

Usos de los recursos hidricos en el Uruguay

Facundo Morini, Mariana Ceresa

2023-06-20

Introducción

A lo largo de los últimos años, el cambio climático y las sequías han impactado significativamente a nuestro país, generando una creciente preocupación en torno al uso del agua. En este contexto, resulta fundamental analizar el aprovechamiento de los recursos hídricos en Uruguay. El presente trabajo tiene como objetivo principal explorar e informar sobre los diversos usos que se le da al agua en nuestro país. Para ello, se utilizarán los datos abiertos disponibles proporcionados por la Dirección Nacional del Agua (DINAGUA), específicamente el registro de empresas que realizan solicitudes para explotar los recursos hídricos.

Dentro de este estudio, nos planteamos responder a preguntas fundamentales, tales como:

1. ¿En qué se utiliza la mayor parte del agua que se solicita extraer?
2. ¿Qué sectores de la actividad requieren mayores volúmenes de agua?
3. ¿Qué zonas del país presentan mayores demanda de uso?

Para abordar estas preguntas, realizaremos un análisis de los datos disponibles utilizando las diversas herramientas que proporciona R para el análisis de datos. Buscaremos utilizar y dar visibilidad a la biblioteca Geoupy, que facilita el manejo de datos geograficos relacionados con Uruguay de manera sencilla. Además, crearemos una aplicación utilizando la biblioteca Shiny, que permitirá visualizar de forma interactiva los resultados obtenidos en el informe. Por último, crearemos un modelo predictivo con el objetivo de predecir el uso del agua utilizando diversas variables de respuesta presentes en los datos.

Datos

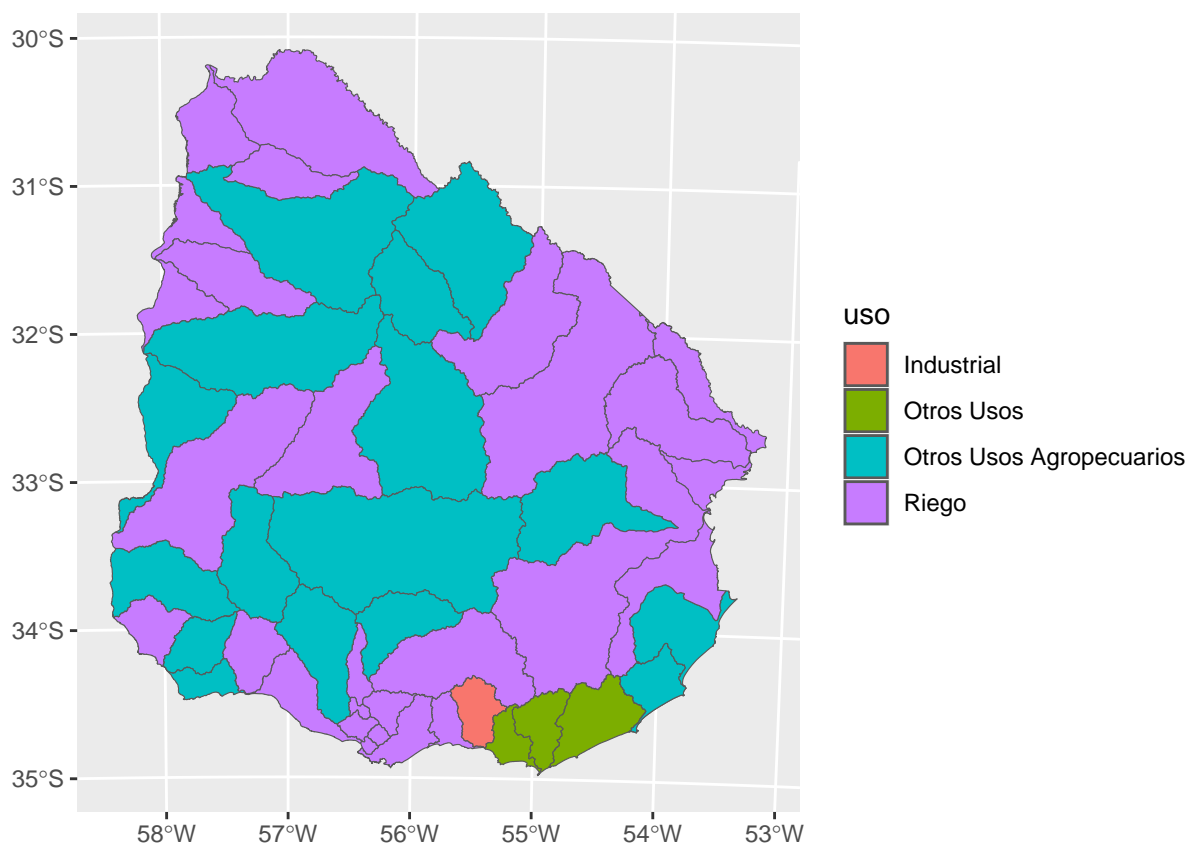
Como se mencionó, los datos utilizados pertenecen a la Dirección Nacional del Agua. Estos datos comprenden el registro de todas las empresas y entes publicos que realizan extracciones de agua para diversos usos productivos o abastecimiento a la poblacion. Los datos abarca un total de 5405 registros que se encuentran actualmente habilitados a la explotacion de recursos hidricos.

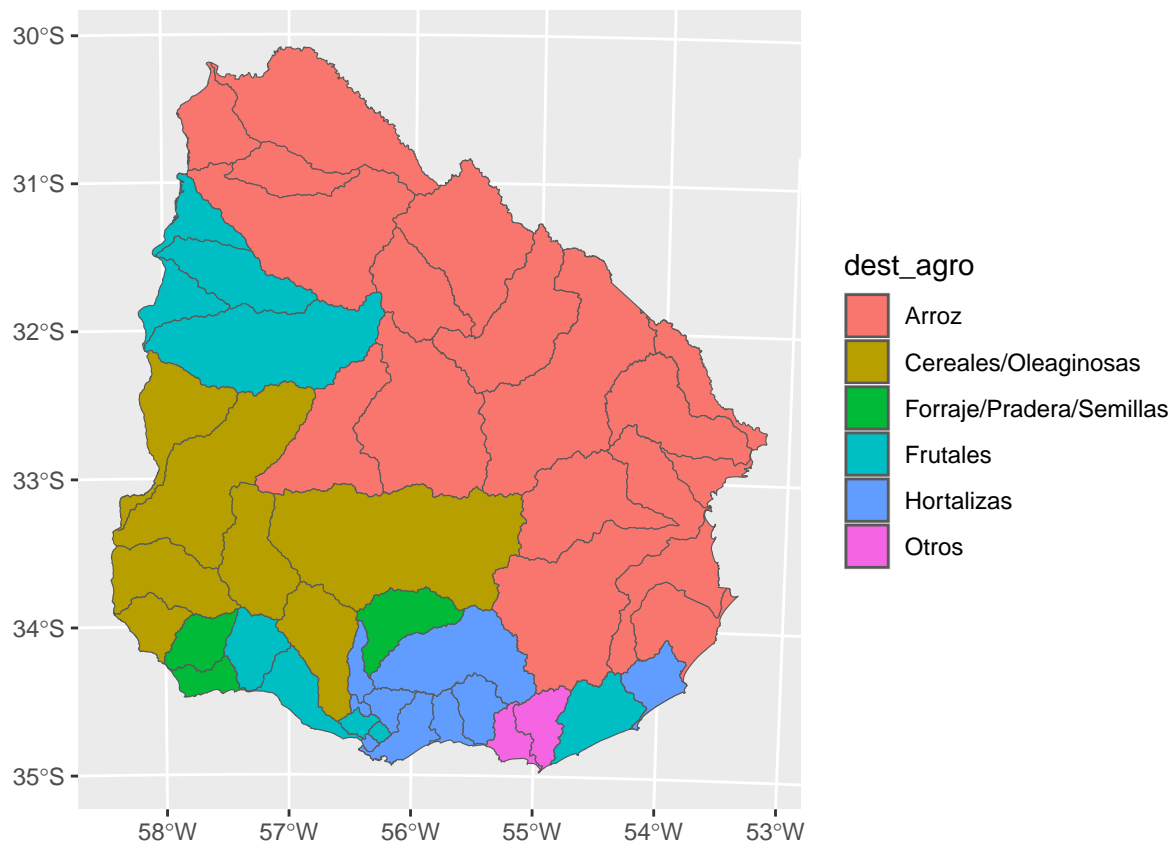
Las principales variables de interes a analizar de los datos son:

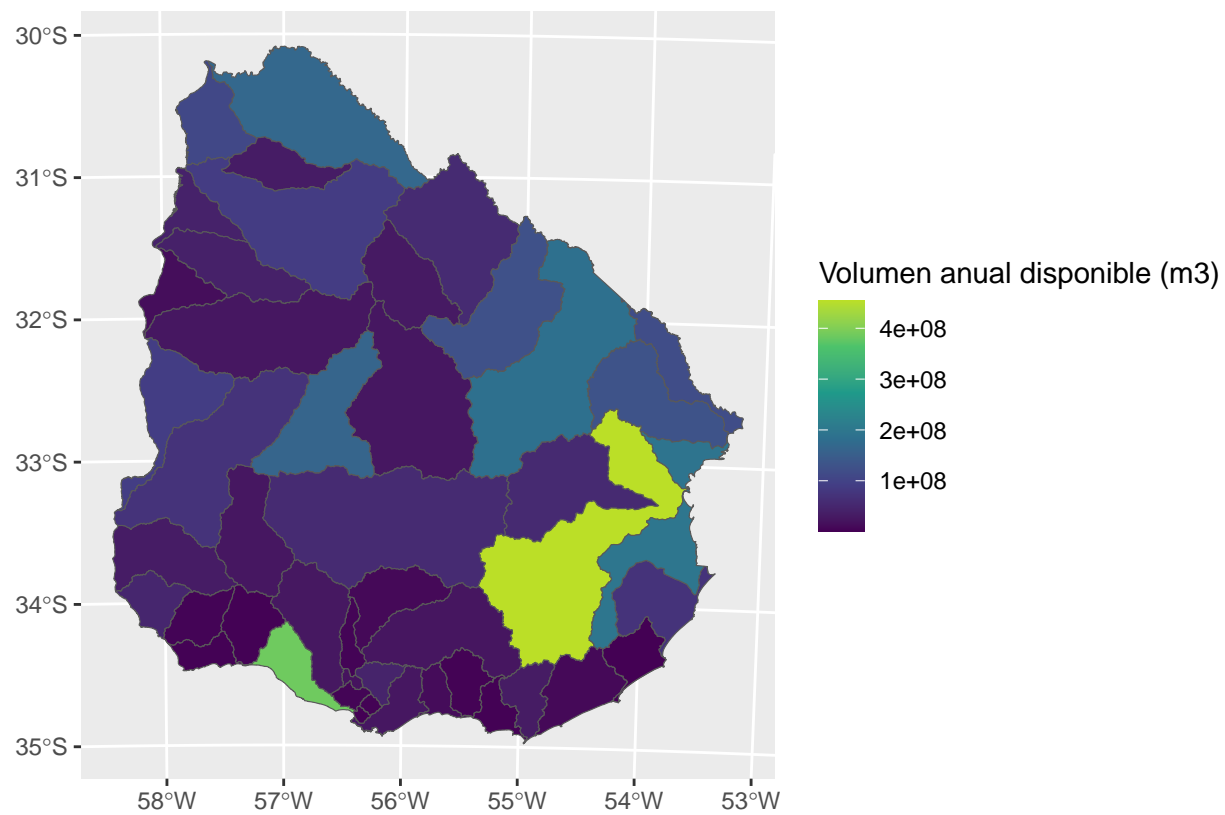
```
## Reading layer 'c098Polygon' from data source
## 'C:\Users\Mariana\AppData\Local\Temp\RtmpMFV2ZX\c098Polygon.shp'
## using driver 'ESRI Shapefile'
## Simple feature collection with 47 features and 5 fields
## Geometry type: MULTIPOLYGON
## Dimension: XY
## Bounding box: xmin: -58.43885 ymin: -34.97394 xmax: -53.18077 ymax: -30.08558
## Geodetic CRS: WGS 84
```

	Nombre Variable	Descripción
1	cod_reg	Refiere al codigo identificador de la solicitud de extraccion frente a la DINAGUA.
2	nom_reg	Nombre de la empresa, ente o persona registrada.
3	dpto	Departamento donde se encuentra la obra de extracción
4	uso	Hace referencia al propósito o destino que se le dará al agua. Los usos se clasifican en consumo humano, uso industrial.
5	destino	Detalla el lugar o fin específico donde se utilizará el agua, por ejemplo, en el caso del riego, se especifica el tipo de cultivos.
6	vol_anual	Indica el volumen máximo de agua permitido para la extracción. Este valor depende en parte, por el tipo de obra utilizada.
7	inicio	Fecha en la que se inicia la solicitud de extracción
8	fin	Fecha en la que caduca el permiso de extracción
9	tipo_ext	Se refiere al tipo de estructura utilizada para la extracción del agua, que puede ser a través de represas, tomas directas, etc.
10	mes_anio	Cantidad de meses en el año que se usa el agua que se extrae
11	codcuenca	Permite identificar la cuenca hidrográfica de nivel 2 a la cual se está realizando la extracción.
12	codcuenca_nv1	Permite identificar la cuenca hidrográfica de nivel 1 a la cual se está realizando la extracción.
13	dif_dias_ini_fin	Diferencia en días entre la fecha inicio y fin
14	lat	Coordenadas de latitud de la obra de extraccion
15	lon	Coordenadas de longitud de la obra de extraccion

Cuencas hidrograficas catalogadas por uso mayoritario

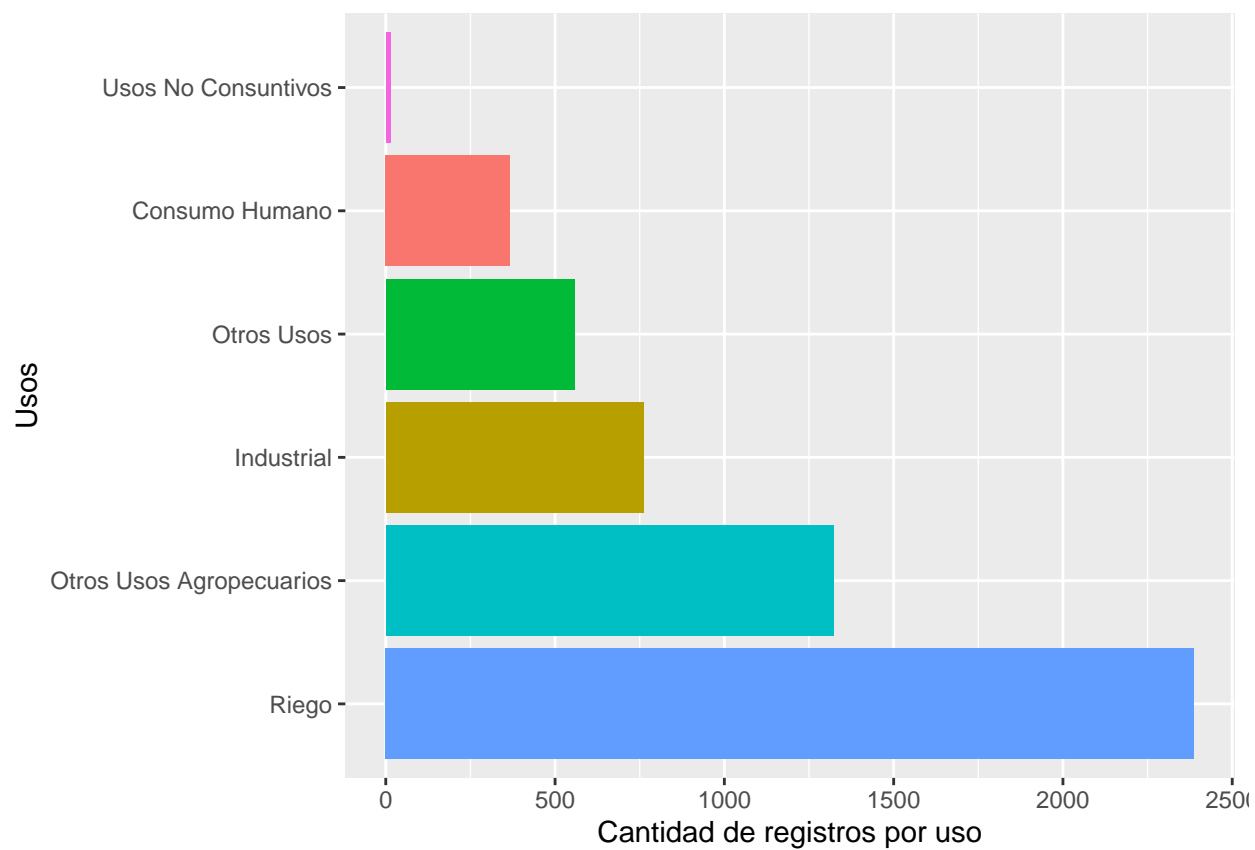






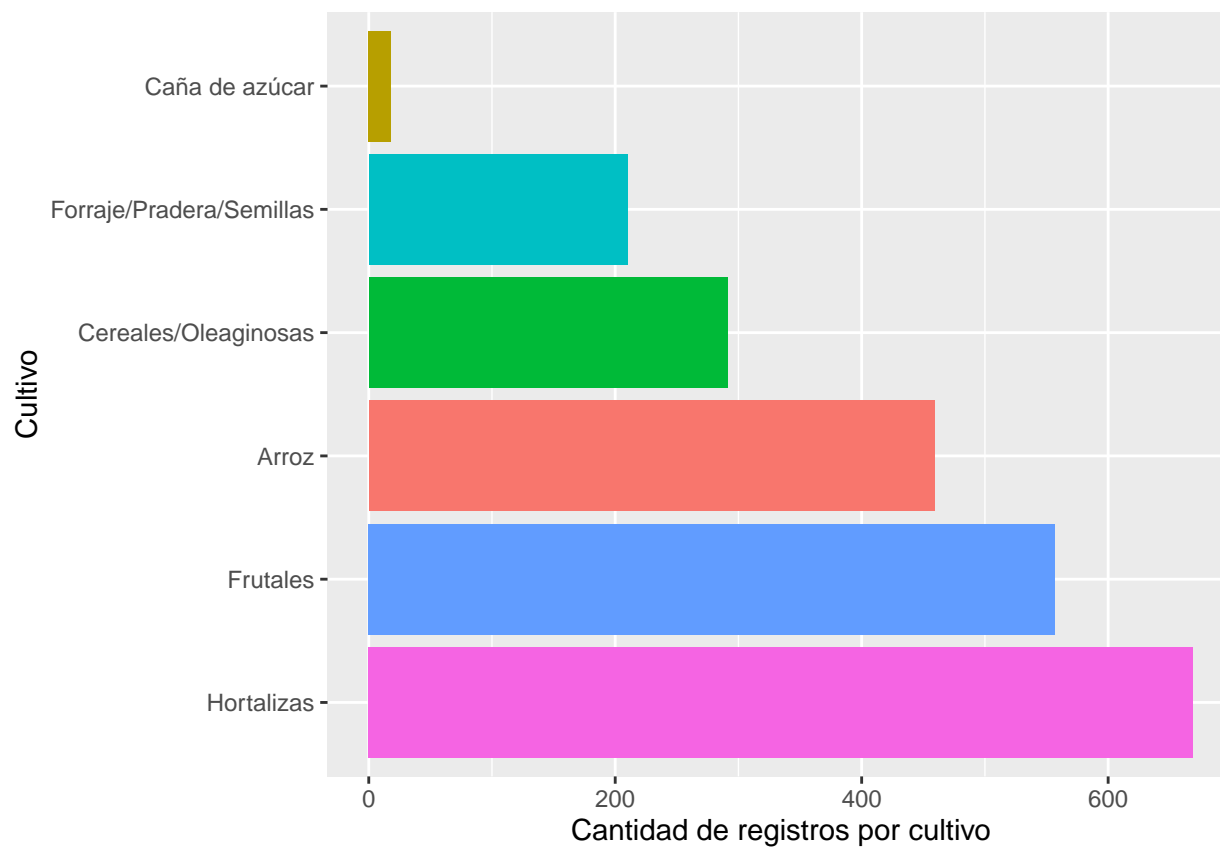
Principales usos

Numero de registros por uso

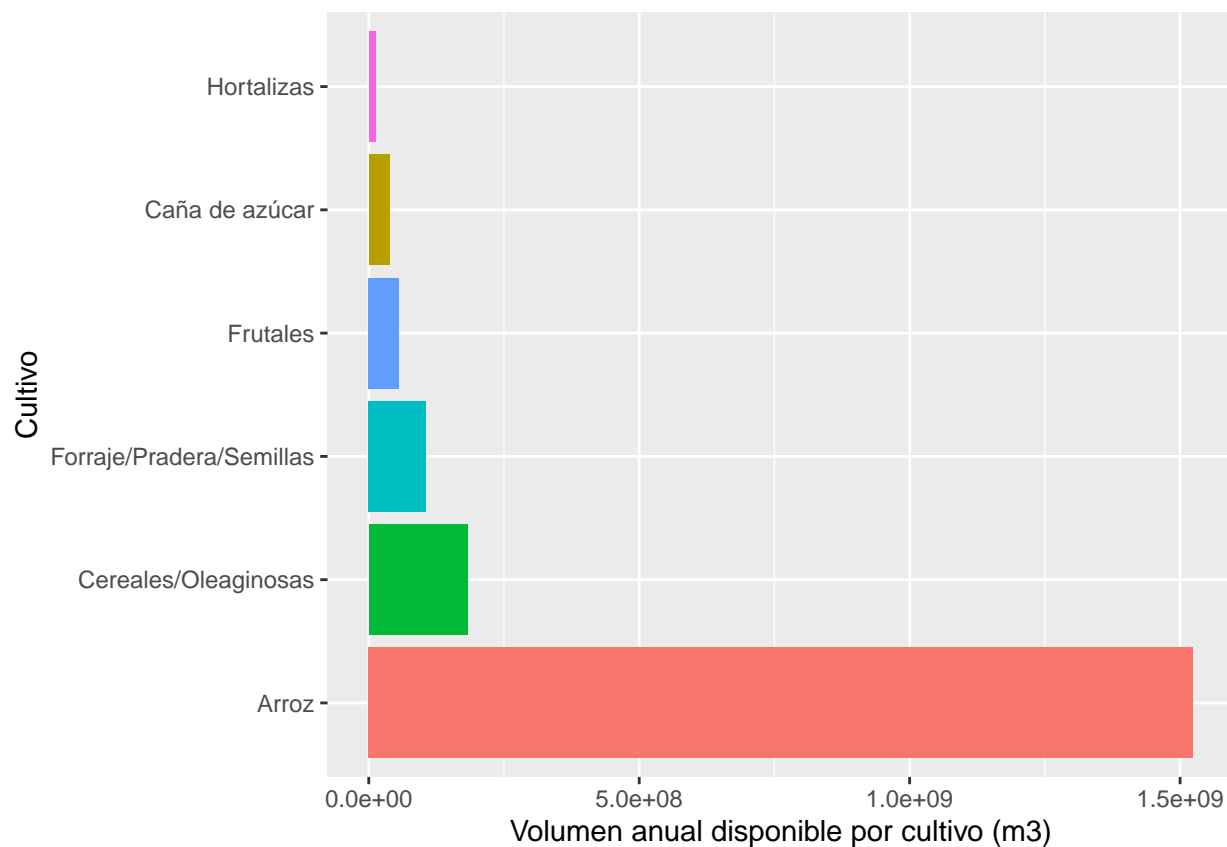


Riego

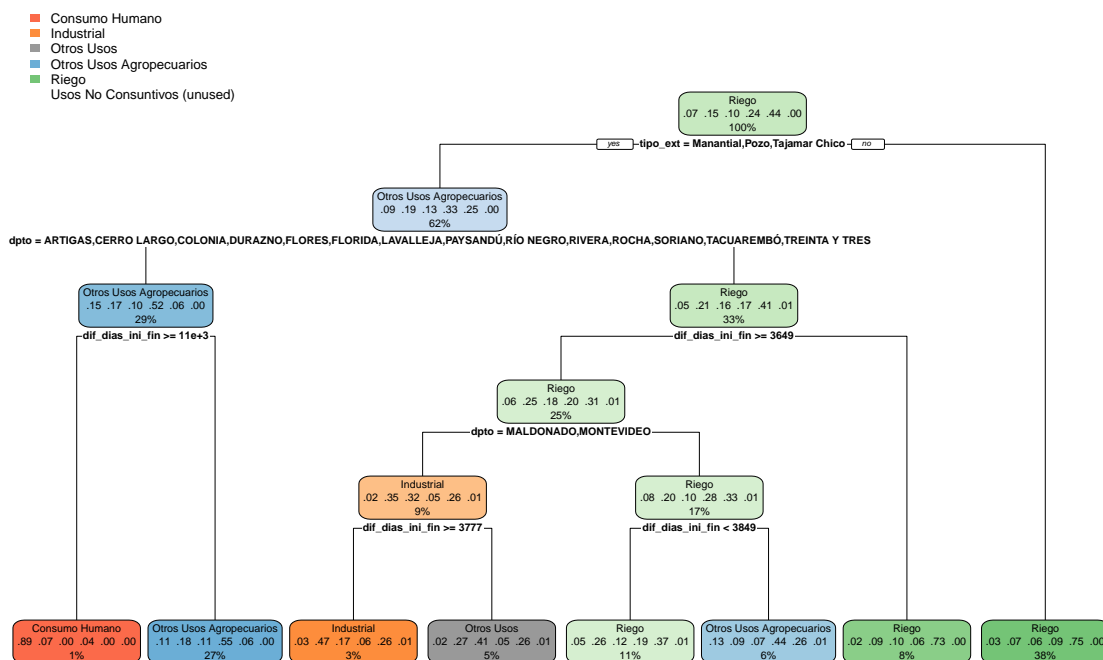
Cantidad de registros por cultivo



Volúmenes anuales por cultivo



```
## # A tibble: 6 x 4
##   uso                Cantidad Porcentaje Proporción
##   <fct>              <int>      <dbl>      <dbl>
## 1 Riego              2387        44.1        0.441
## 2 Otros Usos Agropecuarios 1323        24.5        0.245
## 3 Industrial          761        14.1        0.141
## 4 Otros Usos          557        10.3        0.103
## 5 Consumo Humano       367         6.78       0.0678
## 6 Usos No Consuntivos     14         0.26       0.0026
```



Confusion Matrix and Statistics

##

##

Reference

Prediction

Consumo Humano Industrial Otros Usos

Consumo Humano

15 0 0

Industrial

2 25 9

Otros Usos

6 28 32

Otros Usos Agropecuarios

54 83 67

Riego

31 71 57

Usos No Consuntivos

0 0 0

##

Reference

Prediction

Otros Usos Agropecuarios Riego Usos No Consuntivos

Consumo Humano

0 0 0

Industrial

3 8 1

Otros Usos

3 23 0

Otros Usos Agropecuarios

299 51 0

Riego

100 627 1

Usos No Consuntivos

0 0 0

##

Overall Statistics

##

Accuracy : 0.6253

95% CI : (0.601, 0.6491)

No Information Rate : 0.4442

P-Value [Acc > NIR] : < 2.2e-16

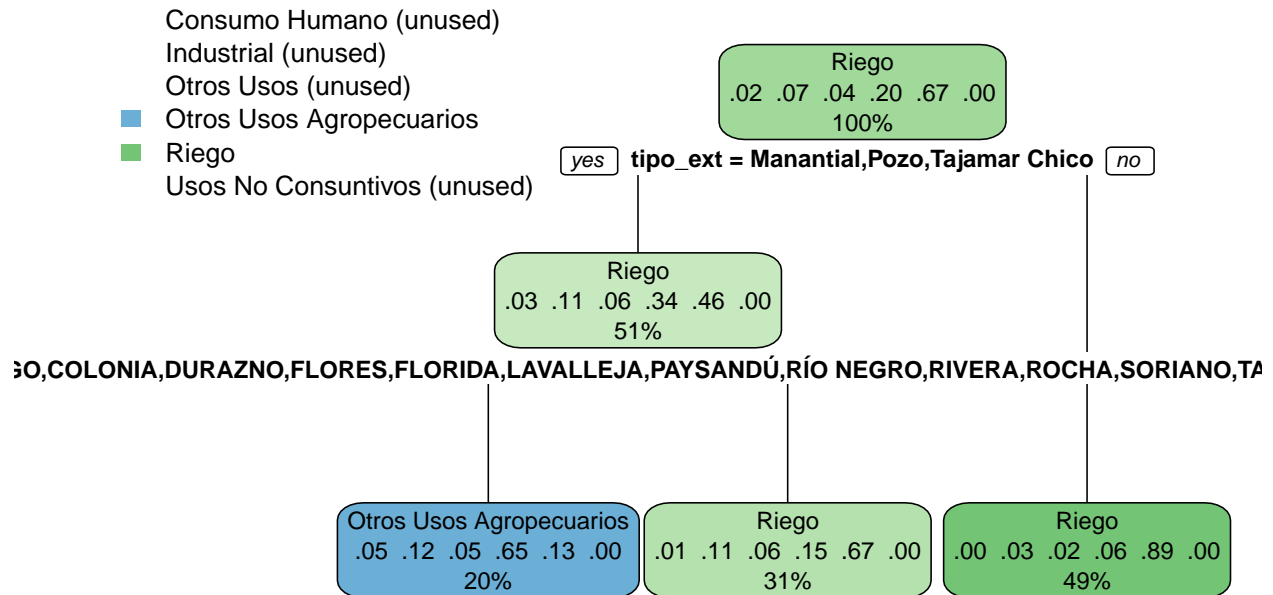
##


```

##                      Kappa : 0.4275
##
## Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Consumo Humano Class: Industrial Class: Otros Usos
## Sensitivity                0.138889                0.12077                0.19394
## Specificity                1.000000                0.98344                0.95807
## Pos Pred Value            1.000000                0.52083                0.34783
## Neg Pred Value            0.941176                0.88243                0.91157
## Prevalence                0.067669                0.12970                0.10338
## Detection Rate            0.009398                0.01566                0.02005
## Detection Prevalence      0.009398                0.03008                0.05764
## Balanced Accuracy          0.569444                0.55211                0.57601
##
##                      Class: Otros Usos Agropecuarios Class: Riego
## Sensitivity                0.7383                0.8843
## Specificity                0.7859                0.7069
## Pos Pred Value            0.5397                0.7069
## Neg Pred Value            0.8983                0.8843
## Prevalence                0.2538                0.4442
## Detection Rate            0.1873                0.3929
## Detection Prevalence      0.3471                0.5558
## Balanced Accuracy          0.7621                0.7956
##
##                      Class: Usos No Consuntivos
## Sensitivity                0.000000
## Specificity                1.000000
## Pos Pred Value            NaN
## Neg Pred Value            0.998747
## Prevalence                0.001253
## Detection Rate            0.000000
## Detection Prevalence      0.000000
## Balanced Accuracy          0.500000

```

Vamos a tratar el desbalance



Confusion Matrix and Statistics

##

##

	Reference		
## Prediction	Consumo Humano	Industrial	Otros Usos
## Consumo Humano	0	0	0
## Industrial	0	0	0
## Otros Usos	0	0	0
## Otros Usos Agropecuarios	62	72	59
## Riego	46	135	106
## Usos No Consuntivos	0	0	0

##

	Reference			
## Prediction	Otros Usos Agropecuarios	Riego	Usos No Consuntivos	
## Consumo Humano		0	0	0
## Industrial		0	0	0
## Otros Usos		0	0	0
## Otros Usos Agropecuarios	262	22		0
## Riego	143	687		2
## Usos No Consuntivos	0	0		0

##

Overall Statistics

##

Accuracy : 0.5946
95% CI : (0.5701, 0.6188)
No Information Rate : 0.4442
P-Value [Acc > NIR] : < 2.2e-16
##

```

##          Kappa : 0.3383
##
## Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##          Class: Consumo Humano Class: Industrial Class: Otros Usos
## Sensitivity          0.00000          0.0000          0.0000
## Specificity          1.00000          1.0000          1.0000
## Pos Pred Value          NaN          NaN          NaN
## Neg Pred Value          0.93233          0.8703          0.8966
## Prevalence          0.06767          0.1297          0.1034
## Detection Rate          0.00000          0.0000          0.0000
## Detection Prevalence          0.00000          0.0000          0.0000
## Balanced Accuracy          0.50000          0.5000          0.5000
##
##          Class: Otros Usos Agropecuarios Class: Riego
## Sensitivity          0.6469          0.9690
## Specificity          0.8195          0.5130
## Pos Pred Value          0.5493          0.6139
## Neg Pred Value          0.8722          0.9539
## Prevalence          0.2538          0.4442
## Detection Rate          0.1642          0.4305
## Detection Prevalence          0.2989          0.7011
## Balanced Accuracy          0.7332          0.7410
##
##          Class: Usos No Consuntivos
## Sensitivity          0.000000
## Specificity          1.000000
## Pos Pred Value          NaN
## Neg Pred Value          0.998747
## Prevalence          0.001253
## Detection Rate          0.000000
## Detection Prevalence          0.000000
## Balanced Accuracy          0.500000

```

No se observa una mejor performance, por lo que vamos a probar con un Random Forest

```
## [1] "Cantidad de registros totales: 5409"
```

```
## [1] "Cantidad de registros sin NA: 5313"
```

```
## Confusion Matrix and Statistics
```

```

##
##          Reference
## Prediction      Consumo Humano Industrial Otros Usos
## Consumo Humano          48           8           4
## Industrial             12          107          25
## Otros Usos              10           13          66
## Otros Usos Agropecuarios  40           36          38
## Riego                   13           39          39
## Usos No Consuntivos       0           1           0
##
##          Reference
## Prediction      Otros Usos Agropecuarios Riego Usos No Consuntivos
## Consumo Humano          6           5           0

```

```

##      Industrial                20      36                1
##      Otros Usos                 9      22                1
##      Otros Usos Agropecuarios   288    36                1
##      Riego                     55    596                1
##      Usos No Consumtivos        1      0                0
##
## Overall Statistics
##
##              Accuracy : 0.7007
##              95% CI : (0.6774, 0.7232)
##      No Information Rate : 0.4407
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.5706
##
## McNemar's Test P-Value : NA
##
## Statistics by Class:
##
##              Class: Consumo Humano Class: Industrial Class: Otros Usos
## Sensitivity                0.39024                0.52451                0.38372
## Specificity                0.98418                0.93154                0.96085
## Pos Pred Value             0.67606                0.53234                0.54545
## Neg Pred Value             0.95020                0.92951                0.92720
## Prevalence                 0.07800                0.12936                0.10907
## Detection Rate             0.03044                0.06785                0.04185
## Detection Prevalence       0.04502                0.12746                0.07673
## Balanced Accuracy           0.68721                0.72802                0.67229
##
##              Class: Otros Usos Agropecuarios Class: Riego
## Sensitivity                0.7599                0.8576
## Specificity                0.8740                0.8333
## Pos Pred Value             0.6560                0.8022
## Neg Pred Value             0.9200                0.8813
## Prevalence                 0.2403                0.4407
## Detection Rate             0.1826                0.3779
## Detection Prevalence       0.2784                0.4711
## Balanced Accuracy           0.8169                0.8454
##
##              Class: Usos No Consumtivos
## Sensitivity                0.000000
## Specificity                0.998729
## Pos Pred Value             0.000000
## Neg Pred Value             0.997460
## Prevalence                 0.002536
## Detection Rate             0.000000
## Detection Prevalence       0.001268
## Balanced Accuracy           0.499364

```

Problemas clasificar en: - Riego - Otros usos agrpecuarios - Industrial - Otros

```

## # A tibble: 4 x 4
##   uso          Cantidad Porcentaje Proporción
##   <fct>         <int>         <dbl>         <dbl>
## 1 Riego          2351          44.2          0.442
## 2 Otros Usos Agropecuarios 1297          24.4          0.244

```

## 3 Otros	927	17.4	0.174
## 4 Industrial	738	13.9	0.139

Confusion Matrix and Statistics

##		Reference			
## Prediction		Industrial	Otros	Otros Usos Agropecuarios	Riego
## Industrial		105	39		20 36
## Otros		28	147		33 31
## Otros Usos Agropecuarios		35	67		277 35
## Riego		36	46		49 593

Overall Statistics

Accuracy : 0.7115
 ## 95% CI : (0.6884, 0.7337)
 ## No Information Rate : 0.4407
 ## P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.5815

McNemar's Test P-Value : 0.0008988

Statistics by Class:

##	Class: Industrial	Class: Otros
## Sensitivity	0.51471	0.49164
## Specificity	0.93081	0.92801
## Pos Pred Value	0.52500	0.61506
## Neg Pred Value	0.92810	0.88640
## Prevalence	0.12936	0.18960
## Detection Rate	0.06658	0.09321
## Detection Prevalence	0.12682	0.15155
## Balanced Accuracy	0.72276	0.70983

##	Class: Otros Usos Agropecuarios	Class: Riego
## Sensitivity	0.7309	0.8532
## Specificity	0.8856	0.8515
## Pos Pred Value	0.6691	0.8191
## Neg Pred Value	0.9123	0.8804
## Prevalence	0.2403	0.4407
## Detection Rate	0.1756	0.3760
## Detection Prevalence	0.2625	0.4591
## Balanced Accuracy	0.8083	0.8524