

**Formula One Race Data**

Maxime THOOR  
Louis TINEL  
Jules ROUSSEAU  
Nils CHARRIER  
Edouard DANSETTE  
Pierre-Maxime COSTA

1. **Présentation de Formula One Data Race**

* La Formule 1

La Formule 1 est la reine des disciplines en sport automobile. Créé en 1950, c’est aujourd’hui un des sports les plus suivi au monde. Le championnat est sous la forme de grands prix qui durent trois jours. Une journée d’essais libre, une journée de qualification et enfin la course. La formule 1 regroupe ce qu’il se fait de mieux en therme d’automobile, et les meilleurs pilotes du monde. Bien sûr, tout cela est encadré par des règles, que les pilotes et les écuries s’efforcent de respecter ou de contourner.

* L’entreprise

Formula One Data Race (FORD) est une société d’études et d’analyses d’écuries et de pilotes de formule 1. En se basant sur l’étude des statistiques des Grand Prix de formule 1, elle délivre des pronostics et conseille les investisseurs et parieurs, particuliers comme professionnels.

Pour pouvoir fournir des pronostics les plus fiables possibles, FORD s’appuie sur un set de données complet et varié et analyse chaque aspect des courses de Formule 1 afin d’anticiper quel pilote ou quelle écurie a le plus de chance de remporter le prochain Grand Prix. Les investisseurs peuvent ainsi investir sur les constructeurs automobiles et les parieurs sur les pilotes.

* Les membres de l’entreprise

Maxime THOOR, chef de projet

Louis TINEL, s’occupe de l’extraction des données et du nettoyage

Jules ROUSSEAU, s’occupe du Data Warehouse

Nils CHARRIER, s’occupe de la partie Datamart

Pierre-Maxime COSTA et Edouard DANSETTE, s’occupe du reporting et de l’analyse des données

1. **Présentation des données**

Les sets de données :

Les différents sets de données sont les suivants :

*circuit.csv drivers.csv results.csv  
 constructor\_results.csv lap\_times.csv seasons.csv  
 constructor\_standing.csv pit\_stops.csv status.csv  
 constructor.csv qualifying.csv  
 driver\_standing.csv races.csv*

Ces jeux de données contiennent toutes les informations nécessaires sur les statistiques des courses des Grand Prix. Les statistiques concernent chaque aspect des courses : les temps réalisés par chaque pilote en course et en qualification, leurs positions, leurs points gagnés. Ces statistiques sont également valables pour les constructeurs.

Avant d’effectuer un éventuel tri, l’ensemble des membres de l’équipe se sont répartis sur différentes études afin d’en tirer des conclusions évidentes pour la suite de notre projet. Grâce à cela, nous avons choisi d’en éliminer 3 d’entre eux qui sont :

*constructor\_standing.csv*

*driver\_standing.csv*

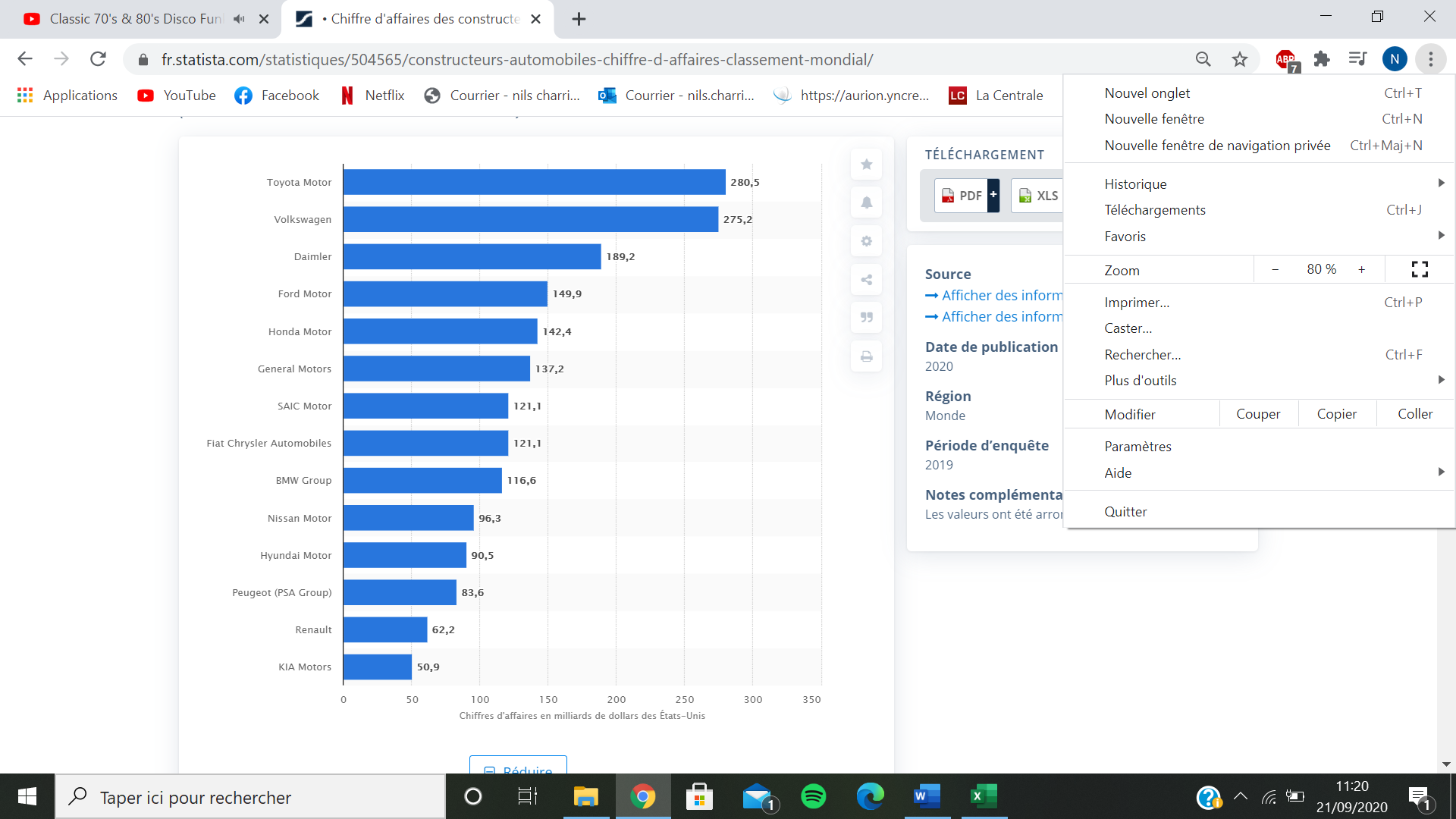
*seasons.csv*

En effet, chacun de ces datasets ne possédaient pas d’éléments majeurs pour délivrer des statistiques importantes pour notre entreprise.

Eléments qu’on ne peut pas prendre en compte (le fait que plusieurs constructeurs collaborent sur une même voiture)

Le marché automobile, indicateur de la santé des écuries :

Chiffre d’affaire des constructeurs en 2019 :



On remarque donc que les écuries les plus performantes sont présentes dans les constructeurs ayant le plus gros chiffre d’affaire, Ferrari qui fait des véhicules en faible quantité sort son épingle du jeu en ayant une grande marge sur chaque véhicule vendu, de même pour McLaren. Ferrari est même la marque qui fait la plus grosse marge par véhicule vendu dans le monde (environ 68 000 par véhicule).

Certains constructeurs dans le top des chiffres d’affaires s’engagent dans d’autre disciplines, comme Toyota et Volkswagen qui sont les leaders en endurance par exemple.

C’est pour cela que dans l’analyse des potentiels nouveaux acteurs en F1, il faudra surveiller aussi le marché automobile mondiale, qui ne cesse d’évoluer. On pourra s’attendre à l’apparition de nouveaux constructeurs en sport automobile comme SAIC Motors, une des plus grosse firme chinoise, ou encore des constructeurs qui décident de changer de discipline ou de se diversifier.

Sur les graphiques précédents, on remarque le détachement d’un trio de tête, avec un énorme budget et qui se partage le podium. Ce podium correspond donc bien au marché automobile actuel qu’on a vu précédemment. On remarque que pour les places suivantes, le budget n’est pas forcément synonyme de victoire. C’est cette partie du classement qui nous intéresse le plus, car c’est sur cette partie du classement, que nous allons essayer de prédire le résultat des courses. C’est à partir de la que nous allons rassembler toutes les données de courses pour rentrer dans le détail et permettre de comprendre et de prédire, qui va gagner la prochaine course.

Les règles en F1 qui renverse les classements :

Surveiller les règles de la F1 et leurs évolutions au fil des années est très important car cela permet à certains de prendre les devants et de bouleverser les classements. Cela permet aussi l’arrivé ou le départ d’écuries ou de constructeurs.

Ex : Audi, maintes fois évoquée, attendue, citée, la marque aux anneaux n'a jamais souhaité sauter le pas, tant les contraintes financières et techniques ne s'adaptaient pas à leur vision du sport automobile de haut niveau.