

Informe cambio climático a partir de registro de temperaturas máximas

El presente informe se ha realizado en base a los datos proporcionados por la Agencia Estatal de Meteorología AEMET, en su página Open Data:

[AEMET OpenData - Agencia Estatal de Meteorología - AEMET. Gobierno de España](#)

Se ha recogido información de los datos disponibles que se enumeran a continuación, según el escenario 23C3M y el proyecto AEMET, con salida en formato txt:

Variable	Escenario	Proyecto	Formato	
Todos Tª máxima Tª mínima Precipitación total acumulada Velocidad del viento a 10m Velocidad máxima del viento a 10m Humedad relativa Percentil 95 de la temperatura máxima diaria Percentil 5 de la temperatura mínima diaria Descargar TODOS los registros de datos de AEMET	Todos HISTORICAL 20C3M C1L RCP4 5 RCP8 5 B1 A1B A2	Todos AEMET ENSEMBLES ESCENA ESTCENA CORDEX	Todos TXT SIG	

Variable	Escenario	Proyecto	Formato	Fichero
Tª máxima	20C3M	AEMET	TXT	Descargar fichero (4 MB)
Tª mínima	20C3M	AEMET	TXT	Descargar fichero (4 MB)
Precipitación total acumulada	20C3M	AEMET	TXT	Descargar fichero (3 MB)
Velocidad del viento a 10m	20C3M	AEMET	TXT	Descargar fichero (3 MB)
Velocidad máxima del viento a 10m	20C3M	AEMET	TXT	Descargar fichero (3 MB)
Humedad relativa	20C3M	AEMET	TXT	Descargar fichero (4 MB)
Percentil 95 de la temperatura máxima diaria	20C3M	AEMET	TXT	Descargar fichero (421 KB)
Percentil 5 de la temperatura mínima diaria	20C3M	AEMET	TXT	Descargar fichero (386 KB)
Percentil 95 de la precipitación diaria	20C3M	AEMET	TXT	Descargar fichero (415 KB)
Precipitación máxima en 24h	20C3M	AEMET	TXT	Descargar fichero (4 MB)
Nº de días con precipitación <1mm	20C3M	AEMET	TXT	Descargar fichero (1 MB)
Nº de días con precipitación >20mm	20C3M	AEMET	TXT	Descargar fichero (612 KB)
Máximo Nº de días consecutivos con precipitación <1mm	20C3M	AEMET	TXT	Descargar fichero (147 KB)

En esta primera fase del estudio, se han descargado los datos de temperatura máxima, que corresponden al valor promedio mensual, y se van a analizar utilizando técnicas de machine learning para analizar la influencia que tiene la temperatura máxima en el cambio climático.

En otras fases del estudio se repetirá este análisis con otros valores registrados como temperatura mínima, precipitación, velocidad del viento o humedad relativa.

El análisis se ha realizado mediante técnicas de machine learning, aplicadas sobre los datos recogidos en más de 2500 estaciones meteorológicas ubicadas en la España peninsular y Baleares desde 1960, y en este documento se presentan los resultados más relevantes.

Ha resultado imposible asociar cada estación meteorológica con una ubicación concreta y sus coordenadas (longitud, latitud, altura) porque el archivo maestro que debe asociar cada ubicación con una estación por medio del código AEMET tiene 1124 filas y las estaciones que aparecen en la tabla de temperaturas máximas por año/mes son 1446.

Alguna estación está repetida, hay estaciones situadas en Canarias cuando la documentación indica que la cuadrícula sólo incluye la península y Baleares, resulta complicado identificar si hay valores repetidos en las estaciones y faltan muchas lecturas que se han corregido a mano de forma que se ha complicado el análisis, pero finalmente se podido realizar el estudio.

La primera visualización correcta de los datos ha sido la que se ve a continuación, se aprecia que falta un encabezado que indique lo que hay en cada una de las columnas, de modo que la primera fila aparece como encabezado.

Antes se ha visualizado de manera errónea y se ha visto que este dataset, de temperatura máxima, estaba enlazado con el fichero maestro de localizaciones.

data.head()	
	196101 16.08 16.33 16.08.1 16.08.2 16.22 16.12 16.01 16.36 16.50 ... 11.79.4 11.79.5 12.54.3 12.54.4 11.36.3 10.84.1 10.26 13.19.2 13.00.1 12.98
0	196102 15.29 15.45 15.56 15.56 15.56 15.31 14.84 16.13 16.09 ... 10.76 10.70 11.37 11.37 10.34 9.81 9.34 12.40 12.23 12.13
1	196103 15.88 16.35 16.64 16.64 17.47 17.03 15.89 17.33 18.14 ... 11.03 11.03 11.69 11.69 10.63 9.96 9.35 12.55 12.35 12.18
2	196104 16.13 17.15 16.57 16.57 17.96 17.82 16.24 17.27 17.71 ... 11.55 11.71 12.47 12.47 11.16 10.46 9.68 12.65 12.36 12.30
3	196105 18.68 20.01 20.31 20.31 23.24 22.28 18.15 22.17 23.90 ... 13.10 13.18 14.29 14.29 12.56 11.63 10.63 14.28 14.13 14.06
4	196106 22.32 22.93 24.75 24.75 27.97 25.82 20.41 27.83 30.61 ... 15.39 15.38 16.39 16.39 14.53 13.57 12.53 16.33 16.13 15.85
5 rows × 1446 columns	

Primera visualización tabla Tª máxima

Identificación de Localidad a partir del Codigo AEMET

data2.shape	
(1124, 5)	
data2.tail()	

	CodigoAEMET		Localidad	Longitud	Latitud	Altura
1119	C449G		ANAGA-SAN ANDRES	-16.1839	28.4967	20
1120	C458A		TACORONTE-A. S.E.A	-16.4092	28.4864	327
1121	C649I	TELDE/AERÓPUERTO DE G.CANARIA (GANDO)		-15.3889	27.9292	24
1122	C659P	LAS PALMAS DE GRAN CANARIA-JUNTA DE OBRAS DEL...		-15.4167	28.15	15
1123	C665I		VALLESECO-CASCO	-15.5736	28.0472	980

Visualización fichero maestro

Se ha añadido a la tabla un encabezado con la fecha y un número de estación meteorológica para poder identificarlas y además se ha dado a todas las celdas formato de número, a excepción de la columna FECHA.

Esta última operación se ha realizado porque se ha visto que muchas celdas tenían formato fecha y aparecían alineadas a la derecha en lugar de alinearse a la izquierda como las celdas con formato número.

TXMM_RCA_ECHAM5_20C3M_196101_200012	
21/09/2022	
A	B
C	D
E	F
G	H
I	J
K	L
M	N
O	P
Q	R
S	T
U	V
W	X
Y	Z
AA	AB
AC	AD
AE	AF
AG	AH
AI	AJ
AK	AL
AM	AN
AO	AP
AQ	AR
AS	AT
AV	AW
AX	AY
AZ	BA
BB	BC
BD	BE
BF	BG
BH	BI
BJ	BL
BM	BN
BO	BP
BQ	BR
BS	BT
BU	BV
BW	BX
BY	BZ
CA	CB
CC	CD
CE	CF
CG	CH
CI	CJ
CK	CL
CM	CN
CO	CP
CQ	CR
CS	CT
CU	CV
CV	CX
CW	CY
CX	CZ
CY	DA
CA	DB
CB	DC
CC	DD
CD	DE
CE	DF
CF	DG
CG	DH
CH	DI
CI	DJ
CJ	DK
CK	DL
CL	DM
CM	DN
CO	DO
CP	DP
CQ	DQ
CR	DR
CS	DS
CT	DT
CU	DU
CV	DV
CW	DW
CX	DX
CY	DY
CZ	EA
DA	EB
DB	EC
DC	ED
DD	EE
DE	EF
DF	EG
DG	EH
DH	EI
DI	EJ
DJ	EL
DK	EM
DL	EN
DM	EO
DN	EP
DO	EQ
DP	ER
DQ	ES
DR	ET
DS	EU
DT	EV
DU	EW
DV	EX
DW	EY
DX	EZ
EY	FA
EZ	FB
FA	FC
FB	FD
FC	FE
FD	FF
FE	FG
FF	FH
FG	FI
FH	FJ
FI	FL
FJ	FM
FK	FN
FL	FO
FM	FP
FN	FQ
FO	FR
FP	FS
FQ	FT
FR	FU
FS	FV
FT	FW
FU	FX
FV	FY
FW	FZ
FX	GA
FY	GB
FZ	GC
GA	GD
GB	GE
GC	GF
GD	GG
GE	GH
GF	GI
GG	GJ
GH	GL
GI	GM
GJ	GN
GL	GO
GM	GP
GN	GQ
GO	GR
GP	GS
GQ	GT
GR	GU
GS	GV
GT	GW
GU	GX
GV	GY
GW	GZ
GX	HA
GY	HB
GZ	HC
HA	HD
HB	HE
HC	HF
HD	HG
HE	HH
HF	HI
HG	HJ
HH	HL
HI	HM
HJ	HN
HK	HO
HL	HP
HM	HQ
HN	HR
HO	HS
HP	HT
HQ	HU
HR	HV
HS	HW
HT	HX
HU	HY
HV	HZ
HW	IA
HX	IB
HY	IC
HZ	ID
IA	IE
IB	IF
IC	IG
ID	IH
IE	II
IF	IJ
IG	IL
IH	IM
II	IN
IJ	IO
IK	IP
IL	IQ
IM	IR
IN	IS
IO	IT
IP	IU
IQ	IV
IR	IW
IS	IX
IT	IY
IU	IZ
IV	JA
IW	JB
IX	JC
IY	JD
IZ	JE
JA	JF
JB	JG
JC	JH
JD	JI
JE	JJ
JF	JK
JG	KL
JH	JM
JI	JN
JJ	JO
JK	JP
KL	JQ
JM	JR
JN	JS
JO	JT
JP	JU
JQ	JV
JR	JW
JS	JX
JT	JY
JU	JZ
JV	KA
JW	KB
JX	KC
JY	KD
JZ	KE
KA	KF
KB	KG
KC	KH
KD	KI
KE	KJ
KF	KK
KG	KL
KH	KM
KI	KN
KJ	KO
KK	KP
KL	KQ
KM	KR
KN	KS
KO	KT
KP	KU
KQ	KV
KR	KW
KS	KX
KT	KY
KU	KZ
KV	LA
KW	LB
KX	LC
KY	LD
KZ	LE
LA	LF
LB	LG
LC	LH
LD	LI
LE	LJ
LF	LK
LG	LL
LH	LM
LI	LN
LJ	LO
LK	LP
LL	LQ
LM	LR
LN	LS
LO	LT
LP	LU
LQ	LV
LR	LW
LS	LX
LT	LY
LU	LZ
LV	MA
LW	MB
LX	MC
LY	MD
LZ	ME
MA	MF
MB	MG
MC	MH
MD	MI
ME	MJ
MF	MK
MG	ML
MH	MM
MI	MN
MJ	MO
MK	MP
ML	MQ
MM	MR
MN	MS
MO	MT
MP	MU
MQ	MV
MR	MW
MS	MX
MT	MY
MU	MZ
MV	NA
MW	NB
MX	NC
MY	ND
MZ	NE
NA	NF
NB	NG
NC	NH
ND	NI
NE	NJ
NF	NK
NG	NL
NH	NM
NI	NN
NJ	NO
NK	NP
NL	NQ
NM	NR
NN	NS
NO	NT
NP	NU
NQ	NV
NR	NW
NS	NX
NT	NY
NU	NZ
NV	OA
NW	OB
NX	OC
NY	OD
NZ	OE
OA	OF
OB	OG
OC	OH
OD	OI
OE	OJ
OF	OK
OG	OL
OH	OM
OI	ON
OJ	OO
OK	OP
OL	OQ
OM	OR
ON	OS
OO	OT
OP	OU
OQ	OV
OR	OW
OS	OX
OT	OY
OU	OZ
OV	PA
OW	PB
OX	PC
OY	PD
OZ	PE
PA	PF
PB	PG
PC	PH
PD	PI
PE	PJ
PF	PK
PG	PL
PH	PM
PI	PN
PJ	PO
PK	PP
PL	PQ
PM	PR
PN	PS
PO	PT
PP	PU
PQ	PV
PR	PW
PS	PX
PT	PY
PU	PZ
PV	QA
PW	QB
PX	QC
PY	QD
PZ	QE
QA	QF
QB	QG
QC	QH
QD	QI
QE	QJ
QF	QK
QG	QL
QH	QM
QI	QN
QJ	QO
QK	QP
QL	QQ
QM	QR
QN	QS
QO	QT
QP	QU
QQ	QV
QR	QW
QS	QX
QT	QY
QU	QZ
QV	RA
QW	RB
QX	RC
QY	RD
QZ	RE
RA	RF
RB	RG
RC	RH
RD	RI
RE	RJ
RF	RK
RG	RL
RH	RM
RI	RN
RJ	RO
RK	RP
RL	RQ
RM	RR
RN	RS
RO	RT
RP	RU
RQ	RV
RR	RW
RS	RX
RT	RY
RU	RZ
RV	SA
RW	SB
RX	SC
RY	SD
RZ	SE
SA	SF
SB	SG
SC	SH
SD	SI
SE	SJ
SF	SK
SG	SL
SH	SM
SI	SN
SJ	SO
SK	SP
SL	SQ
SM	SR
SN	SS
SO	ST
SP	SU
SQ	SV
SR	SW
SS	SX
ST	SY
SU	SZ
SV	TA
SW	TB
SX	TC
SY	TD
SZ	TE
TA	TF
TB	TG
TC	TH
TD	TI
TE	TJ
TF	TK
TG	TL
TH	TM
TI	TN
TJ	TO
TK	TP
TL	TQ
TM	TR
TN	TS
TO	TT
TP	TU
TQ	TV
TR	TW
TS	TX
TT	TY
TU	TZ
TV	UA
TW	UB
TX	UC
TY	UD
TZ	UE
UA	UF
UB	UG
UC	UH
UD	UI
UE	UJ
UF	UK
UG	UL
UH	UM
UI	UN
UJ	UO
UK	UP
UL	UQ
UM	UR
UN	US
UO	UT
UP	UU
UQ	UV
UR	UW
US	UX
UT	UY
UU	UZ
UV	VA
UW	VB
UX	VC
UY	VD
UZ	VE
VA	VF
VB	VG
VC	VH
VD	VI
VE	VJ
VF	VK
VG	VL
VH	VM
VI	VN
VJ	VO
VK	VP
VL	VQ
VM	VR
VN	VS
VO	VT
VP	VU
VQ	VV
VR	VW
VS	VX
VT	VY
VU	VZ
VV	WA
VW	WB
VX	WC
VY	WD
VZ	WE
WA	WF
WB	WG
WC	WH
WD	WI
WE	WJ
WF	WK
WG	WL
WH	WM
WI	WN
WJ	WO
WK	WP
WL	WQ
WM	WR
WN	WS
WO	WT
WP	WU
WQ	WV
WR	WW
WS	WX
WT	WY
WU	WZ
WV	XA
WW	XB
WX	XC
WY	XD
WZ	XE
XA	XF
XB	YG
XC	YH
XD	YI
XE	YJ
XF	YK
YG	YL
YH	YM
YI	YN
YJ	YO
YK	YP
YL	YQ
YM	YR
YN	YS
YO	YT
YP	YU
YQ	YV
YR	YW
YS	YX
YT	YY
YU	YZ
YV	ZA
YW	ZB
YX	ZC
YY	ZD
YZ	ZE
ZA	ZF
ZB	ZG
ZC	ZH
ZD	ZI
ZE	ZJ
ZF	ZK
ZG	ZL
ZH	ZM
ZI	ZN
ZJ	ZO
ZK	ZP
ZL	ZQ
ZM	ZR
ZN	ZS
ZO	ZT
ZP	ZU
ZQ	ZV
ZR	ZW
ZS	ZX
ZT	ZY
ZU	ZZ
ZV	AA
ZW	AB
ZX	AC
ZY	AD
ZZ	AE
AA	AF
AB	AG
AC	AH
AD	AI
AE	AJ
AF	AK
AG	AL
AH	AM
AI	AN
AJ	AO
AK	AP
AL	AQ
AM	AR
AN	AS
AO	AT
AP	AU
AQ	AV
AR	AW
AS	AX
AT	AY
AU	AZ
AV	BA
AW	BB
AX	BC
AY	BD
AZ	BE
BA	BF
BB	BG
BC	BH
BD	BI

Una vez realizados estos cambios se ha obtenido la siguiente visualización en la que se aprecia claramente los valores que han sido añadidos.

data1.head()

	FECHA	500	501	502	503	504	505	506	507	508	...	1935	1936	1937	1938	1939	1940	1941	1942
0	196101		16.33	44.789,00	44.789,00	16.22	44.911,00	44.577,00	16.36	16.50	...	11.79	11.79	12.54	12.54	11.36	10.84	10.26	13.19
1	196102	15.29	15.45	15.56	15.56	15.56	15.31	14.84	16.13	44.820,00	...	10.76	10.70	11.37	11.37	10.34	9.81	9.34	12.40
2	196103	15.88	16.35	16.64	16.64	17.47	44.637,00	15.89	17.33	18.14	...	44.631,00	44.631,00	11.69	11.69	10.63	9.96	9.35	12.55
3	196104	16.13	17.15	16.57	16.57	17.96	17.82	16.24	17.27	17.71	...	11.55	11.71	12.47	12.47	11.16	10.46	9.68	12.65
4	196105	18.68	44.581,00	20.31	20.31	23.24	22.28	18.15	22.17	23.90	...	44.847,00	13.18	14.29	14.29	12.56	11.63	10.63	14.28

5 rows × 1446 columns

data1.shape

(480, 1446)

De hecho, si se abre con un bloc de notas (Notepad++) se ve como esos datos que se han añadidos posteriormente aparecen como datos NaN (Not a Number).

Archivo

Editar

Buscar

Vista

Codificación

Lenguaje

Configuración

Herramientas

Macro

Ejecutar

Plugins

Ventana

?

TM5M_RCA_HADCM300_30CM_196101_200012.csv

150	196110	19.82	19.90	20.90	20.90	21.83	20.63	18.68	22.69	23.68	23.68	22.86	21.57	19.50	18.19	21.27	24.79	25.00	25.00	23.75	2
151	196111	14.00	14.20	14.49	14.49	15.03	14.59	12.97	15.24	15.74	15.74	14.74	14.14	13.35	12.33	14.36	16.20	16.16	16.16	14.67	1
152	196112	11.23	11.78	11.89	11.89	13.11	12.56	10.00	12.76	13.95	13.95	13.04	12.57	11.48	9.25	11.73	14.03	14.50	14.50	13.25	1
153	196201	10.43	10.51	11.21	11.21	11.88	11.24	9.27	12.41	13.63	13.63	12.35	11.55	10.57	8.63	11.10	14.07	14.66	14.66	13.02	1
154	196202	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
155	196203	14.11	14.69	15.22	15.22	16.39	15.72	13.66	16.60	17.72	17.72	16.66	15.46	14.31	12.96	15.38	18.35	18.52	18.52	16.78	1
156	196204	19.29	20.50	20.93	20.93	23.48	22.15	18.65	22.48	24.06	24.06	24.20	23.28	20.22	18.28	20.81	24.61	24.75	24.75	23.84	2
157	196205	21.39	22.63	22.72	22.72	25.27	24.23	21.34	24.39	25.89	25.89	26.13	25.44	22.58	21.05	22.56	26.55	26.78	26.78	26.19	2
158	196206	26.32	27.73	27.82	27.82	30.43	29.31	26.48	29.54	30.98	30.98	31.79	31.18	27.31	25.87	28.15	31.49	31.78	31.78	31.60	3
159	196207	30.98	32.75	32.29	32.29	35.76	34.76	31.35	33.51	35.93	35.93	37.62	37.31	33.10	31.28	31.39	35.59	36.82	36.82	37.71	3
160	196208	27.94	29.27	29.05	29.05	31.20	30.56	28.41	30.33	31.63	31.63	32.38	31.95	29.00	28.03	28.97	32.02	32.31	32.31	32.35	3
161	196209	23.70	24.02	25.20	25.20	26.52	25.35	22.85	27.29	28.01	28.01	28.04	26.57	23.88	22.13	26.26	29.38	29.49	29.49	28.84	2
162	196210	19.72	20.40	20.66	20.66	22.18	21.16	18.49	22.05	23.07	23.07	22.52	21.57	19.57	17.88	20.63	23.65	23.96	23.96	22.83	2
163	196211	14.62	14.93	15.06	15.06	16.01	15.51	13.46	15.88	16.78	16.78	15.68	15.07	14.31	12.81	14.95	17.04	17.29	17.29	15.79	1
164	196212	12.15	12.26	12.70	12.70	13.04	12.27	10.68	13.60	14.06	14.06	12.63	11.66	10.84	9.93	12.62	14.56	14.54	14.54	12.87	1
165	196301	10.37	10.37	10.83	10.83	10.90	10.40	9.28	11.56	11.78	11.78	10.45	9.62	9.14	8.56	10.94	12.28	12.12	12.12	10.59	
166	196302	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
167	196303	15.50	16.11	16.77	16.77	18.52	17.33	14.64	18.02	20.25	20.25	19.08	17.75	16.20	14.10	15.99	19.94	20.72	20.72	19.25	1
168	196304	15.47	16.53	16.14	16.14	17.64	17.32	15.14	17.14	17.97	17.97	17.65	17.00	15.79	14.59	15.89	18.57	18.43	18.43	17.32	1
169	196305	18.56	19.83	19.54	19.54	21.78	21.36	18.64	20.79	22.04	22.04	22.19	20.92	19.42	18.22	19.25	22.39	22.50	22.50	21.06	1
170	196306	23.11	24.23	24.17	24.17	26.03	25.53	23.25	25.76	26.68	26.68	27.04	26.12	24.07	22.97	24.62	27.51	27.57	27.57	26.87	2
171	196307	25.01	27.47	26.95	26.95	29.78	29.36	26.48	28.40	30.06	30.06	31.22	30.95	27.97	26.45	26.84	30.44	31.00	31.00	31.00	2
172	196308	28.47	30.47	30.03	30.03	33.92	32.72	28.90	31.56	34.37	34.37	35.89	35.81	31.22	29.09	29.25	34.00	35.34	35.34	35.93	3
173	196309	24.50	25.41	25.75	25.75	27.81	26.64	23.81	27.47	28.84	28.84	28.78	27.81	24.73	23.42	25.86	29.67	30.00	30.00	29.00	2
174	196310	17.52	17.98	18.30	18.30	19.50	18.78	16.77	19.50	20.50	20.50	19.64	18.70	17.17	16.19	18.30	21.05	21.19	21.19	19.73	1
175	196311	16.12	16.60	16.57	16.57	17.48	17.04	15.03	17.53	18.20	18.20	17.12	16.51	15.65	14.45	16.45	18.64	18.59	18.59	17.20	1
176	196312	14.83	15.47	15.18	15.18	16.10	15.66	13.69	15.83	16.46	16.46	15.57	14.99	14.22	13.04	14.85	16.58	16.50	16.50	15.30	1
177	196401	12.51	12.88	12.95	12.95	13.51	12.98	11.40	13.62	14.10	14.10	13.07	12.40	11.52	10.60	12.82	14.41	14.24	14.24	13.01	1
178	196402	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
179	196403	12.97	13.34	13.25	13.25	13.80	13.47	12.11	13.86	14.12	14.12	13.30	12.67	12.03	11.33	13.08	14.55	14.23	14.23	13.14	1
180	196404	14.74	15.58	15.19	15.19	16.72	16.30	14.24	15.69	16.42	16.42	16.33	15.96	14.72	13.71	14.58	16.47	16.32	16.32	15.74	1
181	196405	17.15	18.07	17.54	17.54	19.23	18.99	17.35	17.91	18.76	18.76	19.13	18.79	17.71	17.02	16.88	18.71	18.63	18.63	18.30	1
182	196406	24.27	25.29	25.40	25.40	27.69	26.82	24.87	26.40	28.18	28.18	28.69	27.12	25.60	24.62	24.99	28.29	28.69	28.69	27.74	2
183	196407	28.64	29.80	30.19	30.19	32.72	31.48	28.59	31.83	33.72	33.72	34.52	33.19	29.59	27.86	30.16	34.23	34.83	34.83	34.38	3
184	196408	27.80	29.75	29.11	29.11	33.13	31.92	28.28	30.10	32.85	32.85	34.42	34.27	30.30	28.14	27.61	31.97	33.22	33.22	33.83	3
185	196409	23.80	24.70	24.89	24.89	26.83	25.75	23.67	26.47	27.67	27.67	27.94	27.06	24.22	22.97	25.03	28.42	28.69	28.69	28.08	2
186	196410	19.83	20.12	20.44	20.44	21.12	20.61	19.09	21.47	22.05	22.05	21.34	20.41	19.33	18.61	20.53	22.70	22.69	22.69	21.55	1
187	196411	14.48	14.92	14.78	14.78	15.59	15.26	13.40	15.48	16.14	16.14	14.64	13.82	12.70	14.48	16.40	16.42	16.42	15.04	1	
188	196412	12.84	12.84	13.49	13.49	13.77	13.16	11.71	14.68	15.37	15.37	13.87	12.91	12.21	11.10	13.57	16.09	16.18	16.18	14.45	1
189	196501	12.58	12.78	13.13	13.13	13.74	13.14	11.30	14.24	15.11	15.11	13.61	12.77	11.91	10.66	13.02	15.70	15.89	15.89	14.02	1
190	196502	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
191	196503	16.27	17.10	17.01	17.01	18.20	17.63	15.56	17.91	18.77	18.77	17.93	17.20	16.01	14.87	16.56	19.20	18.93	18.93	17.81	1

length: 6246240 lines: 481 Ln: 1 Col: 1 Pos: 1 Windows (CR LF) UTF-8 INS

Tiene sentido que falten valores en un dataset porque los datos en crudo se han obtenido de estaciones ubicadas, tanto en lugares institucionales (aeropuertos, colegios, etc....) como particulares y las lecturas no siempre se han podido hacer en tiempo y forma.

Los productos en localidades puntuales ofrecidos en **Escenarios-PNACC Datos mensuales** se refieren a 374 puntos para los productos relacionados con la temperatura y 2.321 puntos para los relacionados con la precipitación, que se distribuyen por toda la geografía de España como se refleja en los mapas de la *Figura 6*. Estos puntos se corresponden a estaciones climatológicas de AEMET que han sido seleccionadas tras un control de calidad basado en criterios de longitud temporal de la serie, número reducido de lagunas y homogeneidad (Brunet *et al.* 2008, Herrera 2011, Herrera *et al.* 2012).

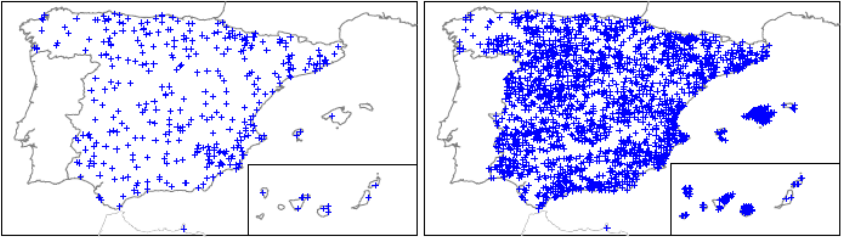


Figura 6. Distribución de las localizaciones de las 374 estaciones con datos de temperatura y las 2.321 estaciones con datos de precipitación para los productos de Escenarios PNACC Datos mensuales

Mapa de estaciones climatológicas ubicadas en España

Una primera conclusión de este estudio se refiere a la detección de errores provocados por desplazamientos de ubicación, cambios en la instrumentación o el entorno, calibraciones, etc... lo que se refleja en las series como cambios artificiales añadidos con formato diferente.

No obstante, la serie que se ha utilizado para estudiar la influencia de la temperatura en el cambio climático tiene la misma tendencia que la señal climática real y los datos climáticos son homogéneos, con largo recorrido temporal y probada calidad de modo que aceptan como válidos para esta investigación de carácter general, sin especificar zonas concretas de la geografía española.

A partir de aquí se trabaja con los datos originales y se va a mostrar mediante diferentes visualizaciones la influencia de la temperatura máxima en el cambio climático.

Plots y visualización de datos

Se ha cargado el archivo original sobre data1 y se van a mostrar las posibles visualizaciones que se pueden obtener a partir de él.

Plots y visualización de los datos

```
: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
: data1.head()
```

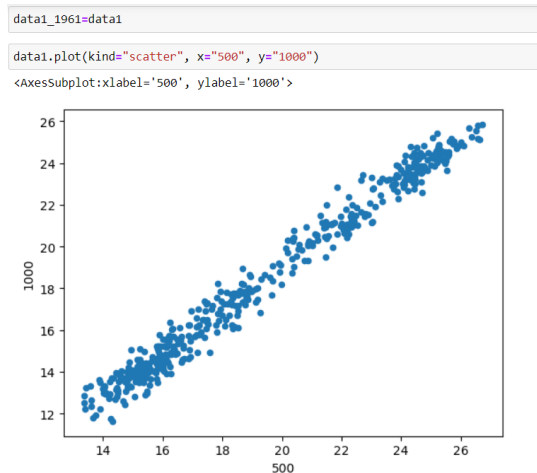
	FECHA	500	501	502	503	504	505	506	507	508	...	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944
0	196101	16.08	16.33	16.08	16.08	16.22	16.12	16.01	16.36	16.50	...	11.79	11.79	12.54	12.54	11.36	10.84	10.26	13.19	13.00	12.98
1	196102	15.29	15.45	15.56	15.56	15.56	15.31	14.84	16.13	16.09	...	10.76	10.70	11.37	11.37	10.34	9.81	9.34	12.40	12.23	12.13
2	196103	15.88	16.35	16.64	16.64	17.47	17.03	15.89	17.33	18.14	...	11.03	11.03	11.69	11.69	10.63	9.96	9.35	12.55	12.35	12.18
3	196104	16.13	17.15	16.57	16.57	17.96	17.82	16.24	17.27	17.71	...	11.55	11.71	12.47	12.47	11.16	10.46	9.68	12.65	12.36	12.30
4	196105	18.68	20.01	20.31	20.31	23.24	22.28	18.15	22.17	23.90	...	13.10	13.18	14.29	14.29	12.56	11.63	10.63	14.28	14.13	14.06

5 rows × 1446 columns

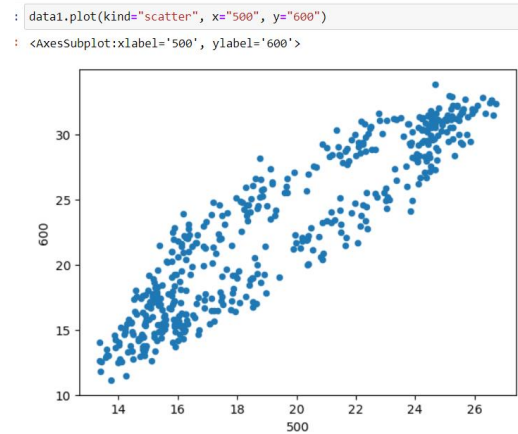
Visualización básica: Scatterplot (nube de dispersión/puntos)

Representación de TODOS los puntos que recoge el dataset, en la estación 500 y la estación 1000 a lo largo del tiempo desde enero de 1961 hasta diciembre del año 2000.

Parece que hay una tendencia ascendente y relativamente similar entre ambas



Sin embargo, al comparar la estación 500 con la estación 1940 a lo largo del tiempo desde enero de 1961 hasta diciembre del año 2000 se ve claramente que la separación entre los puntos es mayor, lo cual se debe seguramente a que las estaciones están más separadas y tienen climas distintos.



Hace falta un análisis más exhaustivo para verificar la representación gráfica del aumento de temperaturas.

Se necesita ajustar más la visualización para ver más claro el análisis, haciendo uso del Data Wranglin (cirugía de datos).

Se puede extraer un subconjunto de datos, por ejemplo, de una sola estación meteorológica y el objeto que se obtiene es de tipo series o vectores.

Crear un subconjunto de datos

```
Estacion500=data1_1961["500"]
Estacion500.head()
0    16.08
1    15.29
2    15.88
3    16.13
4    18.68
Name: 500, dtype: float64
type(Estacion500)
pandas.core.series.Series
```

También se pueden extraer varias columnas, varias estaciones meteorológicas, y el objeto obtenido sería un DataFrame.

```
subset=data1_1961[["500", "600", "700"]]
subset.head()
   500   600   700
0  16.08  18.09  18.86
1  15.29  14.18  16.10
2  15.88  18.35  19.96
3  16.13  21.63  22.58
4  18.68  24.34  24.84
type(subset)
pandas.core.frame.DataFrame
```

Otra opción para esto en especificar las columnas deseadas es utilizar un solo paso mediante la opción `desired_columns`.

```
desired_columns=["500","1000","1500","1944"]
subset=data1[desired_columns]
subset.head()
```

```

      500   1000   1500   1944
0  16.08  14.53  12.42  12.98
1  15.29  13.99  12.15  12.13
2  15.88  14.38  14.66  12.18
3  16.13  15.52  16.75  12.30
4  18.68  17.78  22.48  14.06
```

Así, se haría una comparación entre estaciones, en principio, distribuidas por toda la malla.

Para reducir el peso del archivo se puede escoger una serie de columnas de trabajo y eliminar las demás, lo cual tendría más sentido si se conociera la ubicación exacta de las estaciones, pero aún así se ha hecho escogiendo únicamente las estaciones por centenas.

```
desired_columns=["500","600","700","800","900","1000","1100","1200","1300","1400","1500","1600","1700","1800","1900","2000","2100"]
desired_columns[
```

```
[
'500',
'600',
'700',
'800',
'900',
'1000',
'1100',
'1200',
'1300',
'1400',
'1500',
'1600',
'1700',
'1800',
'1900',
'2000',
'2100',
'2200',
'2300',
'2400',
'2500',
'2600',
'2700',
'2800',
'2900']
```

A continuación, se muestran todos los elementos de la lista del subset:

```
all_columns_list=data1_1961.values.tolist()
```

```
all_columns_list
```

```
[[196101.0,
16.08,
16.33,
16.08,
16.08,
16.22,
16.12,
16.01,
16.36,
16.5,
16.5,
15.39,
14.73,
15.17,
15.48,
16.23,
16.75,
16.56,
16.56,
```

Otra forma de filtrar datos de forma, posiblemente más fácil, es la siguiente:

```

1 | a = set(desired_columns)
2 | b = set(all_columns_list)
3 | sublist = b-a
4 | sublist = list(sublist)
```

Histogramas de frecuencias

Concatenación de datos (función merge)

Tipos de joins

Inner join

Left join