

Práctica 2

Parte 1

- A partir el dataset `Batting.csv`, obtener el siguiente dataframe para `teamID == 'BOS'`:

	yearID	teamID	R	HR	SO	IBB
8123	1901	BOS	2	0	0.0	NaN
8162	1901	BOS	108	6	18.0	NaN
8171	1901	BOS	26	0	9.0	NaN
8181	1901	BOS	4	0	1.0	NaN
8211	1901	BOS	104	3	22.0	NaN

- Reemplazar los NaN de la variable `SO` (*strikeouts*) usando el *IQR*.
- Reemplazar los NaN de la variable `IBB` (*base por bolas intencionales*) usando su media.
- Dibujar la distribución de densidad antes y después de reemplazar los datos para ambas variables.
- Remover los outliers de la variable `HR` y comparar la correlación entre las variables `R`, `HR`, `SO` e `IBB` antes y después removerlos.
- Explicar por qué cambia tanto la correlación entre `HR` e `IBB` después de remover los outliers.

Parte 2

- Extraer la siguiente serie de tiempo del dataset `Batting.csv`:

yearID	avgHR
1871	0.408696
1872	0.235669
1873	0.376000
1874	0.325203
1875	0.183486

- Aplicar el filtro asimétrico de medianas móviles a la serie de tiempo con $q = 20$.
- Aplicar el filtro exponencial a la serie de tiempo con $\theta = 0.1$.
- Graficar la serie original junto con estos filtros.