

TPMineria

Bodean, Emiliano & Ojeda, Zacarías

1 Resumen

2 Introducción

En el trabajo se realiza un estudio sobre sentencias judiciales, en el análisis se pretende evaluar los tiempos de demora de la sentencias. Detectando patrones comunes, o causas, entre las diferentes sentencias que resultan con demoras significativas.

Se considera la demora como el tiempo entre que inicia el proceso y sale la resolución correspondiente.

3 Procesamiento de Datos

3.1 Lectura de Datasets

Cargamos los datasets originales. Contamos con un listado de 25279 sentencias las cuales tiene los siguientes datos de interés:

- Tipo de Proceso.
- Fecha de Inicio y Fecha de Resolución.
- Cantidad de Justiciable.
- Circunscripción.
- Materia.
- Capital Reclamado.
- Organismos.

Este último campo, organismos, esta con un código interno (iep) por lo que se importa otra tabla para traducir este código.

```
sentencias_1c <- read_csv("./sentencias_1c.csv") %>%  
  filter(is.na(mat) | toupper(mat)=="C") %>%  
  # solo incluir materia civil  
  select(-mat)  
  
organismos <- read_csv("./organismos.csv")
```

Table 1. Sentencias Primera Instancia Original

nro	tproc	as	con	finicio	fdesp	fvenc1	fvenc2	fres	tres	justiciales	reccap	lep
12858	RESTRICCIONES A LA CAPACIDAD	S	0	15/09/2015	28/11/2017	22/12/2017	19/02/2018	29/12/2017	7	1	0	jdofam0002gch
11852/5	INCIDENTE	S	1	15/04/2015	19/09/2017	04/10/2017	19/10/2017	29/12/2017	6	3	0	jdofam0002gch
12237	ORDINARIO FILIACION E INDEMNIZACION DE DAÑOS	S	1	03/12/2014	12/10/2017	13/12/2017	19/03/2018	29/12/2017	7	1	0	jdofam0002gch
14440	MEDIDA CAUTELAR (FAMILIA)	S	0	21/04/2017	29/11/2017	18/12/2017	02/02/2018	29/12/2017	7	1	0	jdofam0002gch
11507	ORDINARIO DAÑOS Y PERJUICIOS	S	1	13/03/2014	30/11/2017	02/02/2018	06/04/2018	29/12/2017	7	2	0	jdofam0002gch
8133	ORDINARIO FILIACION E INDEMNIZACION DE DAÑOS	S	1	17/06/2010	16/03/2017	17/05/2017	28/07/2017	29/12/2017	7	2	0	jdofam0002gch

Table 2. Organismos

X1	organismo	organismo_descripcion	email_oficial	fuero	circunscripcion	localidad	categoria	tipo	materia
1	jdocco0000dia	Jdo Ctv y Com Lab	jdccyc-dia@juserterrios.gov.ar	Civil y Comercial	Diamante	Diamante	NA	jdo	cco/lab
2	jdocco0000fed	Jdo Ctv y Com Lab Fam	jdccyc-feion@juserterrios.gov.ar	Civil y Comercial	Federación	Federación	NA	jdo	cco/fam/lab
3	jdocco0000fcl	Jdo Ctv y Com Lab	jdccyc-fcl@juserterrios.gov.ar	Civil y Comercial	Federal	Federal	NA	jdo	cco/lab
4	jdocco0000ssa	Jdo Ctv y Com Lab Fam	jdccyclab-ssdor@juserterrios.gov.ar	Civil y Comercial	San Salvador	San Salvador	NA	jdo	cco/fam/lab
5	jdocco0000tal	Jdo Ctv y Com -comp.Laboral	jdccyc-tala@juserterrios.gov.ar	Civil y Comercial	Tala	Rosario del Tala	NA	jdo	cco/lab
6	jdocco0000vic	Jdo Ctv y Com -comp.Laboral	jdccyc-vic@juserterrios.gov.ar	Civil y Comercial	Victoria	Victoria	NA	jdo	cco/lab

3.2 Inspección y categorización de los datos

A continuación hacemos una inspección de los datos para poder categorizarlos en rangos que nos sean útiles para la generación de reglas de asociación.

En las siguientes tablas se ven fragmentos de las tablas de sentencias y organismos son procesar.

Quitamos los tipos de procesos “Monitorios”, ya que son de mero trámite y no interesan en el análisis. Se eliminan 13687 registros.

```
sentencias_1c <- sentencias_1c %>%
  filter(!grepl("MONITORIO", tproc))
```

Calcula duración como Fecha de Resolución menos Fecha de inicio, generamos una nueva columna “duracion” y mostramos un fragmento de la nueva tabla en la siguiente gráfica.

Table 3. Sentencias con duración

nro	tproc	as	con	finicio	fdesp	fvenc1	fvenc2	fres	tres	justiciales	reccap	lep	duracion
12858	RESTRICCIONES A LA CAPACIDAD	S	0	2015-09-15	28/11/2017	22/12/2017	19/02/2018	2017-12-29	7	1	0	jdofam0002gch	836 days
11852/5	INCIDENTE	S	1	2015-04-15	19/09/2017	04/10/2017	19/10/2017	2017-12-29	6	3	0	jdofam0002gch	989 days
12237	ORDINARIO FILIACION E INDEMNIZACION DE DAÑOS	S	1	2014-12-03	12/10/2017	13/12/2017	19/03/2018	2017-12-29	7	1	0	jdofam0002gch	1122 days
14440	MEDIDA CAUTELAR (FAMILIA)	S	0	2017-04-21	29/11/2017	18/12/2017	02/02/2018	2017-12-29	7	1	0	jdofam0002gch	252 days
11507	ORDINARIO DAÑOS Y PERJUICIOS	S	1	2014-03-13	30/11/2017	02/02/2018	06/04/2018	2017-12-29	7	2	0	jdofam0002gch	1387 days
8133	ORDINARIO FILIACION E INDEMNIZACION DE DAÑOS	S	1	2010-06-17	16/03/2017	17/05/2017	28/07/2017	2017-12-29	7	2	0	jdofam0002gch	2752 days

Eliminamos filas que tienen datos inválidos de fecha (datos nulos o futuros por error de tipo). Se eliminan 16 registros.

```
sentencias_1c <- sentencias_1c %>%
  filter(!is.na(finicio)) %>%
  filter(!is.na(fres)) %>%
  filter(fres < '2018-09-01', finicio < '2018-09-01')
```

Se calculan los cuartiles 25% y 75% de duración por cada tipo de proceso (tproc), y con estos parámetros se clasifican en rápido, normal o demorado si

duración se encuentra en cada uno de los rangos definidos. * Rápido: duración menor al cuartil 25% * Normal: duración entre el cuartil 25% y el 75%. * Demorado: duración mayor al cuartil 75%.

```
demora <- sentencias_1c %>%
  group_by(tproc) %>%
  summarise(techo_rapido=quantile(duracion, probs=0.25),
            piso_demorado=quantile(duracion, probs=0.75))

sentencias_1c <- sentencias_1c %>%
  left_join(demora, by="tproc") %>%
  mutate(rapido = duracion <= techo_rapido) %>%
  mutate(normal = duracion > techo_rapido & duracion
         < piso_demorado) %>%
  mutate(demorado = duracion >= piso_demorado) %>%
  select(-duracion, -techo_rapido, -piso_demorado)
# quitando columnas temporales
```

En al siguiente tabla se muestra un fragmento de los datos con los cambios realizados.

Table 4. Agregando columnas demora

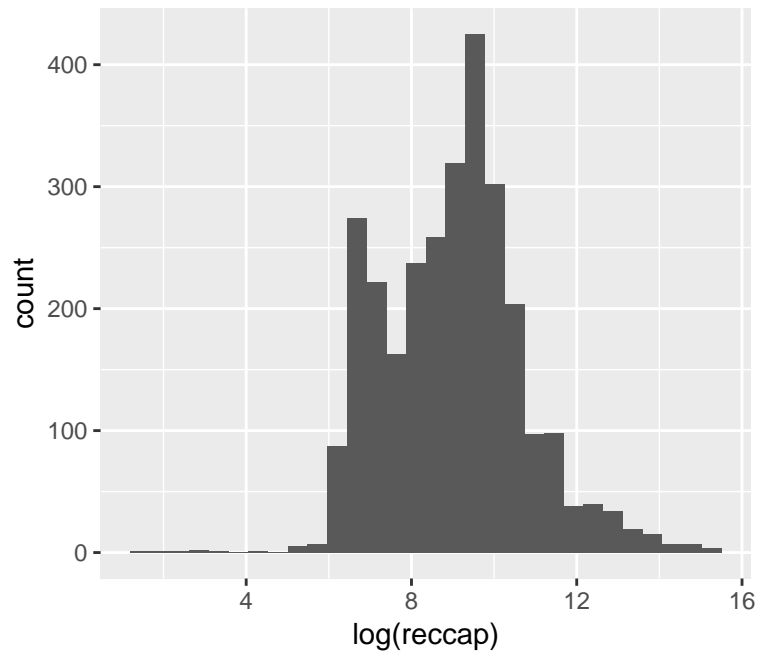
nro	tproc	as	econ	finicio	fdesp	fwenc1	fwenc2	fres	tres	justiciales	reccap	iep	rapido	normal	demorado
12858	RESTRICCIONES A L...	S	0	2015-09-15	28/11/2017	22/12/2017	19/02/2018	2017-12-29	7	1	0	jdofam0002gch	FALSE	TRUE	FALSE
11852/5	INCIDENTE	S	1	2015-04-15	19/09/2017	04/10/2017	19/10/2017	2017-12-29	6	3	0	jdofam0002gch	FALSE	FALSE	TRUE
12237	ORDINARIO FILIACL...	S	1	2014-12-03	12/10/2017	13/12/2017	19/03/2018	2017-12-29	7	1	0	jdofam0002gch	FALSE	TRUE	FALSE
14440	MEDIDA CAUTELAR (...)	S	0	2017-04-21	29/11/2017	18/12/2017	02/02/2018	2017-12-29	7	1	0	jdofam0002gch	FALSE	FALSE	TRUE
11507	ORDINARIO DAÑOS Y...	S	1	2014-03-13	30/11/2017	02/02/2018	06/04/2018	2017-12-29	7	2	0	jdofam0002gch	FALSE	TRUE	FALSE
8133	ORDINARIO FILIACL...	S	1	2010-06-17	16/03/2017	17/05/2017	28/07/2017	2017-12-29	7	2	0	jdofam0002gch	FALSE	FALSE	TRUE

Se agregan los datos de los organismos para tenerlos separados por columna, actualmente el dato se encontraba en columna iep.

```
organismos <- organismos %>%
  select(organismo, circunscripcion, localidad, materia)

sentencias_1c <- sentencias_1c %>%
  left_join(organismos, by = c('iep'='organismo'))
```

Se explora la variable capital reclamado, para definir los rangos y categorizar, graficando un histograma del logaritmo.



Se calculan los cuartiles para evaluar si sirven para parametrizar el capital reclamado (`reccap`).

```
print('1º Curtil:')
## [1] "1º Curtil:"

quantile(pull(sentencias_1c[, 'reccap']), .25, na.rm = TRUE)
## 25%
## 0

print('2º Curtil:')
## [1] "2º Curtil:"

quantile(pull(sentencias_1c[, 'reccap']), .50, na.rm = TRUE)
## 50%
## 0

print('3º Curtil:')
## [1] "3º Curtil:"

quantile(pull(sentencias_1c[, 'reccap']), .75, na.rm = TRUE)
```

```
## 75%
## 0

#View(sentencias_1c)

Como los todos los cuartiles obtenidos son cero, se vuelven a calcular los
cuartiles sin los datos ceros.

#reccap_not_cero <- which(sentencias_1c$reccap != 0)

print('1º Curtil:')
## [1] "1º Curtil:"
quantile(which(sentencias_1c$reccap != 0),.25)
## 25%
## 3181.5
print('2º Curtil:')
## [1] "2º Curtil:"
quantile(which(sentencias_1c$reccap != 0),.50)
## 50%
## 6271.5
print('3º Curtil:')
## [1] "3º Curtil:"
quantile(which(sentencias_1c$reccap != 0),.75)
## 75%
## 8948.75

capmedio <- mean(pull(sentencias_1c[, 'reccap']))

sentencias_1c <- sentencias_1c %>%
  mutate(reccap_0 = reccap == 0) %>%
  mutate(reccap_1 = (reccap < quantile(which(
    sentencias_1c$reccap != 0),.25)) & (reccap!=0)) %>%
  mutate(reccap_2 = (reccap >= quantile(which(
    sentencias_1c$reccap != 0),.25)) & (reccap < quantile(which(
    sentencias_1c$reccap != 0),.50))) %>%
  mutate(reccap_3 = (reccap >= quantile(which(
    sentencias_1c$reccap != 0),.50)) & (reccap < quantile(which(
    sentencias_1c$reccap != 0),.75))) %>%
  mutate(reccap_4 = (reccap >= quantile(which(
    sentencias_1c$reccap != 0),.75)))
```

A partir de los cuartiles obtenidos, se generan las siguientes categorías:

- Capital reclamado igual a cero.
- Capital reclamado distinto de cero y menor al cuartil 25%.
- Capital reclamado entre los cuartiles 25% y 50%.
- Capital reclamado entre los cuartiles 50% y 75%.
- Capital reclamado mayor al cuartil 75%.

Se separa la columna justiciables en los siguientes 6 rangos para categorizar.

- Justiciables igual a 1
- Justiciables igual a 2 o 3
- Justiciables igual a 4 o 5
- Justiciables igual a 6 o 7
- Justiciables igual a 8 o 9
- Justiciables mayor a 9

Se expresan las variables localidad, tipo de proceso, circunscripción y materia como factor, esto se requiere para aplicar apriori.

```
sentencias_1c <- sentencias_1c %>%  
  mutate(localidad = as.factor(localidad))  
  
sentencias_1c <- sentencias_1c %>%  
  mutate(tproc = as.factor(tproc)) %>%  
  mutate(circunscripcion = as.factor(circunscripcion)) %>%  
  mutate(materia = as.factor(materia))
```

Generamos una nueva tabla con las columnas tipo booleanos y categóricas. Se muestra un fragmento en la siguiente tabla.

Table 5. Tabla final a utilizar en el algoritmo apriori

Linea	personal_familia	personal_guerrilla	personal_militar	personal_policia	personal_privado	personal_publico	personal_seguridad	personal_sistema	personal_tercer	personal_otros	personal_otros2	personal_otros3	personal_otros4	personal_otros5
RESTRICCIONES A LA CAPACIDAD INCIDENTE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
ORDINARIO FILIACION E INDEMNIZACION DE DAÑOS	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
MEDIDA CAUTELAR (FAMILIA)	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
ORDINARIO DAÑOS Y PERJUICIOS	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
ORDINARIO DAÑOS Y PERJUICIOS	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

4 Generación de Reglas

Con los datos ya pre-procesados aplicamos apriori para generar las reglas de asociación. Inicialmente tomamos como valores límite un soporte de 0.001 y una confianza de 0.5.

```
rules <- apriori(sentencias_final, parameter = list(
  supp=0.001, conf=0.5, minlen=2), appearance = list(
  rhs=c("demorado", "rapido")))
```

```

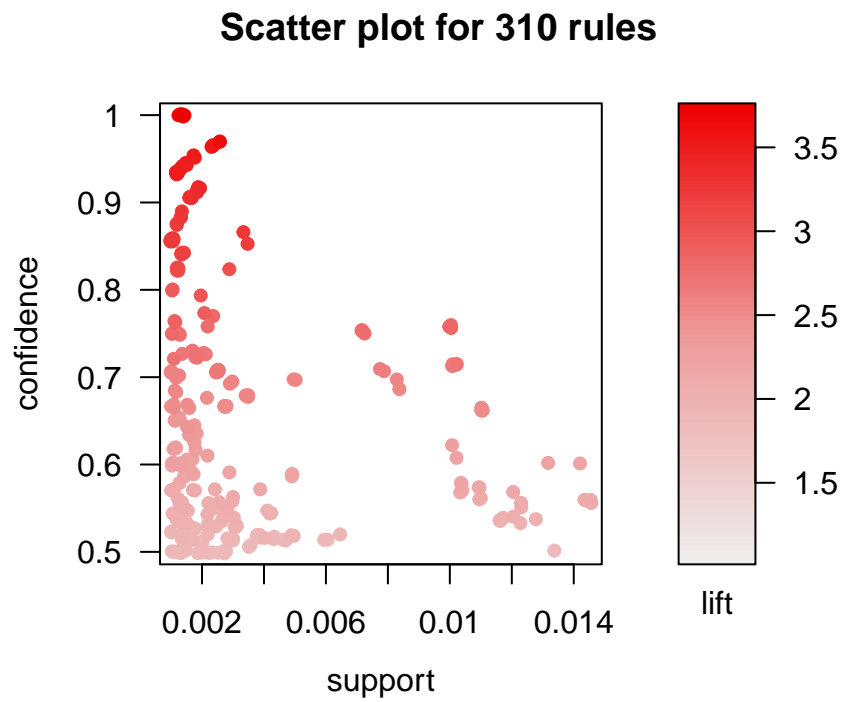
## Apriori
##
## Parameter specification:
## confidence minval smax arem aval originalSupport maxtime support minlen
##          0.5    0.1    1 none FALSE                TRUE      5  0.001      2
## maxlen target  ext
##          10 rules FALSE
##
## Algorithmic control:
## filter tree heap memopt load sort verbose
##      0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 11
##
## set item appearances ...[2 item(s)] done [0.00s].
## set transactions ...[275 item(s), 11576 transaction(s)] done [0.00s].
## sorting and recoding items ... [105 item(s)] done [0.00s].
## creating transaction tree ... done [0.01s].
## checking subsets of size 1 2 3 4 5 6 done [0.00s].
## writing ... [310 rule(s)] done [0.00s].
## creating S4 object ... done [0.00s].

summary(rules)

## set of 310 rules
##
## rule length distribution (lhs + rhs):sizes
##  2   3   4   5   6
##  2  47 126 107  28
##
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  2.000  4.000  4.000  4.361  5.000  6.000
##
## summary of quality measures:
##      support      confidence      lift      count
##  Min.   :0.001037  Min.   :0.5000  Min.   :1.837  Min.   : 12.0
##  1st Qu.:0.001296  1st Qu.:0.5385  1st Qu.:2.004  1st Qu.: 15.0
##  Median :0.001814  Median :0.6197  Median :2.321  Median : 21.0
##  Mean   :0.003352  Mean   :0.6674  Mean   :2.474  Mean   : 38.8
##  3rd Qu.:0.003455  3rd Qu.:0.7582  3rd Qu.:2.840  3rd Qu.: 40.0
##  Max.   :0.014599  Max.   :1.0000  Max.   :3.749  Max.   :169.0
##
## mining info:
##      data ntransactions support confidence
##  sentencias_final      11576  0.001      0.5

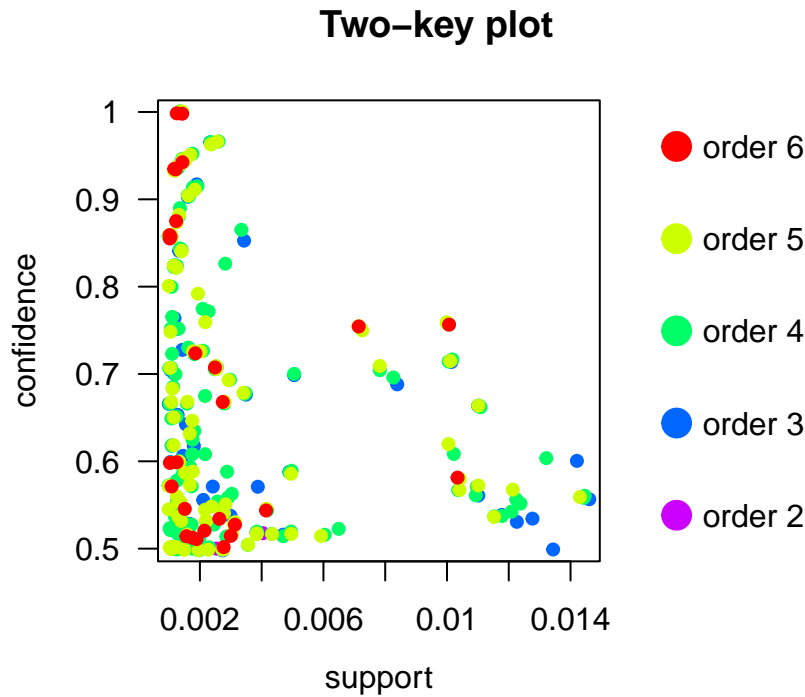
```

Graficamos las reglas para ver como varia el soporte y la confianza.



Se grafica nuevamente pero incluyendo en dato del orden de las reglas con colores.

To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.



Se realiza un inspect de las primeras reglas. Se puede ver que hay reglas que no son de interés por tener baja confianza.

```
inspect(rules[1:8])
```

##	lhs	rhs	support	confidence
## [1]	{circunscripcion=San Salvador}	=> {rapido}	0.002505183	0.5000000 1.
## [2]	{justiciables10_N}	=> {demorado}	0.004060124	0.5164835 1.
## [3]	{tproc=EJECUTIVO,			
##	circunscripcion=San Salvador}	=> {rapido}	0.001295784	0.5555556 2.
## [4]	{circunscripcion=San Salvador,			
##	materia=paz}	=> {rapido}	0.001814098	0.6176471 2.
## [5]	{circunscripcion=San Salvador,			
##	reccap_0}	=> {rapido}	0.002418798	0.5714286 2.
## [6]	{circunscripcion=San Salvador,			
##	justiciables2_3}	=> {rapido}	0.002159641	0.5555556 2.
## [7]	{tproc=DIVORCIO POR MUTUO CONSENTIMIENTO,			
##	circunscripcion=Guaileguay}	=> {rapido}	0.001641327	0.9047619 3.
## [8]	{materia=cco,			
##	justiciables10_N}	=> {demorado}	0.002764340	0.5000000 1.

Se vuelven a visualizar las reglas pero ordenas por soporte y confianza.

```
top.confidence <- sort(rules, decreasing = TRUE,
                        na.last = NA, by = "confidence")
inspect(top.confidence[1:8])
```

	lhs	rhs	support	confidence	lift	count
## [1]	{tproc=ACCION DE AMPARO, circunscripcion=Nogoyá, justiciables0_1}	=> {demorado}	0.001382170	1.0000000	3.748705	16
## [2]	{tproc=ACCION DE AMPARO, circunscripcion=Nogoyá, materia=paz, justiciables0_1}	=> {demorado}	0.001382170	1.0000000	3.748705	16
## [3]	{tproc=ACCION DE AMPARO, circunscripcion=Nogoyá, reccap_0, justiciables0_1}	=> {demorado}	0.001382170	1.0000000	3.748705	16
## [4]	{tproc=EJECUTIVO, circunscripcion=Paraná, reccap_1, justiciables0_1}	=> {rapido}	0.001295784	1.0000000	3.673754	15
## [5]	{tproc=ACCION DE AMPARO, circunscripcion=Nogoyá, materia=paz, reccap_0, justiciables0_1}	=> {demorado}	0.001382170	1.0000000	3.748705	16
## [6]	{tproc=EJECUTIVO, circunscripcion=Paraná, materia=paz, reccap_1, justiciables0_1}	=> {rapido}	0.001295784	1.0000000	3.673754	15
## [7]	{tproc=ACCION DE AMPARO, circunscripcion=Nogoyá, materia=paz}	=> {demorado}	0.002591569	0.9677419	3.627779	30
## [8]	{tproc=ACCION DE AMPARO, circunscripcion=Nogoyá, materia=paz, reccap_0}	=> {demorado}	0.002591569	0.9677419	3.627779	30

```
top.support <- sort(rules, decreasing = TRUE,
                    na.last = NA, by = "support")
inspect(top.support[1:8])
```

	lhs	rhs	support	confidence	lift	count
## [1]	{tproc=APREMIO, reccap_1}	=> {demorado}	0.01459917	0.5577558	2.090862	169
## [2]	{tproc=APREMIO, reccap_1,					

```

##      justiciables2_3}          => {demorado} 0.01451279 0.5600000 2.099275 168
## [3] {tproc=APREMIO,
##      materia=paz,
##      reccap_1}          => {demorado} 0.01442640 0.5585284 2.093758 167
## [4] {tproc=APREMIO,
##      materia=paz,
##      reccap_1,
##      justiciables2_3}      => {demorado} 0.01434001 0.5608108 2.102314 166
## [5] {tproc=APREMIO,
##      reccap_0}          => {rapido} 0.01416724 0.6007326 2.206944 164
## [6] {circunscripcion=Uruguay,
##      justiciables2_3}      => {demorado} 0.01338977 0.5000000 1.874352 155
## [7] {tproc=APREMIO,
##      reccap_0,
##      justiciables2_3}      => {rapido} 0.01321700 0.6023622 2.212931 153
## [8] {tproc=EJECUTIVO,
##      circunscripcion=Gualedguaychú} => {demorado} 0.01278507 0.5362319 2.010175 148

```

Se puede observar que tenemos reglas ...

5 Resultados / Discusión

...

6 Conclusiones

...

7 Referencias

...