

Observations des données

Après avoir nettoyé nos données, nous avons effectué quelques observations dans l'optique de tirer d'avantages informations sur les données. cette phase avait pour objectifs de faire ressortir les réponses aux trois questions posées:

- **Quels sont les noms des chiens les plus populaires?**

Pour répondre à cette question, nous avons effectué un `value_counts()` sur la variable `name` et nous concluons que les noms Charlie, Cooper, Lucy et Oliver sont les noms les plus populaires de cet ensemble de données.

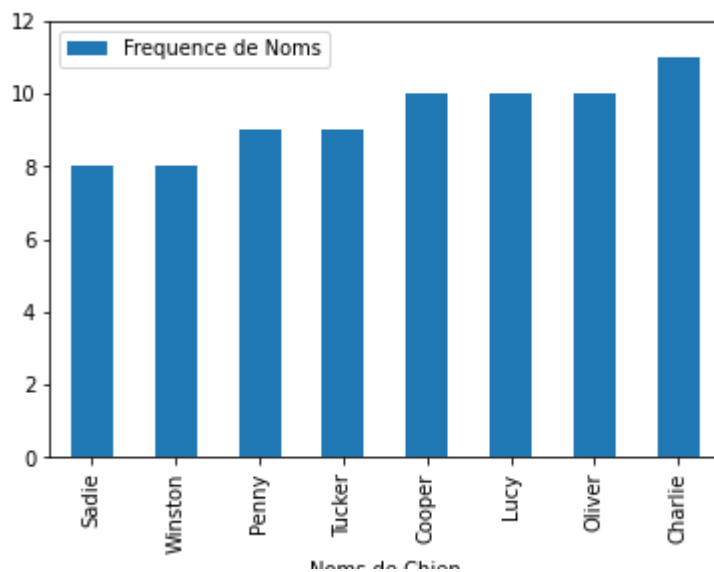


Diagramme représentant les noms les plus populaires en fonction du nombre de tweet

- **Quelle est la source de tweets la plus courante?**

Nous avons groupé les différents tweets en fonction de leurs sources et nous concluons que 98% de tweets inscrits dans le dataset ont été effectués via un iPhone.

- **A quelle classe appartient le chien le plus populaire?**

Il en ressort que le chien le plus populaire('charlie') appartient à la classe Pupper.