Race Data Analysis

Goal: Develop graphs of WR statistics of certain tracks over time

Test the hypothesis: World Record Times have been cut down as new technology in racing becomes better

Import Data and Modules

```
from seaborn import FacetGrid

from DataFrameImport import get_schema_info, list_schemas, schema

import pandas as pd
import seaborn as sns

path = 'resources/pickled_tables/'
extension = '.plk'

race_data_table = 'race_data'
race_table = 'race'
circuit_table = 'circuit'

race_data_file = path + race_data_table + extension
race_file = path + race_table + extension
circuit_file = path + circuit_table + extension

race_data = pd.read_pickle(race_data_file)
race = pd.read_pickle(race_file)
circuit = pd.read_pickle(circuit_file)

race_data
```

```
        race_id                         type  position_display_order  \
0           290     PRE_QUALIFYING_RESULT                          1
1           290     PRE_QUALIFYING_RESULT                          2
2           290     PRE_QUALIFYING_RESULT                          3
3           290     PRE_QUALIFYING_RESULT                          4
4           290     PRE_QUALIFYING_RESULT                          5
...         ...                       ...                     ...
183627     1143  DRIVER_OF_THE_DAY_RESULT                          1
183628     1143  DRIVER_OF_THE_DAY_RESULT                          2
183629     1143  DRIVER_OF_THE_DAY_RESULT                          3
183630     1143  DRIVER_OF_THE_DAY_RESULT                          4
183631     1143  DRIVER_OF_THE_DAY_RESULT                          5

        position_number position_text driver_number
driver_id  \
0                   1.0             1            40   gilles-
```

```
villeneuve
1                    2.0              2            23      patrick-
tambay
2                    3.0              3            34  jean-pierre-
jarier
3                    4.0              4            30        brett-
lunger
4                    5.0              5            38        brian-
henton
...                  ...            ...           ...                ..
.
183627               1.0              1            14     fernando-
alonso
183628               2.0              2            63      george-
russell
183629               3.0              3             1      max-
verstappen
183630               4.0              4             4        lando-
norris
183631               5.0              5            44      lewis-
hamilton

       constructor_id engine_manufacturer_id tyre_manufacturer_id  ...
\
0             mclaren                   ford             goodyear  ...

1              ensign                   ford             goodyear  ...

2              penske                   ford             goodyear  ...

3             mclaren                   ford             goodyear  ...

4               march                   ford             goodyear  ...

...               ...                    ...                  ...  ...

183627   aston-martin               mercedes              pirelli  ...

183628       mercedes               mercedes              pirelli  ...

183629       red-bull             honda-rbpt              pirelli  ...

183630        mclaren               mercedes              pirelli  ...

183631        ferrari                ferrari              pirelli  ...


       fastest_lap_time_millis  fastest_lap_gap fastest_lap_gap_millis
\
0                          NaN             None                    NaN
```

```
1                           NaN              None                       NaN

2                           NaN              None                       NaN

3                           NaN              None                       NaN

4                           NaN              None                       NaN

...                         ...               ...                       ...

183627                      NaN              None                       NaN

183628                      NaN              None                       NaN

183629                      NaN              None                       NaN

183630                      NaN              None                       NaN

183631                      NaN              None                       NaN


        fastest_lap_interval fastest_lap_interval_millis
pit_stop_stop  \
0                       None                          NaN
NaN
1                       None                          NaN
NaN
2                       None                          NaN
NaN
3                       None                          NaN
NaN
4                       None                          NaN
NaN
...                      ...                          ...              ..
.
183627                  None                          NaN
NaN
183628                  None                          NaN
NaN
183629                  None                          NaN
NaN
183630                  None                          NaN
NaN
183631                  None                          NaN
NaN

        pit_stop_lap pit_stop_time  pit_stop_time_millis  \
0                NaN          None                   NaN
1                NaN          None                   NaN
2                NaN          None                   NaN
```

```
3                NaN          None                    NaN
4                NaN          None                    NaN
...              ...           ...                    ...
183627           NaN          None                    NaN
183628           NaN          None                    NaN
183629           NaN          None                    NaN
183630           NaN          None                    NaN
183631           NaN          None                    NaN

       driver_of_the_day_percentage
0                               NaN
1                               NaN
2                               NaN
3                               NaN
4                               NaN
...                             ...
183627                         22.5
183628                         16.4
183629                         14.5
183630                          8.7
183631                          7.6

[183632 rows x 71 columns]

race

        id  year  round        date   time  grand_prix_id  \
0        1  1950      1  1950-05-13   None  great-britain
1        2  1950      2  1950-05-21   None         monaco
2        3  1950      3  1950-05-30   None   indianapolis
3        4  1950      4  1950-06-04   None    switzerland
4        5  1950      5  1950-06-18   None        belgium
...    ...   ...    ...         ...    ...            ...
1144  1145  2025     20  2025-10-26  20:00         mexico
1145  1146  2025     21  2025-11-09  17:00      sao-paulo
1146  1147  2025     22  2025-11-23  04:00      las-vegas
1147  1148  2025     23  2025-11-30  16:00          qatar
1148  1149  2025     24  2025-12-07  13:00      abu-dhabi

                                        official_name
qualifying_format  \
0                          1950 RAC British Grand Prix
TWO_SESSION
1                            Grand Prix de Monaco 1950
TWO_SESSION
2                               1950 Indianapolis 500
FOUR_LAPS
3                      Grosser Preis der Schweiz 1950
TWO_SESSION
4                              1950 Belgian Grand Prix
```

```
TWO_SESSION
...                                                   ...
...
1144  Formula 1 Gran Premio de la Ciudad de México 2025
KNOCKOUT
1145  Formula 1 MSC Cruises Grande Prêmio de São Pau...
KNOCKOUT
1146       Formula 1 Heineken Las Vegas Grand Prix 2025
KNOCKOUT
1147      Formula 1 Qatar Airways Qatar Grand Prix 2025
KNOCKOUT
1148  Formula 1 Etihad Airways Abu Dhabi Grand Prix ...
KNOCKOUT

     sprint_qualifying_format         circuit_id  ...
qualifying_2_date  \
0                        None        silverstone  ...
None
1                        None             monaco  ...
None
2                        None       indianapolis  ...
None
3                        None         bremgarten  ...
None
4                        None  spa-francorchamps  ...
None
...                       ...                ...  ...                ..
.
1144                     None        mexico-city  ...
None
1145          SPRINT_SHOOTOUT         interlagos  ...
None
1146                     None          las-vegas  ...
None
1147          SPRINT_SHOOTOUT              lusail  ...
None
1148                     None         yas-marina  ...
None

     qualifying_2_time  qualifying_date  qualifying_time  \
0                 None             None             None
1                 None             None             None
2                 None             None             None
3                 None             None             None
4                 None             None             None
...                ...              ...              ...
1144              None       2025-10-25            21:00
1145              None       2025-11-08            18:00
1146              None       2025-11-22            04:00
```

```
1147               None       2025-11-29            18:00
1148               None       2025-12-06            14:00

      sprint_qualifying_date  sprint_qualifying_time  sprint_race_date
\
0                       None                    None              None

1                       None                    None              None

2                       None                    None              None

3                       None                    None              None

4                       None                    None              None

...                      ...                     ...               ...

1144                    None                    None              None

1145              2025-11-07                   18:30        2025-11-08

1146                    None                    None              None

1147              2025-11-28                   17:30        2025-11-29

1148                    None                    None              None


      sprint_race_time  warming_up_date  warming_up_time
0                 None             None             None
1                 None             None             None
2                 None             None             None
3                 None             None             None
4                 None             None             None
...                ...              ...              ...
1144              None             None             None
1145             14:00             None             None
1146              None             None             None
1147             14:00             None             None
1148              None             None             None

[1149 rows x 42 columns]

circuit

            id                 name                        full_name  \
0     adelaide            Adelaide        Adelaide Street Circuit
1         aida                Aida  Okayama International Circuit
2     ain-diab            Ain-Diab              Ain-Diab Circuit
3      aintree             Aintree   Aintree Motor Racing Circuit
4   anderstorp  Anderstorp Raceway             Anderstorp Raceway
```

```
..          ...                  ...                              ...
72  yas-marina          Yas Marina             Yas Marina Circuit
73     yeongam               Korea    Korea International Circuit
74   zandvoort           Zandvoort         Circuit Park Zandvoort
75     zeltweg             Zeltweg                         Zeltweg
76      zolder              Zolder                 Circuit Zolder

           previous_names    type       direction      place_name  \
0                    None  STREET       CLOCKWISE        Adelaide
1         TI Circuit Aida    RACE       CLOCKWISE            Aida
2                    None    ROAD       CLOCKWISE      Casablanca
3                    None    ROAD       CLOCKWISE         Aintree
4   Scandinavian Raceway    RACE       CLOCKWISE      Anderstorp
..                   ...     ...             ...             ...
72                   None    RACE  ANTI_CLOCKWISE       Abu Dhabi
73                   None    RACE  ANTI_CLOCKWISE         Yeongam
74                   None    RACE       CLOCKWISE       Zandvoort
75                   None    ROAD       CLOCKWISE         Zeltweg
76                   None    RACE       CLOCKWISE  Heusden-Zolder

              country_id   latitude   longitude  length  turns  \
0              australia -34.927222  138.617222   3.780     16
1                  japan  34.915000  134.221111   3.703     13
2                morocco  33.578611   -7.687500   7.618     18
3         united-kingdom  53.476944   -2.940556   4.828      8
4                 sweden  57.264167   13.601389   4.031      8
..                   ...        ...         ...     ...    ...
72  united-arab-emirates  24.467222   54.603056   5.281     16
73           south-korea  34.733333  126.416667   5.615     18
74           netherlands  52.388819    4.540922   4.259     14
75               austria  47.202222   14.742222   3.186      4
76               belgium  50.988889    5.255556   4.262     15

    total_races_held
0                 11
1                  2
2                  1
3                  5
4                  6
..               ...
72                16
73                 4
74                35
75                 1
76                10

[77 rows x 13 columns]
```

Clean Data

Data Documentation Look-up

Display helper tools that describe the information contained in the table

```
get_schema_info('races')
```

```
races
The list of races, each representing detailed information about
individual races, including results, participants, and statistics.
```

```
get_schema_info('circuits')
```

```
circuits
The list of circuits, each representing a specific racing track,
including geographical location and race history.
```

```
print(schema.Race)
```

```
id = type = integer
description = The unique identifier of the race.

year = type = integer
description = The year of the season.

round = type = integer
description = The round number of the race in the season.

date = type = string
description = The date of the race in UTC.

time = type = ['string', 'null']
description = The start time of the race in UTC.

grandPrixId = type = string
description = The identifier of the Grand Prix associated with the
race.

officialName = type = string
description = The official name of the race.

qualifyingFormat = description = The qualifying format of the race.

sprintQualifyingFormat = description = The sprint qualifying format of
the race.

circuitId = type = string
description = The identifier of the circuit where the race takes
```

```
place.

circuitType = description = The type of the circuit.

direction = description = The direction of the circuit.

courseLength = type = number
description = The length of the circuit (race course) in kilometers.

turns = type = integer
description = The number of turns (corners) in the configuration of
the circuit.

laps = type = integer
description = The total number of laps of the race.

distance = type = number
description = The total distance of the race in kilometers.

scheduledLaps = type = ['integer', 'null']
description = The scheduled number of laps of the race.

scheduledDistance = type = ['number', 'null']
description = The scheduled distance of the race in kilometers.

driversChampionshipDecider = type = boolean
description = Whether this race was the decider of the World Drivers'
Championship.

constructorsChampionshipDecider = type = boolean
description = Whether this race was the decider of the World
Constructors' Championship.

preQualifyingDate = type = ['string', 'null']
description = The date of the pre-qualifying session in UTC.

preQualifyingTime = type = ['string', 'null']
description = The start time of the pre-qualifying session in UTC.

preQualifyingResults = type = ['array', 'null']
description = The results of the pre-qualifying session.

freePractice1Date = type = ['string', 'null']
description = The date of the 1st free practice session in UTC.

freePractice1Time = type = ['string', 'null']
description = The start time of the 1st free practice session in UTC.

freePractice1Results = type = ['array', 'null']
description = The results of the 1st free practice session.
```

```
freePractice2Date = type = ['string', 'null']
description = The date of the 2nd free practice session in UTC.

freePractice2Time = type = ['string', 'null']
description = The start time of the 2nd free practice session in UTC.

freePractice2Results = type = ['array', 'null']
description = The results of the 2nd free practice session.

freePractice3Date = type = ['string', 'null']
description = The date of the 3rd free practice session in UTC.

freePractice3Time = type = ['string', 'null']
description = The start time of the 3rd free practice session in UTC.

freePractice3Results = type = ['array', 'null']
description = The results of the 3rd free practice session.

freePractice4Date = type = ['string', 'null']
description = The date of the 4th free practice session UTC.

freePractice4Time = type = ['string', 'null']
description = The start time of the 4th free practice session in UTC.

freePractice4Results = type = ['array', 'null']
description = The results of the 4th free practice session.

qualifying1Date = type = ['string', 'null']
description = The date of the 1st qualifying session in UTC.

qualifying1Time = type = ['string', 'null']
description = The start time of the 1st qualifying session in UTC.

qualifying1Results = type = ['array', 'null']
description = The results of the 1st qualifying session.

qualifying2Date = type = ['string', 'null']
description = The date of the 2nd qualifying session UTC.

qualifying2Time = type = ['string', 'null']
description = The start time of the 2nd qualifying session UTC.

qualifying2Results = type = ['array', 'null']
description = The results of the 2nd qualifying session.

qualifyingDate = type = ['string', 'null']
description = The date of the qualifying session UTC.

qualifyingTime = type = ['string', 'null']
description = The start time of the qualifying session UTC.
```

```
qualifyingResults = type = ['array', 'null']
description = The results of the qualifying session.

sprintQualifyingDate = type = ['string', 'null']
description = The date of the sprint qualifying session in UTC.

sprintQualifyingTime = type = ['string', 'null']
description = The start time of the sprint qualifying session in UTC.

sprintQualifyingResults = type = ['array', 'null']
description = The results of the sprint qualifying session.

sprintStartingGridPositions = type = ['array', 'null']
description = The starting grid positions for the sprint race.

sprintRaceDate = type = ['string', 'null']
description = The date of the sprint race in UTC.

sprintRaceTime = type = ['string', 'null']
description = The start time of the sprint race in UTC.

sprintRaceResults = type = ['array', 'null']
description = The results of the sprint race.

warmingUpDate = type = ['string', 'null']
description = The date of the warming-up session in UTC.

warmingUpTime = type = ['string', 'null']
description = The start time of the warming-up session in UTC.

warmingUpResults = type = ['array', 'null']
description = The results of the warming-up session.

startingGridPositions = type = ['array', 'null']
description = The starting grid positions for the race.

raceResults = type = ['array', 'null']
description = The results of the race.

fastestLaps = type = ['array', 'null']
description = The fastest laps recorded during the race..

pitStops = type = ['array', 'null']
description = The pit stops made during the race.

driverOfTheDayResults = type = ['array', 'null']
description = The results of the Driver of the Day vote.

driverStandings = type = ['array', 'null']
description = The driver standings after the race.
```

```
constructorStandings = type = ['array', 'null']
description = The constructor standings after the race.
```

Initial Analysis

```
print(schema)

Continent = id = type = string
description = The unique identifier for the continent.

code = type = string
description = The unique code of the continent.

name = type = string
description = The name of the continent.

demonym = type = string
description = The demonym used for people from the continent.


Country = id = type = string
description = The unique identifier for the country.

alpha2Code = type = string
description = The unique ISO 3166-1 alpha-2 code of the country.

alpha3Code = type = string
description = The unique ISO 3166-1 alpha-3 code of the country.

iocCode = type = ['string', 'null']
description = The unique International Olympic Committee (IOC) code of
the country.

name = type = string
description = The name of the country.

demonym = type = ['string', 'null']
description = The demonym for citizens of the country.

continentId = type = string
description = The identifier for the continent where the country is
located.


Driver = id = type = string
description = The unique identifier of the driver.

name = type = string
description = The name of the driver, typically used for display
```

```
purposes.

firstName = type = string
description = The given name or first name of the driver.

lastName = type = string
description = The family name or last name of the driver.

fullName = type = string
description = The full name of the driver, usually a combination of
first, middle and last names.

abbreviation = type = string
description = The three-letter abbreviation for the driver, consisting
of uppercase letters (e.g., 'SEN' for Ayrton Senna).

permanentNumber = type = ['string', 'null']
description = The permanent racing number chosen by the driver.

gender = description = The gender of the driver.

dateOfBirth = type = string
description = The birth date of the driver.

dateOfDeath = type = ['string', 'null']
description = The death date of the driver, if applicable.

placeOfBirth = type = string
description = The place of birth of the driver.

countryOfBirthCountryId = type = string
description = The identifier of the country where the driver was born.

nationalityCountryId = type = string
description = The identifier of the nationality of the driver.

secondNationalityCountryId = type = ['string', 'null']
description = The identifier of the second nationality of the driver,
if applicable.

familyRelationships = type = ['array', 'null']
description = The family relationships involving the driver, such as
parent or sibling relationships.

bestChampionshipPosition = type = ['integer', 'null']
description = The best finishing position achieved by the driver in a
World Drivers' Championship.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the driver
in a race.
```

```
bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the driver in a
race.

totalChampionshipWins = type = integer
description = The total number of World Drivers' Championship titles
won by the driver.

totalRaceEntries = type = integer
description = The total number of races entered by the driver.

totalRaceStarts = type = integer
description = The total number of races started by the driver.

totalRaceWins = type = integer
description = The total number of races won by the driver.

totalRaceLaps = type = integer
description = The total number of laps completed by the driver.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the
driver.

totalPoints = type = number
description = The total number of points accumulated by the driver.

totalChampionshipPoints = type = number
description = The total number of World Drivers' Championship points
accumulated by the driver.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the
driver.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the driver.

totalDriverOfTheDay = type = integer
description = The total number of Driver of the Day awards received by
the driver.

totalGrandSlams = type = integer
description = The total number of Grand Slams achieved by the driver,
defined as pole position, fastest lap, and leading every lap of the
race.


DriverFamilyRelationship = positionDisplayOrder = type = integer
description = The display order of the family relationship relative to
```

```
other relationships of the parent driver.

driverId = type = string
description = The identifier of the related driver.

type = description = The type of the family relationship.


Constructor = id = type = string
description = The unique identifier of the constructor.

name = type = string
description = The name of the constructor, typically used for display
purposes.

fullName = type = string
description = The full name of the constructor.

countryId = type = string
description = The identifier of the country of origin of the
constructor.

chronology = type = ['array', 'null']
description = The chronology of the constructor.

bestChampionshipPosition = type = ['integer', 'null']
description = The best finishing position achieved by the constructor
in a World Constructors' Championship.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the
constructor in a race.

bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the constructor
in a race.

totalChampionshipWins = type = integer
description = The total number of World Constructors' Championship
titles won by the constructor.

totalRaceEntries = type = integer
description = The total number of races entered by the constructor.

totalRaceStarts = type = integer
description = The total number of races started by the constructor.

totalRaceWins = type = integer
description = The total number of races won by the constructor.

total1And2Finishes = type = integer
```

```
description = The total number of races in which the constructor
finished in both 1st and 2nd place.

totalRaceLaps = type = integer
description = The total number of laps completed by the constructor.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the
constructor.

totalPodiumRaces = type = integer
description = The total number of races in which the constructor
finished on the podium.

totalPoints = type = number
description = The total number of points accumulated by the
constructor.

totalChampionshipPoints = type = number
description = The total number of World Constructors' Championship
points accumulated by the constructor.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the
constructor.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the constructor.


ConstructorChronology = positionDisplayOrder = type = integer
description = The display order of the constructor within the
chronological sequence of the parent constructor.

constructorId = type = string
description = The identifier of the constructor.

yearFrom = type = integer
description = The year from.

yearTo = type = ['integer', 'null']
description = The year to, or null if it is still active.


Chassis = id = type = string
description = The unique identifier of the chassis.

constructorId = type = string
description = The identifier of the constructor associated with the
chassis.
```

```
name = type = string
description = The name of the chassis.

fullName = type = string
description = The full name of the chassis.


EngineManufacturer = id = type = string
description = The unique identifier of the engine manufacturer.

name = type = string
description = The name of the engine manufacturer.

countryId = type = string
description = The identifier of the country of origin of the engine
manufacturer.

bestChampionshipPosition = type = ['integer', 'null']
description = The best finishing position achieved by the engine
manufacturer in a World Constructors' Championship.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the engine
manufacturer in a race.

bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the engine
manufacturer in a race.

totalChampionshipWins = type = integer
description = The total number of World Constructors' Championship
titles won by the engine manufacturer.

totalRaceEntries = type = integer
description = The total number of races entered by the engine
manufacturer.

totalRaceStarts = type = integer
description = The total number of races started by the engine
manufacturer.

totalRaceWins = type = integer
description = The total number of races won by the engine
manufacturer.

totalRaceLaps = type = integer
description = The total number of laps completed by the engine
manufacturer.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the
```

```
engine manufacturer.

totalPodiumRaces = type = integer
description = The total number of races in which the engine
manufacturer finished on the podium.

totalPoints = type = number
description = The total number of points accumulated by the engine
manufacturer.

totalChampionshipPoints = type = number
description = The total number of World Constructors' Championship
points accumulated by the engine manufacturer.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the
engine manufacturer.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the engine
manufacturer.


Engine = id = type = string
description = The unique identifier of the engine.

engineManufacturerId = type = string
description = The identifier of the engine manufacturer associated
with the engine.

name = type = string
description = The name of the engine.

fullName = type = string
description = The full name of the engine.

capacity = type = ['number', 'null']
description = The capacity of the engine, measured in liters.

configuration = description = The configuration of the engine, such as
V6, V8, etc.

aspiration = description = The aspiration of the engine, such as
naturally aspirated or turbocharged.


TyreManufacturer = id = type = string
description = The unique identifier of the tyre manufacturer.

name = type = string
description = The name of the tyre manufacturer.
```

```
countryId = type = string
description = The identifier of the country of origin of the tyre
manufacturer.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the tyre
manufacturer in a race.

bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the tyre
manufacturer in a race.

totalRaceEntries = type = integer
description = The total number of races entered by the tyre
manufacturer.

totalRaceStarts = type = integer
description = The total number of races started by the tyre
manufacturer.

totalRaceWins = type = integer
description = The total number of races won by the tyre manufacturer.

totalRaceLaps = type = integer
description = The total number of laps completed by the tyre
manufacturer.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the tyre
manufacturer.

totalPodiumRaces = type = integer
description = The total number of races in which the tyre manufacturer
finished on the podium.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the tyre
manufacturer.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the tyre
manufacturer.


Circuit = id = type = string
description = The unique identifier of the circuit.

name = type = string
description = The name of the circuit, typically used for display
purposes.
```

```
fullName = type = string
description = The full official name of the circuit.

previousNames = type = ['array', 'null']
description = The previous names used for the circuit.

type = description = The type of the circuit.

direction = description = The direction of the current or most
recently used configuration of the circuit.

placeName = type = string
description = The place name where the circuit is located.

countryId = type = string
description = The identifier of the country where the circuit is
located.

latitude = type = number
description = The latitude coordinate where the circuit is located.

longitude = type = number
description = The longitude coordinate where the circuit is located.

length = type = number
description = The length of the current or most recently used
configuration of the circuit in kilometers.

turns = type = integer
description = The number of turns (corners) in the current or most
recently used configuration of the circuit.

totalRacesHeld = type = integer
description = The total number of races held at the circuit.


GrandPrix = id = type = string
description = The unique identifier of the Grand Prix.

name = type = string
description = The name of the Grand Prix, typically used for display
purposes.

fullName = type = string
description = The full name of the Grand Prix.

shortName = type = string
description = The short name of the Grand Prix.

abbreviation = type = string
```

```
description = The three-character abbreviation of the Grand Prix.

countryId = type = ['string', 'null']
description = The identifier of the country where the Grand Prix is
held.

totalRacesHeld = type = integer
description = The total number of races held for this Grand Prix.


SeasonEntrant = entrantId = type = string
description = The identifier of the entrant.

countryId = type = string
description = The identifier of the country of the entrant.

constructors = type = array
description = The constructors associated with the entrant.


Entrant = id = type = string
description = The unique identifier of the entrant.

name = type = string
description = The name of the entrant.


Season = year = type = integer
description = The year of the season.

entrants = type = ['array', 'null']
description = The entrants competing in the season.

constructors = type = ['array', 'null']
description = The constructors competing in the season.

engineManufacturers = type = ['array', 'null']
description = The engine manufacturers competing in the season.

tyreManufacturers = type = ['array', 'null']
description = The tyre manufacturers competing in the season.

drivers = type = ['array', 'null']
description = The drivers competing in the season.

driverStandings = type = ['array', 'null']
description = The driver standings of the season.

constructorStandings = type = ['array', 'null']
description = The constructor standings of the season.
```

```
SeasonEntrantConstructor = constructorId = type = string
description = The identifier of the constructor.

engineManufacturerId = type = string
description = The identifier of the engine manufacturer.

chassis = type = array
description = The chassis used by the constructor.

engines = type = array
description = The engines used by the constructor.

tyreManufacturers = type = array
description = The tyre manufacturers used by the constructor.

drivers = type = ['array', 'null']
description = The drivers who drove for the constructor.


SeasonEntrantChassis = chassisId = type = string
description = The identifier of the chassis.


SeasonEntrantEngine = engineId = type = string
description = The identifier of the engine.


SeasonEntrantTyreManufacturer = tyreManufacturerId = type = string
description = The identifier of the tyre manufacturer.


SeasonEntrantDriver = driverId = type = string
description = The identifier of the driver.

rounds = type = ['array', 'null']
description = The rounds in which the driver participated.

roundsText = type = ['string', 'null']
description = The textual representation of the rounds in which the
driver participated.

testDriver = type = boolean
description = Whether the driver was a test / free practice driver.


SeasonConstructor = year = type = integer
description = The year of the season.

constructorId = type = string
description = The identifier of the constructor.
```

```
positionNumber = type = ['integer', 'null']
description = The numerical position of the constructor in the season
standings.

positionText = type = ['string', 'null']
description = The textual representation of the constructor's position
in the season standings, including special statuses.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the
constructor during the season.

bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the constructor
during the season.

totalRaceEntries = type = integer
description = The total number of races entered by the constructor
during the season.

totalRaceStarts = type = integer
description = The total number of races started by the constructor
during the season.

totalRaceWins = type = integer
description = The total number of races won by the constructor during
the season.

total1And2Finishes = type = integer
description = The total number of races in which the constructor
finished in both 1st and 2nd place during the season.

totalRaceLaps = type = integer
description = The total number of laps completed by the constructor
during the season.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the
constructor during the season.

totalPodiumRaces = type = integer
description = The total number of races in which the constructor
finished on the podium during the season.

totalPoints = type = number
description = The total number of points accumulated by the
constructor during the season.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the
```

```
constructor during the season.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the constructor
during the season.


SeasonEngineManufacturer = year = type = integer
description = The year of the season.

engineManufacturerId = type = string
description = The identifier of the engine manufacturer.

positionNumber = type = ['integer', 'null']
description = The numerical position of the engine manufacturer in the
season standings.

positionText = type = ['string', 'null']
description = The textual representation of the engine manufacturer's
position in the season standings, including special statuses.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the engine
manufacturer during the season.

bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the engine
manufacturer during the season.

totalRaceEntries = type = integer
description = The total number of races entered by the engine
manufacturer during the season.

totalRaceStarts = type = integer
description = The total number of races started by the engine
manufacturer during the season.

totalRaceWins = type = integer
description = The total number of races won by the engine manufacturer
during the season.

totalRaceLaps = type = integer
description = The total number of laps completed by the engine
manufacturer during the season.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the
engine manufacturer during the season.

totalPodiumRaces = type = integer
description = The total number of races in which the engine
```

```
manufacturer finished on the podium during the season.

totalPoints = type = number
description = The total number of points accumulated by the engine
manufacturer during the season.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the
engine manufacturer during the season.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the engine
manufacturer during the season.


SeasonTyreManufacturer = year = type = integer
description = The year of the season.

tyreManufacturerId = type = string
description = The identifier of the tyre manufacturer.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the tyre
manufacturer during the season.

bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the tyre
manufacturer during the season.

totalRaceEntries = type = integer
description = The total number of races entered by the tyre
manufacturer during the season.

totalRaceStarts = type = integer
description = The total number of races started by the tyre
manufacturer during the season.

totalRaceWins = type = integer
description = The total number of races won by the tyre manufacturer
during the season.

totalRaceLaps = type = integer
description = The total number of laps completed by the tyre
manufacturer during the season.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the tyre
manufacturer during the season.

totalPodiumRaces = type = integer
description = The total number of races in which the tyre manufacturer
```

```
finished on the podium during the season.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the tyre
manufacturer during the season.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the tyre
manufacturer during the season.


SeasonDriver = year = type = integer
description = The year of the season.

driverId = type = string
description = The identifier of the driver.

positionNumber = type = ['integer', 'null']
description = The numerical position of the driver in the season
standings.

positionText = type = ['string', 'null']
description = The textual representation of the driver's position in
the season standings, including special statuses.

bestStartingGridPosition = type = ['integer', 'null']
description = The best starting grid position achieved by the driver
during the season.

bestRaceResult = type = ['integer', 'null']
description = The best finishing position achieved by the driver
during the season.

totalRaceEntries = type = integer
description = The total number of races entered by the driver during
the season.

totalRaceStarts = type = integer
description = The total number of races started by the driver during
the season.

totalRaceWins = type = integer
description = The total number of races won by the driver during the
season.

totalRaceLaps = type = integer
description = The total number of laps completed by the driver during
the season.

totalPodiums = type = integer
description = The total number of podium finishes achieved by the
```

```
driver during the season.

totalPoints = type = number
description = The total number of points accumulated by the driver
during the season.

totalPolePositions = type = integer
description = The total number of pole positions achieved by the
driver during the season.

totalFastestLaps = type = integer
description = The total number of fastest laps set by the driver
during the season.

totalDriverOfTheDay = type = integer
description = The total number of Driver of the Day awards received by
the driver during the season.

totalGrandSlams = type = integer
description = The total number of Grand Slams achieved by the driver
during the season, defined as pole position, fastest lap, and leading
every lap of the race.


SeasonDriverStanding = positionDisplayOrder = type = integer
description = The display order of the driver's position in the
standings.

positionNumber = type = ['integer', 'null']
description = The numerical position of the driver in the standings.

positionText = type = string
description = The textual representation of the driver's position in
the standings, including special statuses.

driverId = type = string
description = The identifier of the driver.

points = type = number
description = The total number of points determining the driver's
position in the standings.


SeasonConstructorStanding = positionDisplayOrder = type = integer
description = The display order for the constructor's position in the
standings.

positionNumber = type = ['integer', 'null']
description = The numerical position of the constructor in the
standings.
```

```
positionText = type = string
description = The textual representation of the constructor's position
in the standings, including special statuses.

constructorId = type = string
description = The identifier of the constructor.

engineManufacturerId = type = string
description = The identifier of the engine manufacturer.

points = type = number
description = The total number of points determining the constructor's
position in the standings.


Race = id = type = integer
description = The unique identifier of the race.

year = type = integer
description = The year of the season.

round = type = integer
description = The round number of the race in the season.

date = type = string
description = The date of the race in UTC.

time = type = ['string', 'null']
description = The start time of the race in UTC.

grandPrixId = type = string
description = The identifier of the Grand Prix associated with the
race.

officialName = type = string
description = The official name of the race.

qualifyingFormat = description = The qualifying format of the race.

sprintQualifyingFormat = description = The sprint qualifying format of
the race.

circuitId = type = string
description = The identifier of the circuit where the race takes
place.

circuitType = description = The type of the circuit.

direction = description = The direction of the circuit.

courseLength = type = number
```

```
description = The length of the circuit (race course) in kilometers.

turns = type = integer
description = The number of turns (corners) in the configuration of
the circuit.

laps = type = integer
description = The total number of laps of the race.

distance = type = number
description = The total distance of the race in kilometers.

scheduledLaps = type = ['integer', 'null']
description = The scheduled number of laps of the race.

scheduledDistance = type = ['number', 'null']
description = The scheduled distance of the race in kilometers.

driversChampionshipDecider = type = boolean
description = Whether this race was the decider of the World Drivers'
Championship.

constructorsChampionshipDecider = type = boolean
description = Whether this race was the decider of the World
Constructors' Championship.

preQualifyingDate = type = ['string', 'null']
description = The date of the pre-qualifying session in UTC.

preQualifyingTime = type = ['string', 'null']
description = The start time of the pre-qualifying session in UTC.

preQualifyingResults = type = ['array', 'null']
description = The results of the pre-qualifying session.

freePractice1Date = type = ['string', 'null']
description = The date of the 1st free practice session in UTC.

freePractice1Time = type = ['string', 'null']
description = The start time of the 1st free practice session in UTC.

freePractice1Results = type = ['array', 'null']
description = The results of the 1st free practice session.

freePractice2Date = type = ['string', 'null']
description = The date of the 2nd free practice session in UTC.

freePractice2Time = type = ['string', 'null']
description = The start time of the 2nd free practice session in UTC.

freePractice2Results = type = ['array', 'null']
```

```
description = The results of the 2nd free practice session.

freePractice3Date = type = ['string', 'null']
description = The date of the 3rd free practice session in UTC.

freePractice3Time = type = ['string', 'null']
description = The start time of the 3rd free practice session in UTC.

freePractice3Results = type = ['array', 'null']
description = The results of the 3rd free practice session.

freePractice4Date = type = ['string', 'null']
description = The date of the 4th free practice session UTC.

freePractice4Time = type = ['string', 'null']
description = The start time of the 4th free practice session in UTC.

freePractice4Results = type = ['array', 'null']
description = The results of the 4th free practice session.

qualifying1Date = type = ['string', 'null']
description = The date of the 1st qualifying session in UTC.

qualifying1Time = type = ['string', 'null']
description = The start time of the 1st qualifying session in UTC.

qualifying1Results = type = ['array', 'null']
description = The results of the 1st qualifying session.

qualifying2Date = type = ['string', 'null']
description = The date of the 2nd qualifying session UTC.

qualifying2Time = type = ['string', 'null']
description = The start time of the 2nd qualifying session UTC.

qualifying2Results = type = ['array', 'null']
description = The results of the 2nd qualifying session.

qualifyingDate = type = ['string', 'null']
description = The date of the qualifying session UTC.

qualifyingTime = type = ['string', 'null']
description = The start time of the qualifying session UTC.

qualifyingResults = type = ['array', 'null']
description = The results of the qualifying session.

sprintQualifyingDate = type = ['string', 'null']
description = The date of the sprint qualifying session in UTC.

sprintQualifyingTime = type = ['string', 'null']
```

```
description = The start time of the sprint qualifying session in UTC.

sprintQualifyingResults = type = ['array', 'null']
description = The results of the sprint qualifying session.

sprintStartingGridPositions = type = ['array', 'null']
description = The starting grid positions for the sprint race.

sprintRaceDate = type = ['string', 'null']
description = The date of the sprint race in UTC.

sprintRaceTime = type = ['string', 'null']
description = The start time of the sprint race in UTC.

sprintRaceResults = type = ['array', 'null']
description = The results of the sprint race.

warmingUpDate = type = ['string', 'null']
description = The date of the warming-up session in UTC.

warmingUpTime = type = ['string', 'null']
description = The start time of the warming-up session in UTC.

warmingUpResults = type = ['array', 'null']
description = The results of the warming-up session.

startingGridPositions = type = ['array', 'null']
description = The starting grid positions for the race.

raceResults = type = ['array', 'null']
description = The results of the race.

fastestLaps = type = ['array', 'null']
description = The fastest laps recorded during the race..

pitStops = type = ['array', 'null']
description = The pit stops made during the race.

driverOfTheDayResults = type = ['array', 'null']
description = The results of the Driver of the Day vote.

driverStandings = type = ['array', 'null']
description = The driver standings after the race.

constructorStandings = type = ['array', 'null']
description = The constructor standings after the race.


RaceDriverStanding = positionDisplayOrder = type = integer
description = The display order of the driver's position in the
standings.
```

```
positionNumber = type = ['integer', 'null']
description = The numerical position of the driver in the standings.

positionText = type = string
description = The textual representation of the driver's position in
the standings, including special statuses.

driverId = type = string
description = The identifier of the driver.

points = type = number
description = The total number of points determining the driver's
position in the standings.

positionsGained = type = ['integer', 'null']
description = The positions gained in the standings since the previous
race.


RaceConstructorStanding = positionDisplayOrder = type = integer
description = The display order for the constructor's position in the
standings.

positionNumber = type = ['integer', 'null']
description = The numerical position of the constructor in the
standings.

positionText = type = string
description = The textual representation of the constructor's position
in the standings, including special statuses.

constructorId = type = string
description = The identifier of the constructor.

engineManufacturerId = type = string
description = The identifier of the engine manufacturer.

points = type = number
description = The points.

positionsGained = type = ['integer', 'null']
description = The total number of points determining the constructor's
position in the standings.



race.info(verbose=True, memory_usage='deep', show_counts=True)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1149 entries, 0 to 1148
Data columns (total 42 columns):
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   id                                    1149 non-null   int64
 1   year                                  1149 non-null   int64
 2   round                                 1149 non-null   int64
 3   date                                  1149 non-null   object
 4   time                                  48 non-null     object
 5   grand_prix_id                         1149 non-null   object
 6   official_name                         1149 non-null   object
 7   qualifying_format                     1149 non-null   object
 8   sprint_qualifying_format              18 non-null     object
 9   circuit_id                            1149 non-null   object
 10  circuit_type                          1149 non-null   object
 11  direction                             1149 non-null   object
 12  course_length                         1149 non-null   float64
 13  turns                                 1149 non-null   int64
 14  laps                                  1149 non-null   int64
 15  distance                              1149 non-null   float64
 16  scheduled_laps                        80 non-null     float64
 17  scheduled_distance                    80 non-null     float64
 18  drivers_championship_decider          1149 non-null   bool
 19  constructors_championship_decider     1149 non-null   bool
 20  pre_qualifying_date                   0 non-null      object
 21  pre_qualifying_time                   0 non-null      object
 22  free_practice_1_date                  48 non-null     object
 23  free_practice_1_time                  48 non-null     object
 24  free_practice_2_date                  36 non-null     object
 25  free_practice_2_time                  36 non-null     object
 26  free_practice_3_date                  36 non-null     object
 27  free_practice_3_time                  36 non-null     object
 28  free_practice_4_date                  0 non-null      object
 29  free_practice_4_time                  0 non-null      object
 30  qualifying_1_date                     0 non-null      object
 31  qualifying_1_time                     0 non-null      object
 32  qualifying_2_date                     0 non-null      object
 33  qualifying_2_time                     0 non-null      object
 34  qualifying_date                       48 non-null     object
 35  qualifying_time                       48 non-null     object
 36  sprint_qualifying_date                12 non-null     object
 37  sprint_qualifying_time                12 non-null     object
 38  sprint_race_date                      12 non-null     object
 39  sprint_race_time                      12 non-null     object
 40  warming_up_date                       0 non-null      object
 41  warming_up_time                       0 non-null      object
dtypes: bool(2), float64(4), int64(5), object(31)
memory usage: 1.2 MB
```

```
race_data.info(verbose=True, memory_usage='deep', show_counts=True)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 183632 entries, 0 to 183631
Data columns (total 71 columns):
 #   Column                                                Non-Null
Count   Dtype
---  ------
--------------   -----
 0   race_id                                               183632 non-
null  int64
 1   type                                                  183632 non-
null  object
 2   position_display_order                                183632 non-
null  int64
 3   position_number                                       172468 non-
null  float64
 4   position_text                                         183632 non-
null  object
 5   driver_number                                         183632 non-
null  object
 6   driver_id                                             183632 non-
null  object
 7   constructor_id                                        183632 non-
null  object
 8   engine_manufacturer_id                                183632 non-
null  object
 9   tyre_manufacturer_id                                  183632 non-
null  object
 10  practice_time                                         47260 non-
null   object
 11  practice_time_millis                                  47260 non-
null   float64
 12  practice_gap                                          45124 non-
null   object
 13  practice_gap_millis                                   45124 non-
null   float64
 14  practice_interval                                     45124 non-
null   object
 15  practice_interval_millis                              45124 non-
null   float64
 16  practice_laps                                         38322 non-
null   float64
 17  qualifying_time                                       33926 non-
null   object
 18  qualifying_time_millis                                33926 non-
null   float64
 19  qualifying_q1                                         8470 non-
null    object
 20  qualifying_q1_millis                                  8470 non-
```

```
null    float64
 21  qualifying_q2                                          6216 non-
null    object
 22  qualifying_q2_millis                                   6216 non-
null    float64
 23  qualifying_q3                                          3952 non-
null    object
 24  qualifying_q3_millis                                   3952 non-
null    float64
 25  qualifying_gap                                         36049 non-
null   object
 26  qualifying_gap_millis                                  36049 non-
null   float64
 27  qualifying_interval                                    36036 non-
null   object
 28  qualifying_interval_millis                             36036 non-
null   float64
 29  qualifying_laps                                        17016 non-
null   float64
 30  starting_grid_position_qualification_position_number  25680 non-
null   float64
 31  starting_grid_position_qualification_position_text    25809 non-
null   object
 32  starting_grid_position_grid_penalty                    573 non-
null     object
 33  starting_grid_position_grid_penalty_positions          500 non-
null     float64
 34  starting_grid_position_time                            25258 non-
null   object
 35  starting_grid_position_time_millis                     25258 non-
null   float64
 36  race_shared_car                                        27591 non-
null   object
 37  race_laps                                              25664 non-
null   float64
 38  race_time                                              8318 non-
null    object
 39  race_time_millis                                       8318 non-
null    float64
 40  race_time_penalty                                      274 non-
null     object
 41  race_time_penalty_millis                               274 non-
null     float64
 42  race_gap                                               14822 non-
null   object
 43  race_gap_millis                                        7154 non-
null    float64
 44  race_gap_laps                                          7668 non-
null    float64
```

```
 45  race_interval                                 7136 non-
null    object
 46  race_interval_millis                          7136 non-
null    float64
 47  race_reason_retired                           9998 non-
null    object
 48  race_points                                   8505 non-
null    float64
 49  race_pole_position                            27591 non-
null   object
 50  race_qualification_position_number            26872 non-
null   float64
 51  race_qualification_position_text              27009 non-
null   object
 52  race_grid_position_number                     25584 non-
null   float64
 53  race_grid_position_text                       25815 non-
null   object
 54  race_positions_gained                         16626 non-
null   float64
 55  race_pit_stops                                12676 non-
null   float64
 56  race_fastest_lap                              27571 non-
null   object
 57  race_driver_of_the_day                        4601 non-
null    object
 58  race_grand_slam                               27591 non-
null   object
 59  fastest_lap_lap                               16689 non-
null   float64
 60  fastest_lap_time                              16736 non-
null   object
 61  fastest_lap_time_millis                       16736 non-
null   float64
 62  fastest_lap_gap                               15593 non-
null   object
 63  fastest_lap_gap_millis                        15593 non-
null   float64
 64  fastest_lap_interval                          15593 non-
null   object
 65  fastest_lap_interval_millis                   15593 non-
null   float64
 66  pit_stop_stop                                 21889 non-
null   float64
 67  pit_stop_lap                                  21889 non-
null   float64
 68  pit_stop_time                                 21888 non-
null   object
 69  pit_stop_time_millis                          21888 non-
```

```
null   float64
 70  driver_of_the_day_percentage                              720 non-
null     float64
dtypes: float64(34), int64(2), object(35)
memory usage: 253.4 MB
```

```
circuit.info(verbose=True, memory_usage='deep', show_counts=True)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 77 entries, 0 to 76
Data columns (total 13 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   id                77 non-null     object
 1   name              77 non-null     object
 2   full_name         77 non-null     object
 3   previous_names    15 non-null     object
 4   type              77 non-null     object
 5   direction         77 non-null     object
 6   place_name        77 non-null     object
 7   country_id        77 non-null     object
 8   latitude          77 non-null     float64
 9   longitude         77 non-null     float64
 10  length            77 non-null     float64
 11  turns             77 non-null     int64
 12  total_races_held  77 non-null     int64
dtypes: float64(3), int64(2), object(8)
memory usage: 37.4 KB
```

```
circuit.id.unique()
```

```
array(['adelaide', 'aida', 'ain-diab', 'aintree', 'anderstorp',
'austin',
       'avus', 'bahrain', 'baku', 'brands-hatch', 'bremgarten',
'buddh',
       'buenos-aires', 'bugatti', 'caesars-palace', 'catalunya',
       'clermont-ferrand', 'dallas', 'detroit', 'dijon', 'donington',
       'east-london', 'estoril', 'fuji', 'hockenheimring',
'hungaroring',
       'imola', 'indianapolis', 'interlagos', 'istanbul',
'jacarepagua',
       'jarama', 'jeddah', 'jerez', 'kyalami', 'las-vegas', 'long-
beach',
       'lusail', 'magny-cours', 'marina-bay', 'melbourne', 'mexico-
city',
       'miami', 'monaco', 'monsanto', 'mont-tremblant', 'montjuic',
       'montreal', 'monza', 'mosport', 'mugello', 'nivelles',
       'nurburgring', 'paul-ricard', 'pedralbes', 'pescara',
'phoenix',
       'portimao', 'porto', 'reims', 'riverside', 'rouen', 'sebring',
```

```
       'sepang', 'shanghai', 'silverstone', 'sochi', 'spa-
francorchamps',
       'spielberg', 'suzuka', 'valencia', 'watkins-glen', 'yas-
marina',
       'yeongam', 'zandvoort', 'zeltweg', 'zolder'], dtype=object)

circuit.sort_values(by='total_races_held', ascending=False)
[['id','total_races_held']].head()

# Circuits that have been repetitively raced on will show the most
conclusive results
```

	id	total_races_held
48	monza	75
43	monaco	71
65	silverstone	60
67	spa-francorchamps	58
47	montreal	44

Initial Notes:

- constructors_championship_decider is used for postgres functionality, and therefore can be dropped from the dataframe for analysis

Circuit.previous_name is arbitrary info and can be dropped from the dataframe Circuit.name and index are the same values, names can be dropped Circuit.total_races_held can be simplified to be called 'races_count'

```
circuit =
circuit.drop(columns=['name','previous_names','direction','latitude','
longitude'])
circuit = circuit.rename(columns = {
    'total_races_held' : 'race_count'
})
```

Preparing the Data

```
tracks = ['monza', 'monaco', 'silverstone', 'spa-francorchamps',
'montreal'] # Circuits of interest; have the most amount of racing
data to analyze

race[['date']] = race[['date']].apply(pd.to_datetime, format="%Y-%m-
%d")


def query_circuit(circuit_name: str, agg_method : str = 'min') ->
pd.DataFrame:

    selected_data = (race.query(f'circuit_id == "{circuit_name}"')
                     .rename(columns={'id' : 'race_id'}))
```

```
# Change id name to match with race_data for merge

    selected_data = (
        selected_data
            .drop(columns=['time','grand_prix_id',
'circuit_id' ,'circuit_type', 'direction'])      # Remove redundant
information (same for all rows)
            .dropna(axis=1, how='all')
# Remove non-applicable information (null for all rows)

            .merge(right=race_data, on='race_id', how='left')
# Merge selected data and race_data to obtain lap time data
            .rename(columns={'fastest_lap_time_millis' :
'fastest_time'})

            .groupby('race_id')[['fastest_time']]
            .agg(agg_method)
# apply aggregation

            .merge(right=selected_data[['race_id', 'date',
'circuit_id']].copy(), on='race_id', how='left' )      # Merge
selected data back to maintain date column
            .drop(columns='race_id')
        )

    selected_data['fastest_time'] = selected_data['fastest_time'] /
60000                           # Convert to minutes
    selected_data = selected_data.sort_values(by='date')
    # selected_data =
selected_data.set_index('date').resample('YE').agg(agg_method).reset_i
ndex()
    selected_data['record'] =
selected_data['fastest_time'].expanding().min()
# Calculate running minimum

    selected_data = selected_data.melt(id_vars=['circuit_id', 'date'],
value_vars=['fastest_time','record'])        # melt data for plotting

    return selected_data

lap_data_min = pd.concat([query_circuit(table) for table in tracks])
# lap_data_avg = pd.concat([query_circuit(table, 'mean') for table in
tracks])

lap_data_min.head()
```

	circuit_id	date	variable	value
0	monza	1950-09-03	fastest_time	2.000000
1	monza	1951-09-16	fastest_time	1.941667
2	monza	1952-09-07	fastest_time	2.101667

```
3      monza 1953-09-13  fastest_time  2.075000
4      monza 1954-09-05  fastest_time  2.013333
```

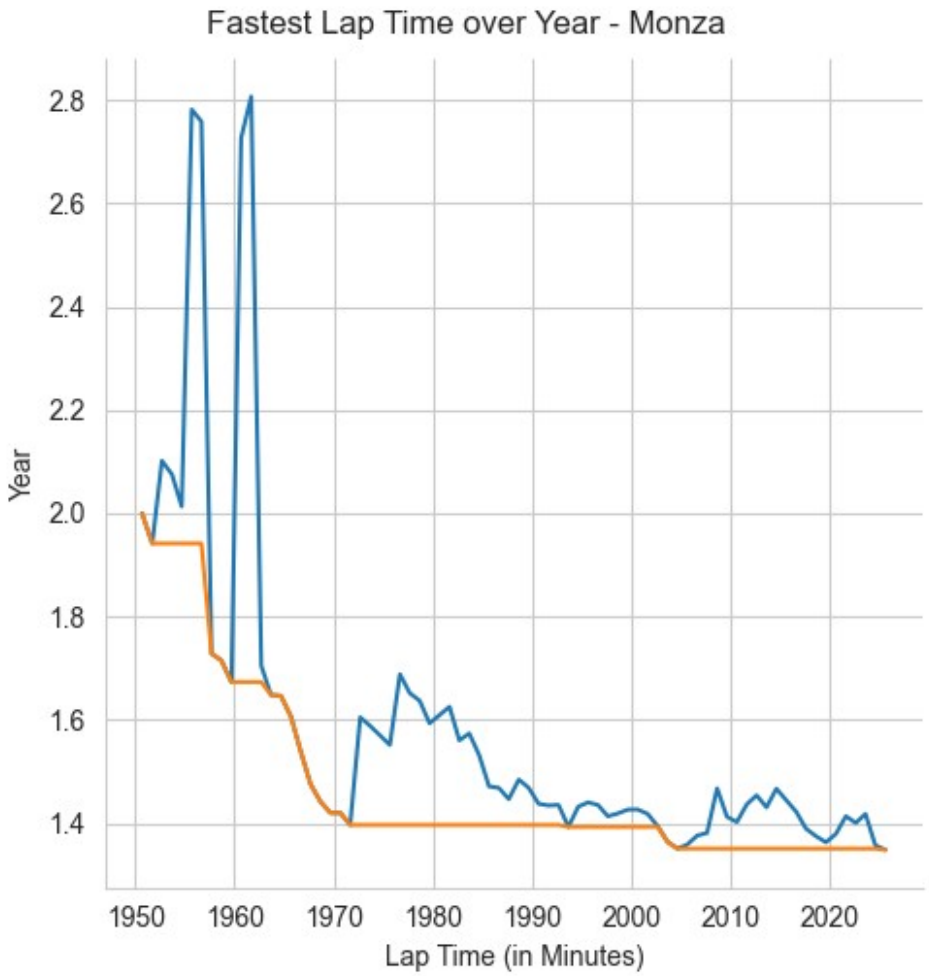
```
# lap_data_avg.head()

data_percent_change = (
    lap_data_min
    .query('variable == "record"')
    .set_index(['circuit_id', 'date'])
    .groupby(level=0)['value']
    .pct_change()                          # Apply percentage
    .apply(lambda x: -1 * x)               # Reverse so all data is
positive
    .to_frame()
    .reset_index()
    .set_index('date',)
)

data_percent_change.head()

df = []

# Downsize data to yearly for clarity

for track in tracks:
    temp = data_percent_change.query(f'circuit_id == "{track}"')
    temp = temp.value.resample('YS').mean().to_frame()
    temp['circuit_id'] = track
    temp = temp.reset_index()
    df.append(temp)

data_percent_change = pd.concat(df) # Recompile into a single
DataFrame (so graphing functions can be used)
data_percent_change.head()
```

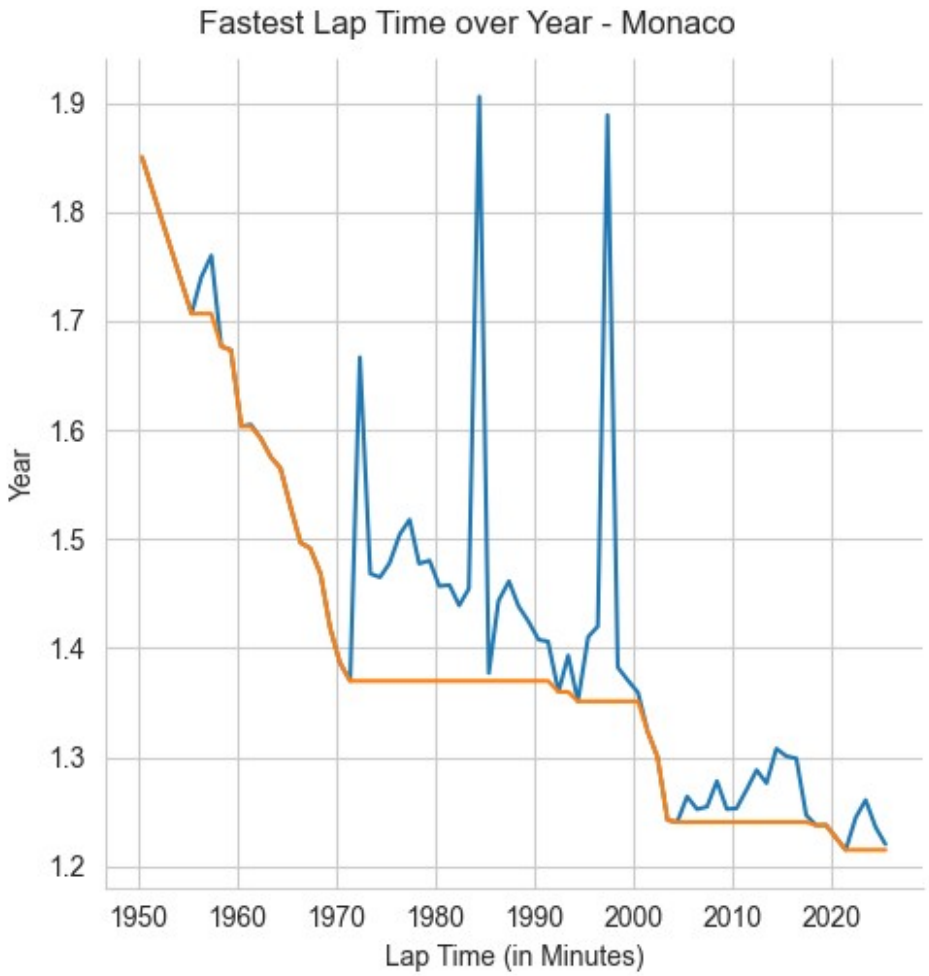
```
        date     value circuit_id
0 1950-01-01       NaN      monza
1 1951-01-01  0.029167      monza
2 1952-01-01  0.000000      monza
3 1953-01-01  0.000000      monza
4 1954-01-01  0.000000      monza
```

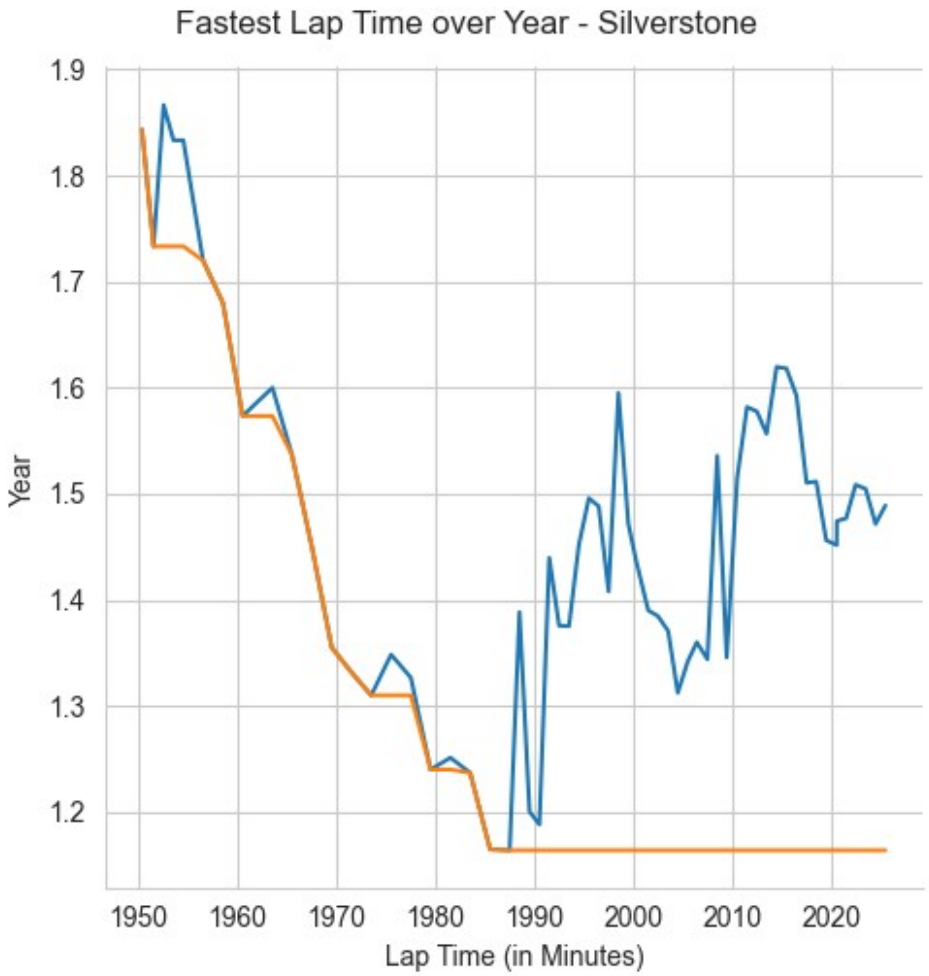
Plotting the data

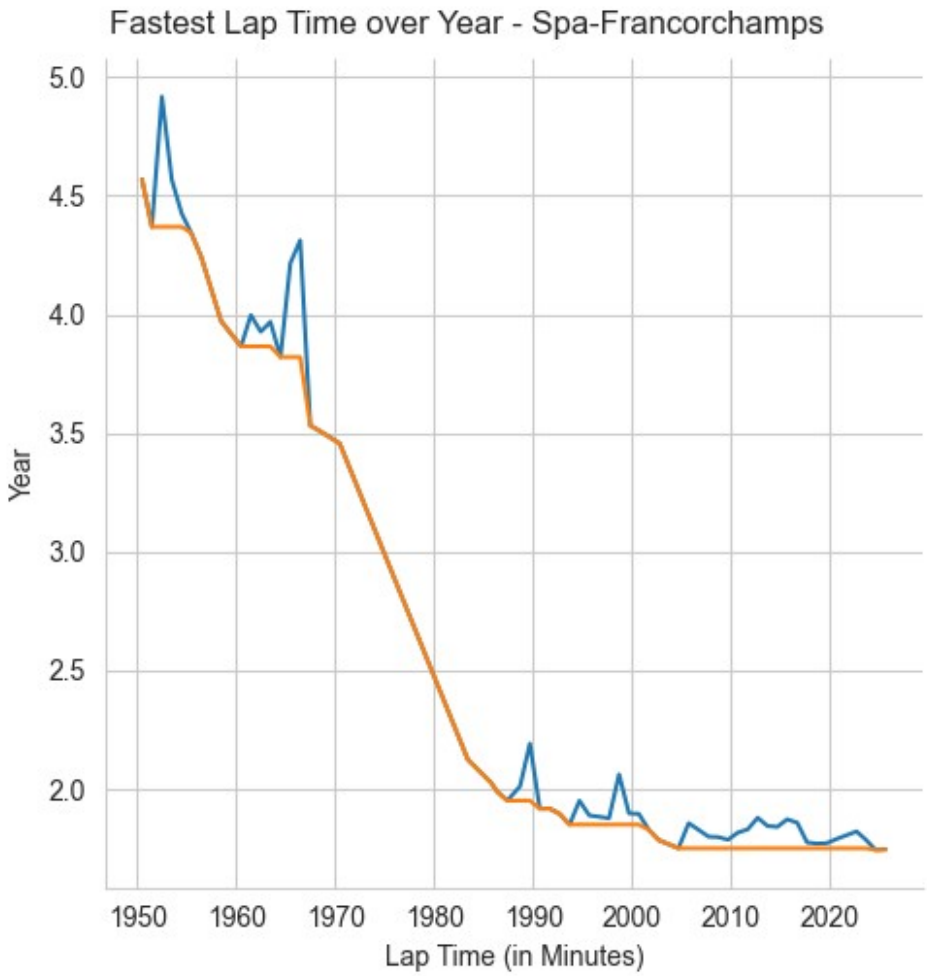
```
def create_plot ( data_set : pd.DataFrame, super_title : str ) ->
sns.FacetGrid:
    # Create single figure with multiple plots
    g = sns.relplot(data=data_set,
                kind="line", x='date', y='value', hue='variable',
                col='circuit_id', facet_kws={ 'sharex' : False,
```

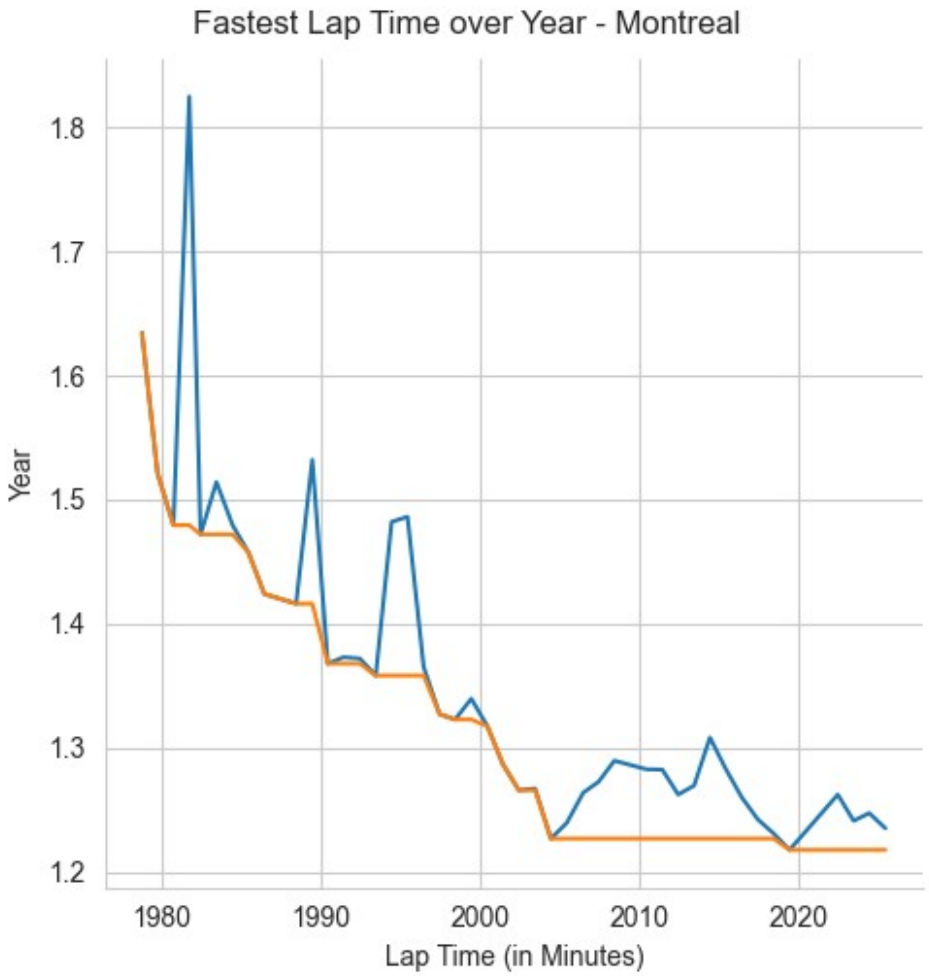
```
'sharey' : False }
    )
    g.set_axis_labels('Year', 'Lap Time (in minutes)')
    g.figure.suptitle(super_title, y = 1.02)

    for ax in g.axes.flat:
            ax.set_title(ax.get_title()[13:].title())
    return g

path = 'charts/circuit_data/'
extension = '.png'

def save_plot (
        data_set : pd.DataFrame,
        super_title : str,
        subfolder : str = '',
        x_label : str = "Lap Time (in Minutes)",
        y_label : str = "Year",
        y_lim : float = None,
):
    # Save individual plots to directory

    sns.set_style('whitegrid')

    for cir in lap_data_min.circuit_id.unique().tolist():

        g = sns.relplot(
            data=data_set.query(f'circuit_id == "{cir}"'),
            kind='line',
            x='date',
            y='value',
            hue= 'variable' if 'variable' in data_set.columns else
None,
            # legend=False      # Used for presentation
        )

        g.figure.suptitle(f'{super_title} - {cir.title()}', y = 1.02)

        g.set_axis_labels(
            x_label,
            y_label
        )

        for ax in g.axes.flat:
            ax.set(ylim = (0, y_lim) if y_lim is not None else None)

        file = path + subfolder + cir + extension
        g.savefig(file)

save_plot(lap_data_min, "Fastest Lap Time over Year")
```


Fastest Lap Time over Year - Monza
Year
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1950
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)


Fastest Lap Time over Year - Monaco
Year
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1950
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)

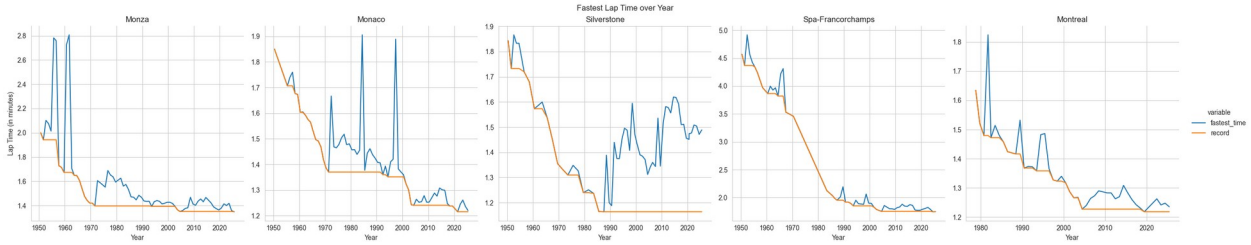

Fastest Lap Time over Year - Silverstone
Year
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1950
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)


Fastest Lap Time over Year - Spa-Francorchamps
5.0
4.5
4.0
3.5
3.0
2.5
2.0
Year
1950
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)

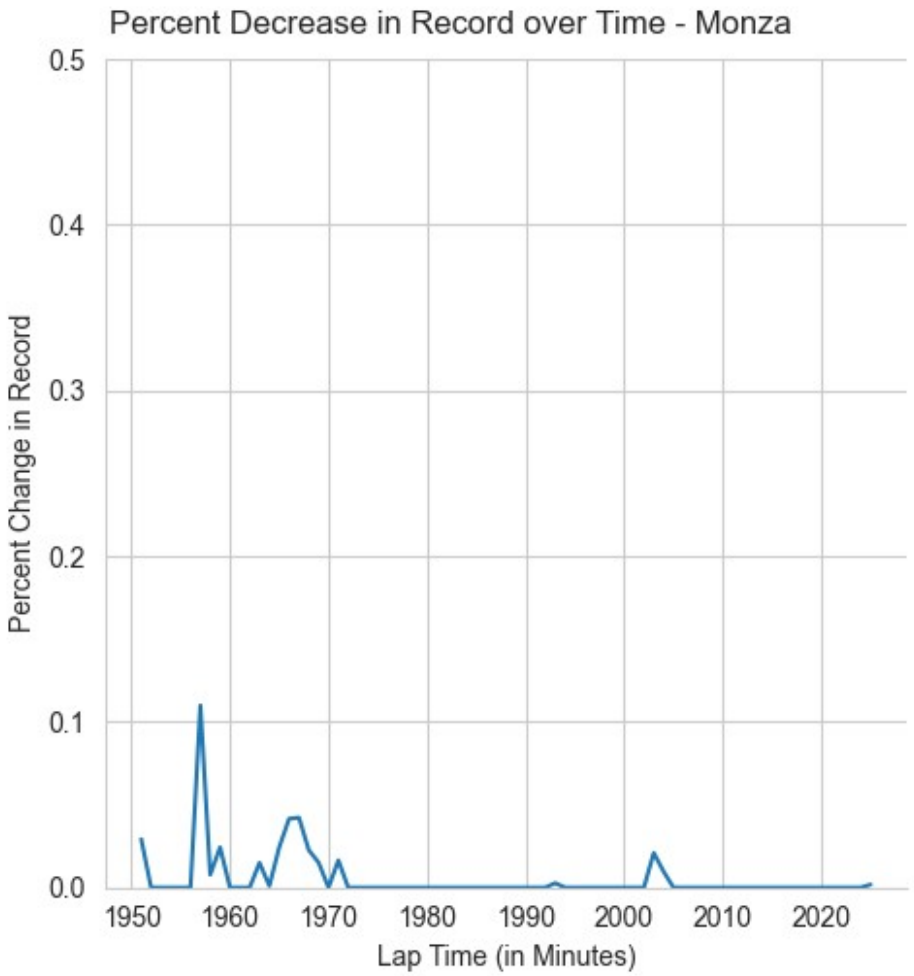


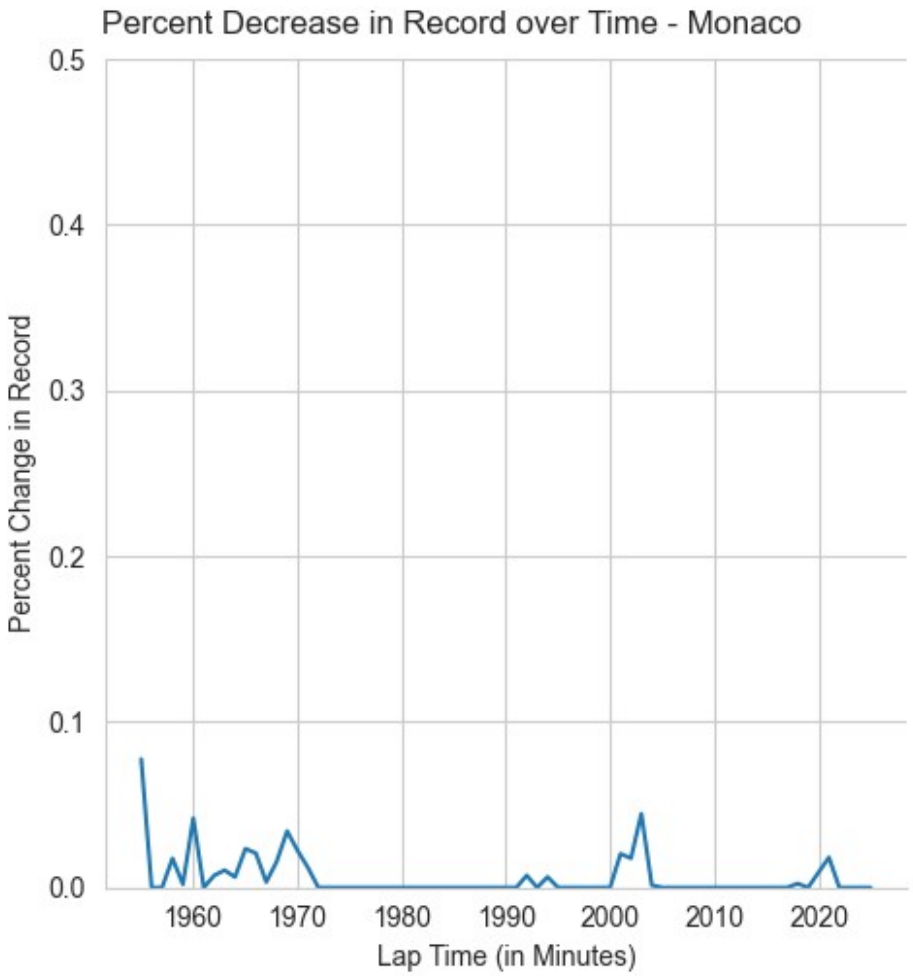

```
create_plot(lap_data_min, "Fastest Lap Time over Year")
```

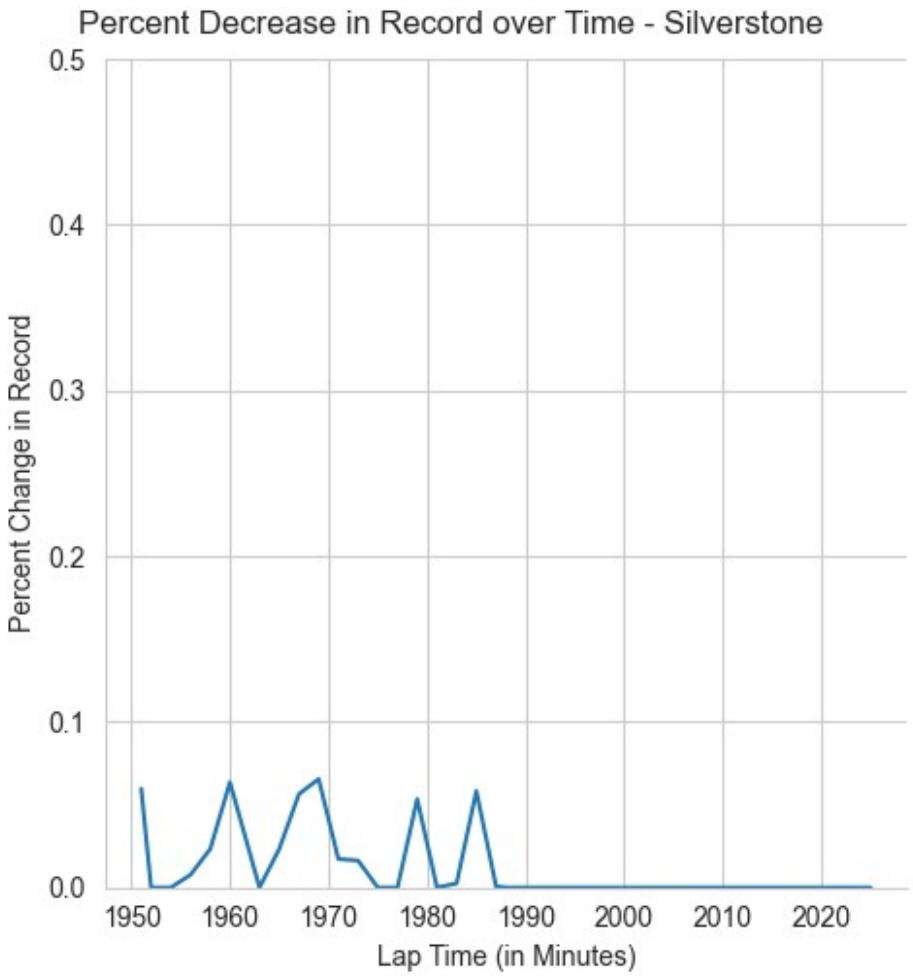
```
<seaborn.axisgrid.FacetGrid at 0x16b3a3750>
```

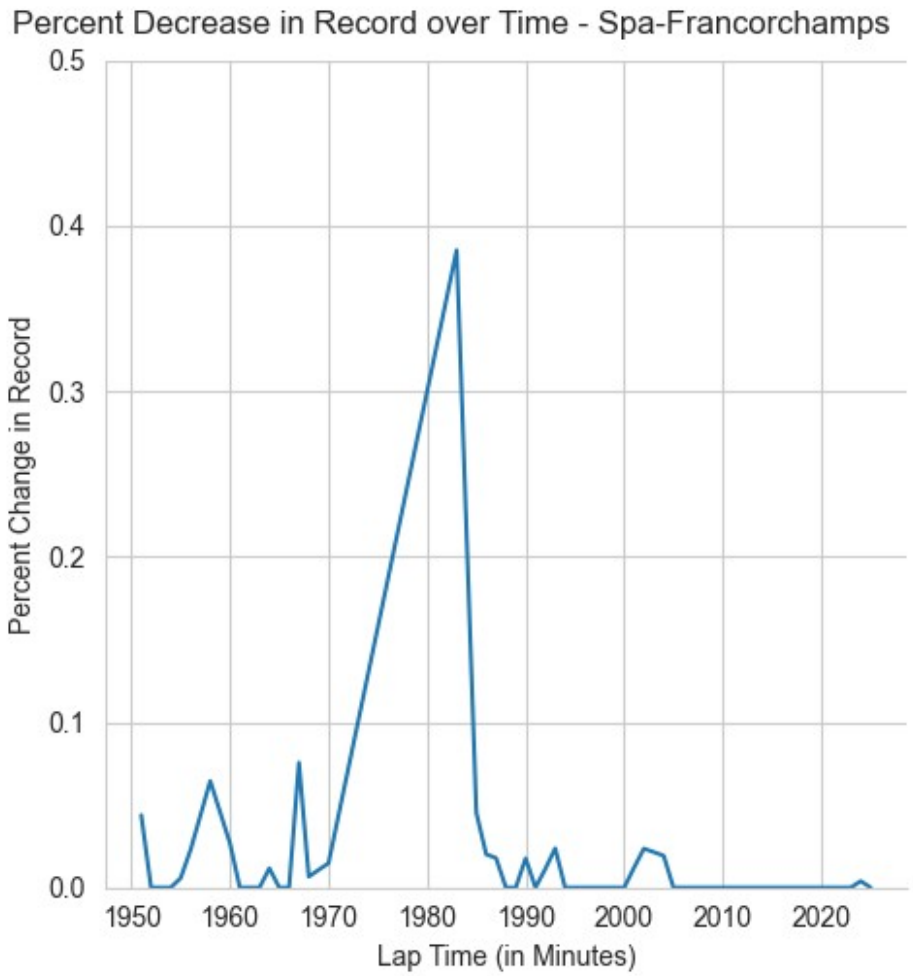


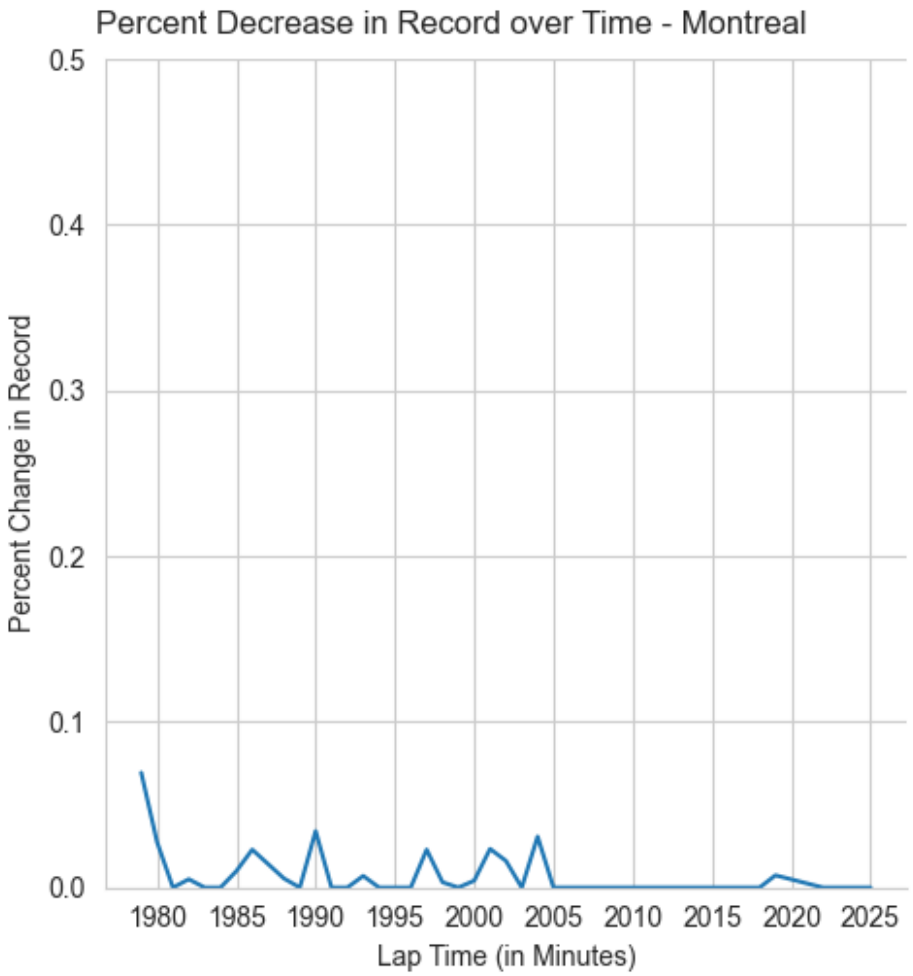

```
save_plot(
    data_set=data_percent_change,
    super_title="Percent Decrease in Record over Time",
    subfolder='pct_change/',
    y_label='Percent Change in Record',
    y_lim = .5
)
```


Percent Decrease in Record over Time - Monza
Percent Change in Record
0.5
0.4
0.3
0.2
0.1
0.0
1950
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)


Percent Decrease in Record over Time - Monaco
Percent Change in Record
0.5
0.4
0.3
0.2
0.1
0.0
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)


Percent Decrease in Record over Time - Silverstone
Percent Change in Record
0.5
0.4
0.3
0.2
0.1
0.0
1950
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)


Percent Decrease in Record over Time - Spa-Francorchamps
Percent Change in Record
0.5
0.4
0.3
0.2
0.1
0.0
1950
1960
1970
1980
1990
2000
2010
2020
Lap Time (in Minutes)


Percent Decrease in Record over Time - Montreal
Percent Change in Record
0.5
0.4
0.3
0.2
0.1
0.0
1980
1985
1990
1995
2000
2005
2010
2015
2020
2025
Lap Time (in Minutes)