

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

import scipy.stats as stats

def t_test(df1, df2, target="SOG"):
    t_stat, p_value = stats.ttest_ind(df1[target].dropna(), df2[target].dropna())
    print(f"T-statistic: {t_stat:.3f}, p-value: {p_value:.15f}")

    # If p-value is less than 0.05, the difference is statistically significant
    if p_value < 0.05:
        print("The difference is statistically significant, keeping data split.")
    else:
        print("The difference is not statistically significant, keeping data combined.")
```

```
In [2]: df = pd.read_csv("all_data.csv")
```

```
In [3]: # Filtrer les données de Senseboard le 9 juin
senseboard_9juin = df[
    (df["boat_name"] == "SenseBoard") &
    (df["opponent_name"] == "Karl Maeder") &
    (df["ISODateTimeUTC"].str.startswith("2025-06-09"))]

# Compter le nombre de poids différents
nb_poids_différents = senseboard_9juin["boat_weight"].nunique()

# Afficher le résultat
print(f"Nombre de poids différents utilisés par Senseboard (Gian) le 9 juin : {nb_poids_différents}")
```

Nombre de poids différents utilisés par Senseboard (Gian) le 9 juin : 2

```
In [4]: poids_uniques = senseboard_9juin["boat_weight"].unique()

if len(poids_uniques) != 2:
    print("Erreur : il faut exactement deux poids pour effectuer le t-test.")
else:
    poids1, poids2 = poids_uniques

    groupe1_df = senseboard_9juin[senseboard_9juin["boat_weight"] == poids1]
    groupe2_df = senseboard_9juin[senseboard_9juin["boat_weight"] == poids2]
    t_test(groupe1_df, groupe2_df)
```

T-statistic: -2.676, p-value: 0.007467354374147
The difference is statistically significant, keeping data split.

```
In [5]: print(len(groupe1_df), len(groupe2_df))
```

3940 4622

```
In [ ]:
```