

## Opis charakteru danych

Dane prezentują losowe wyniki poszczególnych kierowców reprezentujących w każdym wyścigu 1 z 24 producentów w ramach danego amatorskiego Grand Prix. Na podstawie określonego rezultatu (miejsca od 1 do 20) kierowcy zdobywają za wyścig punkty, które mogą posłużyć do stworzenia klasyfikacji generalnej kierowców bądź producentów marek samochodowych. Etykieta czasowa określa losowy czas zakończenia wyścigu odbytego w przeciągu ostatniego roku.

## Opis atrybutów

Atrybuty w każdym zdarzeniu mają następujące znaczenie:

- car
  - typ: string
  - znaczenie: nazwa marki samochodu
  - kategoria: atrybut lub atrybuty dodatkowe, opisujące zdarzenie
- team
  - typ: string
  - znaczenie: nazwa zespołu (producenta danego samochodu)
  - kategoria: atrybut lub atrybuty dodatkowe, opisujące zdarzenie
- driver
  - typ: string
  - znaczenie: imię i nazwisko kierowcy pojazdu w czasie określonego wyścigu
  - kategoria: atrybut lub atrybuty dodatkowe, opisujące zdarzenie
- car\_type
  - typ: string
  - znaczenie: rodzaj nadwozia pojazdu
  - kategoria: atrybut, lub atrybuty, po których można dane grupować (zawierający ograniczoną liczbę możliwych wartości (5-10))
- fuel\_type
  - typ: string
  - znaczenie: rodzaj paliwa wykorzystywanego przez dany pojazd
  - kategoria: atrybut, lub atrybuty, po których można dane grupować (zawierający ograniczoną liczbę możliwych wartości (5-10))
- result
  - typ: int
  - znaczenie: zdobyta pozycja przez kierowcę w wyścigu
  - kategoria: atrybut lub atrybuty dodatkowe, opisujące zdarzenie
- points
  - typ: int
  - znaczenie: liczba zdobytych punktów za zajęcie określonego miejsca przez kierowcę w wyścigu
  - kategoria: atrybut lub atrybuty, których wartości można agregować (wartości numeryczne)
- place
  - typ: string
  - znaczenie: nazwa określonego wyścigu zawierająca miejsce odbywania się danego Grand Prix

- kategoria: atrybut lub atrybuty dodatkowe, opisujące zdarzenie
- ts
  - typ: string
  - znaczenie: czas zakończenia wyścigu (występu danego kierowcy w określonym Grand Prix)
  - kategoria: kategoria: znacznik czasowy zdarzeń (inny niż czas systemowy)

## Opis trzech przykładowych analiz

Propozycje możliwych analiz opartych na tym generatorze danych

### Agregacja

Utrzymuj informacje o sumie punktów zdobytych przez poszczególne zespoły (producentów) w całym roku (sezonie) lub określonym przedziale czasu.

### Wykrywanie anomalii

Wykrywaj przypadki zdobycia przez kierowców liczby większej niż 10 za dany wyścig.

### Wykrywanie anomalii oparte na agregacji

Znajduj przypadki, w których liczba punktów zdobytych przez danego producenta samochodami elektrycznymi będzie stanowiła co najmniej 20% wszystkich zdobytych punktów przez tego producenta w całym sezonie lub określonym okresie czasu.