

#### Práctica 1 (35% nota final)

1. Contexto. Explicar en qué contexto se ha recolectado la información. Explique por qué el sitio web elegido proporciona dicha información.

Los datos han sido extraidos de la pagina web "https://www.motorflash.com/" con el objetivo de analizar la variación de precios de coches de segunda mano con el objetivo principal de visualizar las diferencias de precios entre vendedores profesionales y particulares. La web de motor flash ha sido elegida en primera instancia al poder usar el scrapper elegido para el proyecto (Scrapy), en segundo lugar proporciona una amplia colección de diferentes marcas y modelos de coches así como la diferenciación de venta particular y profesional. Otros sitios web de venta de coches de segunda mano tienen sistemas anti-scrapping como es por "https://www.coches.net/". Otro ejemplo de sitio web "https://www.coches.com/coches-segunda-mano/" ya que aunque permitia el uso del scraper y anunciaba un número mayor de modelos y marcas al realizar la prueba de scraping se vio que el csv resultante sólo contenía 9000 datos en vez de los 90000 anunciados en la página.

2. Definir un título para el dataset. Elegir un título que sea descriptivo.

El título seleccionado para el dataset es: "motorflash.csv" debido a que el dataset contiene toda la información de cada vehiculo proveniente de la web de "https://www.motorflash.com/"

3. Descripción del dataset. Desarrollar una descripción breve del conjunto de datos que se ha extraído (es necesario que esta descripción tenga sentido con el título elegido).

El dataset generado contiene una representación temporal de las características principales de los coches de segunda mano, se compone de 32 campos en los cuales están incluidos todos los datos tecnicos del vehiculo asi como los datos importantes para la realización del estudio como son el "Modelo", "Descripcion", "Precio Oferta", "Precio Nuevo" y el "Anunciante"



#### 4. Representación gráfica. Presentar una imagen o esquema que identifique el dataset visualmente

			V.2 620 mm Cam				color_t	Gasolina (	consum_can	e consum_cor 5.6 l/100km									Peugeot : 5 Puerl		peso_ 0 1.750		as potenci	a pr		recio_oferta		ubicacion Valencia	velocidad
								Gasoina	4.1 1/100km																				
),3 segu 1.556 mm							. T.				5,9 U100km								Peugeot 25 Puert		(g 1.690			(110 CV, 22				Barcelona	
9 segun 1.445 mm					66 litros		Marra*	n Gasolna	5,1 l/100km		8,9 l/100km						4.686 mm		Mercedes 4 Puerl		(g 2.135			(258 C176				Barcelona	
			B 2 620 mm Cam			Bianco		DiAtisel	3,1 l/100km	4,6 l/100km							4.253 mm		Peugeot 35 Puerl		g 1.810			(130 CV, 27				Barcelona	
),8 segu 1.624 mm						Gris		Gasolina	4,5 l/100km	5,1 l/100km									Peugeot 55 Puert		g 1.910			(130 CV, 29					188 km/h
3,3 segui 1.542 mm					33 litros	Negro	Negro	Gasoina	4.0 l/100km	4,5 l/100km	5.2 l/100km	Combinado 4.5	I Coupe Pass	sion 104 gr/km	12 meses 1	10000 km	2.695 mm	06-2010	Smart For 3 Puert		1.020	Kg 2	52 kW	(71 CV, §12	2.143 3	890,00	3 puertas	Madrid	145 km/h
1,6 segui 1.556 mm	1.739 mm	<b>PSARETAI</b>	. B 2 538 mm Cam	bio au 1.400 litros	50 litros	Blanco		Gasolna	6,2 l/100km	5,7 l/100km	6,7 l/100km	Combinado 5.7	I 1.2 Pure lect	h S 126 gr/km	12 meses 1	5368 km	4.159 mm	07-2018	Peugeot 26 Puert	tas 1.225 R	g 1.690	Kg 5	81 KW	(110 CV, 22	2.900 15	5490,00	5 puertas	Barcelona	187 km/h
6 seguni 1.454 mm					70 litros	Blanco	Gris	DiAdsel	4,6 l/100km	5,5 l/100km	7,3 V100km								SEAT Exe 5 Puerl	tas 1.605 k	(g 2.090	Kg 5	105 kW	(143 C13)	1.821 7	900,000	5 puertas	Guipuzcoa	
1 segun 1.476 mm	1.828 mm	<b>PSARETAI</b>	A 2.817 mm Cam	bio ma 1.598 litros	72 litros	Gris		DiÃOsel	3,8 l/100km	4,2 V100km	5,0 V100km	Combinado 4.2	I 2.0 BlueHDI.	All: 110 gr/km	24 meses 90	5351 km	4.829 mm	07-2017	Peugeot 55 Puert	tas 1.520 k	G 2.200	Kg 5	110 KW	(150 CV36	3.300 13	3900,00	5 puertas	CoruĂta	210 km/h
3,3 segu 1.474 mm	1.749 mm	<b>PSARETAI</b>	A 2.540 mm Cam	bio ma 922 litros	45 litros	Negro		Gasolina		4,2 l/100km		Combinado 4.2	1 Pure Tech 82	2 Fc96 gr/km	24 meses 11	5366 km	3.996 mm	02-2020	Citroen C:5 Puert	tas 980 Kg	1.555	Kg 5	60 kW	82 CV, 115	5.500 11	1900,00	5 puertas	CoruAsa	169 km/h
9 seguni 1.435 mm	1.745 mm	<b>PSARETAII</b>	A 2.538 mm Cam	bio ma 1.118 litros	44 litros	Gris		Gasolina	3,6 l/100km	18,2 V100km	5.2 l/100km	Combinado 18:	2 1.2 Turbo XI	HL 196 gr/km	24 meses 11	1753 km	4.060 mm	02-2020	Opel Cors 5 Puert	tas 1.165 k	(g 1.620	Kg 5	74 KW	100 CV, 19	700 1	3900.00	5 puertas	Coruña	188 km/h
9 seguni 1.435 mm	1.745 mm	PSA RETAIL	A 2.538 mm Cam	bio ma 1,118 litros	44 litros	Gris		Gasolina	3,6 l/100km	18,2 l/100km	5,2 V100km	Combinado 18:	2 1.2 Turbo XI	HL 196 gr/km	24 meses 13	2337 km	4.060 mm	02-2020	Opel Cors 5 Puerl	tas 1,165 R	(g 1.620	Kg 5	74 KW	(100 CV, 19	700 13	3900,00	5 puertas	CoruA±a	188 km/t
1.457 mm	1.804 mm	<b>PSARETAI</b>	A 2.620 mm Cam	bio ma 1.309 litros	53 litros	Blanco		Gasolina	3.9 V100km	5.6 l/100km	5.3 V100km	Combinado 5.6	PureTech 11	10 £101 gr/km	24 meses 7	156 km	4.253 mm	11-2019	Peugeot 35 Puert	tas 1.165 H	a 1.750	Kg 5	81 kW	110 CV. 25	5.400 16	6490.00	5 puertas	AlmerÃ-a	186 km/h
),5 segu 1.647 mm	1.819 mm	<b>PSARETAI</b>	V.2 636 mm Cam	no au 1.251 litros	60 litros	Blanco		DiAppet	5.1 V100km	5.7 V100km	6.6 l/100km	Combinado 5.7	1 2.0 Mutuet L	Lon 148 gr/km	12 meses 2	8550 km	4.394 mm	07-2018	Jeep Com 5 Puerl	tas 1.619 N	g 2.080	Kg 5	103 KW	(140 C135	330 1	9900.00	5 puertas	Valencia	190 km/tr
			V.2.620 mm Cam			Azul		Gasolina	3.9 V100km	5.6 V100km	5.3 V100km						4.253 mm		Peugeot 35 Puerf		G 1.750	Kg 5		110 CV 21			5 puertas		186 km/t
1.820 mm	1.848 mm	<b>PSARETAII</b>	V.2.785 mm Cam	bio mr. 1 355 liftos	50 litros	Azul		Diābset	4.2 V100km	4.4 l/100km	4.8 V100km	Combinado 4.4	I BlueHDi 100	All 112 gr/km	24 meses 2	km	4 403 mm	01-2020	Peugeot F5 Puert	tas 1 424 H	(g 2.100	Ka 5	73 kW	100 CV.26	950 1	7400.00	5 puertas	Valencia	170 km/t
5 segun 1.624 mm								Gasoina	4.8 1/100km	5.3 l/100km	6.1 V100km						4.447 mm	08-2019	Peugeot 35 Puert		G 1.910			130 CV 30			5 puertas		188 km/t
0.5 segu 1.440 mm						Gris	Negro	DiApsel	3.9 V100km	4.2 V100km		Combinado 4.2							Mercedes 5 Puerl		g 1.935			116 CV. 37			5 puertas		202 km/t
6 segun 1.442 mm					66 litros	Gris	Negro	Gasoina	4.4 V100km	5.3 l/100km									Mercedes 4 Puerl		0 2.045			(211 CV4)			4 puertas		244 km/h
0.0 segu 1.710 mm						Rojo	riegro	DiAtisel	4.7 V100km	5.3 ¥100km									Mazda CX5 Puerl		g 2.050			(150 CV36				Madrid	198 km/t
1.5 segui 1.635 mm								Gasolna	5.0 l/100km	5.9 l/100km									Misubishi 6 Puerl		g 1.870			117 CV. 21				Madrid	183 km/h
0 segur 1 633 mm						Blanco		DiAnsel	5.7 V100km	7.0 U100km									Chevrolet 5 Puerl		(q 2.291			(163 C)25				Madrid	195 km/t
9 segun 1.761 mm						Blanco		DIÃOSEI	5.6 V100km	6.6 l/100km									Opel Anta 5 Puerl		0 2.505			(163 C136				Madrid	188 km²t
2 segun 1.710 mm						Rojo		DiAbsel	4.1 V100km	4.6 l/100km									Mazda CX5 Puerl		g 2.000			(150 C\30				Madrid	202 km/h
2 segun 1.710 mm 2.1 segu 1.474 mm						Bianco		DiApsel	3.4 V100km		5,4 U100km	Combinado 4.6 Combinado 3.9							Hyundai C5 Puerl		g 2.050 g 1.690			190 CV 218				Madrid	175 km/t
2,4 segu 1.653 mm					52 litros	Granati	Negro	DIADSEL	3,9 l/100km	4,0 l/100km	4,2 V100km								Renault S 5 Puert		g 2.097			(110 CV, 28				Madrid	184 km/t
2,5 segu 1.640 mm						Negro		DiA0sel	3,9 V100km	4,1 V100km	4,5 l/100km								Renault S 5 Puert		g 1.967			(110 CV, 24				Madrid	180 km/h
),9 segu 1.495 mm					41 litros	Gris		DiAdsel	3,2 1/100km	3,6 l/100km									Ford Fies 5 Puerl		g 1.575			(96 CV, 21)				Madrid	180 km/h
2,1 segui 1.500 mm						Gris		Gasolna	4,1 V100km	4,9 l/100km									Hyundai i 5 Puerl		G 1.450			87 CV, 12				Madrid	175 km²
3,7 segu 1.660 mm						Blanco		DiAttset	4,2 l/100km	4,6 l/100km									Hyundai T5 Puert		(g 2.000			(115 CV, 28				Madrid	176 km/h
1,5 segur 1.680 mm						Blanco		Gasolna	5,5 l/100km	6,2 l/100km	7,4 F100km								Misubish(5 Puerl		(g 1.985			(150 CV2)				Madrid	192 km/t
7 segun 1.471 mm						Gris		Gasolna	4,7 l/100km	5,3 l/100km									Renault M5 Puerl		cg 1.781			(115 CV, 19				Madrid	200 km/t
),9 segu 1.484 mm						Bianco		DIADsel	3,7 l/100km	4,2 l/100km									Ford Foci 5 Puerl		(g 1.900			(115 CV, 20				Madrid	193 km/h
),9 segu 1.605 mm						Blanco		Gasolina	4,3 l/100km	4,7 l/100km									Opel Cros 5 Puerl		g 1.770			(110 CV, 21				Madrid	187 km/h
2 segun 1.426 mm	1.785 mm	MENAL S.L.	2.637 mm Cam	bio au 1.220 litros	50 litros	Gris Pt	M Negro	Gasoina	4,4 l/100km	5,1 l/100km	6,4 l/100km	Combinado 5.2	I S line 35 TF	SI 117 gt/km	24 meses 4	790 km	4.313 mm	07-2019	Audi A3 S 5 Puerl	tas 1.315 H	(g 1.800	Kg 5	110 KW	(150 CV39	9.048 24	4400,00	5 puertas		220 km/h
0 seguni 1.420 mm	1.859 mm	Cobendai C	ca 2.793 mm Cam	bio au 1,780 litros	62 litros	Bianco	Negro	Gasolina	4,4 l/100km	5,4 l/100km	7,1 l/100km	Combinado 5.4	I Puretech 18	10 S 118 gr/km	12 meses 2	301 km	4.790 mm	12-2019	Peugeot 55 Puert	tas 1.430 k	g 2.025	Kg 5	133 KW	(180 C) 40	0.000 33	2490,00	5 puertas	Madrid	226 km/h
1 segun 1.476 mm	1.828 mm	DriveOn	2.817 mm Cam	bio mé 1.598 litros	72 litros	Blanco	Negro	DIÃOsel	3,8 l/100km	4,2 l/100km	5,0 l/100km	Combinado 4.2	I 2.0 BlueHDL	Alli 110 gr/km	12 meses 8	5156 km	4.829 mm	07-2016	Peugeot 55 Puert	tas 1.520 H	(g 2.200	Kg 5	110 KW	(150 CV36	300 13	3499,00	5 puertas	Madrid	210 km/t
1 segun 1.476 mm	1.828 mm	DriveOn	2.817 mm Cam	bio ma 1.598 litros	72 litros	Negro	Negro	DiÃOsel	3,8 l/100km	4,2 l/100km	5,0 l/100km	Combinado 4.2	1 2.0 BlueHDI	GT 110 gr/km	12 meses 90	3863 km	4.829 mm	09-2016	Peugeot 55 Puert	tas 1,520 N	Q 2.200	Kg 5	110 kW	(150 C\36	3.750 13	3799,00	5 puertas	Madrid	210 km/h
1,3 segur 1 476 mm	1.828 mm	AUTOMOVI	LE 2.817 mm Cam	bio ma 1.598 litros	72 litros	Gris		DiAbsel	3,8 1/100km	4,2 l/100km	4,8 l/100km	Combinado 4.2	I 1.6 BlueHDI.	Act 100 gr/km	12 meses 7	7991 km	4.829 mm	01-2017	Peugeot £5 Puert	tas   1.430 K	g 2.120	Kg 5	88 KW	(120 CV, 30	750 11	1400,00	5 puertas	Madrid	199 km/h
0 segun 1.420 mm	1.859 mm	Cobendai C	0c82.793 mm Cam	bio au 1.780 litros	62 litros	Azul	Negro	Gasolna	4,4 l/100km	5,4 l/100km	7,1 V100km	Combinado 5.4	Puretech All	ure 116 gr/km	12 meses 1	00 km	4.790 mm	12-2019	Peugeot 55 Puert	tas 1.430 H	g 2.025	Kg 5	133 KW	(180 C13)	700 2	4890,00	5 puertas	Madrid	226 km/t
1,3 segur 1,476 mm	1.828 mm	Convesa	2.817 mm Cam	bio ma 1.598 litros	72 litros	Gris	Negro	DiAbsel	3.8 V100km	4.2 l/100km	4,8 l/100km	Combinado 4.2	I 1.6 BlueHDL	Act 104 gr/km	12 meses 7	5751 km	4.829 mm	02-2016	Peugeot 55 Puert	tas 1,430 H	(g 2.120	Kg 5	88 KW	(120 CV, 30	0.400 13	2900.00	5 puertas	Badajoz	199 km/t
3.2 segu 1.481 mm	1.746 mm	Alguiber	2.510 mm Cam	bio ma 1.120 litros	45 litros	Gris	Negro	Gasoina	4.2 l/100km	5.2 V100km	6.9 l/100km								Opel Cors 5 Puert		(g 1.615		66 KW	90 CV. 16	690 8	690.00	5 puertas	Madrid	175 km/t
).9 segu 1.454 mm			2 642 mm Cam		70 litros		3.0	DiApsel	4.1 V100km	5.3 l/100km									SEAT Exe 5 Puerl		g 2.050			120 CV.25			5 puertas		196 km/t
1.8 segu 1.481 mm				bio ma 1 120 litror	45 litros	Blanco	Negro	DiAnsel	3.3 V100km	3.8 l/100km		Combinado 3 8	I 13 CDTI Ex	pre 100 gr/km	12 meses 4	9488 km	4 021 mm	06-2016	Opel Cors 5 Puerf		G 1.695			75 CV : 16			5 puertas		164 km/h
				bio au 1.482 litros			9.0	DiÃOsel				Combinado 5.2							Peugeot 35 Puert		g 2.000			130 CV.30			5 puertas		192 km/t

Imagen del dataset recuperado de la web de motorflash, mostrando los diferentes campos obtenidos.

# 5. Contenido. Explicar los campos que incluye el dataset, el periodo de tiempo de los datos y cómo se ha recogido.

Por cada anuncio, el cual se corresponde con un registro en el conjunto de datos, se almacenan las siguientes características:

- 1. modelo : modelo del vehículo anunciado
- 2. descripcion : descripción del anuncio
- 3. anunciante : propietario del vehículo del anuncio
- 4. precio\_oferta : precio ofertado por el anunciante.
- 5. precio\_nuevo : precio del vehículo 0 kilómetros.
- 6. imagen : url con la imagen del vehículo.
- 7. combustible : combustible del vehículo anunciado.
- 8. kilometros : kilometraje del vehículo al momento de anunciarlo. (km)
- 9. potencia : potencia máxima del vehículo. (kw)
- 10. color : color de la chapa del vehículo..
- 11. matriculación : fecha de matriculación del vehículo.
- 12. color tap : color de la tapicería del vehículo.
- 13. plazas : número de plazas del vehículo
- 14. ubicacion : localización del vehículo.
- 15. cambio: tipo de cambio del vehículo.
- 16. garantia : tiempo de garantía ofrecido por el anunciante.
- 17. consumo: consumo combinado mixto (l/100km).
- 18. altura : altura del vehículo (mm)
- 19. anchura : anchura del vehículo (mm)



- 20. longitud : longitud del vehículo (mm)
- 21. cap\_maletero : capacidad del maletero (I)
- 22. batalla : distancia entre ejes del vehículo (mm)
- 23. peso: peso del vehículo (kd)
- 24. peso\_maximo : peso máximo admitido (kg)
- 25. num\_puertas : número de puertas del vehículo
- 26. capacidad dep : capacidad del depósito (I)
- 27. velocidad\_max : velocidad máxima (km/h)
- 28. aceleracion: aceleracion maxima (seg)
- 29. emision\_co2 : nivel de emisión de dióxido de carbono (gr/km)
- 30. consum\_carretera : consumo en carretera (l/100km)
- 31. consum\_urbano : consumo urbano (l/100km).
- 32. consum\_comb\_mix: consumo combinado mixto (l/100km).
- 33. num\_cilindros : número de cilindros
- 34. cilindrada : cilindrada del vehículo (cm3)
- 35. par motor: par motor (Nm)
- 36. est emisiones : estándar de emisiones
- 37. tracción : tracción del vehículo.
- 38. neumaticos del : medidas de los neumáticos delanteros
- 39. neumaticos tra: medidas de los neumáticos traseros.

El periodo de matriculación de los vehículos anunciados va desde 1955 hasta la actualidad. Los datos se han recogido mediante la librería scrapy. Primero hemos revisado el tamaño de la página para ver que dispone de suficientes datos, también el fichero robots.txt para ver que las limitaciones y restricciones a los robots de scraping no nos afectan.

Mediante reglas definimos el comportamiento para rastrear el sitio web, una regla para la paginación con la que recorremos todas las páginas y otra regla para acceder a todos los anuncios de cada página. Dentro de cada anuncio está toda la información que almacenamos.

Finalmente realizamos el mapeo mediante xpath y css de la localización de los campos en el dom del html con el objeto serializado que genera el fichero de salida csv.

El tiempo aproximado de realización de web scraping en toda la web fue de unas 4 horas aproximadamente, obteniendo un total de 50656 registros por 39 columnas de características.

Se ha tenido en cuenta las restricciones que algunas páginas web realizan sobre los robots para evitar el bloqueo de los mismos, pero en este caso concreto no fue necesario.



# 6. Agradecimientos. Presentar al propietario del conjunto de datos. Es necesario incluir citas de investigación o análisis anteriores (si los hay).

Los datos han sido recopilados desde la página web de anuncios de coches de segunda mano <a href="https://www.motorflash.com/">https://www.motorflash.com/</a>. Para ello, se ha hecho uso del lenguaje de programación python y de técnicas de web scraping para extraer información alojada en las páginas HTML. Agradecemos a la compañía de motorflash por poner a disposición de los usuarios dichos datos de manera pública en su página web así como a todos los concesionarios de vehículos que suben los datos de sus vehículos al sitio web.

# 7. Inspiración. Explique por qué es interesante este conjunto de datos y qué preguntas se pretenden responder.

El presente conjunto de datos permite realizar un análisis estadístico para identificar patrones y tendencias en los datos.

Identificar las ciudades que más coches tienen publicados y los modelos mayoritarios en cada una de ellas.

Analizar si el precio de un coche varía por provincias y cuales son los anunciantes con los precios más caros o más baratos.

Detectar los modelos de coches que más se devalúan con respecto a su precio y los que menos.

Identificar los modelos que mejor relación calidad precio tienen en base a sus características. Predecir el precio de un coche de segunda mano dadas sus características para tener una estimación del precio en el mercado de segunda mano.

Clasificar coches por características, en base al precio, potencia, velocidad, etc. para agrupar coches por gama alta, media, baja.

Son algunas de las preguntas que podríamos responder mediante un análisis de datos realizado sobre este conjunto de datos, que podría utilizarse en diversos ámbitos. Uno de ellos podría ser en el periodístico con fines informativos sobre el mercado de segunda mano. Así mismo podría ser de gran utilidad en el campo de la minería de datos, a la hora de elaborar modelos predictivos como los mencionados anteriormente.

### 8. Licencia. Seleccione una de estas licencias para su dataset y explique el motivo de su selección:

- o Released Under CC0: Public Domain License
- Released Under CC BY-NC-SA 4.0 License
- o Released Under CC BY-SA 4.0 License
- Database released under Open Database License, individual contents



- under Database Contents License
- Other (specified above)
- o Unknown License

Teniendo en cuenta que nuestro proyecto es meramente académico, consideramos que debe ser público para que pueda ser utilizado por cualquier persona incluido para una finalidad comercial, ya que el origen de estos datos tiene fines comerciales.

La licencia escogida para la publicación de este conjunto de datos ha sido CC BY-SA 4.0 License. Los motivos que han llevado a la elección de esta licencia va acorde con la idoneidad del trabajo realizado:

- Se debe reconocer adecuadamente la autoría e indicar si se han realizado cambios. De manera que se reconoce el trabajo ajeno y permite realizar aportaciones sobre el mismo
- Se debe compartir igual si se mezcla, transforma o crea a partir del material y deberá difundir las contribuciones bajo la misma licencia que el original.
- Se puede utilizar para una finalidad comercial. Aportaría valor a la empresa que utilice estos datos para generar análisis de calidad reconociendo la autoría original.

## 9. Código. Adjuntar el código con el que se ha generado el dataset, preferiblemente en Python o, alternativamente, en R.

El código se ha realizado íntegramente en python y se ha sincronizado mediante el versionado de git.

Adicionalmente se tiene la carpeta docs, con este mismo documento, el fichero readme.md que contiene una pagina de bienvenida y descripcion del proyecto.

La url de github donde se aloja el proyecto es

https://github.com/CarlosRea/MotorflashScraper

# 10. Dataset. Publicación del dataset en formato CSV en Zenodo (obtención del DOI) con una breve descripción.

En Zenodo se ha añadido una breve descripción del proyecto que incluye el enlace al repositorio Github para quien desee ampliar información.

El dataset publicado se puede consultar en el siguiente enlace <a href="https://doi.org/10.5281/zenodo.4262983">https://doi.org/10.5281/zenodo.4262983</a>



- Lawson, R. (2015). Web Scraping with Python. Packt Publishing Ltd. Chapter 2.
  Scraping the Data.
- Simon Munzert, Christian Rubba, Peter Meißner, Dominic Nyhuis. (2015).
  Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. John Wiley & Sons.
- Tutorial de Github <a href="https://quides.github.com/activities/hello-world">https://quides.github.com/activities/hello-world</a>.

#### Criterios de valoración

Todos los apartados son obligatorios. La ponderación de los ejercicios es la siguiente:

- Los apartados 1, 2, 3 y 4 valen 0,25 puntos cada uno.
- Los apartados 5, 6, 7 y 8 valen 1 punto cada uno.
- Los apartados 9 y 10 valen 2,5 puntos cada uno.

Otros criterios que se tomarán en cuenta para la evaluación son:

- Idoneidad de las respuestas (deberán ser claras y completas).
- Complejidad del sitio web elegido para la extracción.
- Síntesis y claridad, a través del uso de comentarios, del código resultante.
- Presentación adecuada de los datos.
- Organización y claridad de los documentos de entrega final.
- Completitud de los documentos requeridos para la entrega final.

#### Formato y fecha de entrega

Durante la semana del 26 al 30 de octubre, el grupo podrá entregar al profesor una entrega parcial opcional. Esta entrega parcial es muy recomendable para recibir asesoramiento sobre la práctica y verificar que la dirección tomada es la correcta. Se entregarán comentarios a los estudiantes que hayan efectuado la entrega parcial pero no contará para la nota de la práctica. En la entrega parcial los estudiantes deberán entregar por correo electrónico, al profesor encargado del aula, el enlace al repositorio Github con lo que hayan avanzado.

En referente a la entrega final, hay que entregar un único fichero que contenga el enlace a Github donde haya:

- 1. Una Wiki donde estén los nombres de los componentes del grupo y una descripción de los ficheros.
- Un documento PDF con las respuestas a las preguntas y los nombres de los componentes del grupo. Además, al final del documento, debe aparecer la siguiente tabla de contribuciones al trabajo, la cual debe firmar cada integrante



del grupo con sus iniciales. Las iniciales representan la confirmación por parte del grupo que el integrante ha participado en dicho apartado. Todos los integrantes deben participar en cada apartado, por lo que, idealmente, los apartados deberían estar firmados por todos los integrantes.

Contribuciones	Firma						
Investigación previa	CR, YN						
Redacción de las respuestas	CR, YN						
Desarrollo código	CR, YN						

- 3. Una carpeta con el código Python o R generado para obtener los datos.
- 4. El DOI a los datos.

Este documento de la entrega final se tiene que entregar en el espacio de Entrega y Registro de AC del aula antes de las **23:59 del día 9 de noviembre**. No se aceptarán entregas fuera de plazo.