

# **Analysis Data**

**WeRateDogs TWITTER**

## INTRODUCTION

Les données du monde réel sont rarement claires. L'ensemble de données utiliser dans le cadre de ce projet est l'archive de tweets de l'utilisateur de Twitter **@dog\_rates**, également connu sous le nom de **WeRateDogs**.

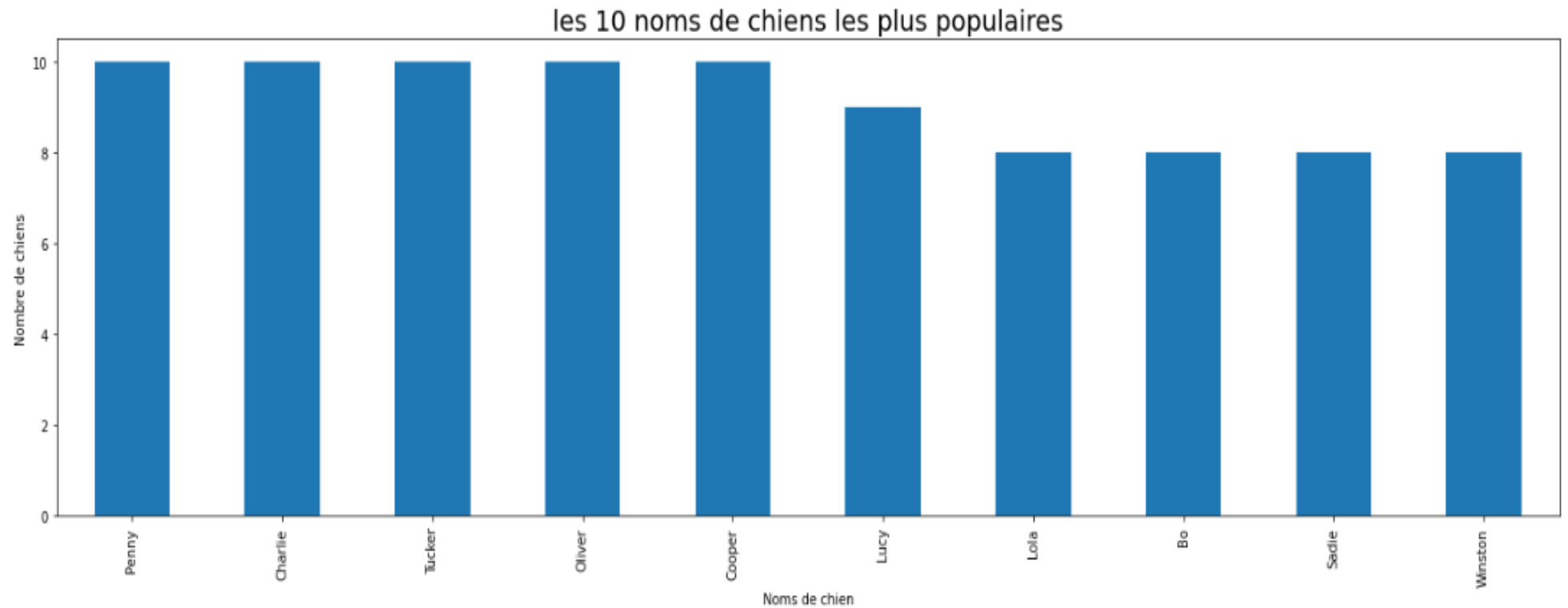
**WeRateDogs** est un compte Twitter qui évalue les chiens des gens avec un commentaire humoristique sur le chien.

À l'aide de Python et de ses bibliothèques, nous avons travaillé sur le processus de préparation des données qui se base sur la collecte, évaluation et nettoyage des données. Cependant, ce rapport présente les résultats de quelque question sur l'analyse et la visualisation réalisées.

Ces questions sont les suivantes :

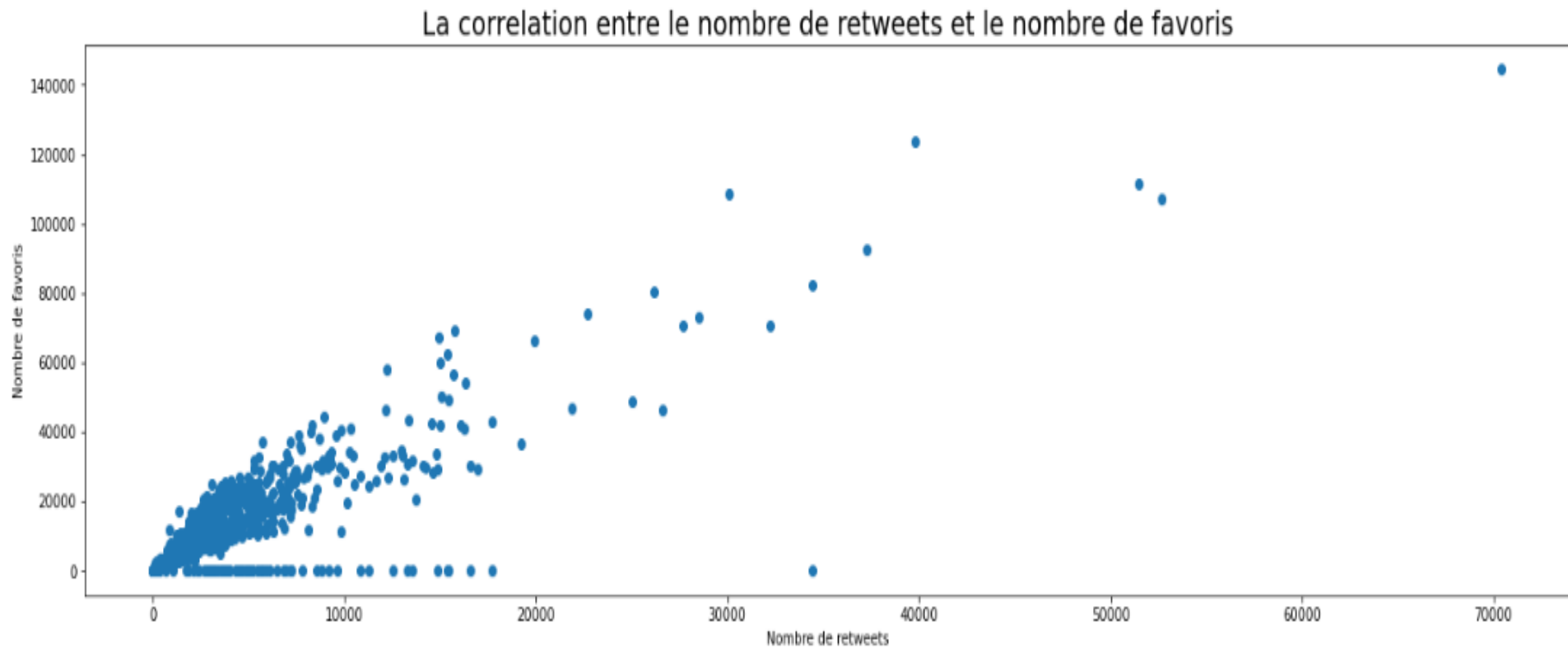
1. Quels sont les noms de chien les plus populaires ?
2. Quelle est la corrélation entre le nombre de retweets et le nombre de favoris ?
3. Quelle sont les stades de chien les plus courant ?
4. Quels sont les stades de chien les plus Likes ?

## 1. Les noms de chien les plus populaires



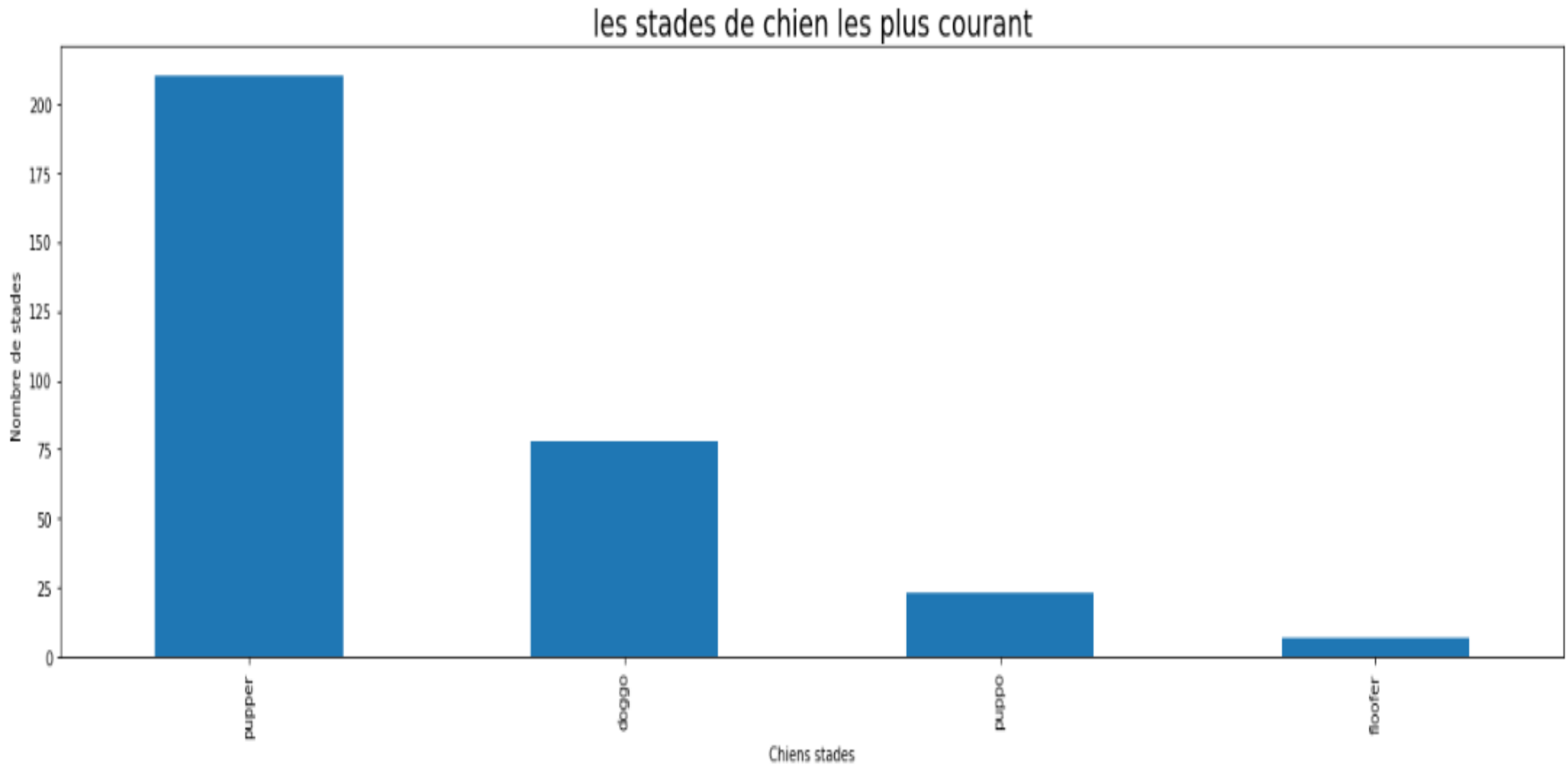
On peut conclure que les 10 noms de chien les plus populaires sont : **Penny, Charlie, Tucker, Oliver, Cooper**. Suivi de **Lucy** et vient ensuite **Lola, Bo, Sadie et Winston**.

## 2. La corrélation entre le nombre de retweets et le nombre de favoris



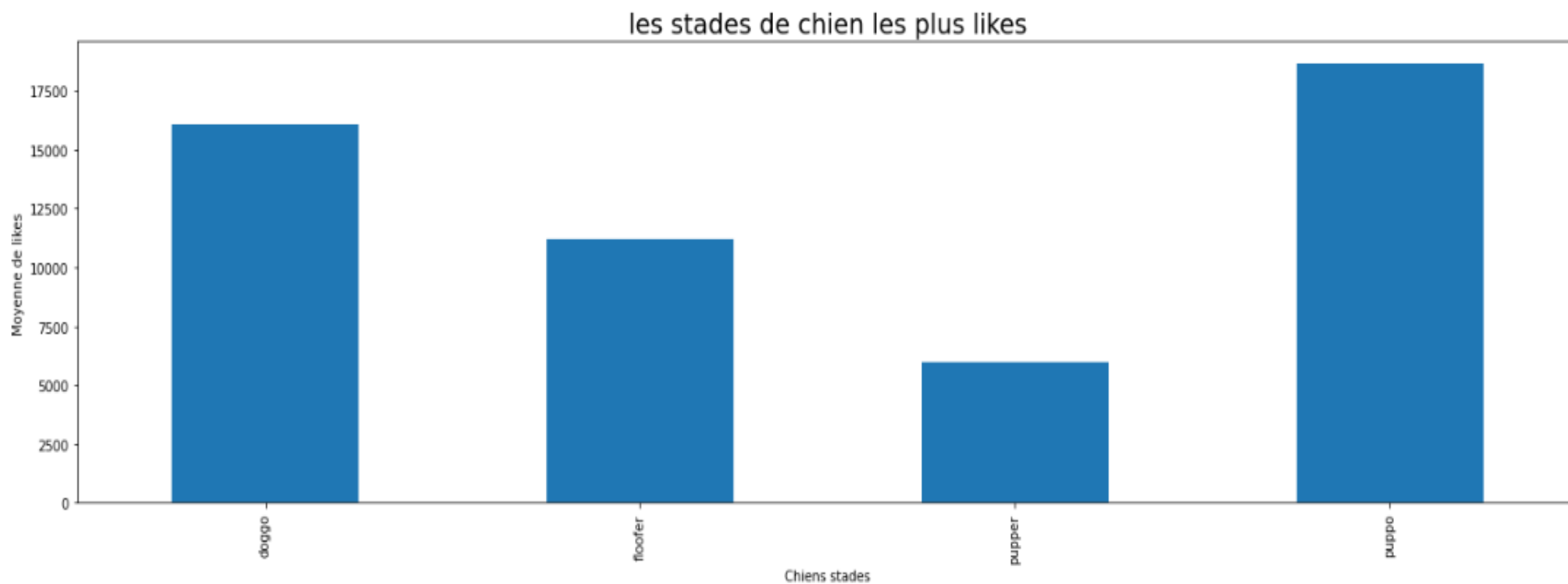
Ce graphique nous montre qu'il existe une corrélation linéaire positive entre le nombre de retweets et le nombre de favoris.

### 3. Les stades de chien les plus courant



On peut voir que le stade de chien le plus courant est **pupper** qui compte **210**, ensuite **doggo** qui compte **78**, **puppo** pour **23** et **flooter** pour **7**

#### 4. Les stades de chien les plus Likes



On peut conclure que **puppo** est le stade de chien le plus likes, suivi de **doggo**, **floofer** et **pupper**

## CONCLUSIONS

Ce projet nous a permis de voir un aperçu sur quelques analyses et visualisations des données sur **WeRateDogs Twitter**. Cependant, il y a tellement d'autres questions qui peuvent être analysées et visualisées de cet ensemble de données.