

Análisis exploratorio de datos2

1. Obtener los datos

1.1. Datos sobre educación:

Vamos a obtener los datos sobre acreditación por departamentos, años y nivel de educación. Nos hemos decidido por la acreditación en cada uno de los niveles educativos en la medida en que da cuenta tanto de la diversidad de la propuesta educativa como de su calidad. Dado que provienen de distintas tablas, es necesario verificar que los años coincidan, los departamentos y su escritura y los niveles educativos:

Una vez juntamos todos los datos, vamos a verificar que estén todos los departamentos y de cuáles años hay registros:

```
## [1] "Tenemos los siguientes departamentos:"

## [1] "AMAZONAS"          "ANTIOQUIA"
## [3] "ARAUCA"            "ATLANTICO"
## [5] "BOGOTA D.C."       "BOLIVAR"
## [7] "BOYACA"            "CALDAS"
## [9] "CAQUETA"           "CASANARE"
## [11] "CAUCA"             "CESAR"
## [13] "CHOCO"             "CORDOBA"
## [15] "CUNDINAMARCA"      "GUAINIA"
## [17] "GUAVIARE"          "HUILA"
## [19] "LA GUAJIRA"        "MAGDALENA"
## [21] "META"              "NARINO"
## [23] "NORTE DE SANTANDER" "PUTUMAYO"
## [25] "QUINDIO"           "RISARALDA"
## [27] "SAN ANDRES Y PROVIDENCIA" "SANTANDER"
## [29] "SUCRE"             "TOLIMA"
## [31] "VALLE DEL CAUCA"   "VAUPES"
## [33] "VICHADA"

## [1] "Para los años:"

## [1] 2010 2011 2012 2013 2014 2015
```

1.2. Obtener los datos sobre PIB:

```
## [1] "los departamentos que quedan son:"

## [1] "AMAZONAS"          "ANTIOQUIA"
## [3] "ARAUCA"            "ATLANTICO"
## [5] "BOGOTA D.C."       "BOLIVAR"
## [7] "BOYACA"            "CALDAS"
## [9] "CAQUETA"           "CASANARE"
## [11] "CAUCA"             "CESAR"
## [13] "CHOCO"             "CORDOBA"
## [15] "CUNDINAMARCA"      "GUAINIA"
## [17] "GUAVIARE"          "HUILA"
## [19] "LA GUAJIRA"        "MAGDALENA"
## [21] "META"              "NARINO"
## [23] "NORTE DE SANTANDER" "PUTUMAYO"
```

```
## [25] "QUINDIO" "RISARALDA"
## [27] "SAN ANDRES Y PROVIDENCIA" "SANTANDER"
## [29] "SUCRE" "TOLIMA"
## [31] "VALLE DEL CAUCA" "VAUPES"
## [33] "VICHADA"

## [1] "Y los años:"

## [1] "2000" "2001" "2002" "2003" "2004" "2005" "2006" "2007" "2008" "2009"
## [11] "2010" "2011" "2012" "2013" "2014"
```

1.3. Obtener los datos sobre proyecciones poblacionales del DANE según censo 2005:

```
##      departamento anio      proy
## 1 NORTE DE SANTANDER 1985  953912
## 2 NORTE DE SANTANDER 1986  967688
## 3 NORTE DE SANTANDER 1987  983168
## 4 NORTE DE SANTANDER 1988  999946
## 5 NORTE DE SANTANDER 1989 1017585
## 6 NORTE DE SANTANDER 1990 1035636

## [1] "los departamentos que quedan son:"

## [1] "AMAZONAS" "ANTIOQUIA"
## [3] "ARAUCA" "ATLANTICO"
## [5] "BOGOTA D.C." "BOLIVAR"
## [7] "BOYACA" "CALDAS"
## [9] "CAQUETA" "CASANARE"
## [11] "CAUCA" "CESAR"
## [13] "CHOCO" "CORDOBA"
## [15] "CUNDINAMARCA" "GUAINIA"
## [17] "GUAVIARE" "HUILA"
## [19] "LA GUAJIRA" "MAGDALENA"
## [21] "META" "NARINO"
## [23] "NORTE DE SANTANDER" "PUTUMAYO"
## [25] "QUINDIO" "RISARALDA"
## [27] "SAN ANDRES Y PROVIDENCIA" "SANTANDER"
## [29] "SUCRE" "TOLIMA"
## [31] "VALLE DEL CAUCA" "VAUPES"
## [33] "VICHADA"

## [1] "Y los años:"

## [1] 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998
## [15] 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
## [29] 2013 2014 2015 2016 2017 2018 2019 2020
```

1.4. Juntar todos los datos:

Al tener para todos los sets de datos los mismos 33 departamentos (BOGOTA D.C. se considera un departamento a parte de CUNDINAMARCA), sólo se vana conservar los datos de los años que coincidan en todos: para el caso, acreditación va a ser el límite de los demás. Y por tanto vamos a tener datos solamente entre 2010 y 2015.

```
## 'data.frame': 165 obs. of 56 variables:
## $ departamento : Factor w/ 33 levels "AMAZONAS","ANTIOQUIA",...: 1 1 1 1 1 2 2 2 2 2 ...
```

```

## $ anio : int 2010 2011 2012 2013 2014 2010 2011 2012 2013 2014 ...
## $ proy : int 45509584 46044601 46581823 47121089 47661787 4383277 4428675 4474369 45...
## $ X1 : int 53 50 51 52 61 4495 4859 4874 5206 5420 ...
## $ X2 : int 0 0 0 0 0 645 754 562 690 761 ...
## $ X3 : int 0 1 0 0 0 2201 2354 2471 2592 2520 ...
## $ X4 : int 2 2 2 2 2 1544 1627 1702 1766 1924 ...
## $ X5 : int 7 7 8 10 9 89 103 114 132 183 ...
## $ X6 : int 44 40 41 40 50 16 21 25 26 32 ...
## $ X7 : int 0 0 0 0 0 1718 2434 2985 2853 2645 ...
## $ X8 : int 0 0 0 0 0 17 43 19 11 25 ...
## $ X9 : int 0 0 0 0 0 749 1155 1432 1553 1232 ...
## $ X10 : int 0 0 0 0 0 697 931 1218 928 927 ...
## $ X11 : int 0 0 0 0 0 255 305 316 361 461 ...
## $ X12 : int 8 9 9 10 10 9916 10950 11827 12100 12330 ...
## $ X13 : int 4 4 4 5 5 2161 2206 2475 2643 2763 ...
## $ X14 : int 4 5 5 5 5 7755 8744 9352 9457 9567 ...
## $ X15 : int 7 7 7 7 7 3981 4427 4748 4849 4845 ...
## $ X16 : int 5 5 5 5 5 2880 3276 3499 3563 3541 ...
## $ X17 : int 0 0 0 0 0 151 166 198 224 222 ...
## $ X18 : int 1 1 1 1 1 493 515 568 569 580 ...
## $ X19 : int 1 1 1 1 1 457 470 483 493 502 ...
## $ X20 : int 0 0 0 0 0 5699 7081 7577 9388 12511 ...
## $ X21 : int 0 0 0 0 0 2701 3532 4138 5119 5938 ...
## $ X22 : int 0 0 0 0 0 2998 3549 3439 4269 6573 ...
## $ X23 : int 73 80 88 96 100 9700 10813 11366 12279 13405 ...
## $ X24 : int 47 52 54 60 62 6577 7414 7523 8035 8763 ...
## $ X25 : int 0 0 0 0 0 915 984 1117 1212 1318 ...
## $ X26 : int 26 28 34 36 38 2208 2415 2726 3032 3324 ...
## $ X27 : int 40 42 42 44 50 4297 4508 4659 5347 5858 ...
## $ X28 : int 0 0 0 0 0 2101 2197 2223 2667 2984 ...
## $ X29 : int 0 0 1 0 0 90 60 53 59 69 ...
## $ X30 : int 18 19 18 19 22 219 228 252 312 348 ...
## $ X31 : int 2 2 2 2 3 260 310 331 365 419 ...
## $ X32 : int 20 21 21 23 25 1627 1713 1800 1944 2038 ...
## $ X33 : int 27 28 33 34 38 15526 17165 18664 19900 21620 ...
## $ X34 : int 18 19 23 24 26 3622 4019 4560 4939 5381 ...
## $ X35 : int 7 7 8 8 10 6015 6398 6845 7307 7780 ...
## $ X36 : int 2 2 2 2 2 5889 6748 7259 7654 8459 ...
## $ X37 : int 158 173 191 210 229 10124 11014 12105 13326 14535 ...
## $ X38 : int 80 87 96 106 116 3111 3406 3746 4263 4657 ...
## $ X39 : int 1 1 1 1 1 1665 1820 1997 2199 2435 ...
## $ X40 : int 49 53 58 65 69 1792 1898 2127 2314 2507 ...
## $ X41 : int 16 18 21 24 28 1617 1766 2020 2261 2534 ...
## $ X42 : int 10 12 13 12 13 1174 1305 1342 1358 1405 ...
## $ X43 : int 1 1 1 1 1 179 184 192 207 223 ...
## $ X44 : int 1 1 1 1 1 586 635 681 724 774 ...
## $ X45 : int 366 389 421 453 495 65456 73251 78805 85248 93169 ...
## $ X46 : int 16 20 22 24 28 5881 7225 7562 7365 8490 ...
## $ X47 : int 382 409 443 477 523 71337 80476 86367 92613 101659 ...
## $ acr_tecpro : int 0 0 0 50 50 532 620 184 705 419 ...
## $ acr_tecnologica : int 4 1 0 1 1 22284 18666 18412 23775 23122 ...
## $ acr_universitaria : int 19 57 65 74 66 78010 81142 93643 103091 105756 ...
## $ acr_especializacion : int 0 9 4 2 2 4908 5200 5103 5986 5923 ...
## $ acr_maestria : int 0 27 27 17 14 2983 3618 3761 4348 4675 ...

```

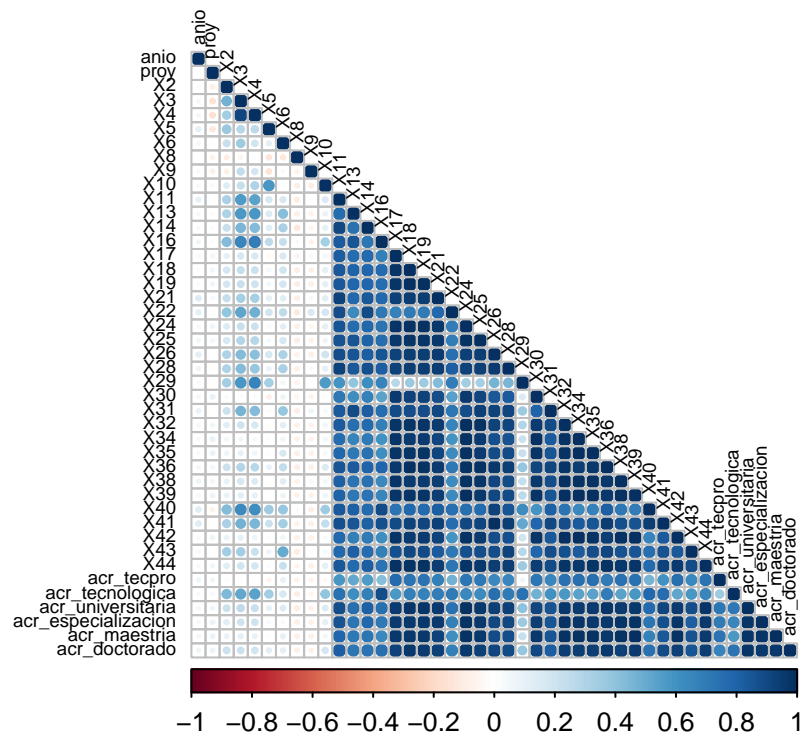
```
## $ acr_doctorado      : int  0 0 0 3 6 610 694 784 958 1129 ...
```

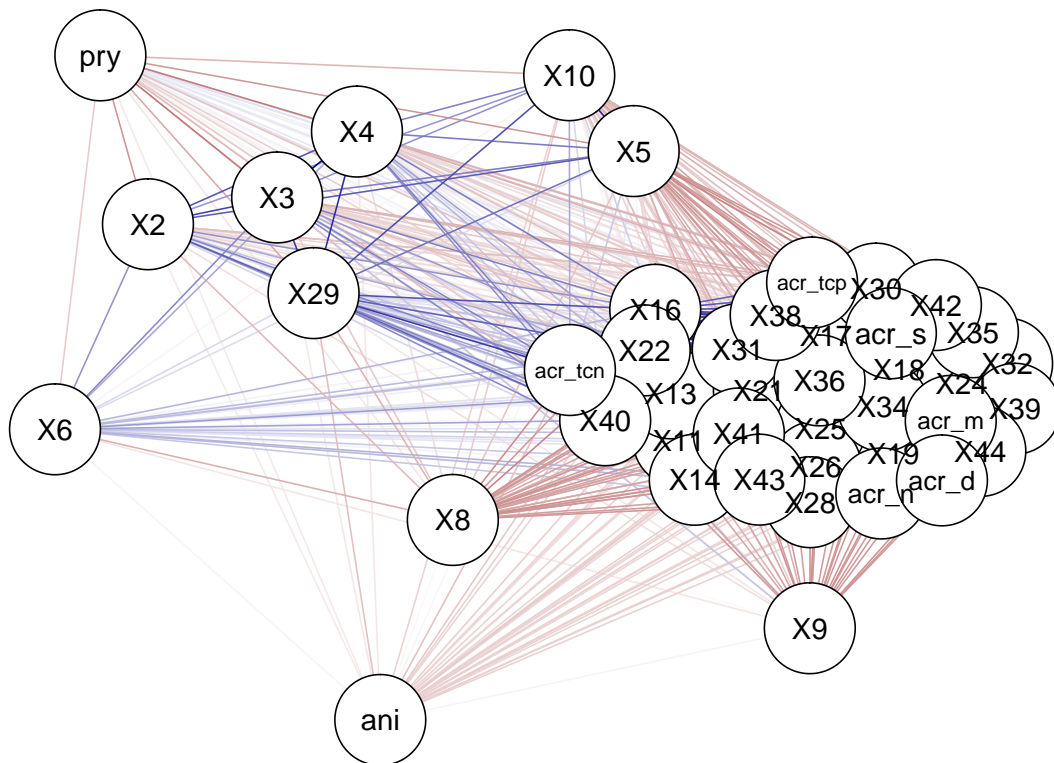
2. Análisis exploratorio de datos:

Ahora que tenemos todos los datos juntos, lo primero que podemos revisar son sus correlaciones (dejando por fuera los datos no numéricos, los nombres de los departamentos).

Pero antes tenemos variables del PIB que son suma de otras; éstas que son sumas deben ser retiradas ya que son redundantes. Y entonces podemos observar las correlaciones.

```
library(corrplot)
corrplot(correlaciones, order = "original", tl.cex = 0.6, tl.col = "black", type = "lower")
```



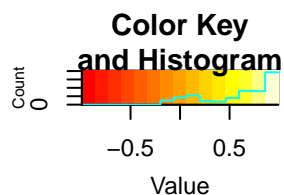


2.1 Reducción del espacio de correlaciones:

Dado que tenemos demasiadas variables para revisar sus correlaciones, vamos a reducir el espacio para que nos muestre las que realmente vamos a revisar: Solo vamos a revisar las correlaciones de los factores productivos vs las variables de educación superior. (blanco es el extremo de correlación positiva y rojo la de negativa)

```
##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##     lowess
```



##	anio	proy	X2	X3	
## acr_tecpro	0.07346082	0.06096417	-0.01733912	0.01028764	
## acr_tecnologica	0.03700365	0.03713761	0.44285604	0.51650966	
## acr_universitaria	0.05861559	0.06688848	0.21755476	0.25522807	
## acr_especializacion	0.04473665	0.08336999	0.12064726	0.20086757	
## acr_maestria	0.05887539	0.08347146	0.09835823	0.11983877	
## acr_doctorado	0.07258512	0.08665516	0.21813665	0.21512213	
##	X4	X5	X6	X8	
## acr_tecpro	-0.001889363	-0.06097334	0.09881419	-0.07401902	
## acr_tecnologica	0.539585028	0.34002719	0.17862916	-0.09537389	
## acr_universitaria	0.228255207	0.10794927	0.14535051	-0.10606516	
## acr_especializacion	0.203242388	0.03692331	0.03797430	-0.08593527	
## acr_maestria	0.097706373	0.01373373	0.10517754	-0.08537518	
## acr_doctorado	0.197647858	0.12459479	0.11464843	-0.08592765	
##	X9	X10	X11	X13	X14
## acr_tecpro	-0.06238422	-0.03994782	0.5839307	0.5146880	0.5689773
## acr_tecnologica	-0.05824440	0.39183401	0.7423403	0.6166575	0.6832924
## acr_universitaria	-0.07598184	0.14963737	0.8879268	0.7525944	0.8149712
## acr_especializacion	-0.07566704	0.11122366	0.8653752	0.7230019	0.7825220
## acr_maestria	-0.08293805	0.05782900	0.7997899	0.7180260	0.7269328
## acr_doctorado	-0.07671714	0.18305480	0.8298424	0.7223131	0.7333010
##	X16	X17	X18	X19	X21
## acr_tecpro	0.4211869	0.7713430	0.7302088	0.7322636	0.6873769
## acr_tecnologica	0.8828883	0.5722778	0.6838506	0.6769843	0.6992654
## acr_universitaria	0.7588508	0.9568053	0.9791681	0.9777987	0.9583047
## acr_especializacion	0.7172540	0.9585457	0.9663521	0.9666263	0.9575978
## acr_maestria	0.6443916	0.9641490	0.9641364	0.9644577	0.9279725
## acr_doctorado	0.7574592	0.9254482	0.9615770	0.9583221	0.9309905

```
##           X22      X24      X25      X26      X28
## acr_tecpro    0.4709339 0.7268049 0.7191197 0.6455846 0.6518248
## acr_tecnologica 0.7086367 0.6458546 0.6352381 0.7514465 0.7088317
## acr_universitaria 0.7418878 0.9808497 0.9610216 0.9365753 0.9375583
## acr_especializacion 0.6852860 0.9892780 0.9683235 0.9088240 0.9080794
## acr_maestria    0.6198805 0.9827838 0.9649410 0.8908697 0.8871980
## acr_doctorado   0.6866480 0.9670852 0.9380159 0.9185339 0.8939005
##           X29      X30      X31      X32      X34
## acr_tecpro    0.07475124 0.7225057 0.6547477 0.7439150 0.7521866
## acr_tecnologica 0.74362631 0.4792212 0.6072180 0.5782948 0.5347228
## acr_universitaria 0.40727875 0.9159427 0.8607277 0.9674019 0.9550068
## acr_especializacion 0.36254939 0.9454556 0.8525570 0.9812148 0.9741064
## acr_maestria    0.23389280 0.9664722 0.8523674 0.9837008 0.9890681
## acr_doctorado   0.38581399 0.9069541 0.8265073 0.9433796 0.9465143
##           X35      X36      X38      X39      X40
## acr_tecpro    0.7556965 0.7065277 0.7422986 0.7400456 0.4693818
## acr_tecnologica 0.5374682 0.6997011 0.5378336 0.5801125 0.7800640
## acr_universitaria 0.9598720 0.9782299 0.9505675 0.9690442 0.7774009
## acr_especializacion 0.9646179 0.9627333 0.9700495 0.9847556 0.7209350
## acr_maestria    0.9869341 0.9626976 0.9802660 0.9913365 0.6701953
## acr_doctorado   0.9415588 0.9653648 0.9313916 0.9595194 0.7481473
##           X41      X42      X43      X44
## acr_tecpro    0.6005697 0.7561000 0.6404676 0.7316552
## acr_tecnologica 0.7699199 0.5246032 0.5855231 0.6501187
## acr_universitaria 0.9108276 0.9524413 0.8840326 0.9711210
## acr_especializacion 0.8711185 0.9721451 0.8283960 0.9615659
## acr_maestria    0.8489680 0.9856200 0.8650814 0.9595195
## acr_doctorado   0.8936394 0.9343866 0.8530790 0.9429277
```

```
acr_tecpro <- df_limpio[, c(1:39)]
model <- lm(acr_tecpro ~ ., data = acr_tecpro)
summary(model)
```

```
##
## Call:
## lm(formula = acr_tecpro ~ ., data = acr_tecpro)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -394.41  -67.76    0.99   80.64  605.34
##
## Coefficients:
##              Estimate Std. Error t value
## (Intercept)    -6.048e+03  4.585e+04  -0.132
## departamentoANTIOQUIA    7.352e+03  6.464e+03   1.137
## departamentoARAUCA    1.076e+03  5.344e+03   0.201
## departamentoATLANTICO   -5.748e+02  5.343e+03  -0.108
## departamentoBOGOTA D.C.    2.794e+04  7.678e+03   3.639
## departamentoBOLIVAR   -4.671e+02  5.196e+03  -0.090
## departamentoBOYACA    1.437e+03  5.378e+03   0.267
## departamentoCALDAS    2.604e+03  5.251e+03   0.496
## departamentoCAQUETA    1.161e+03  5.315e+03   0.218
## departamentoCASANARE    8.717e+02  5.339e+03   0.163
## departamentoCAUCA    4.681e+03  5.221e+03   0.897
## departamentoCESAR    7.611e+02  5.262e+03   0.145
```

## departamentoCHOCO	3.021e+03	5.373e+03	0.562
## departamentoCORDOBA	2.052e+03	5.195e+03	0.395
## departamentoCUNDINAMARCA	6.084e+03	5.501e+03	1.106
## departamentoGUAINIA	7.408e+02	5.272e+03	0.141
## departamentoGUAVIARE	6.684e+02	5.392e+03	0.124
## departamentoHUILA	1.288e+03	5.292e+03	0.243
## departamentoLA GUAJIRA	7.819e+02	5.261e+03	0.149
## departamentoMAGDALENA	1.246e+03	5.212e+03	0.239
## departamentoMETA	1.872e+03	5.243e+03	0.357
## departamentoNARINO	1.305e+03	5.167e+03	0.253
## departamentoNORTE DE SANTANDER	4.136e+03	5.196e+03	0.796
## departamentoPUTUMAYO	1.416e+03	5.351e+03	0.265
## departamentoQUINDIO	4.658e+02	5.304e+03	0.088
## departamentoRISARALDA	2.699e+03	5.248e+03	0.514
## departamentoSAN ANDRES Y PROVIDENCIA	5.210e+02	5.381e+03	0.097
## departamentoSANTANDER	1.944e+02	5.280e+03	0.037
## departamentoSUCRE	3.122e+02	5.269e+03	0.059
## departamentoTOLIMA	2.536e+03	5.039e+03	0.503
## departamentoVALLE DEL CAUCA	6.024e+03	6.093e+03	0.989
## departamentoVAUPES	7.964e+02	5.403e+03	0.147
## departamentoVICHADA	6.029e+02	4.684e+03	0.129
## anio	2.713e+00	2.371e+01	0.114
## proy	3.192e-05	1.166e-04	0.274
## X2	7.111e-02	6.010e-01	0.118
## X3	-4.589e-03	3.989e-01	-0.012
## X4	-2.824e-02	5.268e-01	-0.054
## X5	-6.114e+00	3.618e+00	-1.690
## X6	-9.249e+00	4.580e+00	-2.020
## X8	-3.792e-02	1.116e-01	-0.340
## X9	-2.270e-02	2.605e-02	-0.872
## X10	-5.852e-01	2.319e-01	-2.524
## X11	1.767e+01	6.416e+00	2.755
## X13	-1.143e-01	7.003e-01	-0.163
## X14	8.695e-02	9.256e-02	0.939
## X16	4.709e-01	7.744e-01	0.608
## X17	4.516e+00	5.789e+00	0.780
## X18	-3.767e+01	8.824e+00	-4.269
## X19	2.945e+01	1.266e+01	2.326
## X21	-6.090e-02	2.661e-01	-0.229
## X22	-7.619e-01	2.601e-01	-2.930
## X24	1.535e+00	5.592e-01	2.745
## X25	2.881e+00	3.421e+00	0.842
## X26	7.201e+00	2.156e+00	3.340
## X28	-2.371e-02	4.010e-01	-0.059
## X29	1.451e+01	1.003e+01	1.446
## X30	-1.834e+01	3.772e+00	-4.862
## X31	7.181e-01	4.752e+00	0.151
## X32	-1.243e+01	2.626e+00	-4.735
## X34	6.290e+00	8.492e-01	7.407
## X35	8.060e-01	8.267e-01	0.975
## X36	-2.234e+00	4.193e-01	-5.326
## X38	1.518e+00	5.377e-01	2.823
## X39	-1.772e+01	2.628e+00	-6.741
## X40	-7.015e+00	1.497e+00	-4.685

## X41	3.297e+00	2.178e+00	1.514
## X42	-6.330e+00	1.502e+00	-4.215
## X43	3.174e+00	1.292e+01	0.246
## X44	2.351e+01	7.072e+00	3.325
##	Pr(> t)		
## (Intercept)	0.895340		
## departamentoANTIOQUIA	0.258278		
## departamentoARAUCA	0.840899		
## departamentoATLANTICO	0.914553		
## departamentoBOGOTA D.C.	0.000446	***	
## departamentoBOLIVAR	0.928559		
## departamentoBOYACA	0.789963		
## departamentoCALDAS	0.621190		
## departamentoCAQUETA	0.827520		
## departamentoCASANARE	0.870655		
## departamentoCAUCA	0.372223		
## departamentoCESAR	0.885307		
## departamentoCHOCO	0.575250		
## departamentoCORDOBA	0.693765		
## departamentoCUNDINAMARCA	0.271535		
## departamentoGUAINIA	0.888549		
## departamentoGUAVIARE	0.901593		
## departamentoHUILA	0.808169		
## departamentoLA GUAJIRA	0.882174		
## departamentoMAGDALENA	0.811635		
## departamentoMETA	0.721886		
## departamentoNARINO	0.801127		
## departamentoNORTE DE SANTANDER	0.427986		
## departamentoPUTUMAYO	0.791839		
## departamentoQUINDIO	0.930205		
## departamentoRISARALDA	0.608168		
## departamentoSAN ANDRES Y PROVIDENCIA	0.923062		
## departamentoSANTANDER	0.970716		
## departamentoSUCRE	0.952872		
## departamentoTOLIMA	0.615868		
## departamentoVALLE DEL CAUCA	0.325381		
## departamentoVAUPES	0.883131		
## departamentoVICHADA	0.897856		
## anio	0.909114		
## proy	0.784918		
## X2	0.906068		
## X3	0.990845		
## X4	0.957370		
## X5	0.094349	.	
## X6	0.046233	*	
## X8	0.734627		
## X9	0.385606		
## X10	0.013258	*	
## X11	0.007041	**	
## X13	0.870655		
## X14	0.349913		
## X16	0.544625		
## X17	0.437200		
## X18	4.64e-05	***	

```
## X19                0.022138 *
## X21                0.819476
## X22                0.004250 **
## X24                0.007231 **
## X25                0.401784
## X26                0.001197 **
## X28                0.952962
## X29                0.151552
## X30                4.59e-06 ***
## X31                0.880202
## X32                7.64e-06 ***
## X34                5.25e-11 ***
## X35                0.332057
## X36                6.71e-07 ***
## X38                0.005795 **
## X39                1.21e-09 ***
## X40                9.34e-06 ***
## X41                0.133305
## X42                5.69e-05 ***
## X43                0.806497
## X44                0.001259 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 183.1 on 95 degrees of freedom
## Multiple R-squared:  0.9804, Adjusted R-squared:  0.9661
## F-statistic: 68.73 on 69 and 95 DF,  p-value: < 2.2e-16
```

```
acr_tecpro <- df_limpio[, c(2:39)]
model <- lm(acr_tecpro ~ ., data = acr_tecpro)
summary(model)
```

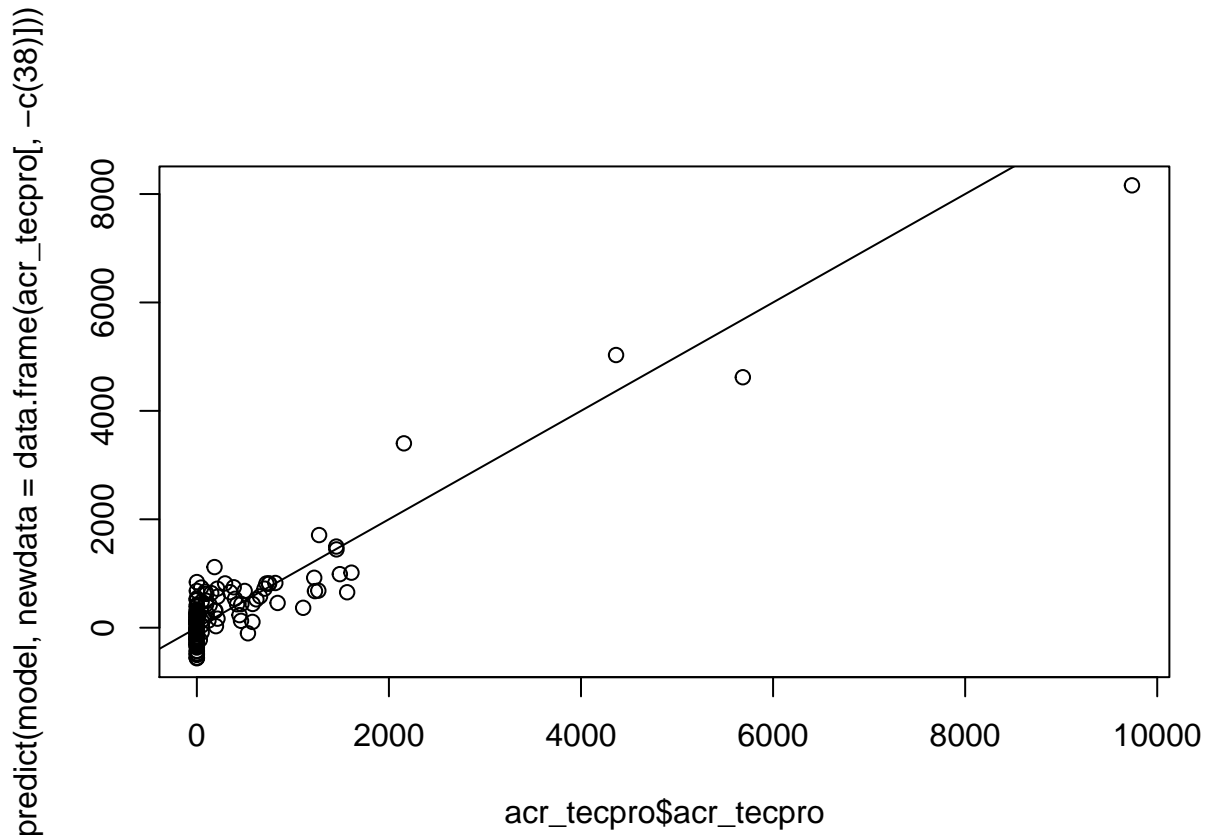
```
##
## Call:
## lm(formula = acr_tecpro ~ ., data = acr_tecpro)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1244.90  -187.45    -7.61   192.69  1576.83
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.326e+04  6.310e+04   0.527 0.599036
## anio         -1.660e+01  3.138e+01  -0.529 0.597767
## proy          1.038e-05  6.364e-06   1.630 0.105490
## X2           -2.080e+00  6.404e-01  -3.247 0.001490 **
## X3              9.872e-01  4.661e-01   2.118 0.036101 *
## X4            -8.547e-01  4.760e-01  -1.796 0.074926 .
## X5              2.119e+00  1.785e+00   1.187 0.237396
## X6            -8.986e+00  4.438e+00  -2.025 0.045009 *
## X8            -1.089e-01  6.017e-02  -1.810 0.072695 .
## X9            -8.046e-02  2.152e-02  -3.738 0.000279 ***
## X10           -9.366e-01  2.662e-01  -3.519 0.000602 ***
## X11              1.687e+00  3.725e+00   0.453 0.651496
## X13              1.040e+00  5.396e-01   1.927 0.056262 .
```

```

## X14      -2.091e-01  1.013e-01  -2.065  0.040960  *
## X16      9.265e-01  5.433e-01   1.705  0.090564  .
## X17     -3.588e+00  6.583e+00  -0.545  0.586742
## X18     -1.025e+01  1.456e+01  -0.704  0.482819
## X19      1.424e+01  1.493e+01   0.954  0.342132
## X21      9.675e-01  2.528e-01   3.827  0.000202  ***
## X22      3.301e-01  2.145e-01   1.539  0.126396
## X24      3.511e-02  4.870e-01   0.072  0.942638
## X25     -8.262e-01  2.193e+00  -0.377  0.706977
## X26      2.682e+00  1.286e+00   2.086  0.039021  *
## X28     -1.362e-01  4.845e-01  -0.281  0.779090
## X29      7.533e+00  1.092e+01   0.690  0.491577
## X30     -8.738e+00  3.413e+00  -2.560  0.011626  *
## X31      4.276e+00  2.539e+00   1.684  0.094574  .
## X32     -6.290e+00  2.411e+00  -2.608  0.010192  *
## X34      6.824e+00  1.017e+00   6.713  5.72e-10  ***
## X35      4.445e-01  5.176e-01   0.859  0.392008
## X36      4.915e-02  3.265e-01   0.151  0.880574
## X38      8.182e-01  5.151e-01   1.589  0.114642
## X39     -1.874e+01  2.620e+00  -7.149  6.09e-11  ***
## X40      2.537e+00  1.267e+00   2.003  0.047295  *
## X41     -7.796e+00  1.319e+00  -5.909  2.96e-08  ***
## X42      7.294e-01  1.614e+00   0.452  0.652120
## X43     -5.441e-01  4.958e+00  -0.110  0.912786
## X44      1.089e+01  4.388e+00   2.483  0.014350  *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 405.7 on 127 degrees of freedom
## Multiple R-squared:  0.8711, Adjusted R-squared:  0.8336
## F-statistic: 23.21 on 37 and 127 DF,  p-value: < 2.2e-16

plot(acr_tecpro$acr_tecpro, predict(model, newdata = data.frame(acr_tecpro[, -c(38)])))
abline(coef = c(0, 1))

```



```
library(caret)

## Loading required package: lattice
## Loading required package: ggplot2
modelrf <- train(acr_tecpro ~ ., data = acr_tecpro, method = "rf")

## Loading required package: randomForest
## randomForest 4.6-12
## Type rfNews() to see new features/changes/bug fixes.
##
## Attaching package: 'randomForest'
## The following object is masked from 'package:ggplot2':
##
##     margin
plot(acr_tecpro$acr_tecpro, predict(modelrf, newdata = data.frame(acr_tecpro[, -c(38)])))
abline(coef = c(0, 1))
```

```
predict(modelrf, newdata = data.frame(acr_tecpro[, -c(38)]))
```

