

F1WP – Formula 1 Winner Prediction

Progetto per il corso di Fondamenti di Intelligenza Artificiale A.A.

2021/2022 di Giorgio Angelo Esposito, matricola 0512107389

<https://github.com/giorgio-angelo-esposito/F1WP.git>

1. INTRODUZIONE

La Formula Uno è uno sport automobilistico nato ufficialmente nel 1950 ed è attualmente lo sport automobilistico di più alta categoria per quanto riguarda le vetture monoposto a ruota scoperta da corsa su circuito.

Il termine “*Formula*” si riferisce all’insieme di regole che i partecipanti (team e piloti) devono rispettare.

Nel tempo le vetture si sono evolute molto diventando, nell’ultimo decennio, l’apice della tecnologia nelle corse automobilistiche.

Una gara di Formula Uno non comprende però, solo la gara stessa: un *Gran Premio* occupa un intero *weekend*: si inizia il giovedì con le interviste ai piloti, prove libere il venerdì e il sabato mattina, qualifiche il sabato pomeriggio e gara la domenica.

Data la natura estremamente dinamica dello sport, fatta di sorpassi, incidenti e guasti, viene naturale provare a prevedere chi tra i partecipanti alla gara sarà il vincitore, cosa non sempre scontata.

2. DESCRIZIONE DELL’AGENTE

L’obiettivo del progetto è quindi quello di realizzare un agente capace di determinare quale pilota vincerà un Gran Premio.

Andiamo ora a definire la misura PEAS (**P**erformance, **E**nvironment, **A**ctuators, **S**ensors) dell’agente:

TABELLA DELLA MISURA PEAS	
PERFORMANCE	La misura di performance dell'agente è la sua capacità di predire correttamente il vincitore di un Gran Premio di Formula Uno
ENVIRONMENT	L'ambiente in cui opera l'agente è: <ul style="list-style-type: none"> • OSSERVABILE: l'agente ha sempre accesso a tutti i dati che ha a disposizione • DISCRETO: l'agente ha un numero limitato di dati da cui apprendere • AGENTE SINGOLO • STATICO: l'ambiente non cambia mentre l'agente sta apprendendo
ACTUATORS	L'attuatore dell'agente corrisponde alla predizione effettuata
SENSORS	I sensori dell'agente sono l'insieme di dati passati in input al machine learner per apprendere

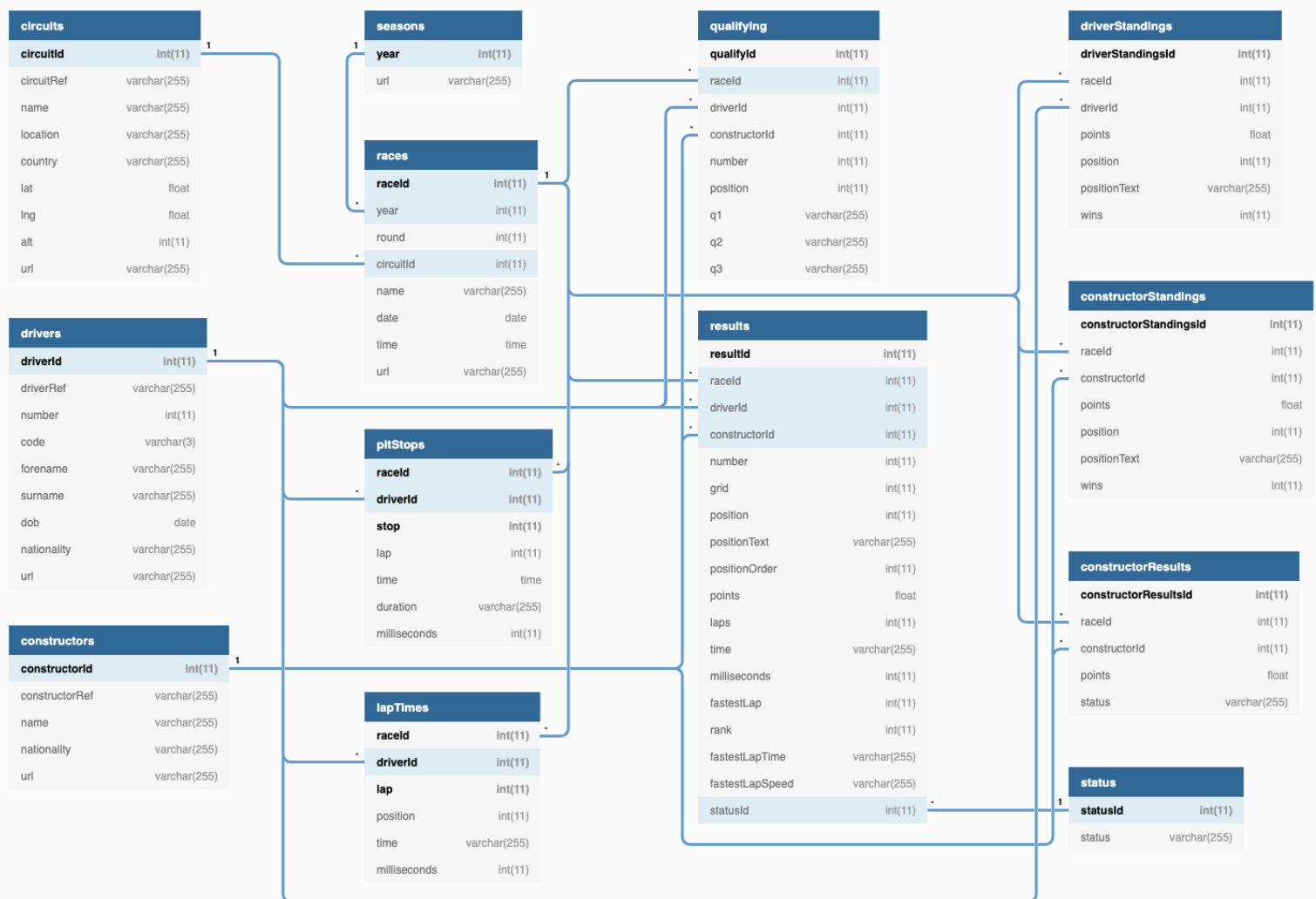
3. RACCOLTA E ANALISI DEI DATI

3.1 SCELTA DEL DATASET

Passiamo ora al dataset che andremo a utilizzare per i nostri scopi. Dopo aver cercato sui vari siti dedicati (Kaggle, Google Dataset Search, ecc...), la scelta è ricaduta su un *web service* chiamato Ergast Developer API (<http://ergast.com/mrd/>). Ergast fornisce dati storici relativi alle corse automobilistiche e, oltre a poter interrogare il database mediante query parametriche, permette di scaricare i dati in formato .csv (comma separated value).

Trattandosi di un database relazionale, avremo diverse tabelle su cui eseguire le tecniche di Data Analysis.

Di seguito, la struttura del database.



Procediamo quindi ad importare i vari file .csv e a descrivere i dati che contengono: per importare i file in Colab utilizziamo il metodo `upload()` di `files`, nativo di Colab.

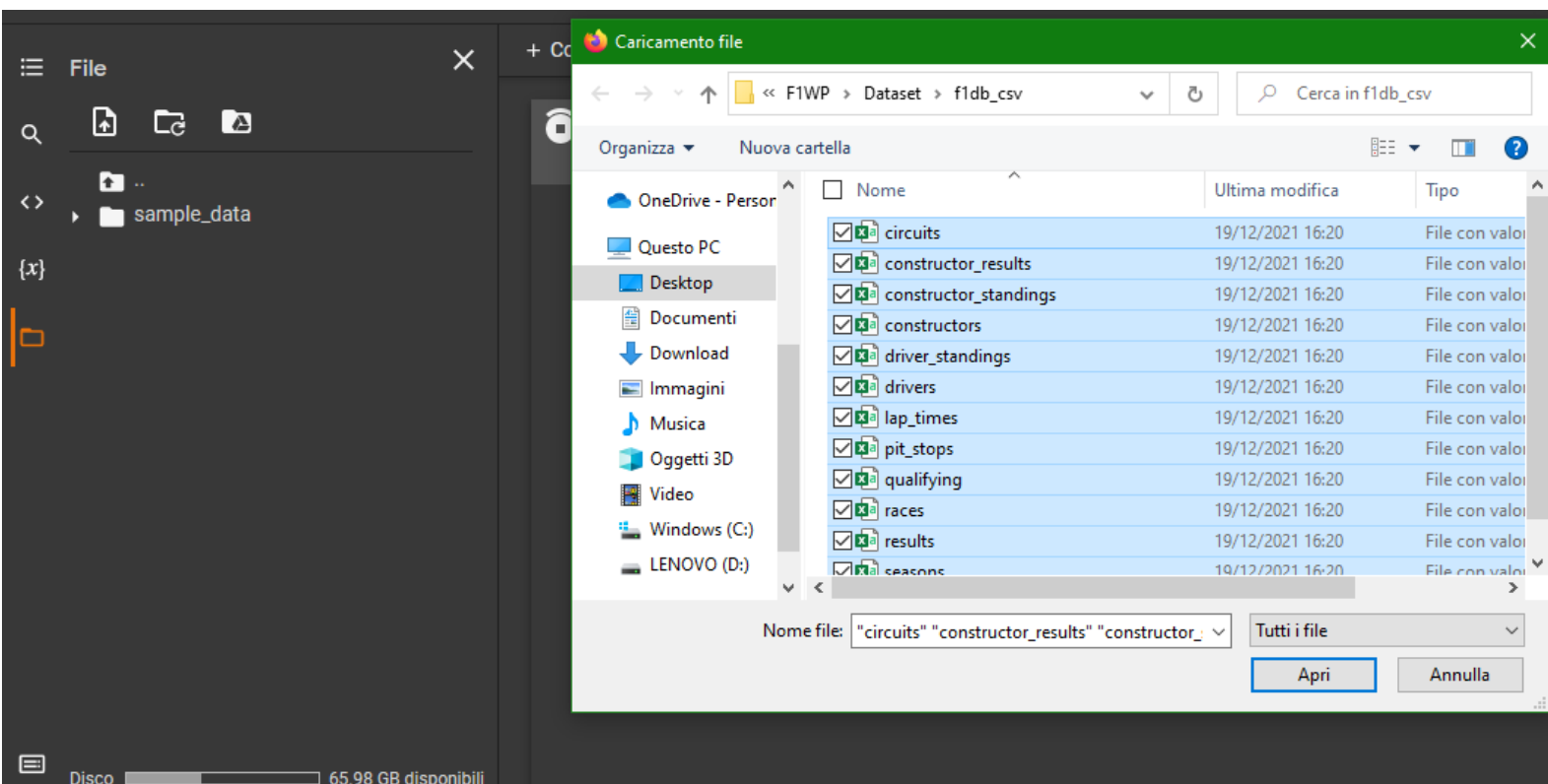
▼ Import del dataset

```
from google.colab import files
file = files.upload() #il metodo upload() permette di caricare un file in Colab
```

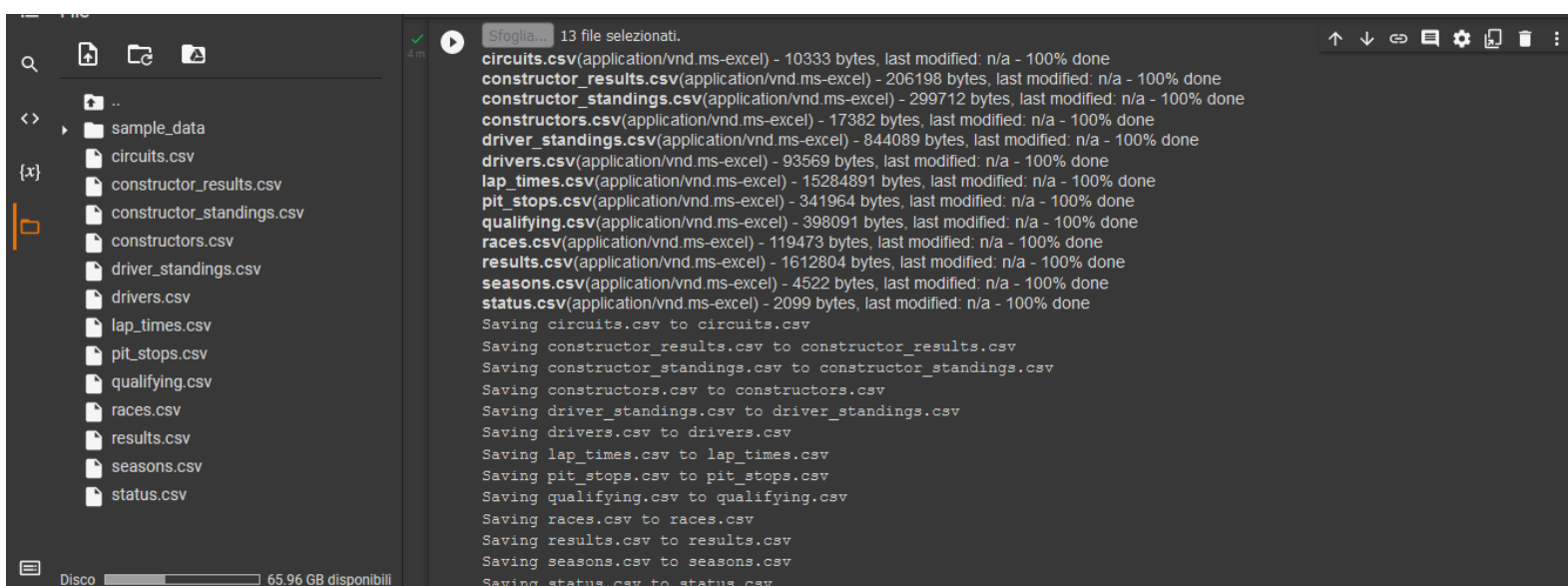
Sfogliala... Nessun file selezionato.

Cancel upload

Selezioniamo quindi i file.



Infine, i file sono importati correttamente.



Procediamo ora ad analizzare tabella per tabella, esaminando i campi in esse contenuti, in maniera tale da scegliere quali sono utili ai nostri scopi.

3.2 DESCRIZIONE TABELLE

Ergast fornisce una descrizione delle tabelle che compongono il suo database.

TABELLA CIRCUITS

circuits table							
Field	Type	Null	Key	Default	Extra	Description	
circuitId	int(11)	NO	PRI	NULL	auto_increment	Primary key	
circuitRef	varchar(255)	NO				Unique circuit identifier	
name	varchar(255)	NO				Circuit name	
location	varchar(255)	YES		NULL		Location name	
country	varchar(255)	YES		NULL		Country name	
lat	float	YES		NULL		Latitude	
lng	float	YES		NULL		Longitude	
alt	int(11)	YES		NULL		Altitude (metres)	
url	varchar(255)	NO	UNI			Circuit Wikipedia page	

La tabella `circuits` riporta le informazioni sui circuiti.

TABELLA CONSTRUCTOR

constructors table							
Field	Type	Null	Key	Default	Extra	Description	
constructorId	int(11)	NO	PRI	NULL	auto_increment	Primary key	
constructorRef	varchar(255)	NO				Unique constructor identifier	
name	varchar(255)	NO	UNI			Constructor name	
nationality	varchar(255)	YES		NULL		Constructor nationality	
url	varchar(255)	NO				Constructor Wikipedia page	

La tabella `constructor` riporta le informazioni sui costruttori, ovvero i team.

TABELLA CONSTRUCTOR RESULTS

constructor_results table							
Field	Type	Null	Key	Default	Extra	Description	
constructorResultsId	int(11)	NO	PRI	NULL	auto_increment	Primary key	
raceId	int(11)	NO		0		Foreign key link to races table	
constructorId	int(11)	NO		0		Foreign key link to constructors table	
points	float	YES		NULL		Constructor points for race	
status	varchar(255)	YES		NULL		"D" for disqualified (or null)	

La tabella `constructor_results` riporta le informazioni sui risultati ottenuti dai costruttori nelle gare.

TABELLA CONSTRUCTOR STANDINGS

```
constructor_standings table
```

Field	Type	Null	Key	Default	Extra	Description
constructorStandingsId	int(11)	NO	PRI	NULL	auto_increment	Primary key
raceId	int(11)	NO		0		Foreign key link to races table
constructorId	int(11)	NO		0		Foreign key link to constructors table
points	float	NO		0		Constructor points for season
position	int(11)	YES		NULL		Constructor standings position (integer)
positionText	varchar(255)	YES		NULL		Constructor standings position (string)
wins	int(11)	NO		0		Season win count

La tabella `constructor_standings` riporta le informazioni sulla classifica costruttori.

TABELLA DRIVERS

```
drivers table
```

Field	Type	Null	Key	Default	Extra	Description
driverId	int(11)	NO	PRI	NULL	auto_increment	Primary key
driverRef	varchar(255)	NO				Unique driver identifier
number	int(11)	YES		NULL		Permanent driver number
code	varchar(3)	YES		NULL		Driver code e.g. "ALO"
forename	varchar(255)	NO				Driver forename
surname	varchar(255)	NO				Driver surname
dob	date	YES		NULL		Driver date of birth
nationality	varchar(255)	YES		NULL		Driver nationality
url	varchar(255)	NO	UNI			Driver Wikipedia page

La tabella `drivers` riporta le informazioni sui piloti.

TABELLA DRIVER STANDINGS

```
driver_standings table
```

Field	Type	Null	Key	Default	Extra	Description
driverStandingsId	int(11)	NO	PRI	NULL	auto_increment	Primary key
raceId	int(11)	NO		0		Foreign key link to races table
driverId	int(11)	NO		0		Foreign key link to drivers table
points	float	NO		0		Driver points for season
position	int(11)	YES		NULL		Driver standings position (integer)
positionText	varchar(255)	YES		NULL		Driver standings position (string)
wins	int(11)	NO		0		Season win count

La tabella `driver_standings` riporta le informazioni sulla classifica ottenuta dai piloti in una gara.

TABELLA LAP TIMES

lap_times table						
Field	Type	Null	Key	Default	Extra	Description
raceId	int(11)	NO	PRI	NULL		Foreign key link to races table
driverId	int(11)	NO	PRI	NULL		Foreign key link to drivers table
lap	int(11)	NO	PRI	NULL		Lap number
position	int(11)	YES		NULL		Driver race position
time	varchar(255)	YES		NULL		Lap time e.g. "1:43.762"
milliseconds	int(11)	YES		NULL		Lap time in milliseconds

La tabella lap_times riporta le informazioni sui tempi ottenuti dai piloti durante le gare.

TABELLA PIT STOPS

pit_stops table						
Field	Type	Null	Key	Default	Extra	Description
raceId	int(11)	NO	PRI	NULL		Foreign key link to races table
driverId	int(11)	NO	PRI	NULL		Foreign key link to drivers table
stop	int(11)	NO	PRI	NULL		Stop number
lap	int(11)	NO		NULL		Lap number
time	time	NO		NULL		Time of stop e.g. "13:52:25"
duration	varchar(255)	YES		NULL		Duration of stop e.g. "21.783"
milliseconds	int(11)	YES		NULL		Duration of stop in milliseconds

La tabella pit_stops riporta le informazioni sui pit stops effettuati dai piloti durante le gare.

TABELLA QUALIFYING

qualifying table						
Field	Type	Null	Key	Default	Extra	Description
qualifyId	int(11)	NO	PRI	NULL	auto_increment	Primary key
raceId	int(11)	NO		0		Foreign key link to races table
driverId	int(11)	NO		0		Foreign key link to drivers table
constructorId	int(11)	NO		0		Foreign key link to constructors table
number	int(11)	NO		0		Driver number
position	int(11)	YES		NULL		Qualifying position
q1	varchar(255)	YES		NULL		Q1 lap time e.g. "1:21.374"
q2	varchar(255)	YES		NULL		Q2 lap time
q3	varchar(255)	YES		NULL		Q3 lap time

La tabella qualifying riporta le informazioni sulle qualifiche dei Gran Premi.

TABELLA RACES

Field	Type	Null	Key	Default	Extra	Description
raceId	int(11)	NO	PRI	NULL	auto_increment	Primary key
year	int(11)	NO		0		Foreign key link to seasons table
round	int(11)	NO		0		Round number
circuitId	int(11)	NO		0		Foreign key link to circuits table
name	varchar(255)	NO				Race name
date	date	NO		0000-00-00		Race date e.g. "1950-05-13"
time	time	YES		NULL		Race start time e.g. "13:00:00"
url	varchar(255)	YES	UNI	NULL		Race Wikipedia page

La tabella `races` riporta le informazioni sui Gran Premi.

TABELLA RESULTS

Field	Type	Null	Key	Default	Extra	Description
resultId	int(11)	NO	PRI	NULL	auto_increment	Primary key
raceId	int(11)	NO		0		Foreign key link to races table
driverId	int(11)	NO		0		Foreign key link to drivers table
constructorId	int(11)	NO		0		Foreign key link to constructors table
number	int(11)	YES		NULL		Driver number
grid	int(11)	NO		0		Starting grid position
position	int(11)	YES		NULL		Official classification, if applicable
positionText	varchar(255)	NO				Driver position string e.g. "1" or "R"
positionOrder	int(11)	NO		0		Driver position for ordering purposes
points	float	NO		0		Driver points for race
laps	int(11)	NO		0		Number of completed laps
time	varchar(255)	YES		NULL		Finishing time or gap
milliseconds	int(11)	YES		NULL		Finishing time in milliseconds
fastestLap	int(11)	YES		NULL		Lap number of fastest lap
rank	int(11)	YES		0		Fastest lap rank, compared to other drivers
fastestLapTime	varchar(255)	YES		NULL		Fastest lap time e.g. "1:27.453"
fastestLapSpeed	varchar(255)	YES		NULL		Fastest lap speed (km/h) e.g. "213.874"
statusId	int(11)	NO		0		Foreign key link to status table

La tabella `results` riporta le informazioni sui risultati dei Gran Premi.

TABELLA SEASON

Field	Type	Null	Key	Default	Extra	Description
year	int(11)	NO	PRI	0		Primary key e.g. 1950
url	varchar(255)	NO	UNI			Season Wikipedia page

La tabella `season` riporta le informazioni sulla singola stagione.

TABELLA STATUS

status table						
Field	Type	Null	Key	Default	Extra	Description
statusId	int(11)	NO	PRI	NULL	auto_increment	Primary key
status	varchar(255)	NO				Finishing status e.g. "Retired"

La tabella status riporta le informazioni riguardanti lo stato dei piloti a fine gara.

3.3 ANALISI TABELLE

Com'è facilmente intuibile, andremo adesso ad analizzare le tabelle per poter decidere quali caratteristiche potranno essere più rilevanti per risolvere il nostro problema. Fatto ciò andremo ad unire tutte le tabelle in una sola, che sarà il nostro dataset da cui il modello potrà apprendere.

4. ALGORITMI UTILIZZATI

5. VALUTAZIONI

6. SVILUPPI FUTURI

7. GLOSSARIO

- **Formula:** si riferisce all'insieme di regole che i partecipanti (team e piloti) devono rispettare.
- **Prove Libere:** le prove libere sono tre sessioni durante le quali i piloti possono prendere confidenza con la pista e gli ingegneri aggiustare e adattare meglio l'assetto della vettura. Attualmente ogni sessione dura 60 minuti, contro i 90 e i 45 adottati nelle precedenti stagioni.
- **Qualifiche:** le qualifiche servono a stabilire l'ordine di partenza della gara della domenica, la griglia di partenza. Attualmente la qualifica si compone di tre sessioni, denominate Q1, Q2 e Q3 della durata di 18, 15 e 12 minuti rispettivamente, con un sistema *knock-out*: le 5 vetture più lente al termine della Q1 sono eliminate, allo stesso modo le ultime 5, tra le vetture rimaste, al termine delle Q2 sono eliminate e le restanti combattono per la *Pole Position* nella Q3.

Il formato delle qualifiche è quello che ha subito più variazioni nel corso della storia della Formula Uno.

- **Pole Position:** è il termine che indica la prima posizione nella griglia di partenza.
- **Gran Premio:** erroneamente confuso con la gara della domenica, il Gran Premio indica tutti gli eventi che si svolgono dal giovedì alla domenica.
- **Costruttore:** un Costruttore può essere visto come il team stesso. Il termine deriva dal fatto che, inizialmente, un team costruiva sia telaio che motore per le monoposto di Formula Uno. Col tempo però, a causa delle grandi quantità di denaro da dover spendere per poter sviluppare una vettura, molti team hanno deciso di utilizzare un motore costruito da altri team, costruendo però da se il telaio, così da poter prendere parte al Mondiale Costruttori.
- **Mondiale Piloti:** il Mondiale Piloti è il titolo che si contendono ogni anno i piloti di Formula Uno. Alla fine di ogni gara, i primi dieci classificati ottengono un punteggio proporzionato alla posizione (1° = 25 punti, 2° = 18 punti, ..., 10° = 1 punto, secondo l'assegnazione attuale). Il pilota che a fine Campionato avrà totalizzato più punti si aggiudicherà il titolo.
- **Mondiale Costruttori:** il Mondiale Costruttori è l'analogo del Mondiale Piloti per i team di Formula Uno. I punteggi sono calcolati come la somma dei punti ottenuti dai piloti di un team al termine di una gara. Il team che avrà totalizzato più punti a fine Campionato avrà totalizzato più punti si aggiudicherà il titolo.