполадки, навыки гонщика, события потенциально опасные для жизни пилота, непредвиденные обстоятельства и человеческий фактор.

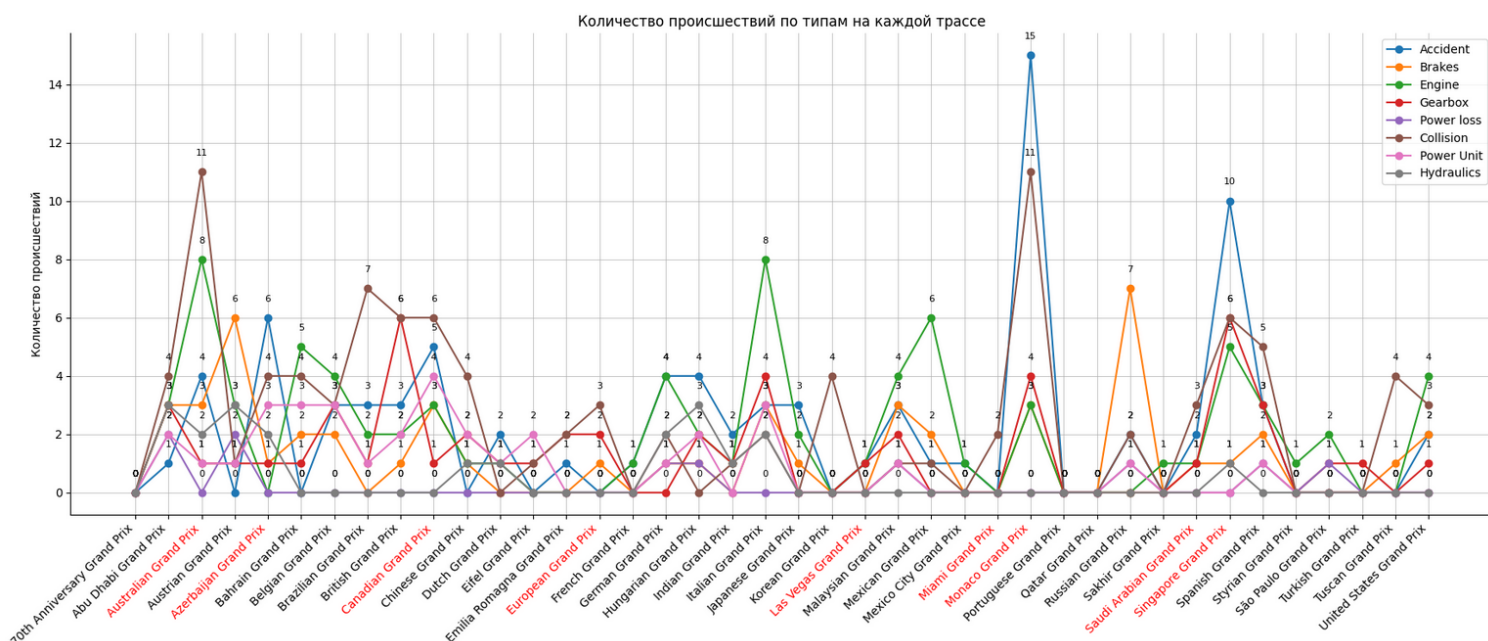
Данные обработаны и представлены в виде подобной сводки по трассам:

Трасса: Albert Park Grand Prix Circuit

	status	status_count
2	Collision	11
7	Engine	8
0	Accident	4
20	Wheel	4
1	Brakes	3

Трасса: Autodromo Enzo e Dino Ferrari

	status	status_count
22	Collision	2
24	Gearbox	2
21	Accident	1
23	Collision damage	1
25	Illness	1



На основании этих данных были выделены трассы и те происшествия, которые наиболее для них характерны. Также было учтено разбиение на городские и гоночные треки. Самые высокие значения на графике в большинстве случаев соотносятся с уличными треками, как наиболее требовательными к технической надежности болида и сложными в пилотировании.

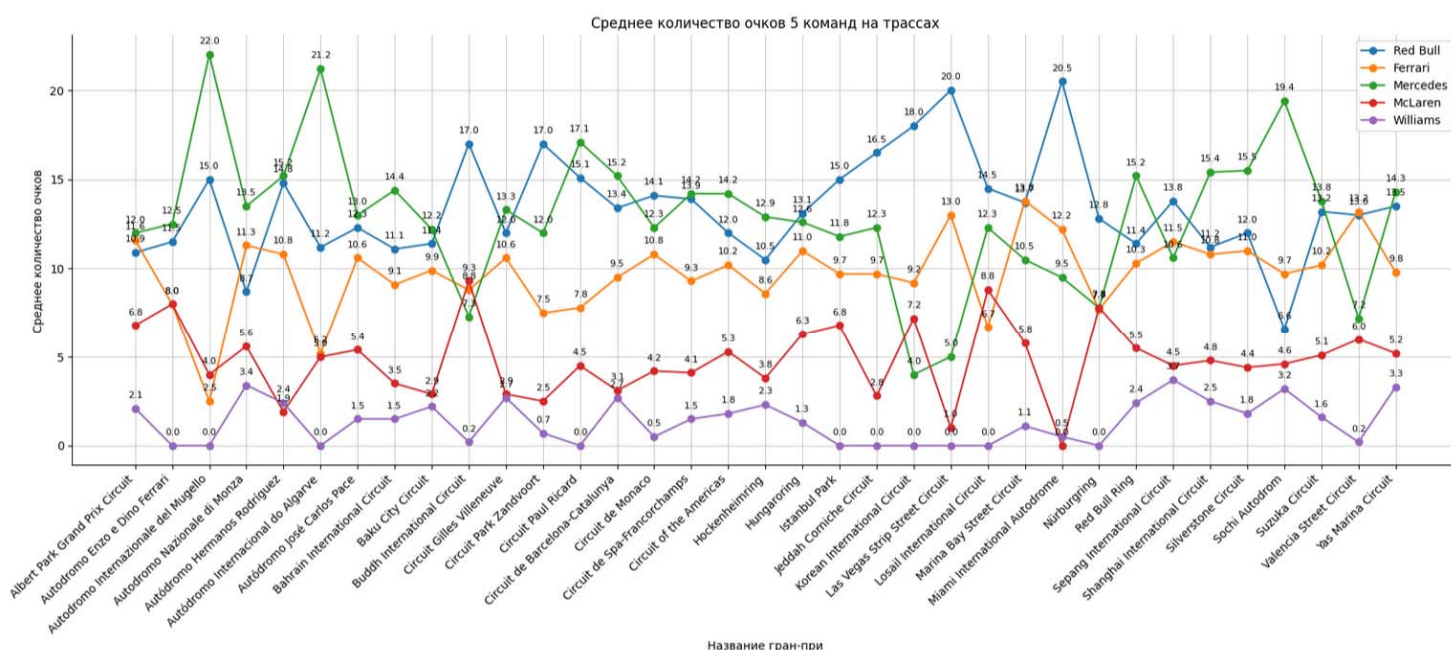
V ЗАКЛЮЧЕНИЕ

В ходе аналитической работы были изучены данные гонок Формулы 1 с акцентом на влияние характеристик трасс на результаты и поведение болидов. Были проанализированы такие параметры, как высота над уровнем моря, количество поворотов, длина прямых и их связь с техническими проблемами, частотой смены позиций и надежность двигателей.

Анализ показал, что трассы с большим количеством поворотов и узкими участками характеризуются меньшим числом обгонов и повышенной важностью стартовой позиции. На высокогорных трассах наблюдается рост количества технических проблем из-за повышенной нагрузки на двигатели и системы охлаждения. Обгоны и смена позиций чаще происходят на трассах с длинными прямыми и эффективными зонами DRS, что позволяет лучше планировать стратегии гонщиков и команд.

Полученные результаты позволяют лучше понять влияние конфигурации трасс на динамику гонок и прогнозировать поведение болидов в различных условиях. Это может помочь командам оптимизировать стратегии и снизить вероятность технических проблем в будущем. Оценить, каких целей удалось достичь.

Команды могут показывать стабильные результаты на определённых типах трасс. На трассах с большим количеством резких поворотов (Монако, Сингапур) чаще случаются аварии и технические проблемы, связанные с тормозами и подвеской.

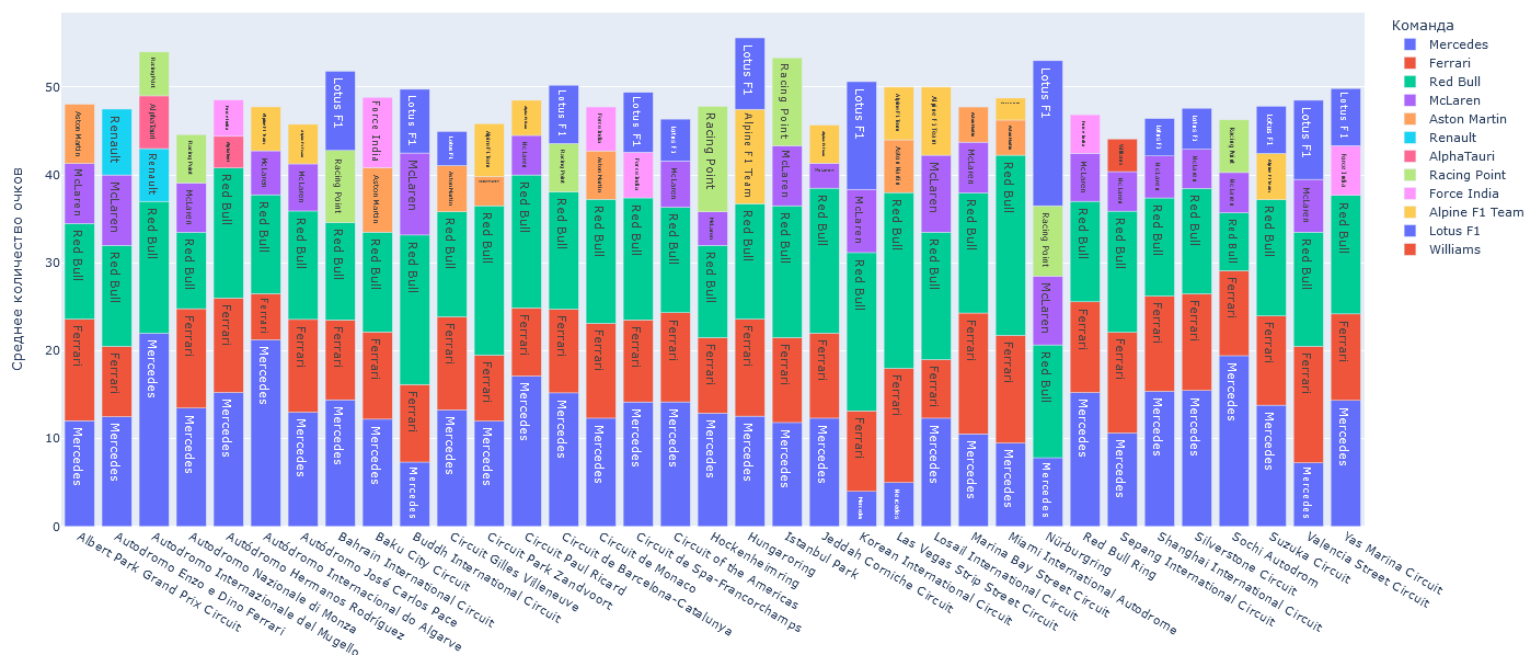


Высокогорные трассы (например, Мехико) значительно увеличивают вероятность перегрева двигателей и отказов гидравлических систем.

Определение трасс, где отдельные команды чаще сталкиваются с техническими проблемами или авариями. Это позволит командам лучше готовиться к таким трассам, устраняя слабые места и минимизируя риски.

Ответом на последний вопрос будет этот и предыдущий график (какие команды уверенно чувствуют себя на любом треке?): последнее десятилетие это эра доминирования Mercedes и Red Bull

Топ-5 команд на каждом гоночном треке (среднее количество очков)



BOX, BOX

