

# Cambio estructural

Antonio Huerta Montellano

7 de agosto del 2022

Importando las librerías:

## Bases de datos

Veamos cuál es el directorio de trabajo

```
getwd()
```

```
## [1] "/home/usuario/Documentos/Github/Proyectos/MLB_HN/Models/Linear_models/Free_agent/Test"
```

Cambiamos el directorio de trabajo y carguemos las bases de datos para el modelo lineal en el mismo chunk:

```
setwd("~/Documentos/Github/Proyectos/MLB_HN/")
catchers_panel_data <- read.csv('Data/New_Data/Models/Article/article_catchers_panel_data.csv')
relief_pitchers_panel_data <- read.csv('Data/New_Data/Models/Article/article_relief_pitchers_panel_data.csv')
starting_pitchers_panel_data <- read.csv('Data/New_Data/Models/Article/article_starting_pitchers_panel_data.csv')
remain_hitters_panel_data <- read.csv('Data/New_Data/Models/Article/article_remain_hitters_panel_data.csv')
```

Observemos el contenido de las bases de datos de los catchers:

```
head(catchers_panel_data)
```

```
##           Jugador Anio_t Anios_contrato_t Valor_contrato_t
## 1 A.J. Pierzynski      4              1         500000
## 2 A.J. Pierzynski      5              1        2000000
## 3 A.J. Pierzynski      6              1        3000000
## 4   Alex Avila        8              2        8250000
## 5   Alex Avila       11              1        1500000
## 6   Alex Avila        7              1        2000000
## Valor_promedio_contrato_t Juegos_t Porcetrnaje_juegos_t At_bats_t Bateos_t
## 1              500000      102         0.630      338      85
## 2             2000000      113         0.698      407     123
## 3             3000000       81         0.503      247      55
## 4             4125000       81         0.500      194      32
## 5             1500000       34         0.210       89      17
## 6             2000000      112         0.691      311      81
## Home_runs_t RBI_t Porcentaje_bateo_t OPS_t ln_Sueldo_base_t ln_Sueldo_t
## 1           5    37          0.252 0.625          12.02897    12.02897
```

## 2	9	49	0.302	0.774	14.50866	14.69098
## 3	2	23	0.223	0.559	14.50866	14.93065
## 4	7	20	0.165	0.603	15.20180	15.20180
## 5	1	9	0.191	0.694	14.22098	14.22098
## 6	14	49	0.261	0.828	14.50866	14.50866
##	ln_Sueldo_regular_t	Juegos_2_t	Porcetnaje_juegos_2_t	At_bats_2_t	Bateos_2_t	
## 1	12.02897	10404		0.396900	114244	7225
## 2	14.69098	12769		0.487204	165649	15129
## 3	14.93065	6561		0.253009	61009	3025
## 4	15.20180	6561		0.250000	37636	1024
## 5	14.22098	1156		0.044100	7921	289
## 6	13.44231	12544		0.477481	96721	6561
##	Home_runs_2_t	RBI_2_t	Porcentaje_bateo_2_t	OPS_2_t	Anio_t_1	
## 1	25	1369	0.063504	0.390625	3	
## 2	81	2401	0.091204	0.599076	4	
## 3	4	529	0.049729	0.312481	5	
## 4	49	400	0.027225	0.363609	7	
## 5	1	81	0.036481	0.481636	10	
## 6	196	2401	0.068121	0.685584	6	
##	Anios_contrato_t_1	Valor_contrato_t_1	Valor_promedio_contrato_t_1	Juegos_t_1		
## 1	1	7500000		7500000	134	
## 2	1	500000		500000	102	
## 3	1	2000000		2000000	113	
## 4	1	2000000		2000000	112	
## 5	1	4250000		4250000	23	
## 6	1	2500000		2500000	57	
##	Porcetnaje_juegos_t_1	At_bats_t_1	Bateos_t_1	Home_runs_t_1	RBI_t_1	
## 1	0.822	503	137	17	70	
## 2	0.630	338	85	5	37	
## 3	0.698	407	123	9	49	
## 4	0.691	311	81	14	49	
## 5	0.383	49	9	1	2	
## 6	0.352	169	36	7	11	
##	Porcentaje_bateo_t_1	OPS_t_1	Juegos_2_t_1	Porcetnaje_juegos_2_t_1		
## 1	0.272	0.722	17956	0.675684		
## 2	0.252	0.625	10404	0.396900		
## 3	0.302	0.774	12769	0.487204		
## 4	0.261	0.828	12544	0.477481		
## 5	0.184	0.641	529	0.146689		
## 6	0.213	0.732	3249	0.123904		
##	At_bats_2_t_1	Bateos_2_t_1	Home_runs_2_t_1	RBI_2_t_1	Porcentaje_bateo_2_t_1	
## 1	253009	18769	289	4900	0.073984	
## 2	114244	7225	25	1369	0.063504	
## 3	165649	15129	81	2401	0.091204	
## 4	96721	6561	196	2401	0.068121	
## 5	2401	81	1	4	0.033856	
## 6	28561	1296	49	121	0.045369	
##	OPS_2_t_1					
## 1	0.521284					
## 2	0.390625					
## 3	0.599076					
## 4	0.685584					
## 5	0.410881					
## 6	0.535824					

Ahora de los que son relief pitchers:

```
head(relief_pitchers_panel_data)
```

```
##      Jugador Anio_t Anios_contrato_t Valor_contrato_t
## 1 Blake Treinen    11          2      17500000
## 2 Brad Brach      10          2      2100000
## 3 Brandon Morrow  6          1          0
## 4 Brian Duensing   7          1      2000000
## 5 Brian Duensing   8          2      7000000
## 6 Brian Wilson     4          2     19000000
## Valor_promedio_contrato_t Juegos_t Juegos_iniciados_t Inning_pitched_t
## 1      8750000      72          0      72.3
## 2     1050000      14          0      12.3
## 3          0      18          0      16.0
## 4     2000000      68          0      62.3
## 5     3500000      49          0      37.7
## 6     9500000      61          0      48.3
## Bateos_pitcher_t Carreras_t Carreras_ganadas_t Walks_t Strike_outs_t Wins_t
## 1          46          20          16      25      85      6
## 2           8           8           8      14      14      1
## 3          19           4           3      3       8      1
## 4          58          19          19      18      61      1
## 5          42          33          32      29      24      3
## 6          49          26          25      29      54      2
## Losses_t Saves_t WHIP_t ERA_t Equipo_t ln_Sueldo_base_t ln_Sueldo_t
## 1          5          7  0.98  1.99    LAD      15.60727      15.89495
## 2          0          0  1.78  5.84    NYM      13.65299      13.79019
## 3          0          0  1.38  1.69    SD       13.13725      13.31711
## 4          1          0  1.22  2.74    CHC      14.50866      14.50866
## 5          0          1  1.88  7.65    CHC      15.06827      15.06827
## 6          4          1  1.61  4.66    LAD      16.11810      16.11810
## ln_Sueldo_regular_t Dominio_t Control_t Comando_t Juegos_2_t
## 1      15.89495  0.13062855  0.03842016  3.4000000      5184
## 2      12.65974  0.12646793  0.12646793  1.0000000      196
## 3      11.83971  0.05555556  0.02083333  2.6666667      324
## 4      14.50866  0.10879258  0.03210273  3.3888889     4624
## 5      15.06827  0.07073386  0.08547009  0.8275862     2401
## 6      16.11810  0.12422360  0.06671268  1.8620690     3721
## Juegos_iniciados_2_t Inning_pitched_2_t Bateos_pitcher_2_t Carreras_2_t
## 1          0          5227.29      2116      400
## 2          0          151.29       64      64
## 3          0          256.00      361      16
## 4          0          3881.29     3364     361
## 5          0          1421.29     1764    1089
## 6          0          2332.89     2401     676
## Carreras_ganadas_2_t Walks_2_t Strike_outs_2_t Wins_2_t Losses_2_t Saves_2_t
## 1      256      625      7225      36      25      49
## 2       64      196      196       1       0       0
## 3        9        9       64       1       0       0
## 4      361      324      3721       1       1       0
## 5     1024      841       576       9       0       1
## 6      625      841      2916       4      16       1
## WHIP_2_t ERA_2_t Control_2_t Comando_2_t Anio_t_1 Anios_contrato_t_1
```

## 1	0.9604	3.9601	0.0014761089	11.5600000	10	1	
## 2	3.1684	34.1056	0.0159941377	1.0000000	9	0	
## 3	1.9044	2.8561	0.0004340278	7.1111111	5	1	
## 4	1.4884	7.5076	0.0010305852	11.4845679	6	1	
## 5	3.5344	58.5225	0.0073051355	0.6848989	7	1	
## 6	2.5921	21.7156	0.0044505811	3.4673008	3	1	
##	Valor_contrato_t_1	Valor_promedio_contrato_t_1	Juegos_t_1				
## 1	10000000	10000000	27				
## 2	0	0	59				
## 3	2500000	2500000	5				
## 4	0	0	14				
## 5	2000000	2000000	68				
## 6	1000000	1000000	18				
##	Juegos_iniciados_t_1	Inning_pitched_t_1	Bateos_pitcher_t_1	Carreras_t_1			
## 1	0	25.7	23	15			
## 2	0	54.3	57	33			
## 3	5	33.0	29	10			
## 4	0	13.3	13	6			
## 5	0	62.3	58	19			
## 6	0	13.7	8	1			
##	Carreras_ganadas_t_1	Walks_t_1	Strike_outs_t_1	Wins_t_1	Losses_t_1	Saves_t_1	
## 1	11	8	22	3	3	1	
## 2	33	31	60	5	4	0	
## 3	10	7	23	2	0	0	
## 4	6	3	10	1	0	0	
## 5	19	18	61	1	1	0	
## 6	1	4	13	2	1	0	
##	WHIP_t_1	ERA_t_1	Equipo_t_1	Dominio_t_1	Control_t_1	Comando_t_1	Juegos_2_t_1
## 1	1.21	3.86	LAD	0.09511457	0.03458712	2.750000	729
## 2	1.62	5.47	NYM	0.12277471	0.06343360	1.935484	3481
## 3	1.09	2.73	SD	0.07744108	0.02356902	3.285714	25
## 4	1.20	4.05	BAL	0.08354219	0.02506266	3.333333	196
## 5	1.22	2.74	CHC	0.10879258	0.03210273	3.388889	4624
## 6	0.88	0.66	LAD	0.10543390	0.03244120	3.250000	324
##	Juegos_iniciados_2_t_1	Inning_pitched_2_t_1	Bateos_pitcher_2_t_1				
## 1	0	660.49	529				
## 2	0	2948.49	3249				
## 3	25	1089.00	841				
## 4	0	176.89	169				
## 5	0	3881.29	3364				
## 6	0	187.69	64				
##	Carreras_2_t_1	Carreras_ganadas_2_t_1	Walks_2_t_1	Strike_outs_2_t_1			
## 1	225	121	64	484			
## 2	1089	1089	961	3600			
## 3	100	100	49	529			
## 4	36	36	9	100			
## 5	361	361	324	3721			
## 6	1	1	16	169			
##	Wins_2_t_1	Losses_2_t_1	Saves_2_t_1	WHIP_2_t_1	ERA_2_t_1	Control_2_t_1	
## 1	9	9	1	1.4641	14.8996	0.0011962686	
## 2	25	16	0	2.6244	29.9209	0.0040238215	
## 3	4	0	0	1.1881	7.4529	0.0005554989	
## 4	1	0	0	1.4400	16.4025	0.0006281368	
## 5	1	1	0	1.4884	7.5076	0.0010305852	

```
## 6      4      1      0      0.7744      0.4356      0.0010524315
## Comando_2_t_1
## 1      7.562500
## 2      3.746098
## 3      10.795918
## 4      11.111111
## 5      11.484568
## 6      10.562500
```

Starting pitchers:

```
head(starting_pitchers_panel_data)
```

```
##      Jugador Anio_t Anios_contrato_t Valor_contrato_t
## 1      A.J. Burnett      5      1      8500000
## 2 Adam Wainwright     11      1      8000000
## 3 Andrew Cashner      8      2     16000000
## 4 Anibal Sanchez      9      2     19000000
## 5 Bartolo Colon       4      2     20000000
## 6 Bartolo Colon       7      1     12500000
## Valor_promedio_contrato_t Juegos_t Juegos_iniciados_t Inning_pitched_t
## 1      8500000      26      26      164.0
## 2      8000000      34      32      206.3
## 3      8000000      30      28      153.0
## 4      9500000      30      30      166.0
## 5     10000000      31      31      202.3
## 6     12500000      28      28      143.0
## Bateos_pitcher_t Carreras_t Carreras_ganadas_t Walks_t Strike_outs_t Wins_t
## 1      174      64      58      49      143      9
## 2      168      72      70      50      174     17
## 3      177      97      90      65      99      5
## 4      153      77      71      58     134     11
## 5      218      97      92      30     151     15
## 6      192     112     103     35      89      7
## Losses_t Saves_t WHIP_t ERA_t Equipo_t ln_Sueldo_base_t ln_Sueldo_t
## 1      7      0      1.36      3.18      PIT      15.95558      15.95558
## 2      7      0      1.06      3.05      STL      15.89495      15.89495
## 3     16      0      1.58      5.29      BAL      15.42495      16.06680
## 4      8      0      1.27      3.85      WSH      15.89495      16.11810
## 5     13      0      1.23      4.09      NYM      16.01274      16.01274
## 6     14      0      1.59      6.48      MIN      13.19002      13.19002
## ln_Sueldo_regular_t Dominio_t Control_t Comando_t Juegos_2_t
## 1     15.95558      0.09688347      0.03319783      2.918367      676
## 2     15.89495      0.09371466      0.02692950      3.480000     1156
## 3     16.06680      0.07189542      0.04720407      1.523077      900
## 4     15.60727      0.08969210      0.03882195      2.310345      900
## 5     16.01274      0.08293513      0.01647718      5.033333      961
## 6     12.31110      0.06915307      0.02719503      2.542857      784
## Juegos_iniciados_2_t Inning_pitched_2_t Bateos_pitcher_2_t Carreras_2_t
## 1      676      26896.00      30276      4096
## 2     1024      42559.69      28224      5184
## 3      784      23409.00      31329      9409
## 4      900      27556.00      23409      5929
```

## 5	961	40925.29	47524	9409		
## 6	784	20449.00	36864	12544		
##	Carreras_ganadas_2_t	Walks_2_t	Strike_outs_2_t	Wins_2_t	Losses_2_t	Saves_2_t
## 1	3364	2401	20449	81	49	0
## 2	4900	2500	30276	289	49	0
## 3	8100	4225	9801	25	256	0
## 4	5041	3364	17956	121	64	0
## 5	8464	900	22801	225	169	0
## 6	10609	1225	7921	49	196	0
##	WHIP_2_t	ERA_2_t	Control_2_t	Comando_2_t	Anio_t_1	Anios_contrato_t_1
## 1	1.8496	10.1124	0.0011020960	8.516868	4	1
## 2	1.1236	9.3025	0.0007251979	12.110400	10	1
## 3	2.4964	27.9841	0.002282239	2.319763	7	1
## 4	1.6129	14.8225	0.0015071442	5.337693	8	1
## 5	1.5129	16.7281	0.0002714974	25.334444	3	1
## 6	2.5281	41.9904	0.0007395695	6.466122	6	1
##	Valor_contrato_t_1	Valor_promedio_contrato_t_1	Juegos_t_1			
## 1	16000000	16000000	34			
## 2	5000000	5000000	10			
## 3	10000000	10000000	28			
## 4	2500000	2500000	25			
## 5	3000000	3000000	30			
## 6	7250000	7250000	34			
##	Juegos_iniciados_t_1	Inning_pitched_t_1	Bateos_pitcher_t_1	Carreras_t_1		
## 1	34	213.7	205	122		
## 2	10	65.7	54	25		
## 3	28	166.7	156	75		
## 4	24	136.7	106	48		
## 5	30	190.3	193	60		
## 6	33	191.7	200	81		
##	Carreras_ganadas_t_1	Walks_t_1	Strike_outs_t_1	Wins_t_1	Losses_t_1	Saves_t_1
## 1	109	96	190	8	18	0
## 2	23	15	54	5	3	0
## 3	63	64	86	11	11	0
## 4	43	42	135	7	6	0
## 5	56	29	117	18	6	0
## 6	73	32	128	15	8	0
##	WHIP_t_1	ERA_t_1	Equipo_t_1	Dominio_t_1	Control_t_1	Comando_t_1
## 1	1.41	4.59	PHI	0.09878854	0.04991421	1.979167
## 2	1.05	3.15	STL	0.09132420	0.02536783	3.600000
## 3	1.32	3.40	TEX	0.05732187	0.04265814	1.343750
## 4	1.08	2.83	ATL	0.10972933	0.03413802	3.214286
## 5	1.17	2.65	OAK	0.06831319	0.01693233	4.034483
## 6	1.21	3.43	NYM	0.07419000	0.01854750	4.000000
##	Juegos_iniciados_2_t_1	Inning_pitched_2_t_1	Bateos_pitcher_2_t_1			
## 1	1156	45667.69	42025			
## 2	100	4316.49	2916			
## 3	784	27788.89	24336			
## 4	576	18686.89	11236			
## 5	900	36214.09	37249			
## 6	1089	36748.89	40000			
##	Carreras_2_t_1	Carreras_ganadas_2_t_1	Walks_2_t_1	Strike_outs_2_t_1		
## 1	14884	11881	9216	36100		
## 2	625	529	225	2916		

```
## 3      5625      3969      4096      7396
## 4      2304      1849      1764      18225
## 5      3600      3136      841      13689
## 6      6561      5329      1024      16384
## Wins_2_t_1 Losses_2_t_1 Saves_2_t_1 WHIP_2_t_1 ERA_2_t_1 Control_2_t_1
## 1      64      324      0      1.9881      21.0681      0.0024914284
## 2      25      9      0      1.1025      9.9225      0.0006435270
## 3      121      121      0      1.7424      11.5600      0.0018197165
## 4      49      36      0      1.1664      8.0089      0.0011654041
## 5      324      36      0      1.3689      7.0225      0.0002867038
## 6      225      64      0      1.4641      11.7649      0.0003440097
## Comando_2_t_1
## 1      3.917101
## 2      12.960000
## 3      1.805664
## 4      10.331633
## 5      16.277051
## 6      16.000000
```

Remain hitters:

```
head(remain_hitters_panel_data)
```

```
##      Jugador Anio_t Anios_contrato_t Valor_contrato_t
## 1 Adeiny Hechavarria      10      1      1000000
## 2 Asdrubal Cabrera      11      1      1750000
## 3 Asdrubal Cabrera      10      1      2500000
## 4 Asdrubal Cabrera      6      2      18500000
## 5 Avisail Garcia      10      2      20000000
## 6 Billy Butler      6      1      49914
## Valor_promedio_contrato_t Juegos_t Porcetnaje_juegos_t At_bats_t Bateos_t
## 1      1000000      27      0.450      59      15
## 2      1750000      110      0.679      309      71
## 3      2500000      52      0.867      190      46
## 4      9250000      141      0.870      521      146
## 5      10000000      53      0.883      181      43
## 6      49914      97      0.599      250      71
## Home_runs_t RBI_t Porcentaje_bateo_t OPS_t ln_Sueldo_base_t ln_Sueldo_t
## 1      0      2      0.254 0.607      13.81551      13.81551
## 2      7      42      0.230 0.681      14.37513      14.37513
## 3      8      31      0.242 0.753      14.73180      14.80319
## 4      24      62      0.280 0.816      15.92572      15.92572
## 5      2      15      0.238 0.659      15.76142      15.79651
## 6      5      35      0.284 0.752      13.13725      13.13725
## ln_Sueldo_regular_t Juegos_2_t Porcetnaje_juegos_2_t At_bats_2_t Bateos_2_t
## 1      12.82226      729      0.202500      3481      225
## 2      12.76034      12100      0.461041      95481      5041
## 3      13.92070      2704      0.751689      36100      2116
## 4      15.92572      19881      0.756900      271441      21316
## 5      14.86023      2809      0.779689      32761      1849
## 6      10.81806      9409      0.358801      62500      5041
## Home_runs_2_t RBI_2_t Porcentaje_bateo_2_t OPS_2_t Anio_t_1
## 1      0      4      0.064516 0.368449      9
```

```

## 2      49    1764      0.052900 0.463761      10
## 3      64     961      0.058564 0.567009       9
## 4     576   3844      0.078400 0.665856       5
## 5       4    225      0.056644 0.434281       9
## 6      25   1225      0.080656 0.565504       5
##   Anios_contrato_t_1 Valor_contrato_t_1 Valor_promedio_contrato_t_1 Juegos_t_1
## 1              0              0              0              84
## 2              1          2500000          2500000              52
## 3              1          3500000          3500000             132
## 4              1          7500000          7500000             143
## 5              1          3500000          3500000             126
## 6              3         30000000         10000000             151
##   Porcetnaje_juegos_t_1 At_bats_t_1 Bateos_t_1 Home_runs_t_1 RBI_t_1
## 1              0.519           203           49           9          33
## 2              0.867           190           46           8          31
## 3              0.815           447          116          18          91
## 4              0.883           505          134          15          58
## 5              0.778           489          138          20          72
## 6              0.932           538          135          15          65
##   Porcentaje_bateo_t_1 OPS_t_1 Juegos_2_t_1 Porcetnaje_juegos_2_t_1
## 1              0.241    0.742           7056          0.269361
## 2              0.242    0.753           2704          0.751689
## 3              0.260    0.783          17424          0.664225
## 4              0.265    0.744          20449          0.779689
## 5              0.282    0.796          15876          0.605284
## 6              0.251    0.713          22801          0.868624
##   At_bats_2_t_1 Bateos_2_t_1 Home_runs_2_t_1 RBI_2_t_1 Porcentaje_bateo_2_t_1
## 1          41209          2401           81          1089          0.058081
## 2          36100          2116           64           961          0.058564
## 3         199809         13456          324          8281          0.067600
## 4         255025         17956          225          3364          0.070225
## 5         239121         19044          400          5184          0.079524
## 6         289444         18225          225          4225          0.063001
##   OPS_2_t_1
## 1  0.550564
## 2  0.567009
## 3  0.613089
## 4  0.553536
## 5  0.633616
## 6  0.508369

```

## Creación del modelo

Para ello y evitar errores, lo haremos con los nombres explícitos, obtengamos los nombres de las columnas

```
colnames(catchers_panel_data)
```

```

## [1] "Jugador"          "Anio_t"
## [3] "Anios_contrato_t" "Valor_contrato_t"
## [5] "Valor_promedio_contrato_t" "Juegos_t"
## [7] "Porcetnaje_juegos_t" "At_bats_t"
## [9] "Bateos_t"         "Home_runs_t"

```



```
## [11] "RBI_t" "Porcentaje_bateo_t"
## [13] "OPS_t" "ln_Sueldo_base_t"
## [15] "ln_Sueldo_t" "ln_Sueldo_regular_t"
## [17] "Juegos_2_t" "Porcetnaje_juegos_2_t"
## [19] "At_bats_2_t" "Bateos_2_t"
## [21] "Home_runs_2_t" "RBI_2_t"
## [23] "Porcentaje_bateo_2_t" "OPS_2_t"
## [25] "Anio_t_1" "Anios_contrato_t_1"
## [27] "Valor_contrato_t_1" "Valor_promedio_contrato_t_1"
## [29] "Juegos_t_1" "Porcetnaje_juegos_t_1"
## [31] "At_bats_t_1" "Bateos_t_1"
## [33] "Home_runs_t_1" "RBI_t_1"
## [35] "Porcentaje_bateo_t_1" "OPS_t_1"
## [37] "Juegos_2_t_1" "Porcetnaje_juegos_2_t_1"
## [39] "At_bats_2_t_1" "Bateos_2_t_1"
## [41] "Home_runs_2_t_1" "RBI_2_t_1"
## [43] "Porcentaje_bateo_2_t_1" "OPS_2_t_1"
```

```
colnames(relief_pitchers_panel_data)
```

```
## [1] "Jugador" "Anio_t"
## [3] "Anios_contrato_t" "Valor_contrato_t"
## [5] "Valor_promedio_contrato_t" "Juegos_t"
## [7] "Juegos_iniciados_t" "Inning_pitched_t"
## [9] "Bateos_pitcher_t" "Carreras_t"
## [11] "Carreras_ganadas_t" "Walks_t"
## [13] "Strike_outs_t" "Wins_t"
## [15] "Losses_t" "Saves_t"
## [17] "WHIP_t" "ERA_t"
## [19] "Equipo_t" "ln_Sueldo_base_t"
## [21] "ln_Sueldo_t" "ln_Sueldo_regular_t"
## [23] "Dominio_t" "Control_t"
## [25] "Comando_t" "Juegos_2_t"
## [27] "Juegos_iniciados_2_t" "Inning_pitched_2_t"
## [29] "Bateos_pitcher_2_t" "Carreras_2_t"
## [31] "Carreras_ganadas_2_t" "Walks_2_t"
## [33] "Strike_outs_2_t" "Wins_2_t"
## [35] "Losses_2_t" "Saves_2_t"
## [37] "WHIP_2_t" "ERA_2_t"
## [39] "Control_2_t" "Comando_2_t"
## [41] "Anio_t_1" "Anios_contrato_t_1"
## [43] "Valor_contrato_t_1" "Valor_promedio_contrato_t_1"
## [45] "Juegos_t_1" "Juegos_iniciados_t_1"
## [47] "Inning_pitched_t_1" "Bateos_pitcher_t_1"
## [49] "Carreras_t_1" "Carreras_ganadas_t_1"
## [51] "Walks_t_1" "Strike_outs_t_1"
## [53] "Wins_t_1" "Losses_t_1"
## [55] "Saves_t_1" "WHIP_t_1"
## [57] "ERA_t_1" "Equipo_t_1"
## [59] "Dominio_t_1" "Control_t_1"
## [61] "Comando_t_1" "Juegos_2_t_1"
## [63] "Juegos_iniciados_2_t_1" "Inning_pitched_2_t_1"
## [65] "Bateos_pitcher_2_t_1" "Carreras_2_t_1"
## [67] "Carreras_ganadas_2_t_1" "Walks_2_t_1"
```

```
## [69] "Strike_outs_2_t_1"      "Wins_2_t_1"
## [71] "Losses_2_t_1"          "Saves_2_t_1"
## [73] "WHIP_2_t_1"            "ERA_2_t_1"
## [75] "Control_2_t_1"         "Comando_2_t_1"
```

```
colnames(starting_pitchers_panel_data)
```

```
## [1] "Jugador"                "Anio_t"
## [3] "Anios_contrato_t"       "Valor_contrato_t"
## [5] "Valor_promedio_contrato_t" "Juegos_t"
## [7] "Juegos_iniciados_t"     "Inning_pitched_t"
## [9] "Bateos_pitcher_t"       "Carreras_t"
## [11] "Carreras_ganadas_t"     "Walks_t"
## [13] "Strike_outs_t"          "Wins_t"
## [15] "Losses_t"               "Saves_t"
## [17] "WHIP_t"                 "ERA_t"
## [19] "Equipo_t"               "ln_Sueldo_base_t"
## [21] "ln_Sueldo_t"            "ln_Sueldo_regular_t"
## [23] "Dominio_t"              "Control_t"
## [25] "Comando_t"              "Juegos_2_t"
## [27] "Juegos_iniciados_2_t"   "Inning_pitched_2_t"
## [29] "Bateos_pitcher_2_t"     "Carreras_2_t"
## [31] "Carreras_ganadas_2_t"   "Walks_2_t"
## [33] "Strike_outs_2_t"        "Wins_2_t"
## [35] "Losses_2_t"             "Saves_2_t"
## [37] "WHIP_2_t"               "ERA_2_t"
## [39] "Control_2_t"            "Comando_2_t"
## [41] "Anio_t_1"               "Anios_contrato_t_1"
## [43] "Valor_contrato_t_1"     "Valor_promedio_contrato_t_1"
## [45] "Juegos_t_1"             "Juegos_iniciados_t_1"
## [47] "Inning_pitched_t_1"     "Bateos_pitcher_t_1"
## [49] "Carreras_t_1"           "Carreras_ganadas_t_1"
## [51] "Walks_t_1"              "Strike_outs_t_1"
## [53] "Wins_t_1"               "Losses_t_1"
## [55] "Saves_t_1"              "WHIP_t_1"
## [57] "ERA_t_1"                "Equipo_t_1"
## [59] "Dominio_t_1"            "Control_t_1"
## [61] "Comando_t_1"            "Juegos_2_t_1"
## [63] "Juegos_iniciados_2_t_1" "Inning_pitched_2_t_1"
## [65] "Bateos_pitcher_2_t_1"   "Carreras_2_t_1"
## [67] "Carreras_ganadas_2_t_1" "Walks_2_t_1"
## [69] "Strike_outs_2_t_1"      "Wins_2_t_1"
## [71] "Losses_2_t_1"           "Saves_2_t_1"
## [73] "WHIP_2_t_1"             "ERA_2_t_1"
## [75] "Control_2_t_1"          "Comando_2_t_1"
```

```
colnames(remain_hitters_panel_data)
```

```
## [1] "Jugador"                "Anio_t"
## [3] "Anios_contrato_t"       "Valor_contrato_t"
## [5] "Valor_promedio_contrato_t" "Juegos_t"
## [7] "Porcentaje_juegos_t"    "At_bats_t"
## [9] "Bateos_t"               "Home_runs_t"
```

```
## [11] "RBI_t" "Porcentaje_bateo_t"
## [13] "OPS_t" "ln_Sueldo_base_t"
## [15] "ln_Sueldo_t" "ln_Sueldo_regular_t"
## [17] "Juegos_2_t" "Porcentaje_juegos_2_t"
## [19] "At_bats_2_t" "Bateos_2_t"
## [21] "Home_runs_2_t" "RBI_2_t"
## [23] "Porcentaje_bateo_2_t" "OPS_2_t"
## [25] "Anio_t_1" "Anios_contrato_t_1"
## [27] "Valor_contrato_t_1" "Valor_promedio_contrato_t_1"
## [29] "Juegos_t_1" "Porcentaje_juegos_t_1"
## [31] "At_bats_t_1" "Bateos_t_1"
## [33] "Home_runs_t_1" "RBI_t_1"
## [35] "Porcentaje_bateo_t_1" "OPS_t_1"
## [37] "Juegos_2_t_1" "Porcentaje_juegos_2_t_1"
## [39] "At_bats_2_t_1" "Bateos_2_t_1"
## [41] "Home_runs_2_t_1" "RBI_2_t_1"
## [43] "Porcentaje_bateo_2_t_1" "OPS_2_t_1"
```

Creemos ahora las formulas para las regresiones lineales sobre  $\ln(\text{Salario}_t)$  donde solo se considerarán las variables que presenten una correlación mayor o igual al 50% y en algunos casos, solo con que sean mayores o iguales a 45%.

## Catchers

En este caso, consideramos las variables que tuvieron un coeficiente de correlación mayor al 50% y después de obtener la primera versión del modelo, solo se dejaron las variables que son estadísticamente significativas en otro modelo.

```
catcher_corr_model <- lm(ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Home_runs_t +
                          Home_runs_t_1 + Home_runs_2_t_1 + OPS_2_t_1,
                          data = catchers_panel_data)
summary(catcher_corr_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Home_runs_t +
##   Home_runs_t_1 + Home_runs_2_t_1 + OPS_2_t_1, data = catchers_panel_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.90246 -0.19074 -0.02702  0.46572  0.77552
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   11.9203975   0.7223853   16.501 0.000000000000102 ***
## Anio_t         0.1366094   0.0495951    2.754   0.0126 *
## Anios_contrato_t 0.5632961   0.2720796    2.070   0.0523 .
## Home_runs_t     0.0101423   0.0312618    0.324   0.7492
## Home_runs_t_1   0.0092846   0.0701906    0.132   0.8962
## Home_runs_2_t_1 -0.0004798   0.0028672   -0.167   0.8689
## OPS_2_t_1       1.5949149   1.4630920    1.090   0.2893
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.657 on 19 degrees of freedom
## Multiple R-squared:  0.6282, Adjusted R-squared:  0.5108
## F-statistic:  5.35 on 6 and 19 DF,  p-value: 0.002207
```

Ahora, se obtendrá el modelo que contiene solo las variables más significativas y que no estuvieran correlacionadas para este

```
catcher_sig_model <- lm(ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Home_runs_t + OPS_t +
                        Home_runs_t_1 + Home_runs_2_t_1 + OPS_2_t_1,
                        data = catchers_panel_data)
summary(catcher_sig_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Home_runs_t +
##     OPS_t + Home_runs_t_1 + Home_runs_2_t_1 + OPS_2_t_1, data = catchers_panel_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.85490 -0.32666  0.08015  0.33821  0.85855
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   10.6235005   1.3842519   7.675 0.00000044 ***
## Anio_t         0.1356620   0.0493417   2.749  0.0132 *
## Anios_contrato_t 0.5153997   0.2741522   1.880  0.0764 .
## Home_runs_t    -0.0231629   0.0434764  -0.533  0.6007
## OPS_t          1.8967369   1.7303469   1.096  0.2875
## Home_runs_t_1   0.0225410   0.0708608   0.318  0.7541
## Home_runs_2_t_1 -0.0009183   0.0028800  -0.319  0.7535
## OPS_2_t_1       1.9468809   1.4903912   1.306  0.2079
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6535 on 18 degrees of freedom
## Multiple R-squared:  0.6514, Adjusted R-squared:  0.5159
## F-statistic: 4.806 on 7 and 18 DF,  p-value: 0.003375
```

## Remain hitters

Se siguió un proceso análogo, primero se dará el modelo con las variables que estén correlacionadas con 45% o más con  $\ln(\text{Sueldo}_t)$

```
remain_hitter_corr_model <- lm(ln_Sueldo_t ~ Anios_contrato_t + RBI_t_1,
                               data = remain_hitters_panel_data)
summary(remain_hitter_corr_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anios_contrato_t + RBI_t_1, data = remain_hitters_panel_data)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6119 -0.4604  0.1419  0.4519  1.7055
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)    13.481050   0.201078  67.044 < 0.0000000000000002 ***
## Anios_contrato_t  0.539445   0.101113   5.335   0.000000779 ***
## RBI_t_1         0.010980   0.004173   2.631   0.0101 *
```

---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7748 on 85 degrees of freedom
## Multiple R-squared: 0.4235, Adjusted R-squared: 0.41
## F-statistic: 31.22 on 2 and 85 DF, p-value: 0.00000000006815

Luego, obtenemos el modelo que solo contenga las variables estadísticamente significativas

```
remain_hitter_sig_model <- lm(ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Home_runs_t +
                             Home_runs_t_1 + Home_runs_2_t_1,
                             data = remain_hitters_panel_data)
summary(remain_hitter_sig_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Home_runs_t +
##      Home_runs_t_1 + Home_runs_2_t_1, data = remain_hitters_panel_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.73499 -0.40508  0.00148  0.41644  2.08108
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)    12.5579412  0.3175891  39.541 < 0.0000000000000002 ***
## Anio_t         0.0634613  0.0300763   2.110   0.03791 *
## Anios_contrato_t  0.5604062  0.0940574   5.958   0.0000000613 ***
## Home_runs_t     0.0309141  0.0105647   2.926   0.00444 **
## Home_runs_t_1    0.0883725  0.0284053   3.111   0.00256 **
## Home_runs_2_t_1 -0.0021114  0.0007883  -2.678   0.00894 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7101 on 82 degrees of freedom
## Multiple R-squared:  0.5329, Adjusted R-squared:  0.5044
## F-statistic: 18.71 on 5 and 82 DF,  p-value: 0.00000000000231
```

## Starting pitchers

Se hará totalmente análogo, el modelo por la matriz de correlación para un umbral del 45% es el siguiente

```
starting_pitchers_corr_model <- lm(ln_Sueldo_t ~ Anios_contrato_t + Inning_pitched_2_t + Inning_pitched_t_1, data = starting_pitchers_panel_data)
summary(starting_pitchers_corr_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anios_contrato_t + Inning_pitched_2_t +
##     Inning_pitched_t_1, data = starting_pitchers_panel_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1291 -0.3933  0.1330  0.4370  1.1226
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.660431750  0.334513286  40.837 < 0.0000000000000002 ***
## Anios_contrato_t    0.488316163  0.128418926   3.803  0.000568 ***
## Inning_pitched_2_t  0.000017515  0.000008885   1.971  0.056891 .
## Inning_pitched_t_1  0.004236738  0.002158079   1.963  0.057845 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7144 on 34 degrees of freedom
## Multiple R-squared:  0.5098, Adjusted R-squared:  0.4665
## F-statistic: 11.79 on 3 and 34 DF, p-value: 0.00001899
```

Por otro lado, el modelo con las variables estadísticamente significativas se muestra a continuación:

```
starting_pitchers_sig_model <- lm(ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Inning_pitched_t_1 +
    Dominio_t + Control_t + Comando_t + Juegos_t_1 + Juegos_iniciados_t_1 + Saves_2_t_1,
    data = starting_pitchers_panel_data)
summary(starting_pitchers_sig_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anio_t + Anios_contrato_t + Inning_pitched_t_1 +
##     Dominio_t + Control_t + Comando_t + Juegos_t_1 + Juegos_iniciados_t_1 +
##     Saves_2_t_1, data = starting_pitchers_panel_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.80833 -0.15281  0.01486  0.33675  0.97744
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    10.97256    1.27613   8.598 0.000000000241 ***
## Anio_t           0.17005    0.05755   2.955  0.006278 **
## Anios_contrato_t  0.46291    0.11877   3.898  0.000552 ***
## Inning_pitched_t_1  0.05255    0.01415   3.713  0.000902 ***
## Dominio_t        4.85287   10.53621   0.461  0.648651
## Control_t       25.49878   24.84358   1.026  0.313502
```

```
## Comando_t          0.22287    0.27148    0.821      0.418613
## Juegos_t_1         -0.11180    0.02749   -4.066      0.000352 ***
## Juegos_iniciados_t_1 -0.16655    0.06996   -2.381      0.024328 *
## Saves_2_t_1        1.77566    0.67708    2.623      0.013961 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5961 on 28 degrees of freedom
## Multiple R-squared:  0.7189, Adjusted R-squared:  0.6285
## F-statistic: 7.957 on 9 and 28 DF,  p-value: 0.000009823
```

## Relief pitchers

De igual manera, el modelo por la matriz de correlación para un umbral del 30% es el siguiente

```
relief_pitchers_corr_model <- lm(ln_Sueldo_t ~ Anios_contrato_t + Losses_t + WHIP_2_t_1,
                                data = relief_pitchers_panel_data)
summary(relief_pitchers_corr_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anios_contrato_t + Losses_t + WHIP_2_t_1,
##     data = relief_pitchers_panel_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32837 -0.30268  0.05194  0.38590  1.28610
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)   14.04495    0.31667  44.353 < 0.0000000000000002 ***
## Anios_contrato_t  0.61128    0.11888   5.142   0.00000255 ***
## Losses_t        0.06156    0.02777   2.216    0.0301 *
## WHIP_2_t_1     -0.28828    0.13565  -2.125    0.0373 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6094 on 67 degrees of freedom
## Multiple R-squared:  0.4168, Adjusted R-squared:  0.3907
## F-statistic: 15.96 on 3 and 67 DF,  p-value: 0.0000000622
```

Por último, el modelo con las variables estadísticamente significativas

```
relief_pitchers_sig_model <- lm(ln_Sueldo_t ~ Anios_contrato_t + Dominio_t +
                                Losses_t + WHIP_2_t_1,
                                data = relief_pitchers_panel_data)
summary(relief_pitchers_sig_model, diagnostics=TRUE)
```

```
##
## Call:
## lm(formula = ln_Sueldo_t ~ Anios_contrato_t + Dominio_t + Losses_t +
##     WHIP_2_t_1, data = relief_pitchers_panel_data)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.40062 -0.28938  0.05351  0.37909  1.25795
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)    13.41222     0.39409   34.033 < 0.0000000000000002 ***
## Anios_contrato_t  0.52943     0.11884    4.455   0.0000332 ***
## Dominio_t        7.08026     2.79778    2.531    0.0138 *
## Losses_t         0.06807     0.02684    2.536    0.0136 *
## WHIP_2_t_1      -0.31006     0.13077   -2.371    0.0207 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5862 on 66 degrees of freedom
## Multiple R-squared:  0.4684, Adjusted R-squared:  0.4361
## F-statistic: 14.54 on 4 and 66 DF,  p-value: 0.00000001451
```