



Instituto Tecnológico  
de Buenos Aires

29/03/2023

**CASO DE ESTUDIO**

**PREDICCIÓN DE LLUVIA EN AUSTRALIA**

# Agenda

01

## Contexto

Introducción a la base.

03

## Outliers

Justificación sobre valores atípicos.

05

## Correlación

Análisis y conclusión sobre Correlación de Spearman.

02

## Missings

Tratamiento de valores faltantes.

04

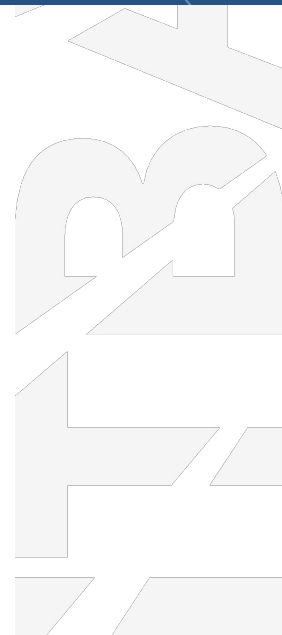
## Gráficos exploratorios

Respondiendo a preguntas sobre la base.

06

## Conclusión

Opinión final + acciones hacia el futuro.



# Contexto

## Las inundaciones en Australia empeoran mientras miles de personas debieron huir de sus hogares en Sidney

Algunas regiones ya acumulan la cantidad de lluvias que deberían recibir en el transcurso de todo el año, pero solamente en tres días

## La lluvia más fuerte en 30 años apaga incendios forestales en Sydney

Las fuertes lluvias en Nueva Gales del Sur, Australia, obligaron a la evacuación de varias ciudades, pero también extinguieron algunos de los persistentes incendios forestales que han devastado gran parte de esa región.

Por [CNN Español](#) Publicado a las 15:04 ET (19:04 GMT) 10 febrero, 2020

## Las fuertes lluvias ayudan a apagar los incendios en Australia pero pueden provocar problemas

COMENTARIOS

Por [Rafael Cerceda](#) • última actualización: 16/01/2020



## Australia faces 'explosive' fire risk after years of rain: Report



Emergency services personnel fighting a fire near Margaret River on Dec 6, 2021. A report is warning that Australia is at the risk of major wildfires in 2023-2024. PHOTO: AFP/WESTERN AUSTRALIA DEPARTMENT OF FIRE AND EMERGENCY SERVICES

## More Flooding Misery for Weather-Weary Australia



A wallaby passes through floodwaters in Fitzroy Crossing, Australia Jan. 3, 2023 in this picture obtained from social media. Callum Lamond/Via REUTERS

# Contexto

kaggle



	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed	WindDir9am	WindDir3pm	WindSpeed9am	WindSpeed3pm	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm	Temp9am	Temp3pm	RainToday	RainTomorrow
1	2008-12-01	Albury	13.4	22.9	0.6	N/A	N/A	W	44	W	WNW	20	24	71	22	1007.7	1007.1	8	N/A	16.9	21.8	No	No
2	2008-12-02	Albury	7.4	25.1	0.0	N/A	N/A	WNW	44	NNW	WSW	4	22	44	25	1010.6	1007.8	N/A	N/A	17.2	24.3	No	No
3	2008-12-03	Albury	12.9	25.7	0.0	N/A	N/A	WSW	46	W	WSW	19	26	38	30	1007.6	1008.7	N/A	2	21.0	23.2	No	No
4	2008-12-04	Albury	9.2	28.0	0.0	N/A	N/A	NE	24	SE	E	11	9	45	16	1017.6	1012.8	N/A	N/A	18.1	26.5	No	No
5	2008-12-05	Albury	17.5	32.3	1.0	N/A	N/A	W	41	ENE	NW	7	20	82	33	1010.8	1006.0	7	8	17.8	29.7	No	No

- Fecha
- Ubicación
- Temperatura (°c)
  - Máxima
  - Mínima
  - 3pm
  - 9am
- Lluvia (mm)
- Evaporación (mm)
- Horas de sol

- Viento
  - Dirección más fuerte
  - Velocidad más fuerte (km/h)
  - Dirección 9am
  - Dirección 3pm
- Humedad (%)
  - 9am
  - 3pm
- Presión atmosférica (hpa)
  - 9am
  - 3pm

- Nubosidad (octas)
  - 9am
  - 3pm
- RainToday ("Yes" si excede 1 mm)
- **Rain Tomorrow**

# Missings

```

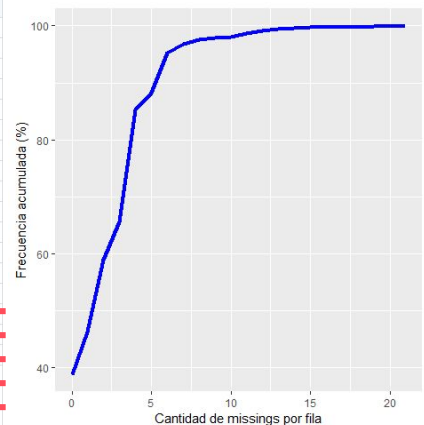
Date                Location                MinTemp                MaxTemp                Rainfall                Evaporation                Sunshine                WindGustDir
Min. : 2007-11-01   Length:145460         Min. : -8.50           Min. : -4.80           Min. : 0.000           Min. : 0.00           Min. : 0.00           Length:145460
1st Qu.: 2011-01-11   Class :character      1st Qu.: 7.60           1st Qu.: 17.90         1st Qu.: 0.000         1st Qu.: 2.60         1st Qu.: 4.80         Class :character
Median : 2013-06-02   Mode :character        Median : 12.00          Median : 22.60         Median : 0.000         Median : 4.80         Median : 8.40         Mode :character
Mean : 2013-04-04     Mean : 12.19           Mean : 23.22           Mean : 2.361           Mean : 5.47           Mean : 7.61
3rd Qu.: 2015-06-14   3rd Qu.: 16.90         3rd Qu.: 28.20         3rd Qu.: 0.800         3rd Qu.: 7.40         3rd Qu.: 10.60
Max. : 2017-06-25     Max. : 33.90           Max. : 48.10           Max. : 371.000         Max. : 145.00         Max. : 34.50
NA's : 1485           NA's : 1261           NA's : 62790          NA's : 69835

WindGustSpeed        WindDir9am                WindDir3pm                WindSpeed9am                WindSpeed3pm                Humidity9am                Humidity3pm                Pressure9am
Min. : 6.00           Length:145460         Length:145460         Min. : 0.00           Min. : 0.00           Min. : 0.00           Min. : 0.00           Min. : 980.5
1st Qu.: 31.00        Class :character      Class :character      1st Qu.: 7.00           1st Qu.: 13.00         1st Qu.: 57.00         1st Qu.: 37.00         1st Qu.: 1012.9
Median : 39.00        Mode :character      Mode :character      Median : 13.00         Median : 19.00         Median : 70.00         Median : 52.00         Median : 1017.6
Mean : 40.03          Mean : 14.04          Mean : 18.66           Mean : 68.88           Mean : 51.54           Mean : 1017.6
3rd Qu.: 48.00        3rd Qu.: 19.00        3rd Qu.: 24.00         3rd Qu.: 83.00         3rd Qu.: 66.00         3rd Qu.: 1022.4
Max. : 135.00         Max. : 130.00         Max. : 87.00           Max. : 100.00         Max. : 100.00         Max. : 1041.0
NA's : 10263          NA's : 1767          NA's : 3062           NA's : 4507           NA's : 15065

Pressure3pm           Cloud9am                Cloud3pm                Temp9am                Temp3pm                RainToday                RainTomorrow
Min. : 977.1          Min. : 0.00           Min. : 0.00           Min. : -7.20         Min. : -5.40           Length:145460         Length:145460
1st Qu.: 1010.4       1st Qu.: 1.00         1st Qu.: 2.00         1st Qu.: 12.30         1st Qu.: 16.60         Class :character      Class :character
Median : 1015.2       Median : 5.00         Median : 5.00         Median : 16.70         Median : 21.10         Mode :character      Mode :character
Mean : 1015.3         Mean : 4.45           Mean : 4.51           Mean : 16.99         Mean : 21.68
3rd Qu.: 1020.0       3rd Qu.: 7.00         3rd Qu.: 7.00         3rd Qu.: 21.60         3rd Qu.: 26.40
Max. : 1039.6         Max. : 9.00           Max. : 9.00           Max. : 40.20         Max. : 46.70
NA's : 15028         NA's : 53888         NA's : 59358         NA's : 1767         NA's : 3609
    
```



missings	cantidad	porcentaje	frecuencia_acumulada
0	56420	38.79	38.8
1	10909	7.50	46.3
2	18346	12.61	58.9
3	9782	6.72	65.6
4	28675	19.71	85.3
5	4141	2.85	88.2
6	10229	7.03	95.2
7	2277	1.57	96.8
8	1072	0.74	97.5
9	492	0.34	97.9
10	328	0.23	98.1
11	719	0.49	98.6
12	781	0.54	99.1
13	617	0.42	99.5
14	75	0.05	99.6
15	82	0.06	99.7
16	59	0.04	99.7
17	107	0.07	99.8
18	56	0.04	99.8
19	96	0.07	99.9
20	15	0.01	99.9
21	182	0.13	100.0



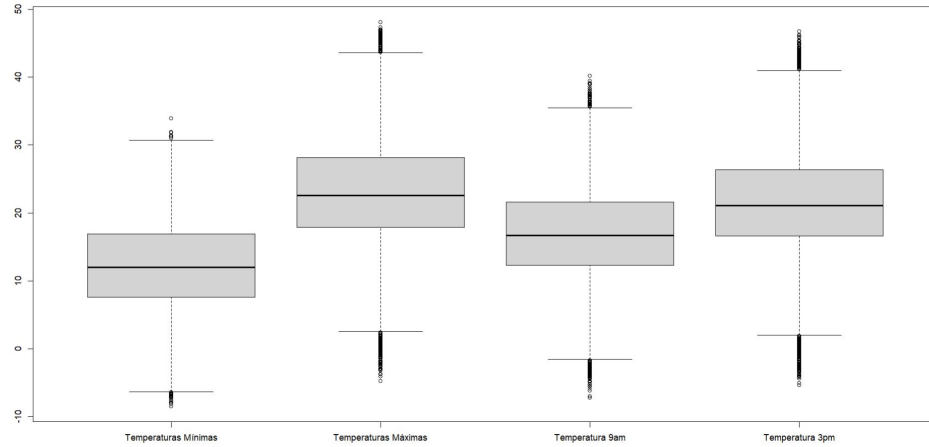
-**21** de las **23** columnas tenían al menos un NA.

-La cantidad total de filas pasó de **145.460** a **144.788** (sólo se redujo un 0,46%)

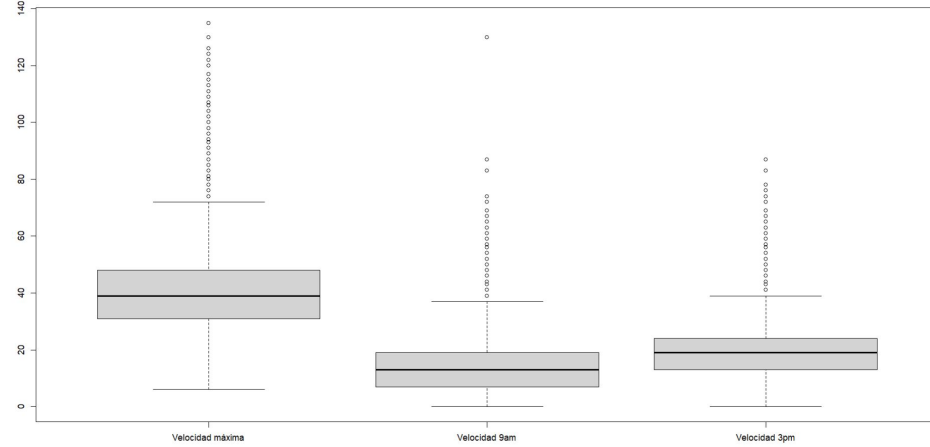
-**No** se detectó algún comportamiento no aleatorio en los datos eliminados (tanto en ubicación como en fecha)

# Outliers

Temperaturas (°c)

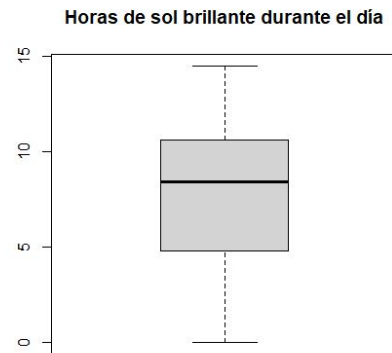
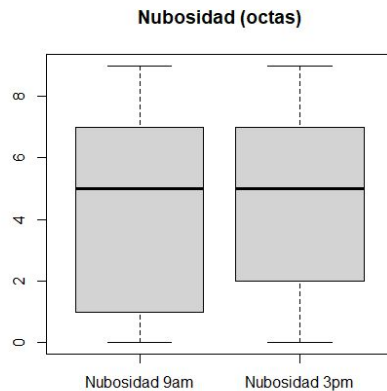
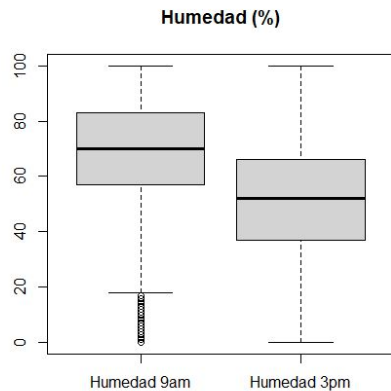


Velocidad del viento (km/h)



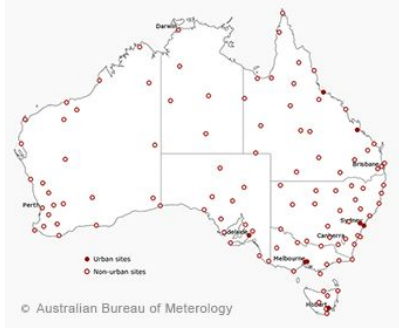
Variables como temperatura, velocidad del viento, presión atmosférica, precipitaciones o la evaporación no tienen un rango de valores (mínimos o máximos) que pueden tomar.

# Outliers

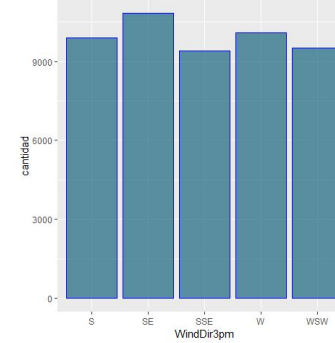
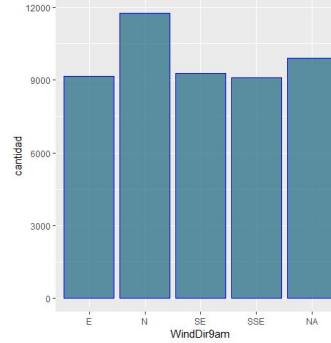


Variables como la humedad, la nubosidad o las horas de sol tienen un rango (mínimo y máximo) de valores que pueden tomar. Esto **no** significa que no puedan haber valores atípicos.

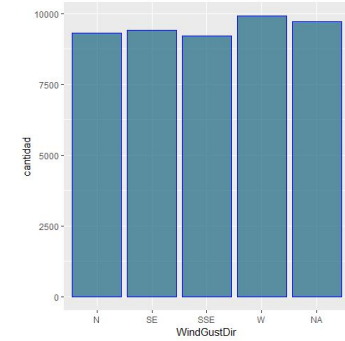
# Variables categóricas



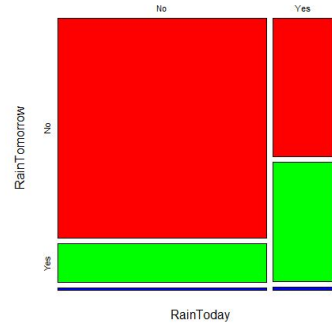
Ubicaciones



Dirección del viento



Rain Today vs. Rain Tomorrow

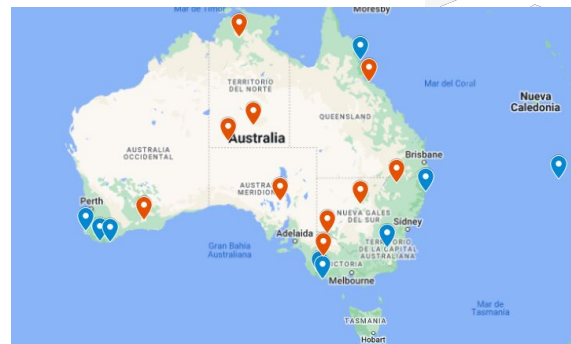
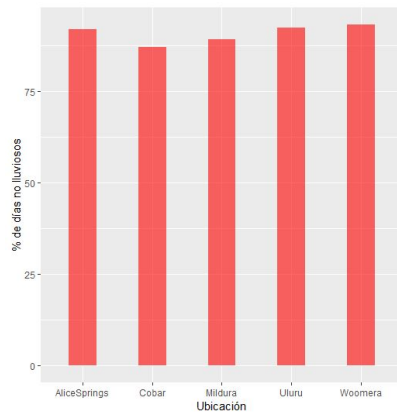
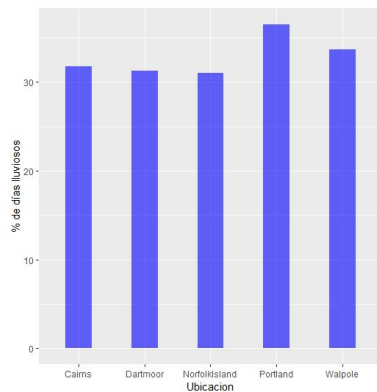


Relación entre RainToday y RainTomorrow



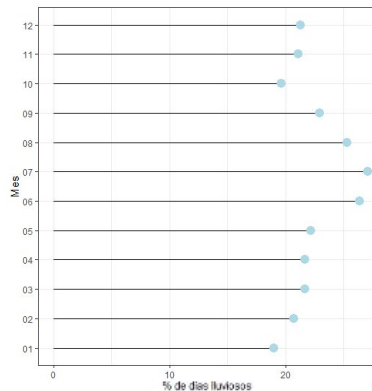
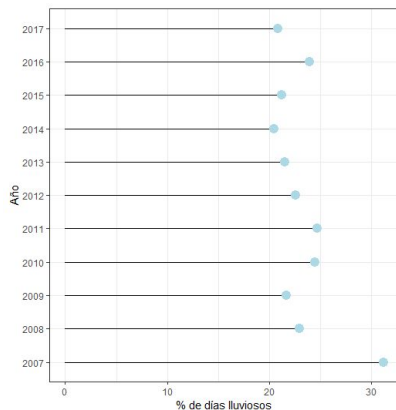
# Gráficos exploratorios

¿En qué lugares llueve más? ¿En cuáles menos?

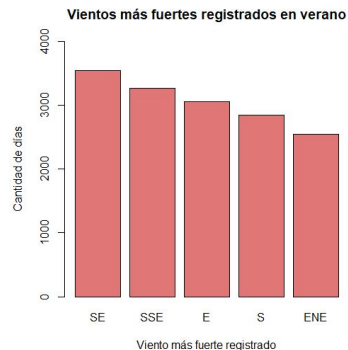


# Gráficos exploratorios

¿Hay alguna época del año donde llueva más? ¿Cómo se vienen manteniendo las lluvias desde el inicio de la base?

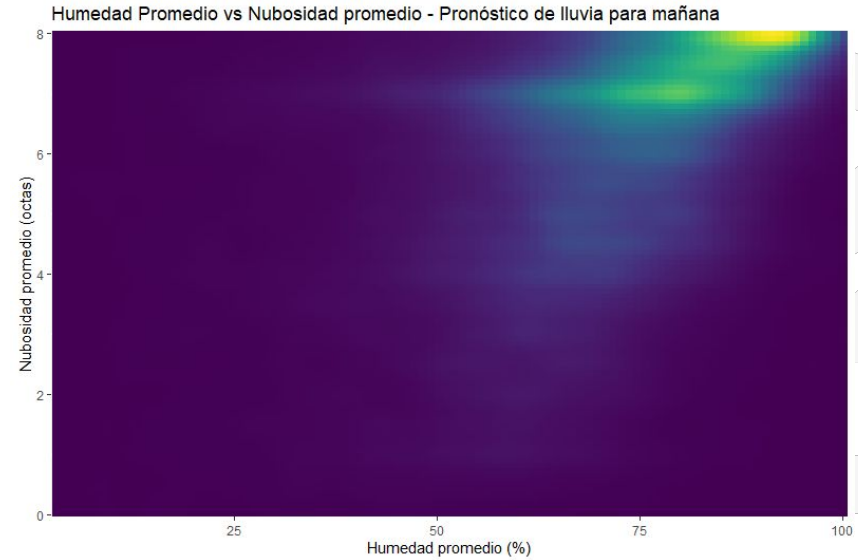
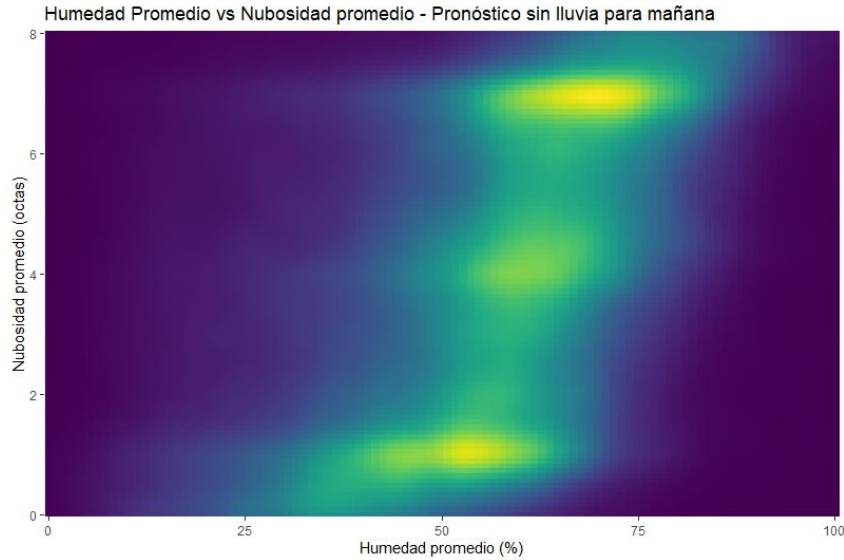


¿Cambia la dirección del viento más fuerte dependiendo de la época del año?



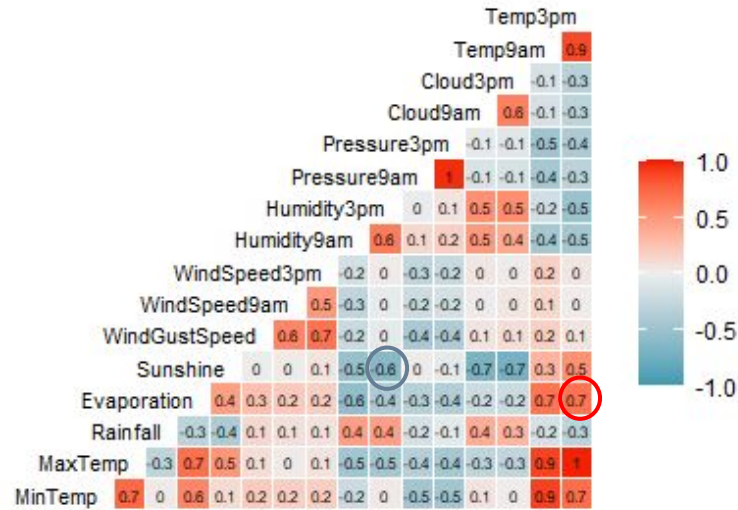
# Gráficos exploratorios

¿Cómo cambia la nubosidad y la humedad promedio dependiendo si se pronostica lluvia para el día siguiente?

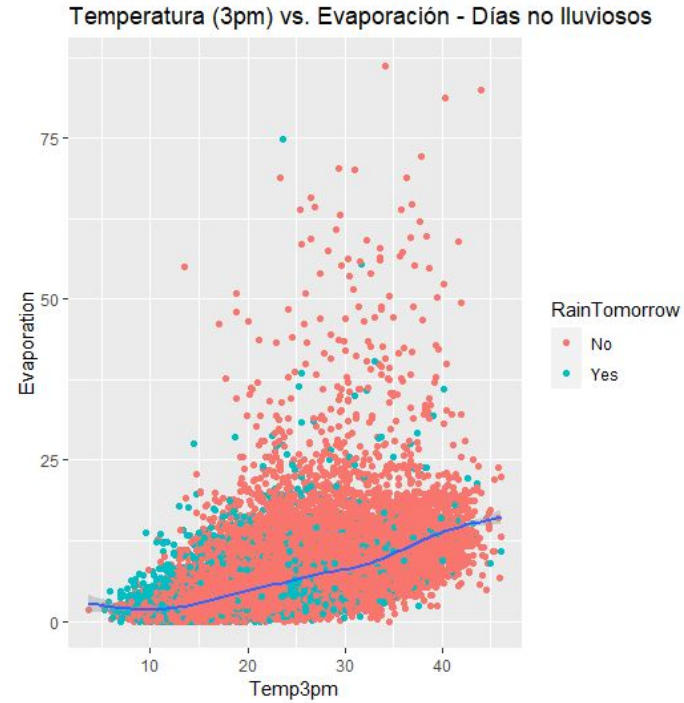
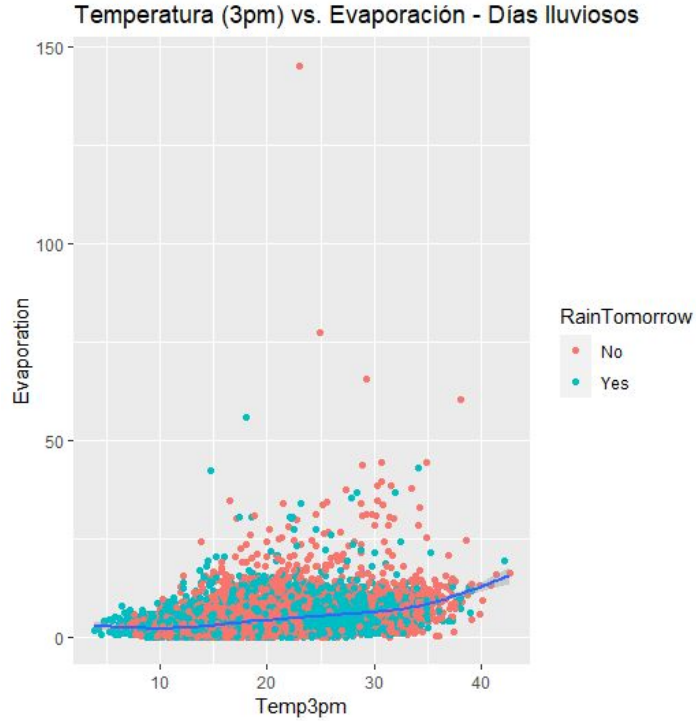


# Análisis de correlación

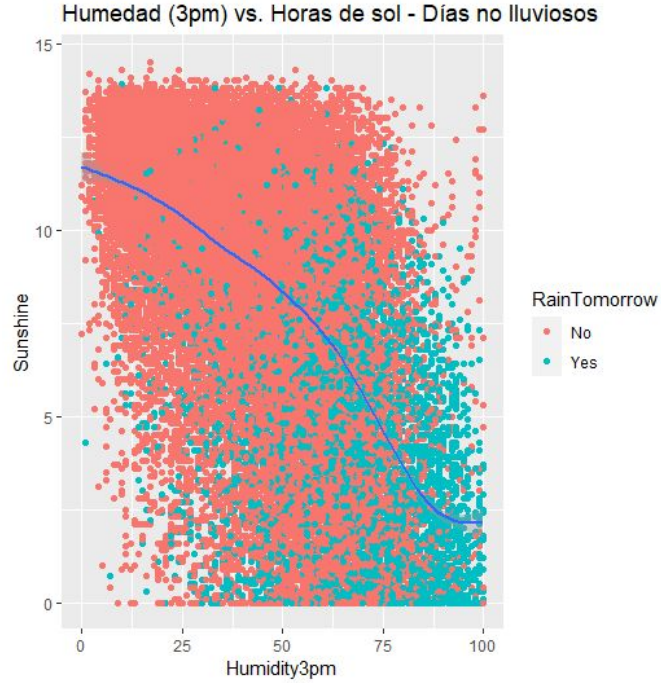
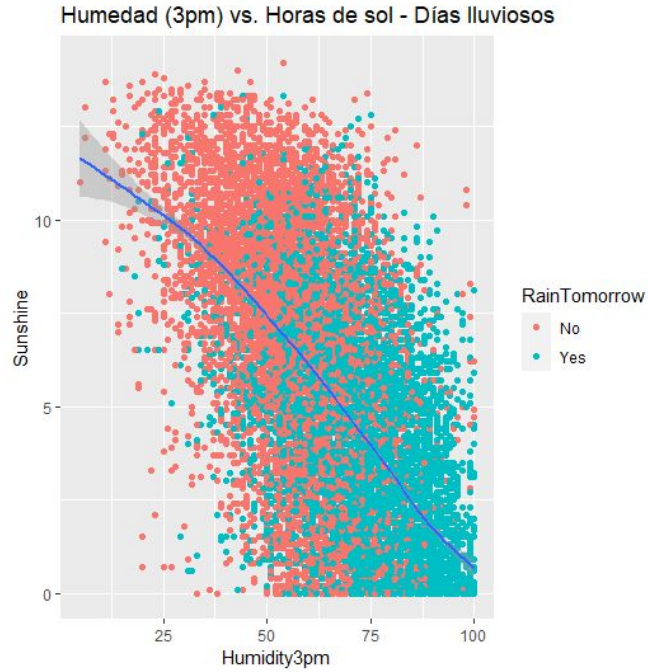
Método de Spearman



# Análisis de correlación



# Análisis de correlación



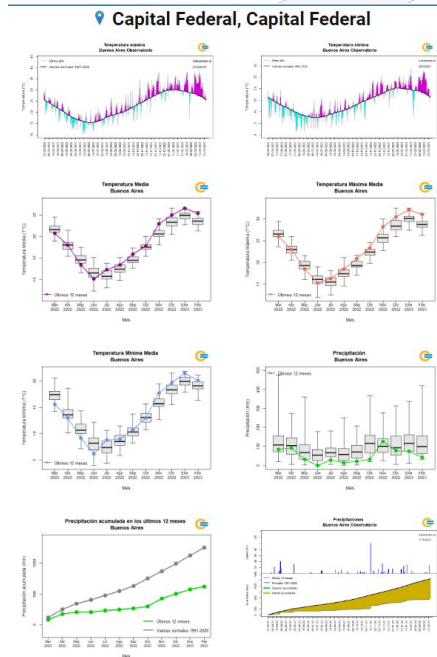
# Conclusión

Sobre la base:

- Hay ciertas variables que pueden incidir más que otras para predecir la columna RainTomorrow (por ejemplo, que ayer haya llovido).
- Puede ser viable crear una columna que contenga más información sobre los últimos días en los que haya llovido (por ejemplo: cantidad de días que llovió en la última semana)

Sobre el objetivo:

- Se podría llevar a cabo (con otras estadísticas) el mismo análisis en nuestro país  
→ <https://www.smn.gov.ar/clima/vigilancia>





Instituto Tecnológico  
de Buenos Aires

¡Gracias!