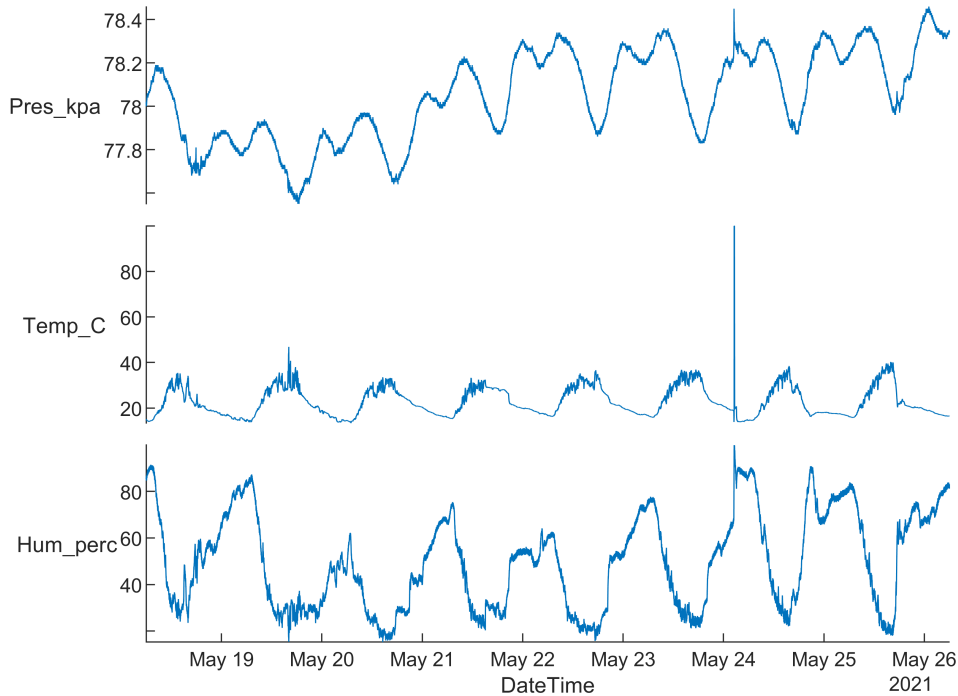


# Estadística la venganza

```
load("../..Clase Github/Utils4SP/Datasets/S5_Estadistica101_LaVenganza.mat");
```

## Plots exploratorios

```
figure  
stackedplot(atmosfera,"XVariable","DateTime")
```



```
summary(atmosfera)
```

Variables:

**Fecha:** 137522×1 cell array of character vectors

**Hora:** 137522×1 cell array of character vectors

**Pres\_kpa:** 137522×1 double

Values:

Min	77.55
Median	78.06
Max	78.46

**Temp\_C:** 137522×1 double

Values:

Min	13.38
Median	21.01
Max	100
NumMissing	11

Hum\_perc: 137522x1 double

Values:

Min	15.27
Median	50.16
Max	99.97
NumMissing	61

DateTime: 137522x1 datetime

Values:

Min	20210518 06:00:03
Median	20210522 05:59:48
Max	20210526 05:59:55

```
% Exploración de números faltantes
% Todas las lecturas de atmosfera
% Tal que existan NaN's en su campo "Humedad"
% Y que sean todos los campos
atmosfera(ismissing(atmosfera.Hum_perc),:)
```

ans = 61x6 table

	Fecha	Hora	Pres_kpa	Temp_C	Hum_perc	DateTime
1	'210519'	'01:55:14'	77.8600	16.5900	NaN	20210519 0...
2	'210523'	'04:27:47'	78.2000	17.2300	NaN	20210523 0...
3	'210523'	'05:17:37'	78.2000	17.2000	NaN	20210523 0...
4	'210524'	'02:32:46'	78.3800	100	NaN	20210524 0...
5	'210524'	'02:32:56'	78.3800	100	NaN	20210524 0...
6	'210524'	'02:33:06'	78.3800	100	NaN	20210524 0...
7	'210524'	'02:33:11'	78.3800	100	NaN	20210524 0...
8	'210524'	'02:33:16'	78.3900	100	NaN	20210524 0...
9	'210524'	'02:33:21'	78.3800	100	NaN	20210524 0...
10	'210524'	'02:33:26'	78.3800	100	NaN	20210524 0...
11	'210524'	'02:33:31'	78.3700	100	NaN	20210524 0...
12	'210524'	'02:33:36'	78.3700	100	NaN	20210524 0...
13	'210524'	'02:33:41'	78.3600	100	NaN	20210524 0...
14	'210524'	'02:33:47'	78.3600	100	NaN	20210524 0...
15	'210524'	'02:33:52'	78.3600	100	NaN	20210524 0...
16	'210524'	'02:33:57'	78.3600	100	NaN	20210524 0...
17	'210524'	'02:34:02'	78.3500	100	NaN	20210524 0...
18	'210524'	'02:34:07'	78.3500	100	NaN	20210524 0...
19	'210524'	'02:34:12'	78.3500	100	NaN	20210524 0...

	Fecha	Hora	Pres_kpa	Temp_C	Hum_perc	DateTime
20	'210524'	'02:34:17'	78.3500	100	NaN	20210524 0...
21	'210524'	'02:34:22'	78.3500	100	NaN	20210524 0...
22	'210524'	'02:34:27'	78.3400	100	NaN	20210524 0...
23	'210524'	'02:34:32'	78.3400	100	NaN	20210524 0...
24	'210524'	'02:34:37'	78.3400	100	NaN	20210524 0...
25	'210524'	'02:34:42'	78.3500	100	NaN	20210524 0...
26	'210524'	'02:34:47'	78.3400	100	NaN	20210524 0...
27	'210524'	'02:34:52'	78.3400	100	NaN	20210524 0...
28	'210524'	'02:34:57'	78.3500	100	NaN	20210524 0...
29	'210524'	'02:35:02'	78.3400	100	NaN	20210524 0...
30	'210524'	'02:35:07'	78.3400	100	NaN	20210524 0...
31	'210524'	'02:35:12'	78.3400	100	NaN	20210524 0...
32	'210524'	'02:35:17'	78.3400	100	NaN	20210524 0...
33	'210524'	'02:35:22'	78.3400	100	NaN	20210524 0...
34	'210524'	'02:35:27'	78.3300	100	NaN	20210524 0...
35	'210524'	'02:35:32'	78.3300	100	NaN	20210524 0...
36	'210524'	'02:35:37'	78.3300	100	NaN	20210524 0...
37	'210524'	'02:35:42'	78.3300	100	NaN	20210524 0...
38	'210524'	'02:35:47'	78.3300	100	NaN	20210524 0...
39	'210524'	'02:35:52'	78.3200	100	NaN	20210524 0...
40	'210524'	'02:35:57'	78.3300	100	NaN	20210524 0...
41	'210524'	'02:36:02'	78.3200	100	NaN	20210524 0...
42	'210524'	'02:36:07'	78.3200	100	NaN	20210524 0...
43	'210524'	'02:36:12'	78.3200	100	NaN	20210524 0...
44	'210524'	'02:36:17'	78.3200	100	NaN	20210524 0...
45	'210524'	'02:36:22'	78.3200	100	NaN	20210524 0...
46	'210524'	'02:36:27'	78.3200	100	NaN	20210524 0...
47	'210524'	'02:36:32'	78.3300	100	NaN	20210524 0...
48	'210524'	'02:36:37'	78.3100	100	NaN	20210524 0...
49	'210524'	'02:36:42'	78.3000	100	NaN	20210524 0...
50	'210524'	'02:36:47'	78.3100	100	NaN	20210524 0...
51	'210524'	'02:36:53'	78.3100	100	NaN	20210524 0...
52	'210524'	'02:36:58'	78.3000	100	NaN	20210524 0...

	Fecha	Hora	Pres_kpa	Temp_C	Hum_perc	DateTime
53	'210524'	'02:37:03'	78.3100	100	NaN	20210524 0...
54	'210524'	'02:37:08'	78.3100	100	NaN	20210524 0...
55	'210524'	'02:37:13'	78.3100	100	NaN	20210524 0...
56	'210524'	'02:37:18'	78.3200	100	NaN	20210524 0...
57	'210524'	'02:37:24'	78.3100	100	NaN	20210524 0...
58	'210524'	'02:37:34'	78.3100	100	NaN	20210524 0...
59	'210524'	'02:37:39'	78.3200	100	NaN	20210524 0...
60	'210524'	'08:33:50'	78.2800	18.3700	NaN	20210524 0...
61	'210526'	'01:52:49'	78.4100	18.4400	NaN	20210526 0...

```
% Exploración de números faltantes
% Todas las lecturas de atmosfera
% Tal que existan NaN's en su campo "Temperatura"
% Y que sean todos los campos
atmosfera(ismissing(atmosfera.Temp_C),:)
```

ans = 11x6 table

	Fecha	Hora	Pres_kpa	Temp_C	Hum_perc	DateTime
1	'210518'	'17:59:58'	77.7100	NaN	49.6800	20210518 1...
2	'210520'	'09:27:14'	77.9500	NaN	38.9700	20210520 0...
3	'210520'	'09:37:27'	77.9500	NaN	39.4700	20210520 0...
4	'210520'	'21:22:43'	77.8400	NaN	44.8400	20210520 2...
5	'210521'	'20:50:27'	78.0400	NaN	48.0700	20210521 2...
6	'210522'	'20:57:34'	78.0700	NaN	51.3400	20210522 2...
7	'210522'	'20:57:44'	78.0800	NaN	51.2100	20210522 2...
8	'210522'	'20:58:39'	78.0700	NaN	51.2900	20210522 2...
9	'210523'	'10:19:08'	78.3200	NaN	53.7100	20210523 1...
10	'210523'	'21:44:54'	78.0100	NaN	52.8200	20210523 2...
11	'210525'	'18:14:48'	78	NaN	61.0600	20210525 1...

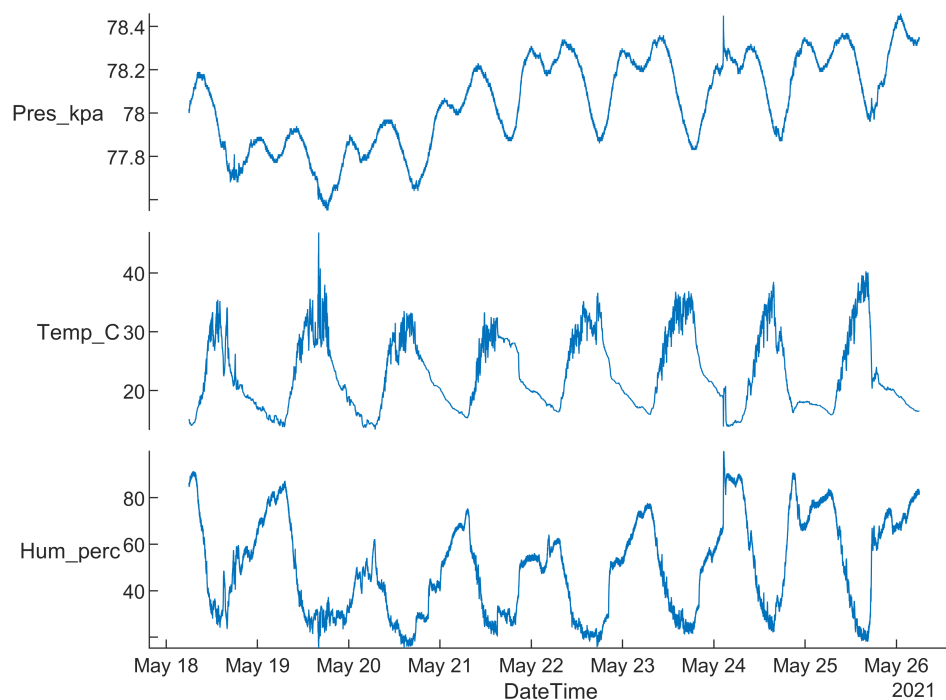
```
% Exploración de números faltantes
% Todas las lecturas de atmosfera
% Tal que existan NaN's en su campo "Temperatura"
% Y que devuelva DateTime
atmosfera(ismissing(atmosfera.Temp_C),["DateTime","Hum_perc"])
```

ans = 11x2 table

	DateTime	Hum_perc
1	20210518 1...	49.6800
2	20210520 0...	38.9700

	DateTime	Hum_perc
3	20210520 0...	39.4700
4	20210520 2...	44.8400
5	20210521 2...	48.0700
6	20210522 2...	51.3400
7	20210522 2...	51.2100
8	20210522 2...	51.2900
9	20210523 1...	53.7100
10	20210523 2...	52.8200
11	20210525 1...	61.0600

```
% Quitar NaN's
atmosfera_clean=rmmissing(atmosfera);
figure
stackedplot(atmosfera_clean,"XVariable","DateTime")
```

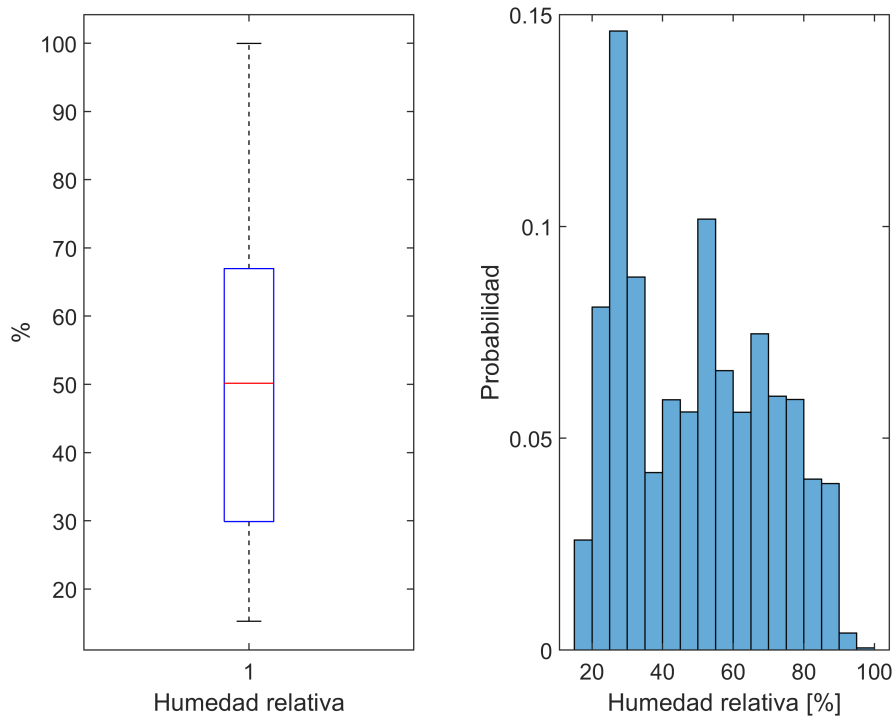


## Métricas de tendencia central

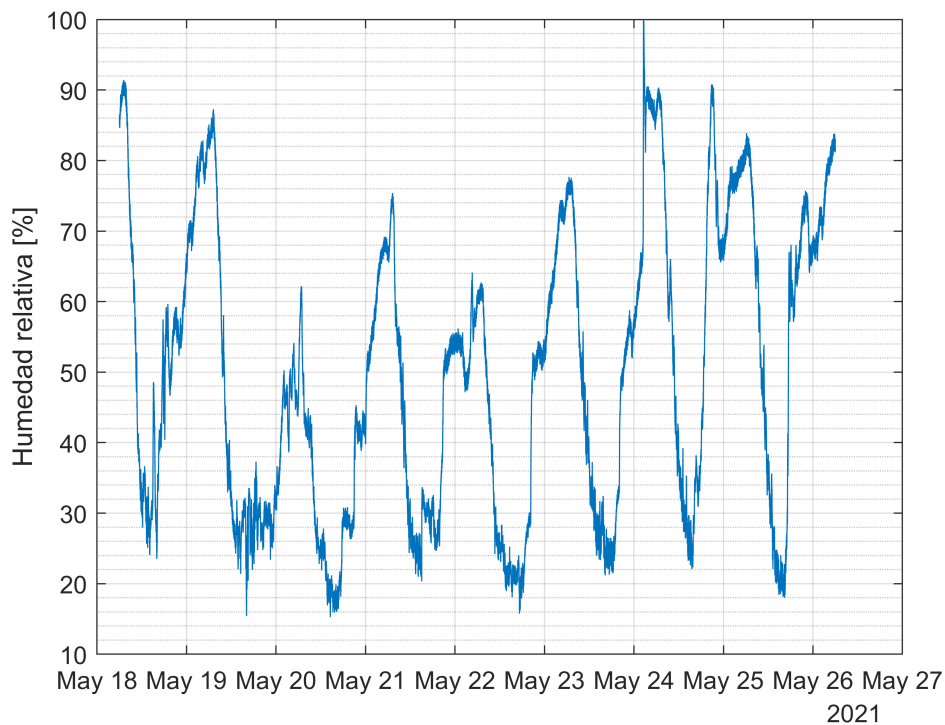
```
% Una gráfica con boxplot e histograma
figure
tiledlayout(1,2)
nexttile
boxplot(atmosfera_clean.Hum_perc)
```

Warning: Unable to set 'Position', 'InnerPosition', 'OuterPosition', or 'ActivePositionProperty' for objects in a TiledChartLayout

```
xlabel("Humedad relativa")
ylabel("%")
nexttile
histogram(atmosfera_clean.Hum_perc,"BinWidth",5,'Normalization',"probability")
xlabel("Humedad relativa [%]")
ylabel("Probabilidad")
```



```
% Serie de tiempo
figure
plot(atmosfera_clean.DateTime,atmosfera_clean.Hum_perc)
ylabel("Humedad relativa [%]")
grid on
grid minor
```



```
% Promedio, mediana y moda
```

```
% Promedio
```

```
hum_mean=mean(atmosfera_clean.Hum_perc)
```

```
hum_mean = 49.6841
```

```
% Mediana
```

```
hum_median=median(atmosfera_clean.Hum_perc)
```

```
hum_median = 50.1550
```

```
% Moda
```

```
hum_mode=mode(atmosfera_clean.Hum_perc)
```

```
hum_mode = 28.9200
```

```
% Cuantiles
```

```
% Primer caurtíl
```

```
hum_1Q=quantile(atmosfera_clean.Hum_perc,0.25)
```

```
hum_1Q = 29.8800
```

```
hum_3Q=quantile(atmosfera_clean.Hum_perc,0.75)
```

```
hum_3Q = 66.9600
```

```
% Añadiendo métricas de tendencia central
```

```
figure
```

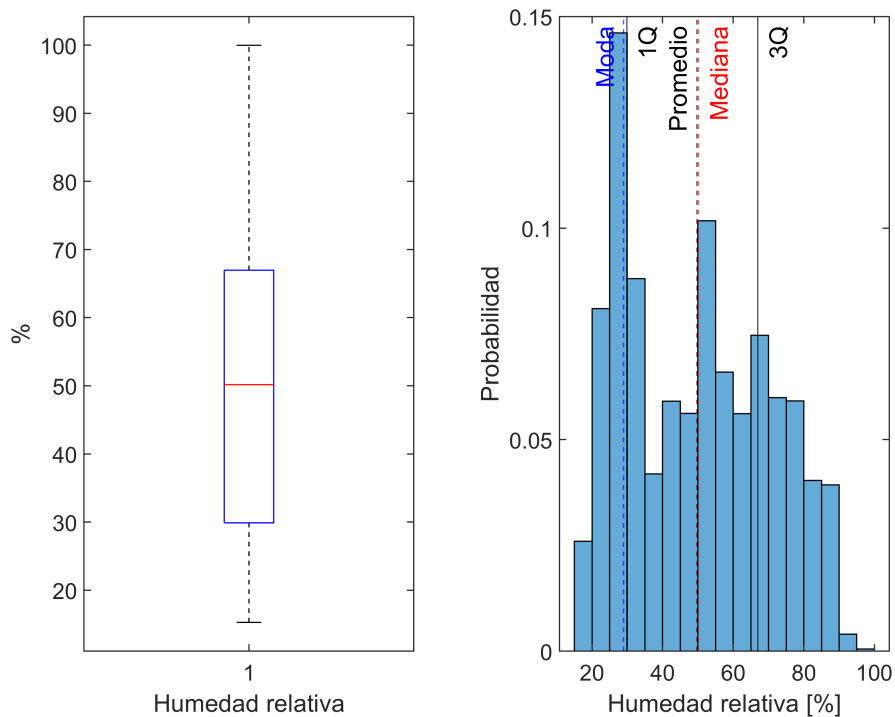
```
tilde layout(1,2)
```

```
nexttile
```

```
boxplot(atmosfera_clean.Hum_perc)
```

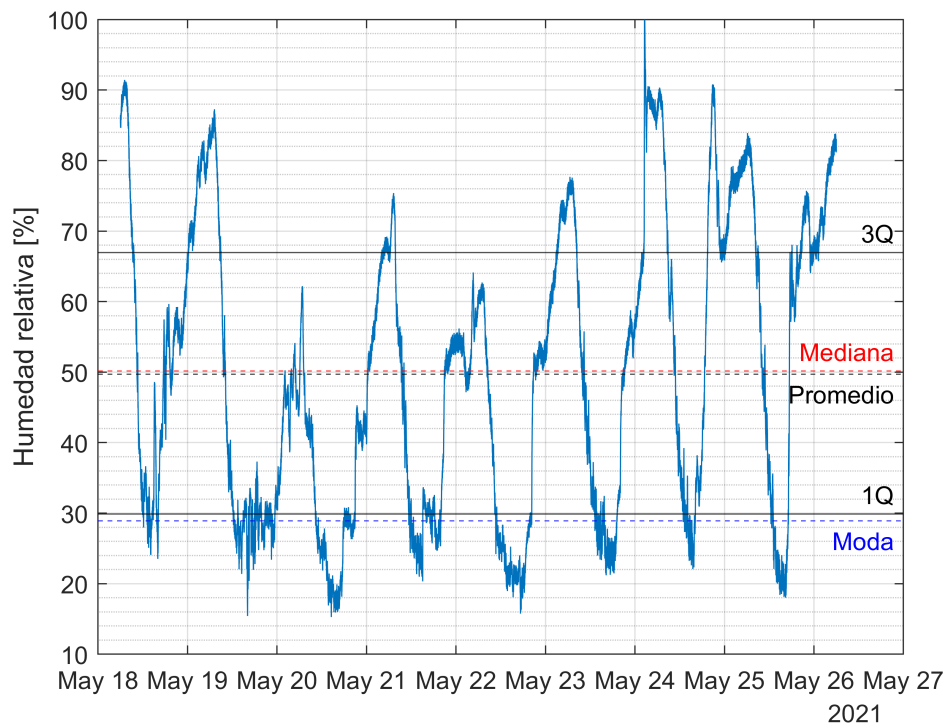
Warning: Unable to set 'Position', 'InnerPosition', 'OuterPosition', or 'ActivePositionProperty' for objects in a TiledChartLayout

```
xlabel("Humedad relativa")
ylabel("%")
nexttile
histogram(atmosfera_clean.Hum_perc,"BinWidth",5,'Normalization',"probability")
hold on
xline(hum_mean,"--k","Promedio","LabelHorizontalAlignment","left")
xline(hum_median,"--r","Mediana")
xline(hum_mode,"--b","Moda","LabelHorizontalAlignment","left")
xline(hum_1Q,"k","1Q")
xline(hum_3Q,"k","3Q")
hold off
xlabel("Humedad relativa [%]")
ylabel("Probabilidad")
```



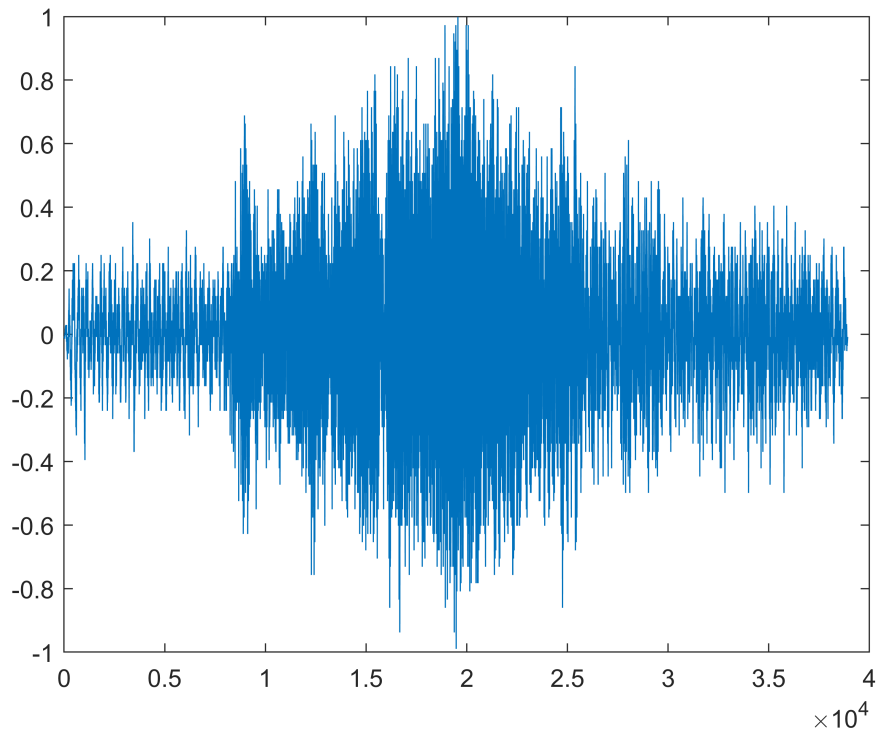
```
% Serie de tiempo
figure
plot(atmosfera_clean.DateTime,atmosfera_clean.Hum_perc)
hold on
yline(hum_mean,"--k","Promedio","LabelVerticalAlignment","bottom")
yline(hum_median,"--r","Mediana")
yline(hum_mode,"--b","Moda","LabelVerticalAlignment","bottom")
yline(hum_1Q,"k","1Q")
yline(hum_3Q,"k","3Q")
hold off
ylabel("Humedad relativa [%]")
grid on
grid minor
```





## Dispersión

```
buho_left=buho(:,1);
plot(buho_left)
```

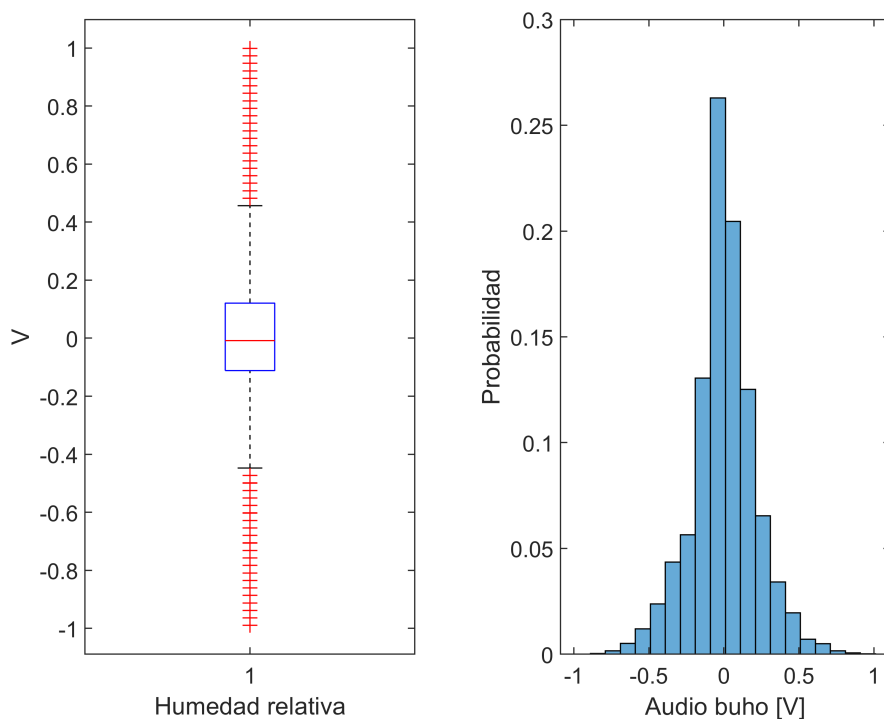


```
% Una gráfica con boxplot e histograma
```

```
figure
tiledlayout(1,2)
nexttile
boxplot(buho_left)
```

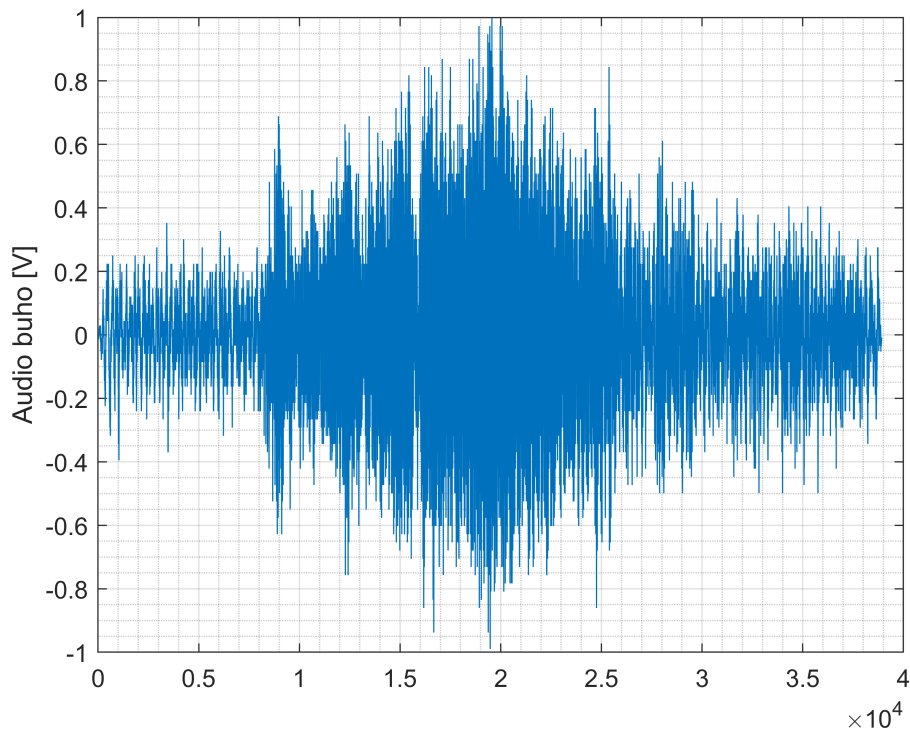
Warning: Unable to set 'Position', 'InnerPosition', 'OuterPosition', or 'ActivePositionProperty' for objects in a TiledChartLayout

```
xlabel("Humedad relativa")
ylabel("V")
nexttile
histogram(buho_left,20,'Normalization','probability')
xlabel("Audio buho [V]")
ylabel("Probabilidad")
```



```
% Serie de tiempo
```

```
figure
plot(buho_left)
ylabel("Audio buho [V]")
grid on
grid minor
```



## Ejercicio

```
buho_desv=std(buho_left)
```

```
buho_desv = 0.2124
```

```
buho_var=var(buho_left)
```

```
buho_var = 0.0451
```

```
buho_mean=mean(buho_left)
```

```
buho_mean = 8.1547e-08
```

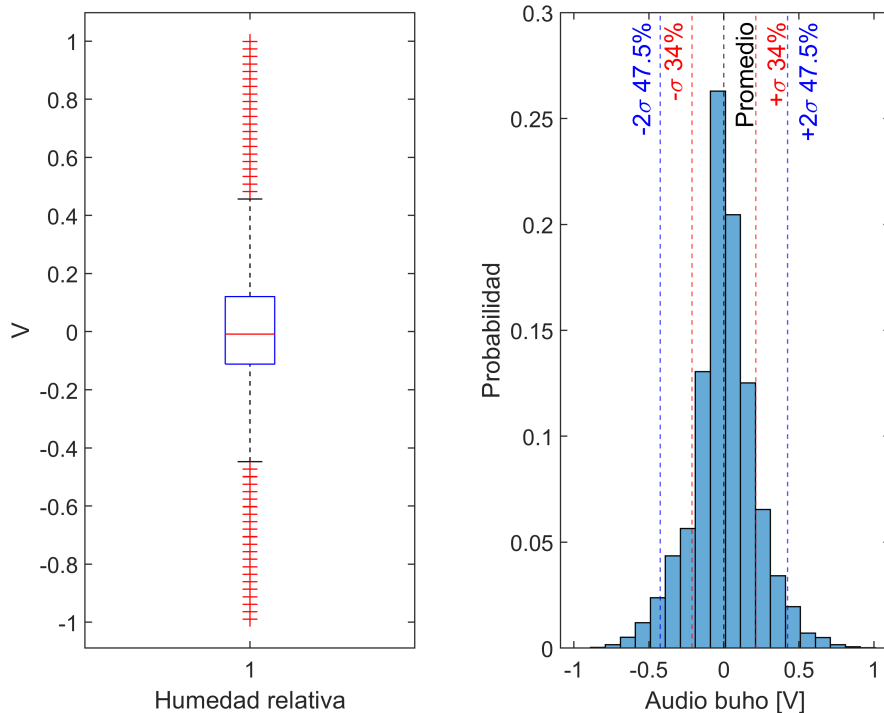
```
% Una gráfica con boxplot e histograma
```

```
figure
tiledlayout(1,2)
nexttile
boxplot(buho_left)
```

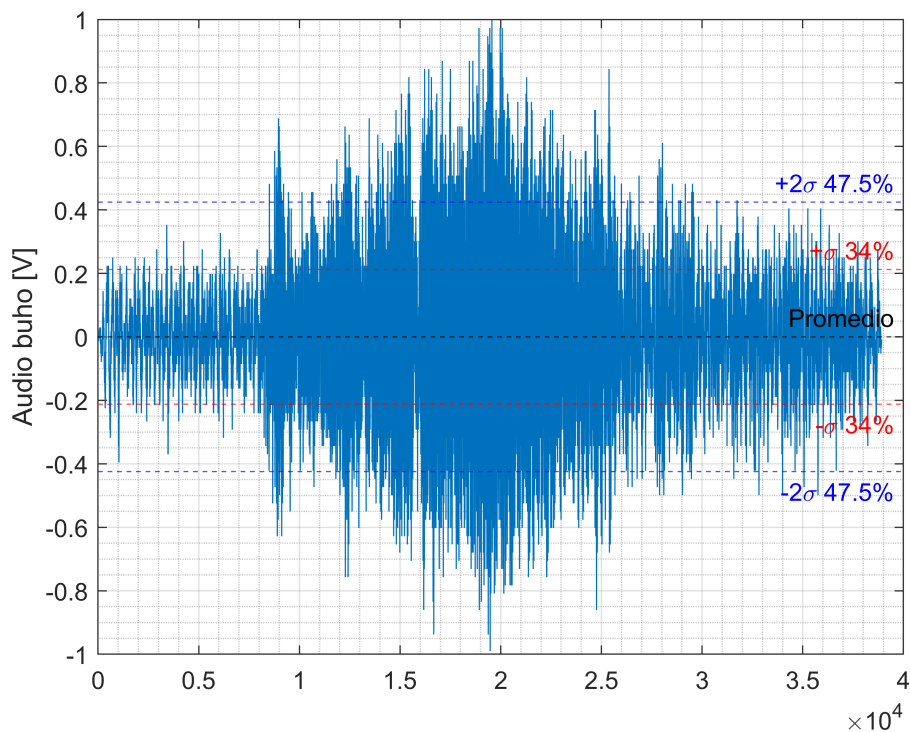
Warning: Unable to set 'Position', 'InnerPosition', 'OuterPosition', or 'ActivePositionProperty' for objects in a TiledChartLayout

```
xlabel("Humedad relativa")
ylabel("V")
nexttile
histogram(buho_left,20,'Normalization','probability')
xlabel("Audio buho [V]")
ylabel("Probabilidad")
hold on
xline(buho_mean,"--k","Promedio")
```

```
xline(buho_mean+buho_desv,"--r","+\\sigma 34%")
xline(buho_mean-buho_desv,"--r","-\\sigma 34%", "LabelHorizontalAlignment","left")
xline(buho_mean+2*buho_desv,"--b","+2\\sigma 47.5%")
xline(buho_mean-2*buho_desv,"--b","-2\\sigma 47.5%", "LabelHorizontalAlignment","left")
hold off
```



```
% Serie de tiempo
figure
plot(buho_left)
hold on
yline(buho_mean,"--k","Promedio")
yline(buho_mean+buho_desv,"--r","+\\sigma 34%")
yline(buho_mean-buho_desv,"--r","-\\sigma 34%", "LabelVerticalAlignment","bottom")
yline(buho_mean+2*buho_desv,"--b","+2\\sigma 47.5%")
yline(buho_mean-2*buho_desv,"--b","-2\\sigma 47.5%", "LabelVerticalAlignment","bottom")
hold off
ylabel("Audio buho [V]")
grid on
grid minor
```



## Ajustes

```
plot(atmosfera_clean.Temp_C,atmosfera_clean.Hum_perc, ".")
xlabel("Temperatura [C]")
ylabel("Humedad [%]")
grid on
[fit_realacion,gof]=fit(atmosfera_clean.Temp_C,atmosfera_clean.Hum_perc,"poly1")
```

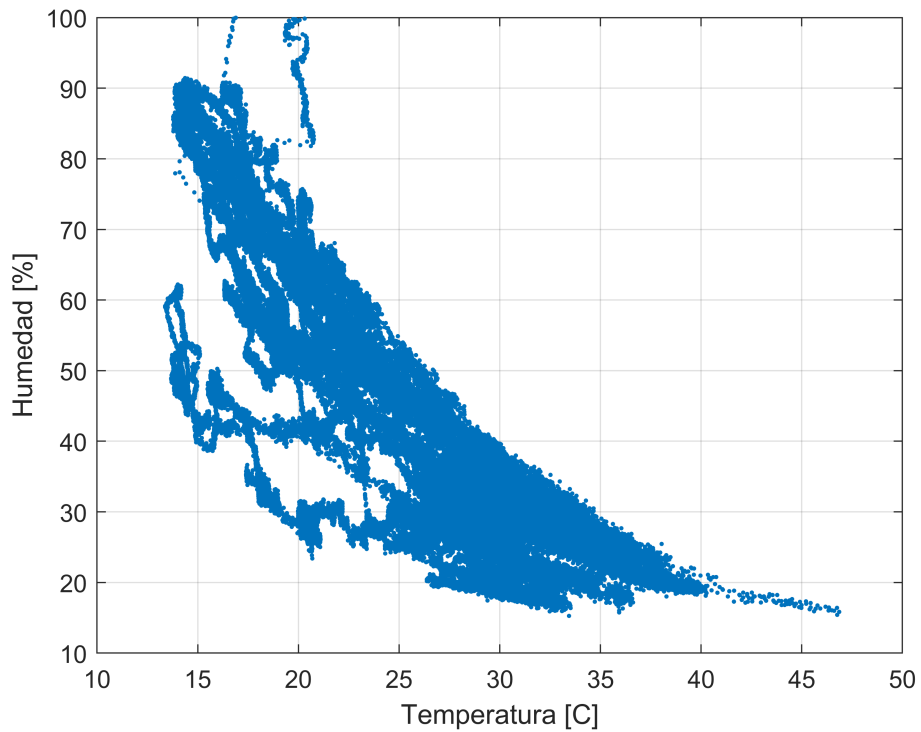
```
fit_realacion =
    Linear model Poly1:
    fit_realacion(x) = p1*x + p2
    Coefficients (with 95% confidence bounds):
        p1 =      -2.772  (-2.781, -2.763)
        p2 =      113.2   (112.9, 113.4)
gof = struct with fields:
    sse: 1.5967e+07
    rsquare: 0.7212
    dfe: 137448
    adjrsquare: 0.7212
    rmse: 10.7783
```

```
[fit_realacion_2,gof]=fit(atmosfera_clean.Temp_C,atmosfera_clean.Hum_perc,"poly2")
```

```
fit_realacion_2 =
    Linear model Poly2:
    fit_realacion_2(x) = p1*x^2 + p2*x + p3
    Coefficients (with 95% confidence bounds):
        p1 =      0.088  (0.08654, 0.08946)
        p2 =     -7.097  (-7.169, -7.024)
        p3 =     162.6  (161.8, 163.5)
gof = struct with fields:
    sse: 1.4498e+07
    rsquare: 0.7468
```

dfe: 137447  
adjrsquare: 0.7468  
rmse: 10.2703

```
figure  
plot(atmosfera_clean.Temp_C,atmosfera_clean.Hum_perc,".")  
hold on  
plot(fit_realacion,'r')  
plot(fit_realacion_2,'b')
```



```
hold off  
xlabel("Temperatura [C]")  
ylabel("Humedad [%]")  
legend("Data","Ajuste de línea","Ajuste polinomio grado 2")  
grid on
```

