

```
import pandas as pd
```

```
df = pd.read_csv("results.csv")
```

```
print(df.info())
```

```
print(df.head())
```

```
print(df.describe())
```

```
print("\nMedia:")
print(df[["home_score", "away_score"]].mean())
```

```
print("\nMediana:")
print(df[["home_score", "away_score"]].median())
```

```
print("\nDesviación estándar:")
print(df[["home_score", "away_score"]].std())
```



```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 41243 entries, 0 to 41242
```

```
Data columns (total 9 columns):
```

#	Column	Non-Null Count	Dtype
0	date	41243 non-null	object
1	home_team	41243 non-null	object
2	away_team	41243 non-null	object
3	home_score	41243 non-null	int64
4	away_score	41243 non-null	int64
5	tournament	41243 non-null	object
6	city	41243 non-null	object
7	country	41243 non-null	object
8	neutral	41243 non-null	bool

```
dtypes: bool(1), int64(2), object(6)
```

```
memory usage: 2.6+ MB
```

```
None
```

	date	home_team	away_team	home_score	away_score	tournament	city	\
0	1872-11-30	Scotland	England	0	0	Friendly	Glasgow	
1	1873-03-08	England	Scotland	4	2	Friendly	London	
2	1874-03-07	Scotland	England	2	1	Friendly	Glasgow	
3	1875-03-06	England	Scotland	2	2	Friendly	London	
4	1876-03-04	Scotland	England	3	0	Friendly	Glasgow	

	country	neutral
0	Scotland	False
1	England	False
2	Scotland	False
3	England	False
4	Scotland	False

	home_score	away_score
count	41243.000000	41243.000000
mean	1.744975	1.189317
std	1.750513	1.406648
min	0.000000	0.000000
25%	1.000000	0.000000
50%	1.000000	1.000000
75%	2.000000	2.000000
max	31.000000	21.000000

```
Media:
```

```
home_score    1.744975
```

```
away_score    1.189317  
dtype: float64
```

```
Mediana:  
home_score    1.0  
away_score    1.0  
dtype: float64
```

```
Desviación estándar:  
home_score    1.750513  
away_score    1.406648  
dtype: float64
```

1. Home Advantage: On average, home teams score more goals (mean = 1.74) compared to away teams (mean = 1.19). This indicates a clear home-field advantage.
2. Median Comparison: The median score for both home and away teams is 1, suggesting that in most matches, teams score around 1 goal regardless of being home or away.
3. Score Variability: The standard deviation is higher for home teams (1.75) than for away teams (1.41), indicating slightly more variability in home team scores.
4. Extreme Values: The maximum number of goals scored by a single team in a match is very high (31 for home teams, 21 for away teams), which could be due to historical or highly unbalanced matches.
5. Balanced Dataset: With over 41,000 matches, the dataset is large and likely to provide meaningful insights into trends over time and across tournaments, countries, and locations.

<https://github.com/Uriel1780/semana-tec>