

Minería de Datos

Bodean, Emiliano - Ojeda, Zacarías

UTN Regional Paraná

Abstract Se presenta un análisis sobre los tiempos entre inicio y sentencia de causas judiciales, en la búsqueda de identificar patrones o asociaciones que lleven a diferentes tiempos de demora en los mismos. Se ha utilizado el lenguaje R y el set de herramientas tidyverse[1]

Palabras Claves: Minería de Datos, sentencia, justicia, demora, reglas de asociación

1 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 asociaciones, que resulten significativas con estos tiempo de demora.

Se considera la demora como el tiempo entre que inicia el proceso y la sentencia correspondiente que pone fin al conflicto.

Se cuenta con una base de datos de resoluciones correspondientes al Superior Tribunal de Justicia de la Provincia de Entre Ríos, que a sido debidamente anonimizada, quitando referencias a las partes intervinientes y carátulas de las causas.

El análisis se realiza analizando demoras según el tipo de proceso, debido a que cada tipo de proceso, por su naturaleza, implica diferente tratamiento por parte de los organismos (juzgados) y por ende supone demoras esperables diferentes.

El procesamiento y análisis de datos se ha realizado con el [lenguaje R](#)[2], utilizando [dplyr](#)[3] para preprocesamiento de datos, [arules](#)[4] para creación de reglas de asociación y [arulesViz](#)[5] para visualizarlas. El presente informe se ha realizado en Rmd para producir resultados reproducibles[6].

Se puede acceder al código fuente del presente trabajo a través de [repositorio git](#), o al archivo generado [PDF](#) o la version [WEB](#).

2 Procesamiento de Datos

2.1 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")
```

2.2 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 sin procesar.

Table 1. Sentencias Primera Instancia Original

nro	tproc	as	con	finicio	fdesp	fvenc1	fvenc2	fres	tres	justiciales	reccap	iep
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

Se quitan 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))
```

Se calcula duración como Fecha de Resolución menos Fecha de inicio, se genera una nueva columna “duracion” y, en la siguiente gráfica, se muestra un fragmento de la nueva tabla.

Se eliminan las 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')
```

Table 2. Datos y Tipos de la tabla de sentencias

variable	tipo	primeros_valores
nro	character	12858, 11852/5
tproc	character	RESTRICCIONES A LA CAPACIDAD, INCIDENTE
as	character	S, S
ccon	integer	0, 1
finicio	character	15/09/2015, 15/04/2015
fdesp	character	28/11/2017, 19/09/2017
fvenc1	character	22/12/2017, 04/10/2017
fvenc2	character	19/02/2018, 19/10/2017
fres	character	29/12/2017, 29/12/2017
tres	integer	7, 6
justiciales	integer	1, 3
reccap	double	0, 0
iep	character	jdofam0002gch, jdofam0002gch

Table 3. Organismos

X1	organismo	organismo_descripcion	email_oficial	fuero	circunscripcion	localidad	categoria	tipo	materia
1	jdocco0000dia	Jdo Civ y Com Lab	jdccyc-dia@jusertreros.gov.ar	Civil y Comercial	Diamante	Diamante	NA jdo	cco	lab
2	jdocco0000fed	Jdo Civ y Com Lab Fam	jdccyc-fcion@jusertreros.gov.ar	Civil y Comercial	Federación	Federación	NA jdo	cco	fam/lab
3	jdocco0000frl	Jdo Civ y Com Lab	jdccyc-fral@jusertreros.gov.ar	Civil y Comercial	Federal	Federal	NA jdo	cco	lab
4	jdocco0000ssa	Jdo Civ y Com Lab Fam	jdccyclab-ssdor@jusertreros.gov.ar	Civil y Comercial	San Salvador	San Salvador	NA jdo	cco	fam/lab
5	jdocco0000tal	Jdo Civ y Com -comp.Laboral	jdccyc-tala@jusertreros.gov.ar	Civil y Comercial	Tala	Rosario del Tala	NA jdo	cco	lab
6	jdocco0000vic	Jdo Civ y Com -comp.Laboral	jdccyc-vic@jusertreros.gov.ar	Civil y Comercial	Victoria	Victoria	NA jdo	cco	lab

Table 4. Sentencias con duración

nro	tproc	as	ccon	finicio	fdesp	fvenc1	fvenc2	fres	tres	justiciales	reccap	iep	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

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 5. Agregando columnas demora

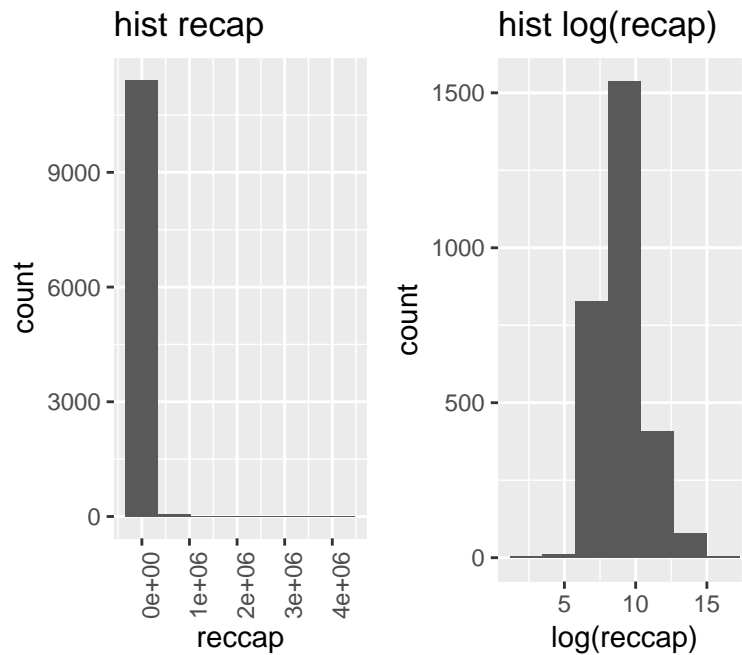
nro	tproc	as	con	finicio	fdesp	fvenc1	fvenc2	fres	tres	justiciales	reccap	iep	rapido	normal	demorado
12858	RESTRICCIONES A L...	S	0	2015-09-15	28/11/2017	22/12/2017	19/02/2018	2017-12-29	7	1	0	jdofam0002gch	FALSE	TRUE	FALSE
11852/5	INCIDENTE	S	1	2015-04-15	19/09/2017	04/10/2017	19/10/2017	2017-12-29	6	3	0	jdofam0002gch	FALSE	FALSE	TRUE
12237	ORDINARIO FILIACI...	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 FILIACI...	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 tanto el histograma del mismo, como el 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 6. Tabla final a utilizar en el algoritmo apriori

tipos	rapido	normal	demorado	circunscripcion	materia	procap_1	procap_2	procap_3	procap_4	justiciables1	justiciables2	justiciables3	justiciables4	justiciables5	justiciables6	justiciables7	justiciables8	justiciables9	justiciables10	N
RESTRICCIONES A LA CAPACIDAD	FALSE	TRUE	FALSE	Gadaguarachi	famipen	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
INCIDENTE	FALSE	FALSE	TRUE	Gadaguarachi	famipen	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
ORDINARIO FILIACION E INDEMNIZACION DE DAÑOS	FALSE	TRUE	FALSE	Gadaguarachi	famipen	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
MEDIDA CAUTELAR (FAMILIA)	FALSE	FALSE	TRUE	Gadaguarachi	famipen	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
ORDINARIO DAÑOS Y PERJUICIOS	FALSE	TRUE	FALSE	Gadaguarachi	famipen	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
ORDINARIO FILIACION E INDEMNIZACION DE DAÑOS	FALSE	FALSE	TRUE	Gadaguarachi	famipen	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

3 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 [0.00s].
## 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

inspect(rules[1:8])

##      lhs      rhs      support confidence      lift count
## [1] {circunscripcion=San Salvador} => {rapido}  0.002505183  0.5000000  1.836877    29
## [2] {justiciables10_N} => {demorado}  0.004060124  0.5164835  1.936144    47
## [3] {tproc=EJECUTIVO,
##      circunscripcion=San Salvador} => {rapido}  0.001295784  0.5555556  2.040975    15
## [4] {circunscripcion=San Salvador,
##      materia=paz} => {rapido}  0.001814098  0.6176471  2.269084    21
## [5] {circunscripcion=San Salvador,
##      reccap_0} => {rapido}  0.002418798  0.5714286  2.099288    28
## [6] {circunscripcion=San Salvador,
##      justiciables2_3} => {rapido}  0.002159641  0.5555556  2.040975    25
## [7] {tproc=DIVORCIO POR MUTUO CONSENTIMIENTO,
##      circunscripcion=Gualeguay} => {rapido}  0.001641327  0.9047619  3.323873    19
## [8] {materia=cco,
##      justiciables10_N} => {demorado}  0.002764340  0.5000000  1.874352    32

```

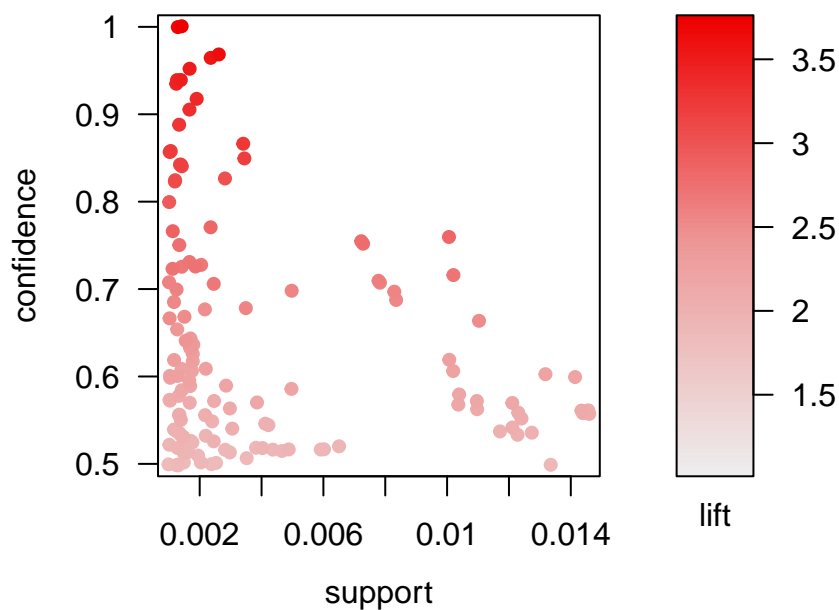
Se eliminan las reglas redundantes y se imprimen las primeras 8. Se puede observar que hay reglas que son similares pero solamente difieren en su nivel de especificidad, por esto deben ser eliminadas para el análisis.


```
rules <- rules[!is.redundant(rules)]
inspect(rules[1:8])
```

	lhs	rhs	support	confidence	lift	count
## [1]	{circunscripcion=San Salvador}	=> {rapido}	0.002505183	0.5000000	1.836877	29
## [2]	{justiciales10_N}	=> {demorado}	0.004060124	0.5164835	1.936144	47
## [3]	{tproc=EJECUTIVO, circunscripcion=San Salvador}	=> {rapido}	0.001295784	0.5555556	2.040975	15
## [4]	{circunscripcion=San Salvador, materia=paz}	=> {rapido}	0.001814098	0.6176471	2.269084	21
## [5]	{circunscripcion=San Salvador, reccap_0}	=> {rapido}	0.002418798	0.5714286	2.099288	28
## [6]	{circunscripcion=San Salvador, justiciales2_3}	=> {rapido}	0.002159641	0.5555556	2.040975	25
## [7]	{tproc=DIVORCIO POR MUTUO CONSENTIMIENTO, circunscripcion=Guaaleguay}	=> {rapido}	0.001641327	0.9047619	3.323873	19
## [8]	{circunscripcion=Paraná, justiciales10_N}	=> {demorado}	0.003023497	0.5384615	2.018533	35

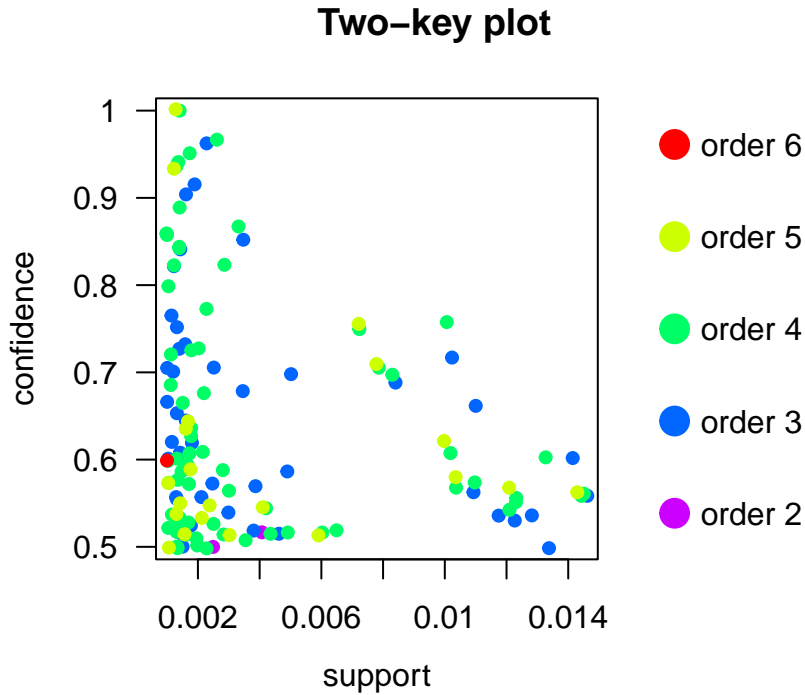
Graficamos las reglas para ver como varía el soporte y la confianza.

Scatter plot for 140 rules



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	lift	count
## [1]	{circunscripcion=San Salvador}	=> {rapido}	0.002505183	0.5000000	1.836877	29
## [2]	{justiciables10_N}	=> {demorado}	0.004060124	0.5164835	1.936144	47
## [3]	{tproc=EJECUTIVO,					
##	circunscripcion=San Salvador}	=> {rapido}	0.001295784	0.5555556	2.040975	15
## [4]	{circunscripcion=San Salvador,					
##	materia=paz}	=> {rapido}	0.001814098	0.6176471	2.269084	21
## [5]	{circunscripcion=San Salvador,					
##	reccap_0}	=> {rapido}	0.002418798	0.5714286	2.099288	28
## [6]	{circunscripcion=San Salvador,					
##	justiciables2_3}	=> {rapido}	0.002159641	0.5555556	2.040975	25
## [7]	{tproc=DIVORCIO POR MUTUO CONSENTIMIENTO,					
##	circunscripcion=Guaaleguay}	=> {rapido}	0.001641327	0.9047619	3.323873	19
## [8]	{circunscripcion=Paraná,					
##	justiciables10_N}	=> {demorado}	0.003023497	0.5384615	2.018533	35

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=EJECUTIVO,
##      circunscripcion=Paraná,
##      reccap_1,
##      justiciables0_1} => {rapido} 0.001295784 1.0000000 3.673754 15
## [3] {tproc=ACCION DE AMPARO,
##      circunscripcion=Nogoyá,
##      materia=paz} => {demorado} 0.002591569 0.9677419 3.627779 30
## [4] {tproc=DIVORCIO,
##      circunscripcion=Gualeguay} => {rapido} 0.002332412 0.9642857 3.542549 27
## [5] {tproc=EJECUTIVO,
##      circunscripcion=Paraná,
##      justiciables0_1} => {rapido} 0.001727713 0.9523810 3.498814 20
## [6] {circunscripcion=Paraná,
##      reccap_1,
##      justiciables0_1} => {rapido} 0.001382170 0.9411765 3.457651 16
## [7] {tproc=EJECUTIVO,
##      reccap_1,
##      justiciables0_1} => {rapido} 0.001295784 0.9375000 3.444145 15
## [8] {tproc=EJECUCION DE HONORARIOS,
##      circunscripcion=Paraná,
##      materia=cco,
##      justiciables0_1} => {rapido} 0.001209399 0.9333333 3.428837 14
```

```
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=Gualeguaychú}	=> {demorado}	0.01278507	0.5362319	2.010175	148

Se toman las reglas ordenandas por confianza para analizar.

```
summary(top.confidence)

## set of 140 rules
##
## rule length distribution (lhs + rhs):sizes
## 2 3 4 5 6
## 2 46 70 21 1
##
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 2.000  3.000  4.000  3.807  4.000  6.000
```

```
##
## summary of quality measures:
##      support      confidence      lift      count
## Min.   :0.001037   Min.   :0.5000   Min.   :1.837   Min.   : 12.00
## 1st Qu.:0.001296   1st Qu.:0.5351   1st Qu.:1.988   1st Qu.: 15.00
## Median :0.001814   Median :0.5888   Median :2.204   Median : 21.00
## Mean   :0.003960   Mean   :0.6396   Mean   :2.369   Mean   : 45.84
## 3rd Qu.:0.004730   3rd Qu.:0.7103   3rd Qu.:2.657   3rd Qu.: 54.75
## Max.   :0.014599   Max.   :1.0000   Max.   :3.749   Max.   :169.00
##
## mining info:
##      data ntransactions support confidence
## sentencias_final      11576    0.001      0.5

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=EJECUTIVO,
##      circunscripcion=Paraná,
##      recap_1,
##      justiciables0_1}      => {rapido} 0.001295784  1.0000000 3.673754    15
## [3] {tproc=ACCION DE AMPARO,
##      circunscripcion=Nogoyá,
##      materia=paz}          => {demorado} 0.002591569  0.9677419 3.627779    30
## [4] {tproc=DIVORCIO,
##      circunscripcion=Guaaleguay}      => {rapido} 0.002332412  0.9642857 3.542549    27
## [5] {tproc=EJECUTIVO,
##      circunscripcion=Paraná,
##      justiciables0_1}      => {rapido} 0.001727713  0.9523810 3.498814    20
## [6] {circunscripcion=Paraná,
##      recap_1,
##      justiciables0_1}      => {rapido} 0.001382170  0.9411765 3.457651    16
## [7] {tproc=EJECUTIVO,
##      recap_1,
##      justiciables0_1}      => {rapido} 0.001295784  0.9375000 3.444145    15
## [8] {tproc=EJECUCION DE HONORARIOS,
##      circunscripcion=Paraná,
##      materia=cco,
##      justiciables0_1}      => {rapido} 0.001209399  0.9333333 3.428837    14
```

4 Resultados / Discusión

4.1 Asociaciones destacadas:

- tipo de proceso *ACCION DE AMPARO* en las circunscripción de **Nogoyá** como *demoradas* con respecto a los valores provinciales.
- *APREMIO* con capital reclamado *recap_1* aparecen en gral *demorados*, sin embargo, cuando el capital reclamado es *recap_0*, se resuelven *rápido*.
- Los procesos *EJECUTIVO* en *Guaaleguaychú* aparecen como *demorados*.

Se han obtenido reglas que eran de esperarse debido, por ejemplo, a que dado un tipo de proceso, al incrementarse el capital reclamado o el nro de justiciables, se puede inferir que todo el proceso se hace más complejo y se elonguen los tiempos de resolución por sentencia.

Sin embargo, *no hay razones* de tipo procesal, para que *diferentes circunscripciones/jurisdicciones presenten diferencias*, las mismas, *son las más significativas para el análisis* y requieren una investigación más profunda. Las reglas generadas pueden poner en evidencia diferencias en dotaciones de personal,

prácticas administrativas y/o alguna circunstancia particular que lleve a estas diferencias.

5 Conclusiones

Las técnicas empleadas aquí han dado como resultado reglas interesantes para investigar, ya que a priori no se esperaban circunscripciones asociadas a diferencias en tiempos de resolución de sentencia, dichas diferencias pueden estar asociadas al modo de trabajo en esas localidades, quizás a la dotación de personal, capacitaciones de los mismos, u otras razones que requieren investigación específica.

En cuanto a las herramientas, el preprocesamiento de datos y la generación de reglas han resultado muy sencillo con las técnicas utilizadas, así mismo el formato seleccionado para la realización del informe, permite hacer evaluaciones interactivas mientras se conforma el documento, como así también hace muy sencillo trabajar de manera colaborativa ya que en el mismo documento está el código que se ha utilizado para manipular los datos y generar las reglas.

Referencias

1. *R packages for data science*, <https://www.tidyverse.org/>.
2. *The R Project for Statistical Computing*, <https://www.r-project.org/about.html>.
3. *dplyr: A Grammar of Data Manipulation*, <https://dplyr.tidyverse.org/>.
4. *arules: Mining Association Rules and Frequent Itemsets*, <https://cran.r-project.org/web/packages/arules/index.html>.
5. *arulesViz: Visualizing Association Rules and Frequent Itemsets*, <https://cran.r-project.org/web/packages/arulesViz/index.html>.
6. Gandrud, C.: Reproducible Research with R and R Studio. (2015).