TP - BigData

Analyse des performances aux championnats du monde de Formule 1.

Wassim BACHA

Nathan SAMOCHVAL

Céleste DUPONT

[**Sujet 3**](#_vohevy185str)

[Projet : 3](#_q3rdbh7rgfp7)

[Objectif : 3](#_sbg5iseffidc)

[Jeux de données : 3](#_rghhks7j08oj)

[**Architecture et traitements : 4**](#_st2wtjmrry88)

[Schéma d’architecture 4](#_xc34c96qgu0s)

[Collecte et stockage des données (Dataset) 4](#_ocynxs8tdq3r)

[Traitement des données 4](#_wxeazl2v58md)

[Nettoyage et transformation de données 4](#_dx9cy3mevo5r)

[**Ajout de champs 4**](#_omlb9hdm09no)

[**Régression linéaire 4**](#_2upnf4xvet06)

# 

# Sujet

## Projet :

Ce projet a pour objectif d'explorer et d'analyser les performances des pilotes, des écuries et des voitures au fil des saisons du Championnat du Monde de Formule 1, de 1950 à 2024. L'étude se concentre sur les données de tours de piste, avec une analyse détaillée des chronos, des écarts entre les pilotes, et des tendances de performance selon divers facteurs tels que le circuit, les conditions météorologiques, et les technologies utilisées.

## Objectif :

L'objectif principal est d'identifier les tendances, les relations et les anomalies dans les performances des pilotes et des écuries grâce à des méthodes d'analyse statistique et de visualisation de données.

Les résultats seront synthétisés sous forme de graphiques et de rapports pour fournir une vue d'ensemble complète des dynamiques de performance en Formule 1.

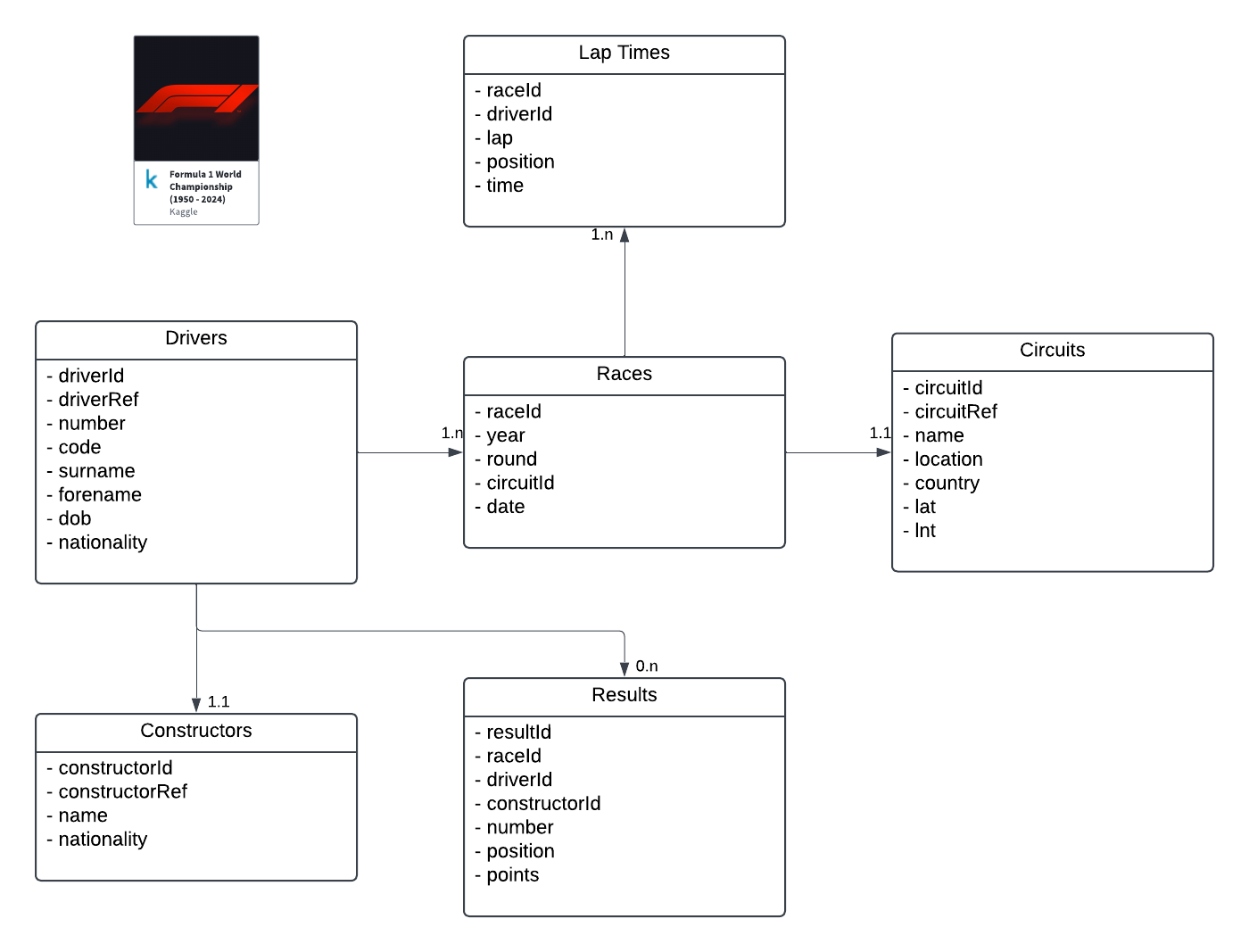
Ces insights peuvent être utilisés pour :

* Comprendre l'évolution historique des performances.
* Mettre en lumière les facteurs clés de succès en compétition.
* Proposer des prédictions basées sur des données historiques.

## Jeux de données :

Pour mener à bien ce projet, le dataset suivant sera utilisé : [Formula 1 World Championship 1950-2024](https://www.kaggle.com/datasets/rohanrao/formula-1-world-championship-1950-2020?select=lap_times.csv).

Le dataset est composé de plusieurs tables de données, afin d’en avoir une vue d’ensemble plus précise, voici le MCD (Modèle Conceptuel de Données) :



# Architecture et traitements :

## Schéma d’architecture

## Collecte et stockage des données (Dataset)

## Traitement des données

### Nettoyage et transformation de données

# Ajout de champs

# Régression linéaire