

Regresion

Tarea 3.

Regresión lineal

Análisis del Problema

El desempeño de un automóvil se puede medir de diferentes formas. Algunas comunes son la cantidad de caballos de fuerza y el rendimiento del mismo, que se puede resumir en cuantas millas puede recorrer el automóvil por cada galón de combustible que consume. Para los clientes, potenciales compradores de un automóvil, este rendimiento es importante pues puede ayudar a tomar una decisión con respecto a cuál automóvil comprar (si, por ejemplo, el cliente quiere un auto que rinda por muchas millas y pueda economizar en la compra de combustible).

Desde este punto de vista, tanto a clientes como a fabricantes de automóviles, les conviene entender cuál es la relación entre diferentes características del automóvil y su rendimiento, pues el conocer estas relaciones les puede ayudar a inferir cuál va a ser la eficiencia del vehículo a partir de ver los valores de otras características. Para fabricantes, puede ser importante conocer estas relaciones para saber cómo hacer cada modelo más eficiente con respecto al anterior.

Entendimiento de los Datos

Con el fin de analizar y tratar de estimar las millas por galón de diferentes modelos de automóviles, se trabajó con un conjunto de datos que contiene 398 observaciones y 9 variables:

- mpg (millas por galón): numérica, con un rango de 9 a 46.60.
- cyl (cilindraje): categórica ordinal, con valores posibles de 3, 4, 5, 6 y 8.
- disp (desplazamiento): numérica, con un rango de 68 a 455.
- hp (caballos de fuerza): numérica, con un rango de 46 a 230 y 6 valores faltantes.
- weight (peso): numérica, con un rango de 1613 a 5140.
- acc (aceleración): numérica, con un rango de 8 a 24.80.
- model year (año): categórica, con 13 valores diferentes representando el año del automóvil.
- origin (origen): categórica, 3 valores posibles: 1, 2, 3.
- model name (nombre del modelo): categórica, con 305 posibles valores.

Ejercicios

1. Cargue el archivo auto-mpg_g.csv en una variable

```
library(GGally)
```

```
## Warning: package 'GGally' was built under R version 3.6.1
```

```
## Loading required package: ggplot2
```

```
## Registered S3 method overwritten by 'GGally':  
##   method from  
##   +.gg      ggplot2
```

```
library(ggplot2)
library(visdat)
library(caTools)
library(lessR)
```

```
## Warning: package 'lessR' was built under R version 3.6.1
```

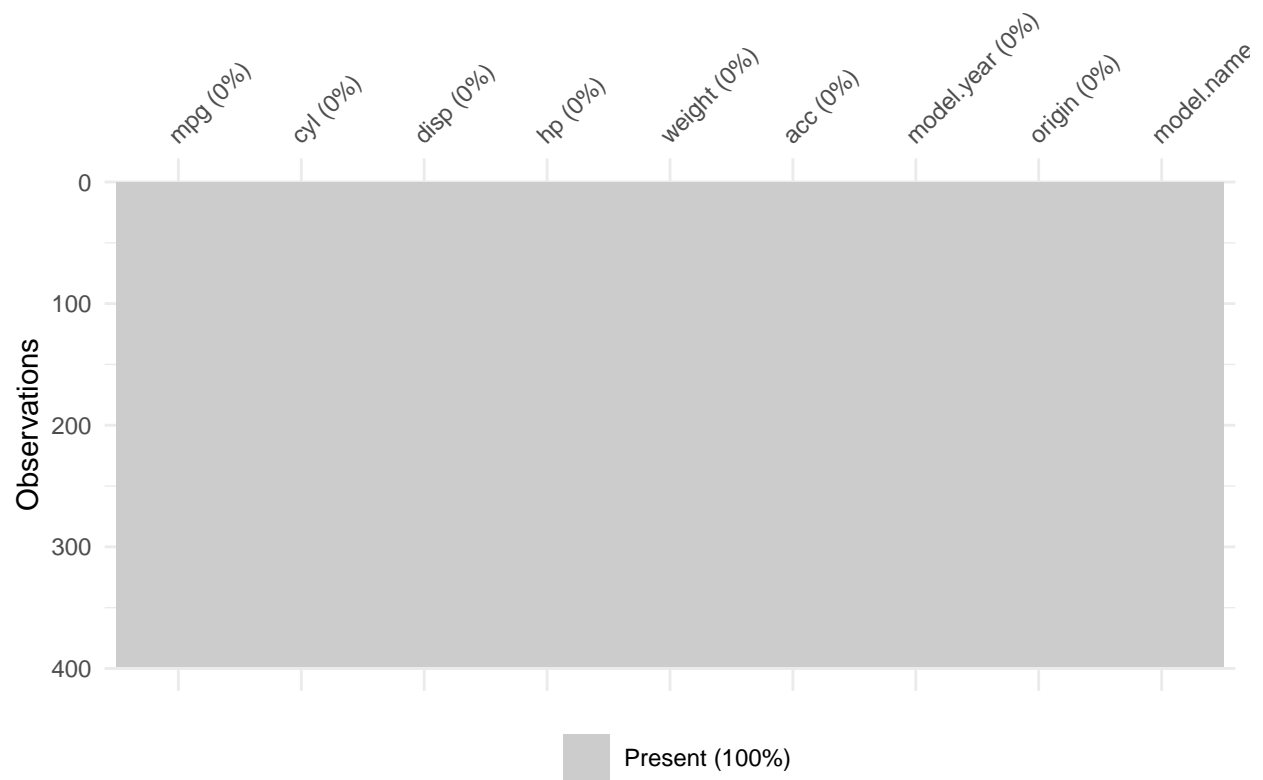
```
##
## lessR 3.8.8      feedback: gerbing@pdx.edu      web: lessRstats.com/new
## -----
## 1. d <- Read("")          Read text, Excel, SPSS, SAS or R data file
##                          d: default data frame, no need for data=
## 2. l <- Read("", var_labels=TRUE) Read variable labels into l,
##                          required name for data frame of labels
## 3. Help()                Get help, and, e.g., Help(Read)
## 4. hs(), bc(), or ca()   All histograms, all bar charts, or both
## 5. Plot(X) or Plot(X,Y)  For continuous and categorical variables
## 6. by1= , by2=          Trellis graphics, a plot for each by1, by2
## 7. reg(Y ~ X, Rmd="eg")  Regression with full interpretative output
## 8. style("gray")        Grayscale theme, + many others available
##   style(show=TRUE)      all color/style options and current values
## 9. getColors()          create many styles of color palettes
##
## lessR parameter names now use _'s. Names with a period are deprecated.
## Ex: bin_width instead of bin.width
```

```
library(bestNormalize)
library(mltools)
```

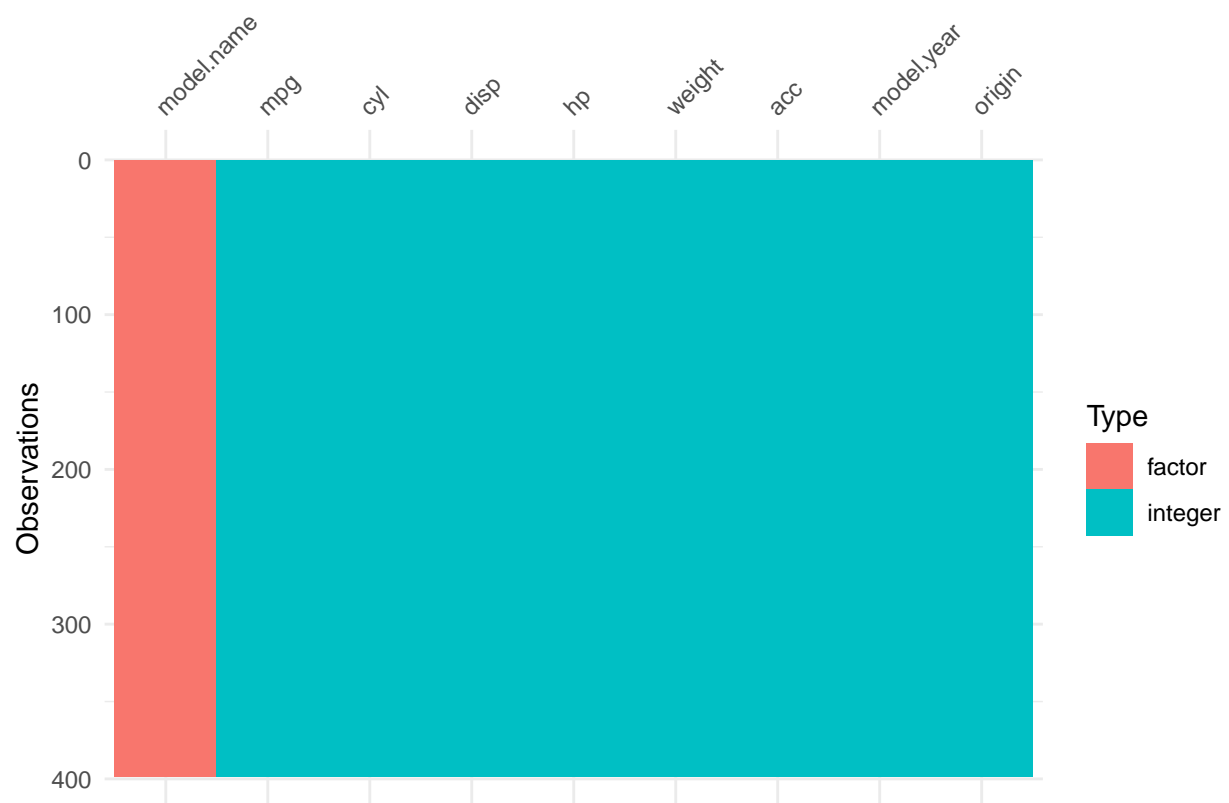
```
autos <- read.csv('auto-mpg_g.csv', header = T, na.strings = '?')
```

2. Utilizando Ggpairs cree un gráfico de los atributos del dataset, observe las correlaciones entre atributos

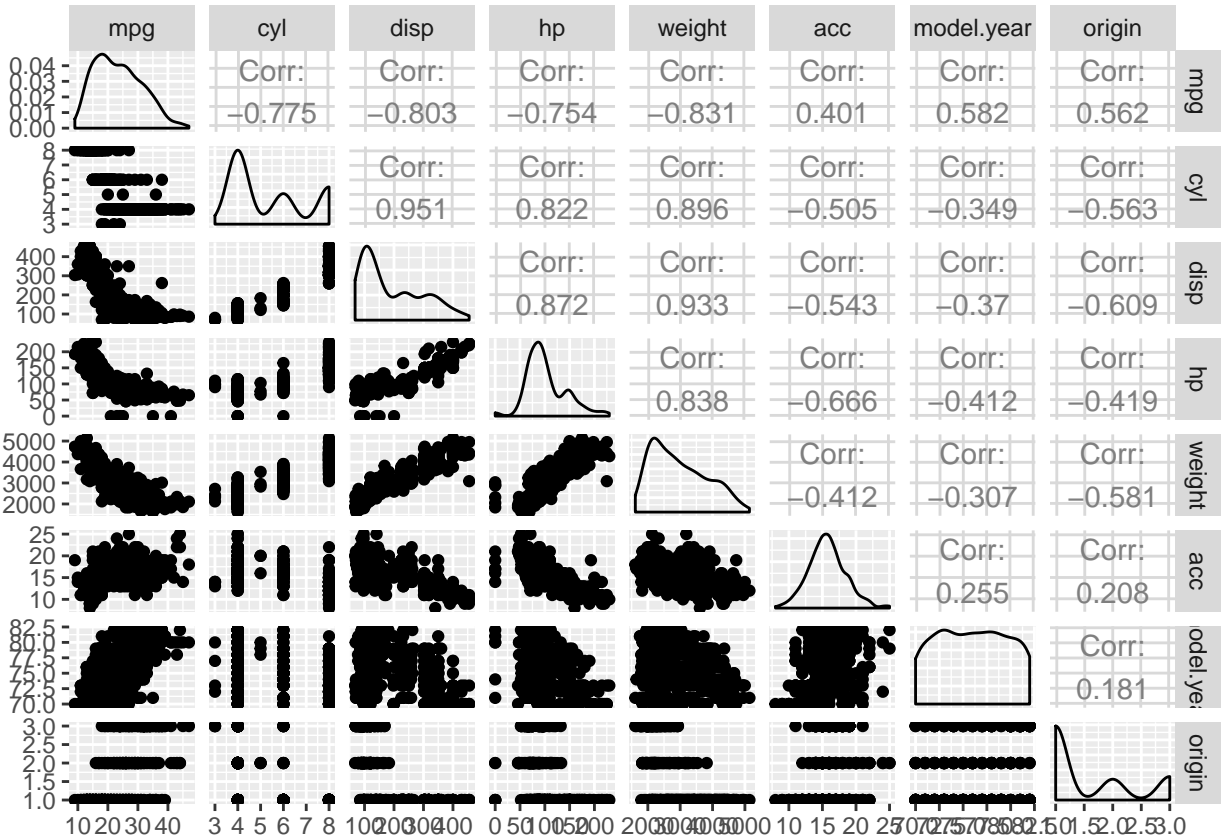
```
vis_miss(autos) # No hay datos faltantes
```



```
vis_dat(autos)
```



```
autos <- autos[,1:8]  
ggpairs(autos, columns = 1:8)
```



3. Separe los datos en 2 conjuntos, uno de entrenamiento y otro de pruebas. Normalmente se trabaja utilizando un 70-80% de los datos para entrenamiento y el resto para pruebas.

Recuerde fijar una semilla para que el documento sea reproducible.

Pista: <https://www.rdocumentation.org/packages/caTools/versions/1.17.1/topics/sample.split>

autos

```
##      mpg cyl disp  hp weight  acc model.year origin
## 1    18   8  307  130  3504   12         70       1
## 2    15   8  350  165  3693   12         70       1
## 3    18   8  318  150  3436   11         70       1
## 4    16   8  304  150  3433   12         70       1
## 5    17   8  302  140  3449   11         70       1
## 6    15   8  429  198  4341   10         70       1
## 7    14   8  454  220  4354    9         70       1
## 8    14   8  440  215  4312    9         70       1
## 9    14   8  455  225  4425   10         70       1
## 10   15   8  390  190  3850    9         70       1
## 11   15   8  383  170  3563   10         70       1
## 12   14   8  340  160  3609    8         70       1
## 13   15   8  400  150  3761   10         70       1
## 14   14   8  455  225  3086   10         70       1
## 15   24   4  113   95  2372   15         70       3
## 16   22   6  198   95  2833   16         70       1
```

## 17	18	6	199	97	2774	16	70	1
## 18	21	6	200	85	2587	16	70	1
## 19	27	4	97	88	2130	15	70	3
## 20	26	4	97	46	1835	21	70	2
## 21	25	4	110	87	2672	18	70	2
## 22	24	4	107	90	2430	15	70	2
## 23	25	4	104	95	2375	18	70	2
## 24	26	4	121	113	2234	13	70	2
## 25	21	6	199	90	2648	15	70	1
## 26	10	8	360	215	4615	14	70	1
## 27	10	8	307	200	4376	15	70	1
## 28	11	8	318	210	4382	14	70	1
## 29	9	8	304	193	4732	19	70	1
## 30	27	4	97	88	2130	15	71	3
## 31	28	4	140	90	2264	16	71	1
## 32	25	4	113	95	2228	14	71	3
## 33	19	6	232	100	2634	13	71	1
## 34	16	6	225	105	3439	16	71	1
## 35	17	6	250	100	3329	16	71	1
## 36	19	6	250	88	3302	16	71	1
## 37	18	6	232	100	3288	16	71	1
## 38	14	8	350	165	4209	12	71	1
## 39	14	8	400	175	4464	12	71	1
## 40	14	8	351	153	4154	14	71	1
## 41	14	8	318	150	4096	13	71	1
## 42	12	8	383	180	4955	12	71	1
## 43	13	8	400	170	4746	12	71	1
## 44	13	8	400	175	5140	12	71	1
## 45	18	6	258	110	2962	14	71	1
## 46	22	4	140	72	2408	19	71	1
## 47	19	6	250	100	3282	15	71	1
## 48	18	6	250	88	3139	15	71	1
## 49	23	4	122	86	2220	14	71	1
## 50	28	4	116	90	2123	14	71	2
## 51	30	4	79	70	2074	20	71	2
## 52	30	4	88	76	2065	15	71	2
## 53	31	4	71	65	1773	19	71	3
## 54	35	4	72	69	1613	18	71	3
## 55	27	4	97	60	1834	19	71	2
## 56	26	4	91	70	1955	21	71	1
## 57	24	4	113	95	2278	16	72	3
## 58	25	4	98	80	2126	17	72	1
## 59	23	4	97	54	2254	24	72	2
## 60	20	4	140	90	2408	20	72	1
## 61	21	4	122	86	2226	17	72	1
## 62	13	8	350	165	4274	12	72	1
## 63	14	8	400	175	4385	12	72	1
## 64	15	8	318	150	4135	14	72	1
## 65	14	8	351	153	4129	13	72	1
## 66	17	8	304	150	3672	12	72	1
## 67	11	8	429	208	4633	11	72	1
## 68	13	8	350	155	4502	14	72	1
## 69	12	8	350	160	4456	14	72	1
## 70	13	8	400	190	4422	13	72	1

## 71	19	3	70	97	2330	14	72	3
## 72	15	8	304	150	3892	13	72	1
## 73	13	8	307	130	4098	14	72	1
## 74	13	8	302	140	4294	16	72	1
## 75	14	8	318	150	4077	14	72	1
## 76	18	4	121	112	2933	15	72	2
## 77	22	4	121	76	2511	18	72	2
## 78	21	4	120	87	2979	20	72	2
## 79	26	4	96	69	2189	18	72	2
## 80	22	4	122	86	2395	16	72	1
## 81	28	4	97	92	2288	17	72	3
## 82	23	4	120	97	2506	15	72	3
## 83	28	4	98	80	2164	15	72	1
## 84	27	4	97	88	2100	17	72	3
## 85	13	8	350	175	4100	13	73	1
## 86	14	8	304	150	3672	12	73	1
## 87	13	8	350	145	3988	13	73	1
## 88	14	8	302	137	4042	15	73	1
## 89	15	8	318	150	3777	13	73	1
## 90	12	8	429	198	4952	12	73	1
## 91	13	8	400	150	4464	12	73	1
## 92	13	8	351	158	4363	13	73	1
## 93	14	8	318	150	4237	15	73	1
## 94	13	8	440	215	4735	11	73	1
## 95	12	8	455	225	4951	11	73	1
## 96	13	8	360	175	3821	11	73	1
## 97	18	6	225	105	3121	17	73	1
## 98	16	6	250	100	3278	18	73	1
## 99	18	6	232	100	2945	16	73	1
## 100	18	6	250	88	3021	17	73	1
## 101	23	6	198	95	2904	16	73	1
## 102	26	4	97	46	1950	21	73	2
## 103	11	8	400	150	4997	14	73	1
## 104	12	8	400	167	4906	13	73	1
## 105	13	8	360	170	4654	13	73	1
## 106	12	8	350	180	4499	13	73	1
## 107	18	6	232	100	2789	15	73	1
## 108	20	4	97	88	2279	19	73	3
## 109	21	4	140	72	2401	20	73	1
## 110	22	4	108	94	2379	17	73	3
## 111	18	3	70	90	2124	14	73	3
## 112	19	4	122	85	2310	19	73	1
## 113	21	6	155	107	2472	14	73	1
## 114	26	4	98	90	2265	16	73	2
## 115	15	8	350	145	4082	13	73	1
## 116	16	8	400	230	4278	10	73	1
## 117	29	4	68	49	1867	20	73	2
## 118	24	4	116	75	2158	16	73	2
## 119	20	4	114	91	2582	14	73	2
## 120	19	4	121	112	2868	16	73	2
## 121	15	8	318	150	3399	11	73	1
## 122	24	4	121	110	2660	14	73	2
## 123	20	6	156	122	2807	14	73	3
## 124	11	8	350	180	3664	11	73	1

##	125	20	6	198	95	3102	17	74	1
##	126	19	6	232	100	2901	16	74	1
##	127	15	6	250	100	3336	17	74	1
##	128	31	4	79	67	1950	19	74	3
##	129	26	4	122	80	2451	17	74	1
##	130	32	4	71	65	1836	21	74	3
##	131	25	4	140	75	2542	17	74	1
##	132	16	6	250	100	3781	17	74	1
##	133	16	6	258	110	3632	18	74	1
##	134	18	6	225	105	3613	17	74	1
##	135	16	8	302	140	4141	14	74	1
##	136	13	8	350	150	4699	15	74	1
##	137	14	8	318	150	4457	14	74	1
##	138	14	8	302	140	4638	16	74	1
##	139	14	8	304	150	4257	16	74	1
##	140	29	4	98	83	2219	17	74	2
##	141	26	4	79	67	1963	16	74	2
##	142	26	4	97	78	2300	15	74	2
##	143	31	4	76	52	1649	17	74	3
##	144	32	4	83	61	2003	19	74	3
##	145	28	4	90	75	2125	15	74	1
##	146	24	4	90	75	2108	16	74	2
##	147	26	4	116	75	2246	14	74	2
##	148	24	4	120	97	2489	15	74	3
##	149	26	4	108	93	2391	16	74	3
##	150	31	4	79	67	2000	16	74	2
##	151	19	6	225	95	3264	16	75	1
##	152	18	6	250	105	3459	16	75	1
##	153	15	6	250	72	3432	21	75	1
##	154	15	6	250	72	3158	20	75	1
##	155	16	8	400	170	4668	12	75	1
##	156	15	8	350	145	4440	14	75	1
##	157	16	8	318	150	4498	15	75	1
##	158	14	8	351	148	4657	14	75	1
##	159	17	6	231	110	3907	21	75	1
##	160	16	6	250	105	3897	19	75	1
##	161	15	6	258	110	3730	19	75	1
##	162	18	6	225	95	3785	19	75	1
##	163	21	6	231	110	3039	15	75	1
##	164	20	8	262	110	3221	14	75	1
##	165	13	8	302	129	3169	12	75	1
##	166	29	4	97	75	2171	16	75	3
##	167	23	4	140	83	2639	17	75	1
##	168	20	6	232	100	2914	16	75	1
##	169	23	4	140	78	2592	19	75	1
##	170	24	4	134	96	2702	14	75	3
##	171	25	4	90	71	2223	17	75	2
##	172	24	4	119	97	2545	17	75	3
##	173	18	6	171	97	2984	15	75	1
##	174	29	4	90	70	1937	14	75	2
##	175	19	6	232	90	3211	17	75	1
##	176	23	4	115	95	2694	15	75	2
##	177	23	4	120	88	2957	17	75	2
##	178	22	4	121	98	2945	15	75	2

## 179	25	4	121	115	2671	14	75	2
## 180	33	4	91	53	1795	18	75	3
## 181	28	4	107	86	2464	16	76	2
## 182	25	4	116	81	2220	17	76	2
## 183	25	4	140	92	2572	15	76	1
## 184	26	4	98	79	2255	18	76	1
## 185	27	4	101	83	2202	15	76	2
## 186	18	8	305	140	4215	13	76	1
## 187	16	8	318	150	4190	13	76	1
## 188	16	8	304	120	3962	14	76	1
## 189	15	8	351	152	4215	13	76	1
## 190	22	6	225	100	3233	15	76	1
## 191	22	6	250	105	3353	15	76	1
## 192	24	6	200	81	3012	18	76	1
## 193	23	6	232	90	3085	18	76	1
## 194	29	4	85	52	2035	22	76	1
## 195	25	4	98	60	2164	22	76	1
## 196	29	4	90	70	1937	14	76	2
## 197	33	4	91	53	1795	17	76	3
## 198	20	6	225	100	3651	18	76	1
## 199	18	6	250	78	3574	21	76	1
## 200	19	6	250	110	3645	16	76	1
## 201	18	6	258	95	3193	18	76	1
## 202	30	4	97	71	1825	12	76	2
## 203	32	4	85	70	1990	17	76	3
## 204	28	4	97	75	2155	16	76	3
## 205	27	4	140	72	2565	14	76	1
## 206	20	4	130	102	3150	16	76	2
## 207	13	8	318	150	3940	13	76	1
## 208	19	4	120	88	3270	22	76	2
## 209	19	6	156	108	2930	16	76	3
## 210	17	6	168	120	3820	17	76	2
## 211	17	8	350	180	4380	12	76	1
## 212	13	8	350	145	4055	12	76	1
## 213	13	8	302	130	3870	15	76	1
## 214	13	8	318	150	3755	14	76	1
## 215	32	4	98	68	2045	19	77	3
## 216	30	4	111	80	2155	15	77	1
## 217	36	4	79	58	1825	19	77	2
## 218	26	4	122	96	2300	16	77	1
## 219	34	4	85	70	1945	17	77	3
## 220	18	8	305	145	3880	13	77	1
## 221	17	8	260	110	4060	19	77	1
## 222	16	8	318	145	4140	14	77	1
## 223	15	8	302	130	4295	15	77	1
## 224	18	6	250	110	3520	16	77	1
## 225	21	6	231	105	3425	17	77	1
## 226	19	6	225	100	3630	18	77	1
## 227	19	6	250	98	3525	19	77	1
## 228	16	8	400	180	4220	11	77	1
## 229	16	8	350	170	4165	11	77	1
## 230	16	8	400	190	4325	12	77	1
## 231	16	8	351	149	4335	15	77	1
## 232	29	4	97	78	1940	15	77	2

##	233	25	4	151	88	2740	16	77	1
##	234	26	4	97	75	2265	18	77	3
##	235	26	4	140	89	2755	16	77	1
##	236	31	4	98	63	2051	17	77	1
##	237	34	4	98	83	2075	16	77	1
##	238	30	4	97	67	1985	16	77	3
##	239	31	4	97	78	2190	14	77	2
##	240	22	6	146	97	2815	15	77	3
##	241	22	4	121	110	2600	13	77	2
##	242	22	3	80	110	2720	14	77	3
##	243	43	4	90	48	1985	22	78	2
##	244	36	4	98	66	1800	14	78	1
##	245	33	4	78	52	1985	19	78	3
##	246	39	4	85	70	2070	19	78	3
##	247	36	4	91	60	1800	16	78	3
##	248	20	8	260	110	3365	16	78	1
##	249	19	8	318	140	3735	13	78	1
##	250	20	8	302	139	3570	13	78	1
##	251	19	6	231	105	3535	19	78	1
##	252	21	6	200	95	3155	18	78	1
##	253	20	6	200	85	2965	16	78	1
##	254	25	4	140	88	2720	15	78	1
##	255	21	6	225	100	3430	17	78	1
##	256	19	6	232	90	3210	17	78	1
##	257	21	6	231	105	3380	16	78	1
##	258	21	6	200	85	3070	17	78	1
##	259	19	6	225	110	3620	19	78	1
##	260	18	6	258	120	3410	15	78	1
##	261	19	8	305	145	3425	13	78	1
##	262	18	6	231	165	3445	13	78	1
##	263	18	8	302	139	3205	11	78	1
##	264	18	8	318	140	4080	14	78	1
##	265	30	4	98	68	2155	17	78	1
##	266	28	4	134	95	2560	14	78	3
##	267	27	4	119	97	2300	15	78	3
##	268	31	4	105	75	2230	15	78	1
##	269	21	4	134	95	2515	15	78	3
##	270	23	4	156	105	2745	17	78	1
##	271	24	4	151	85	2855	18	78	1
##	272	24	4	119	97	2405	15	78	3
##	273	20	5	131	103	2830	16	78	2
##	274	17	6	163	125	3140	14	78	2
##	275	22	4	121	115	2795	16	78	2
##	276	16	6	163	133	3410	16	78	2
##	277	32	4	89	71	1990	15	78	2
##	278	30	4	98	68	2135	17	78	3
##	279	22	6	231	115	3245	15	79	1
##	280	20	6	200	85	2990	18	79	1
##	281	22	4	140	88	2890	17	79	1
##	282	20	6	232	90	3265	18	79	1
##	283	21	6	225	110	3360	17	79	1
##	284	17	8	305	130	3840	15	79	1
##	285	18	8	302	129	3725	13	79	1
##	286	17	8	351	138	3955	13	79	1

##	287	18	8	318	135	3830	15	79	1
##	288	17	8	350	155	4360	15	79	1
##	289	16	8	351	142	4054	14	79	1
##	290	19	8	267	125	3605	15	79	1
##	291	19	8	360	150	3940	13	79	1
##	292	32	4	89	71	1925	14	79	2
##	293	34	4	86	65	1975	15	79	3
##	294	36	4	98	80	1915	14	79	1
##	295	27	4	121	80	2670	15	79	1
##	296	25	5	183	77	3530	20	79	2
##	297	23	8	350	125	3900	17	79	1
##	298	27	4	141	71	3190	25	79	2
##	299	24	8	260	90	3420	22	79	1
##	300	34	4	105	70	2200	13	79	1
##	301	35	4	105	70	2150	15	79	1
##	302	32	4	85	65	2020	19	79	3
##	303	37	4	91	69	2130	15	79	2
##	304	28	4	151	90	2670	16	79	1
##	305	29	6	173	115	2595	11	79	1
##	306	27	6	173	115	2700	13	79	1
##	307	34	4	151	90	2556	13	79	1
##	308	42	4	98	76	2144	15	80	2
##	309	38	4	89	60	1968	19	80	3
##	310	32	4	98	70	2120	16	80	1
##	311	37	4	86	65	2019	16	80	3
##	312	28	4	151	90	2678	17	80	1
##	313	26	4	140	88	2870	18	80	1
##	314	24	4	151	90	3003	20	80	1
##	315	19	6	225	90	3381	19	80	1
##	316	34	4	97	78	2188	16	80	2
##	317	30	4	134	90	2711	16	80	3
##	318	31	4	120	75	2542	18	80	3
##	319	37	4	119	92	2434	15	80	3
##	320	32	4	108	75	2265	15	80	3
##	321	47	4	86	65	2110	18	80	3
##	322	28	4	156	105	2800	14	80	1
##	323	41	4	85	65	2110	19	80	3
##	324	44	4	90	48	2085	22	80	2
##	325	43	4	90	48	2335	24	80	2
##	326	36	5	121	67	2950	20	80	2
##	327	30	4	146	67	3250	22	80	2
##	328	45	4	91	67	1850	14	80	3
##	329	34	4	97	67	2145	18	80	3
##	330	30	4	89	62	1845	15	80	2
##	331	33	6	168	132	2910	11	80	3
##	332	24	3	70	100	2420	13	80	3
##	333	35	4	122	88	2500	15	80	2
##	334	32	4	107	72	2290	17	80	3
##	335	27	4	135	84	2490	16	81	1
##	336	27	4	151	84	2635	16	81	1
##	337	26	4	156	92	2620	14	81	1
##	338	24	6	173	110	2725	13	81	1
##	339	30	4	135	84	2385	13	81	1
##	340	39	4	79	58	1755	17	81	3

##	341	39	4	86	64	1875	16	81	1
##	342	35	4	81	60	1760	16	81	3
##	343	32	4	97	67	2065	18	81	3
##	344	37	4	85	65	1975	19	81	3
##	345	38	4	89	62	2050	17	81	3
##	346	34	4	91	68	1985	16	81	3
##	347	35	4	105	63	2215	15	81	1
##	348	34	4	98	65	2045	16	81	1
##	349	30	4	98	65	2380	21	81	1
##	350	33	4	105	74	2190	14	81	2
##	351	34	4	107	75	2210	14	81	3
##	352	32	4	108	75	2350	17	81	3
##	353	33	4	119	100	2615	15	81	3
##	354	32	4	120	74	2635	18	81	3
##	355	28	4	141	80	3230	20	81	2
##	356	31	6	145	76	3160	20	81	2
##	357	25	6	168	116	2900	13	81	3
##	358	24	6	146	120	2930	14	81	3
##	359	22	6	231	110	3415	16	81	1
##	360	27	8	350	105	3725	19	81	1
##	361	20	6	200	88	3060	17	81	1
##	362	18	6	225	85	3465	17	81	1
##	363	28	4	112	88	2605	20	82	1
##	364	27	4	112	88	2640	19	82	1
##	365	34	4	112	88	2395	18	82	1
##	366	31	4	112	85	2575	16	82	1
##	367	29	4	135	84	2525	16	82	1
##	368	27	4	151	90	2735	18	82	1
##	369	24	4	140	92	2865	16	82	1
##	370	36	4	105	74	1980	15	82	2
##	371	37	4	91	68	2025	18	82	3
##	372	31	4	91	68	1970	18	82	3
##	373	38	4	105	63	2125	15	82	1
##	374	36	4	98	70	2125	17	82	1
##	375	36	4	120	88	2160	15	82	3
##	376	36	4	107	75	2205	15	82	3
##	377	34	4	108	70	2245	17	82	3
##	378	38	4	91	67	1965	15	82	3
##	379	32	4	91	67	1965	16	82	3
##	380	38	4	91	67	1995	16	82	3
##	381	25	6	181	110	2945	16	82	1
##	382	38	6	262	85	3015	17	82	1
##	383	26	4	156	92	2585	15	82	1
##	384	22	6	232	112	2835	15	82	1
##	385	32	4	144	96	2665	14	82	3
##	386	36	4	135	84	2370	13	82	1
##	387	27	4	151	90	2950	17	82	1
##	388	27	4	140	86	2790	16	82	1
##	389	44	4	97	52	2130	25	82	2
##	390	32	4	135	84	2295	12	82	1
##	391	28	4	120	79	2625	19	82	1
##	392	31	4	119	82	2720	19	82	1
##	393	25	4	98	0	2046	19	71	1
##	394	21	6	200	0	2875	17	74	1

```
## 395 41 4 85 0 1835 17 80 2
## 396 24 4 140 0 2905 14 80 1
## 397 35 4 100 0 2320 16 81 2
## 398 23 4 151 0 3035 21 82 1
```

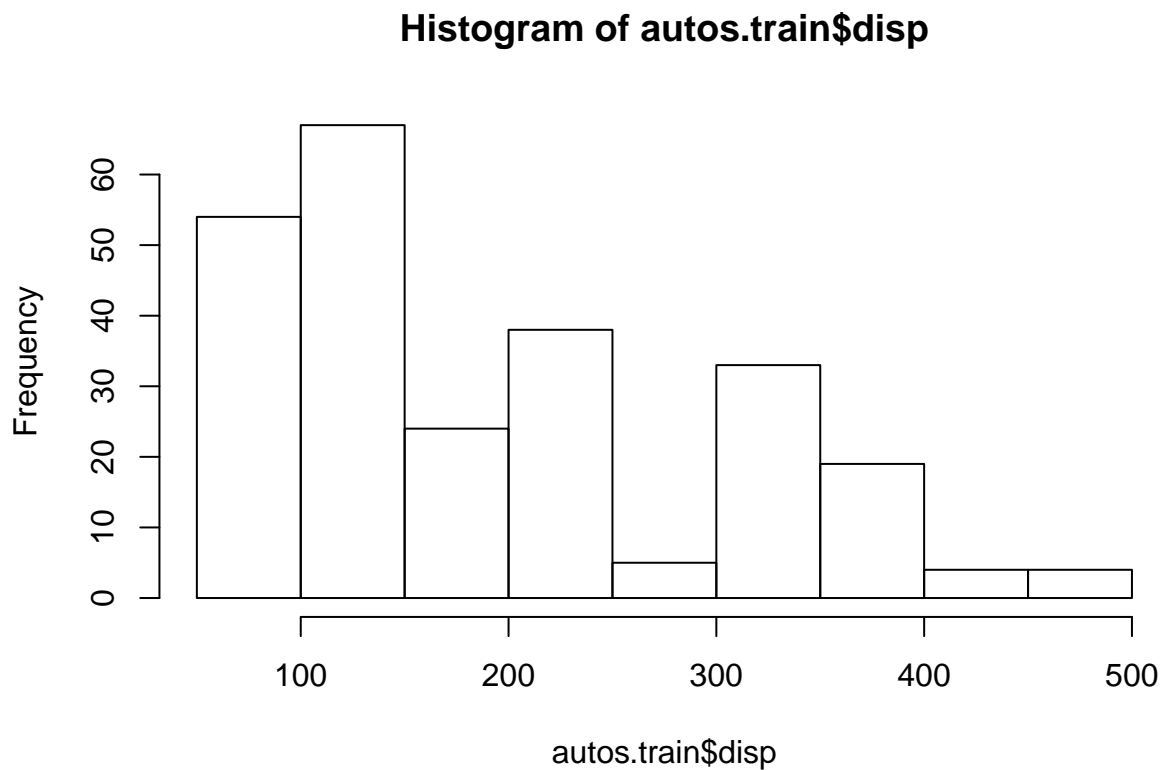
```
result <- sample.split(autos, SplitRatio = .7)
autos.train <- autos[result==TRUE, ]
autos.test <- autos[result==FALSE, ]
```

4. Cree un modelo de regresión lineal utilizando el atributo mpg como la variable objetivo y en base a las correlaciones observadas en el gráfico del punto 2 escoja al menos dos atributos para usarlos como variables predictoras para el modelo.

Pista: <https://www.rdocumentation.org/packages/lessR/versions/1.9.8/topics/reg>

Nota: Al crear el modelo utilice el conjunto de datos de entrenamiento definido en el punto 3.

```
#hist(log10(autos.train$disp))
#hist(log(autos.train$disp))
#hist(sqrt(autos.train$disp))
#hist(orderNorm(autos.train$disp)$x.t) # Ordered Quantile Normalization
hist(autos.train$disp)
```

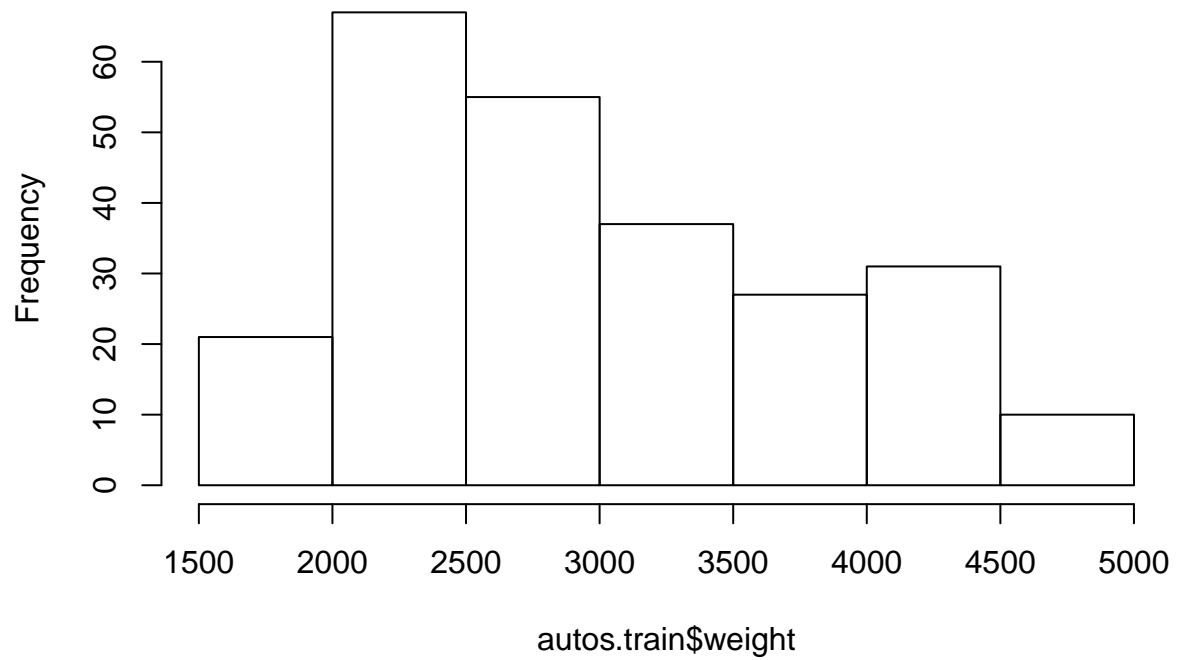


```
autos.train$disp <- orderNorm(autos.train$disp)$x.t  
hist(autos.train$disp)
```



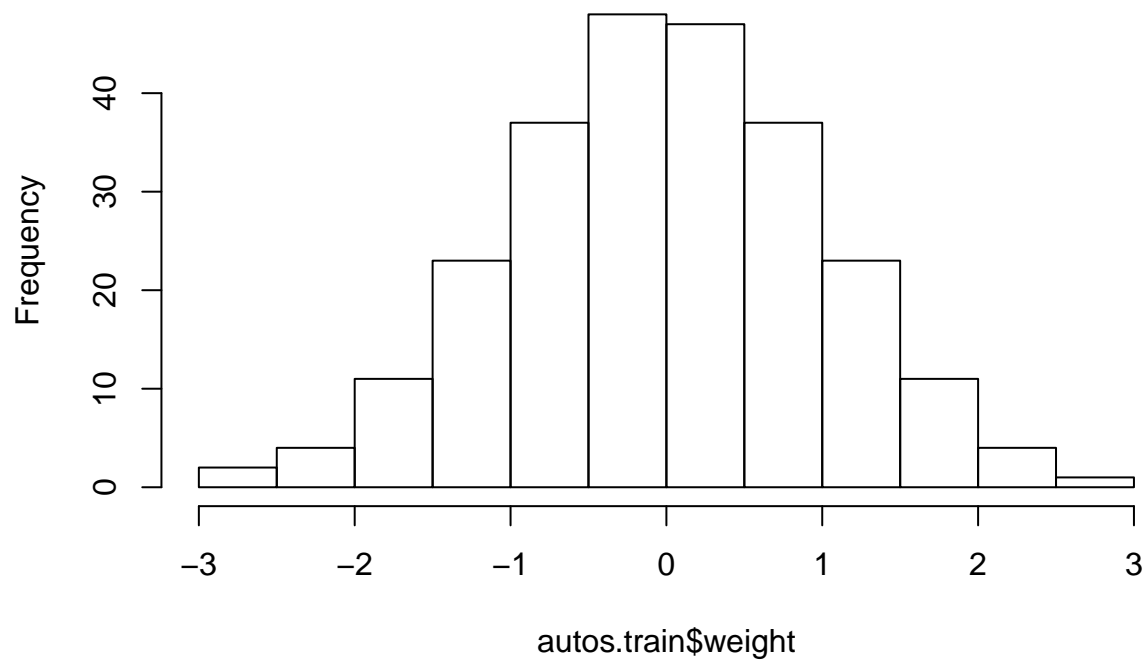
```
#hist(log10(autos.train$weight))  
#hist(log(autos.train$weight))  
#hist(sqrt(autos.train$weight))  
#hist(orderNorm(autos.train$weight)$x.t) # Ordered Quantile Normalization  
hist(autos.train$weight)
```

Histogram of autos.train\$weight



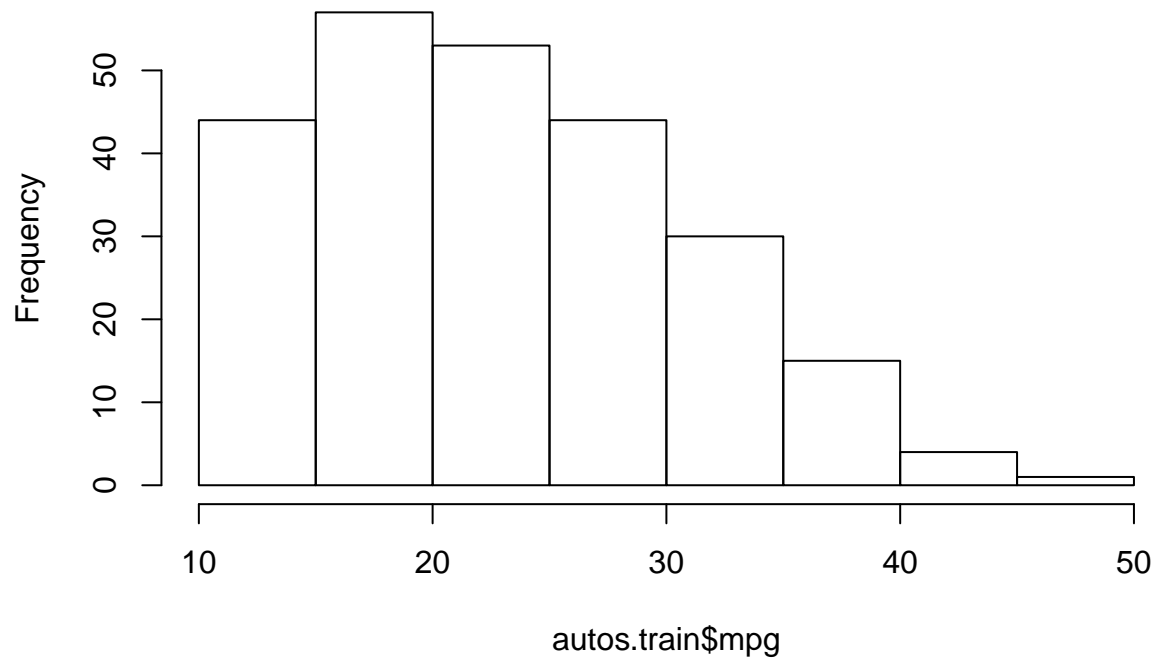
```
autos.train$weight <- orderNorm(autos.train$weight)$x.t  
hist(autos.train$weight)
```

Histogram of autos.train\$weight



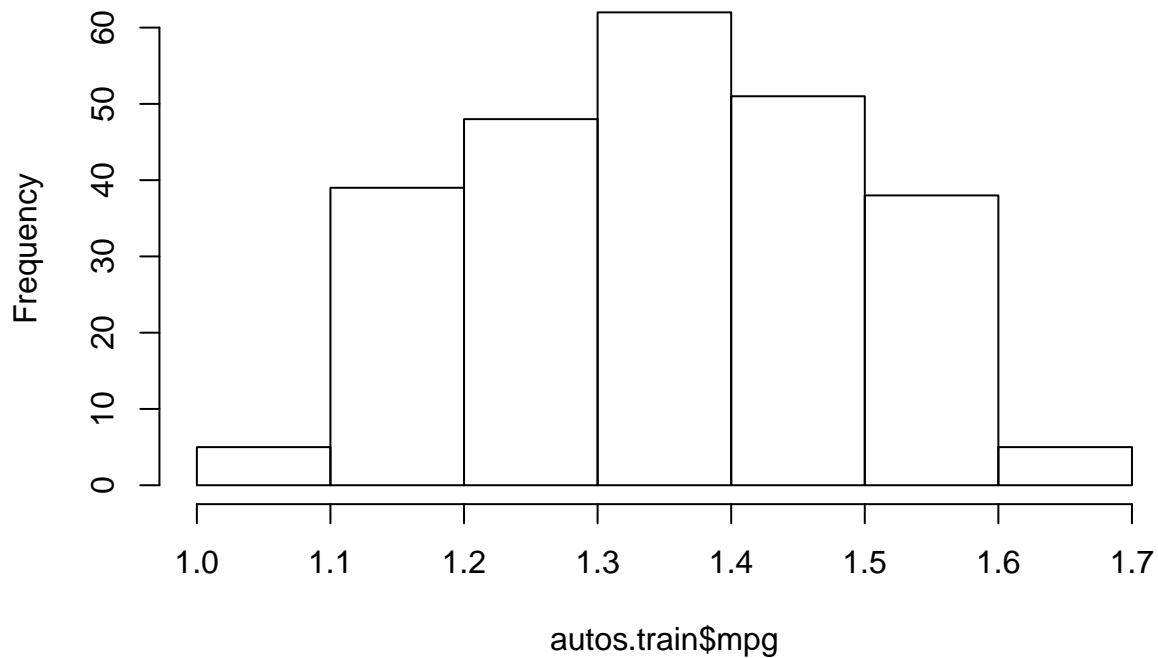
```
#hist(log10(autos.train$mpg))  
#hist(log(autos.train$mpg))  
#hist(sqrt(autos.train$mpg))  
#hist(orderNorm(autos.train$mpg)$x.t) # Ordered Quantile Normalization  
hist(autos.train$mpg)
```


Histogram of autos.train\$mpg



```
autos.train$mpg <- log10(autos.train$mpg)
hist(autos.train$mpg)
```

Histogram of autos.train\$mpg



```
autos.test$disp <- orderNorm(autos.test$disp)$x.t
autos.test$weight <- orderNorm(autos.test$weight)$x.t
autos.test$mpg <- log10(autos.test$mpg)
```

```
# reg <- lessR::reg(mpg ~ disp + weight, data=autos.train)
reg <- lm(formula = mpg ~ disp + weight, data = autos.train) # https://github.com/cran/lessR/blob/master
```

5. Realice predicciones utilizando el conjunto de pruebas y evalúe el resultado con la métrica MSE.

Pista: <https://www.rdocumentation.org/packages/mltools/versions/0.3.5/topics/mse>

```
library(Metrics)
```

```
##
## Attaching package: 'Metrics'
```

```
## The following objects are masked from 'package:mltools':
##
##     mse, msle, rmse, rmsle
```

```
prediccion <- predict(reg, autos.test)
prediccion[1]
```

```
##          2
## 1.239454
```

```
autos.test$mpg[1]
```

```
## [1] 1.176091
```

```
mltools::mse(preds = prediccion, actuals=autos.test$mpg)
```

```
## [1] 0.006382263
```

6. Opcional

6.a Pruebe varios modelos que utilicen diferentes variables y comparar los resultados obtenidos

6.b Investigar como implementar en R las técnicas de preprocesado y normalización vistas en clase y aplicarlas a los datos antes de pasarlos al modelo.