

# Usos de los recursos hídricos en el Uruguay

Facundo Morini, Mariana Ceresa

2023-07-04

## 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 geográficos 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 públicos que realizan extracciones de agua para diversos usos productivos o abastecimiento a la población. Los datos abarcan un total de 5405 registros que se encuentran actualmente habilitados a la explotación de recursos hídricos.

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

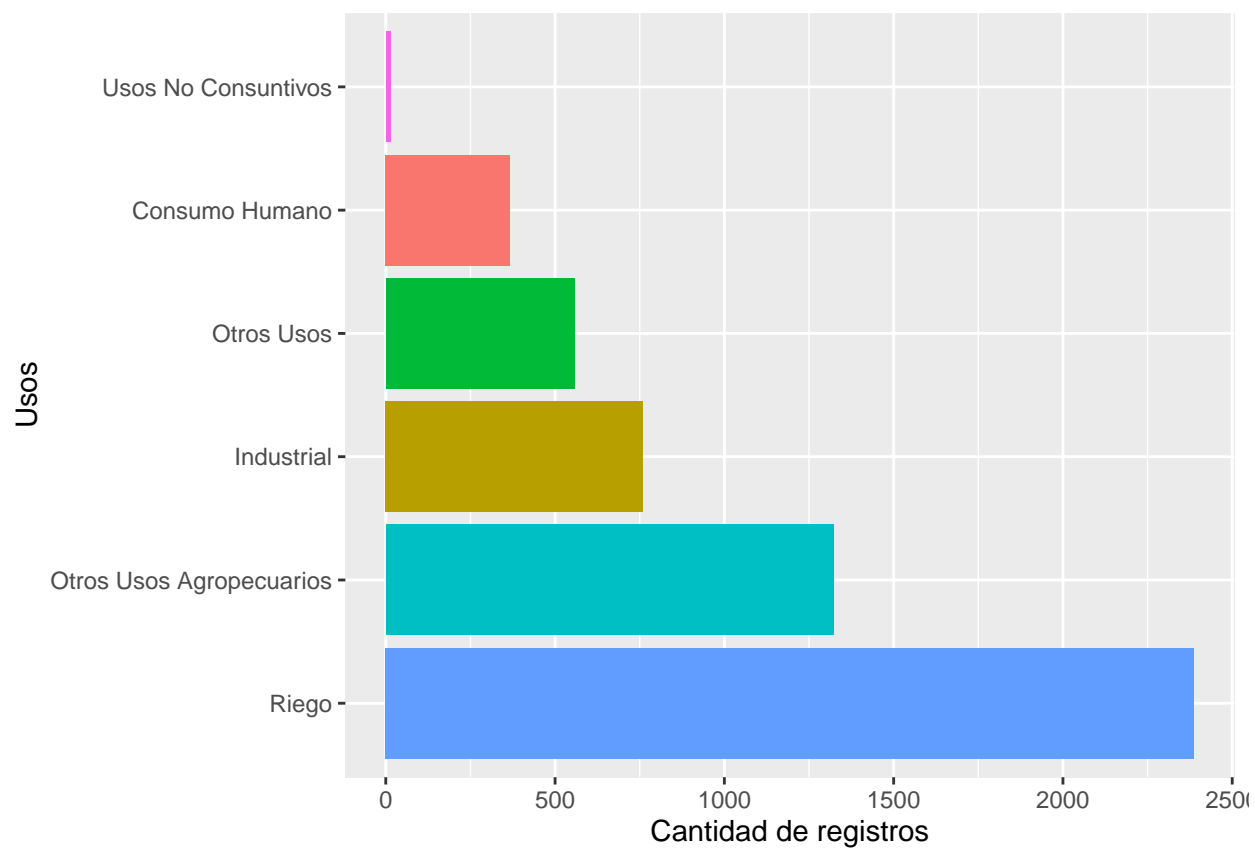
Nombre	
Vari- able	Descripción
cod_reg	Refiere al código identificador de la solicitud de extracción frente a la DINAGUA.
nom_reg	Nombre de la empresa, ente o persona registrada.
dpto	Departamento donde se encuentra la obra de extracción

Nombre	
Vari- able	Descripción
uso	Hace referencia al propósito o destino que se le dará al agua. Los usos se clasifican en consumo humano, uso industrial, riego, otros usos agropecuarios y otros usos adicionales (turismo y recreación, control de incendios, comercio y servicios).
destino	Detalla el lugar o fin específico donde se utilizará el agua, por ejemplo, en el caso del riego, se especifica el tipo de cultivo.
vol_anual	Indica el volumen máximo de agua permitido para la extracción. Este valor depende en parte, por el tipo de obra utilizada para la extracción.
inicio	Fecha en la que se inicia la solicitud de extracción
fin	Fecha en la que caduca el permiso de extracción
tipo_ext	Se refiere al tipo de estructura utilizada para la extracción del agua, que puede ser a través de represas, tomas directas en ríos o pozos, entre otras opciones comunes.
mes_año	Cantidad de meses en el año que se usa el agua que se extrae
codcuenca2	Permite identificar la cuenca hidrográfica de nivel 2 a la cual se está realizando la extracción.
codcuenca1	Permite identificar la cuenca hidrográfica de nivel 1 a la cual se está realizando la extracción.
dias_permiso	Diferencia en días entre la fecha inicio y fin
lat	Coordenadas de latitud de la obra de extraccion
lon	Coordenadas de longitud de la obra de extraccion

```
## Reading layer 'c098Polygon' from data source
## 'C:\Users\Mariana\AppData\Local\Temp\RtmpGIV2k6\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
```

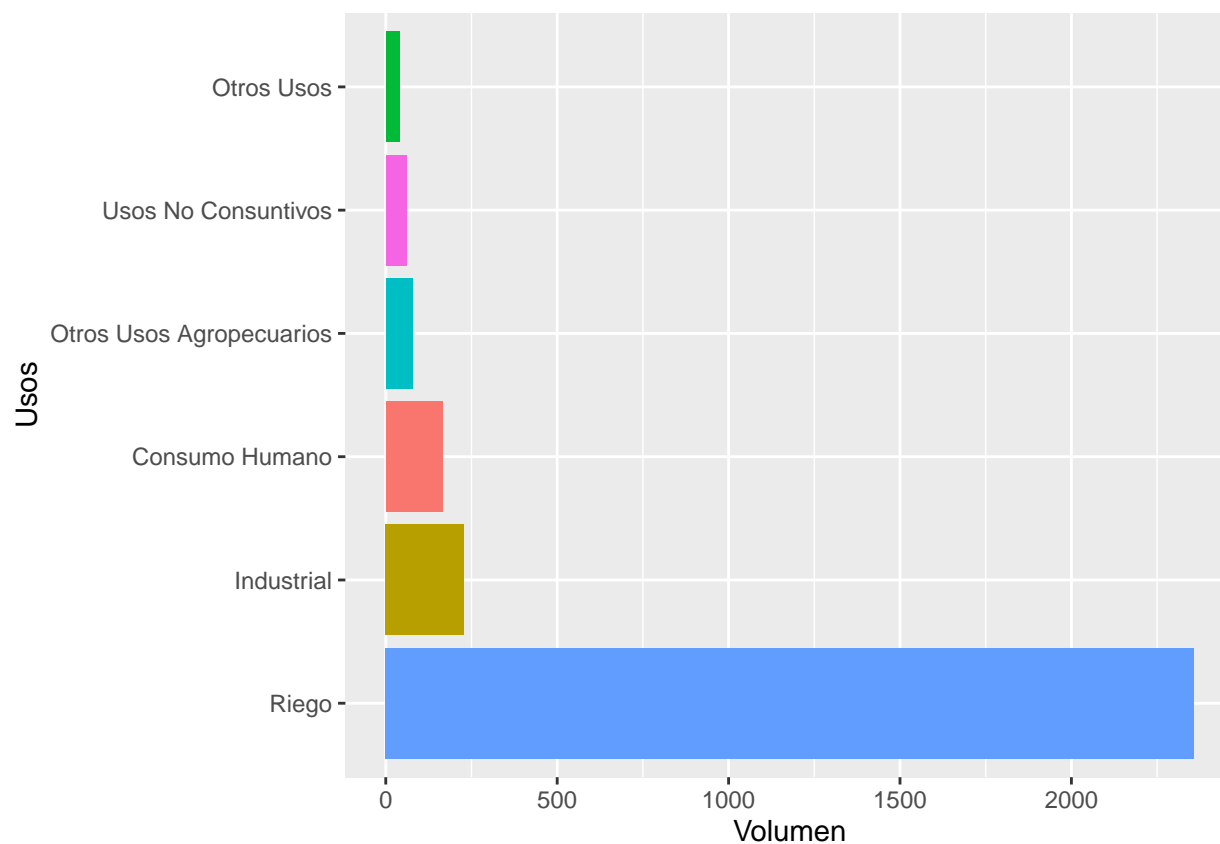
## Cantidad de registros por uso

El siguiente gráfico muestra la cantidad total de solicitudes de extracción agrupadas por el Uso de la misma.



## Volumen total por uso

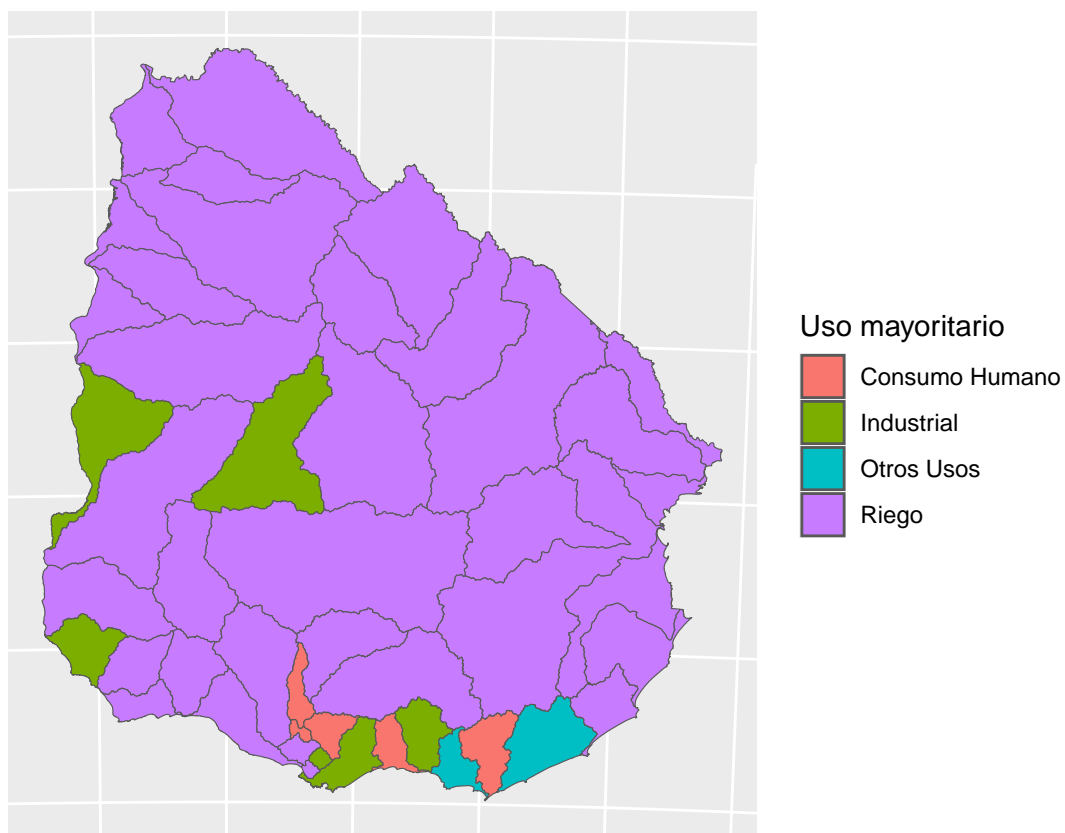
De forma análoga al gráfico anterior, el siguiente muestra el volumen total autorizado agrupados por Uso.



En los dos gráficos anteriores puede visualizarse cómo el Riego es el uso que posee tanto mayor cantidad de solicitudes, como de volumen autorizado para extraer.

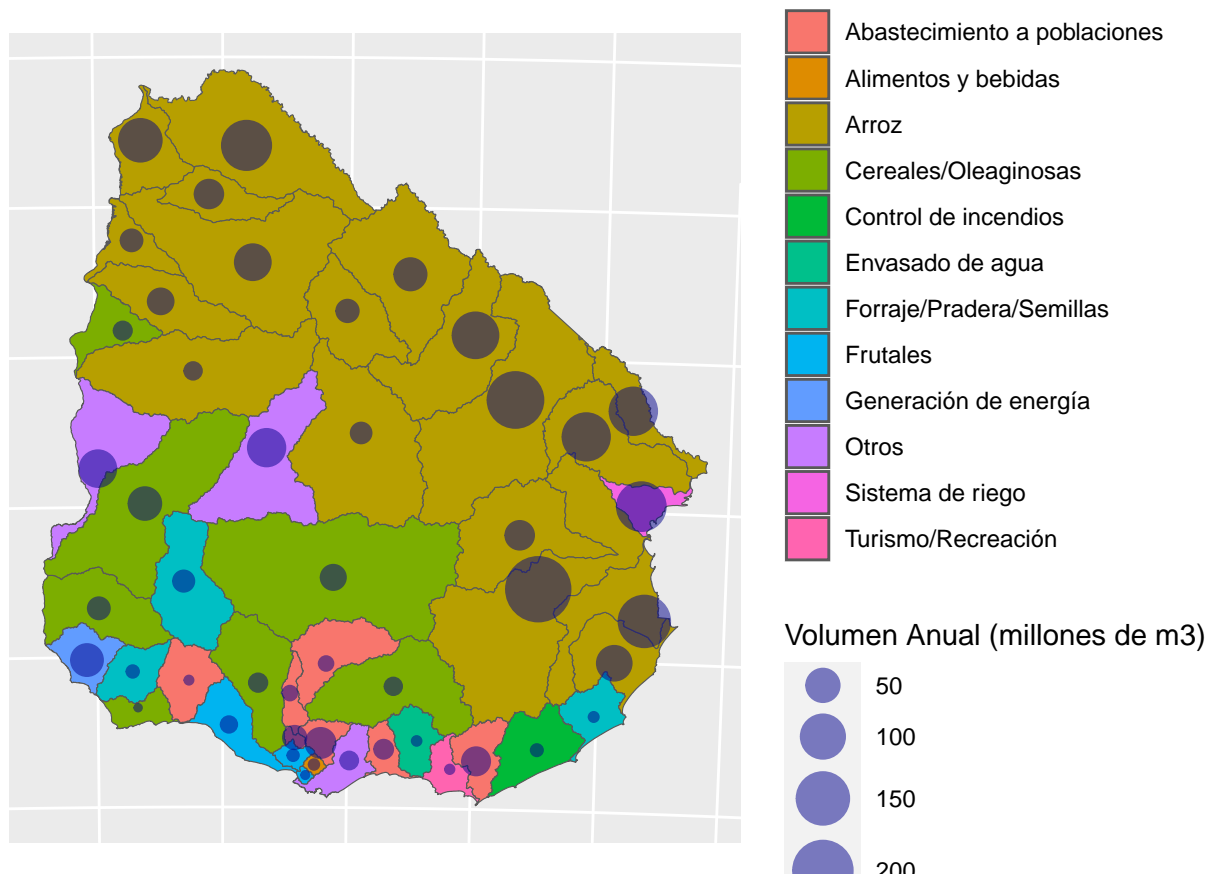
## Cuencas hidrograficas catalogadas por uso mayoritario

El siguiente gráfico nos muestra, para cada cuenca de nivel 2, cuál es el Uso con mayor volumen autorizado. Una vez más, Riego es el que encabeza el listado para la mayoría de las cuencas.



## Cuencas hidrográficas catalogadas por destino mayoritario y volumen anual disponible

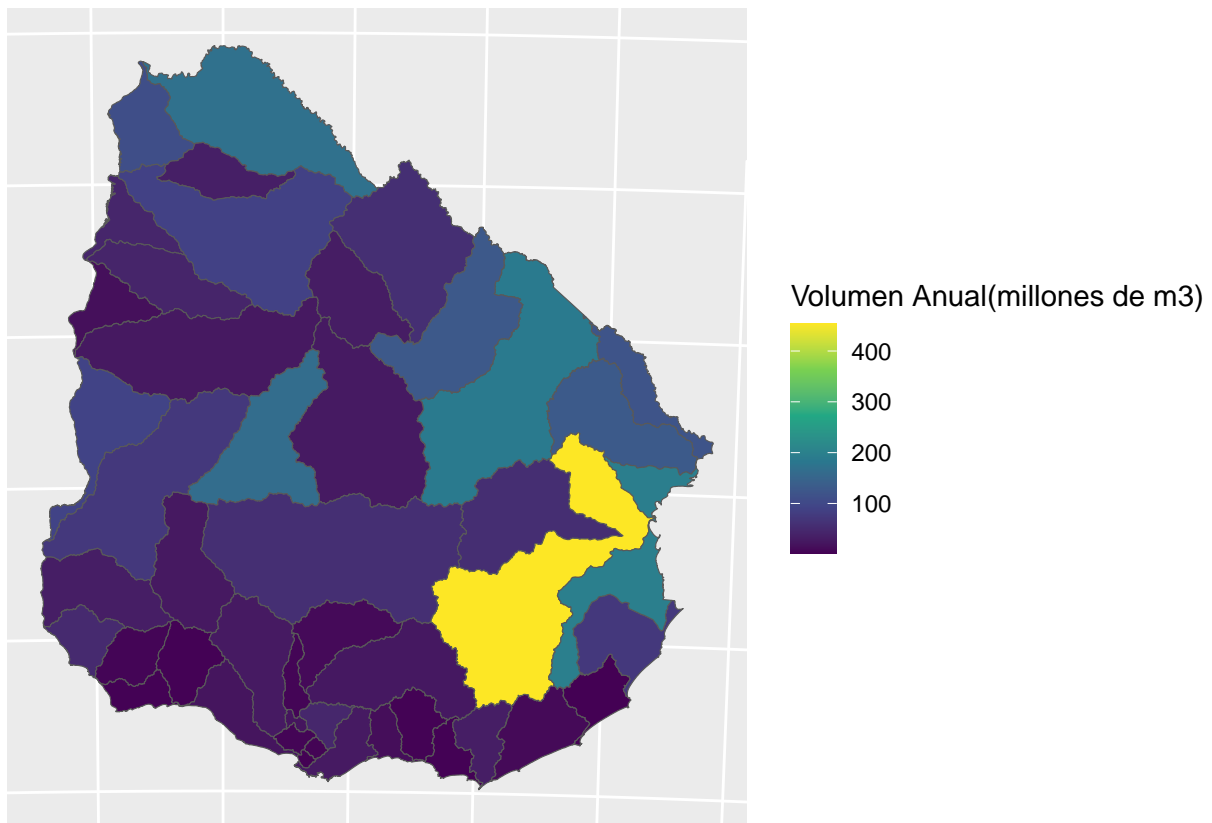
Aquí podemos ver cuál es el destino con mayor volumen para cada cuenca nivel 2. Además también se muestra el rango en el que se encuentra el volumen total para todos los usos y destinos.



En el gráfico puede observarse cómo en la mayoría de las cuencas, el destino con mayores volúmenes de agua consumidos es el Arroz, seguido por “Cereales/Oleaginosas”.

## Cuencas hidrograficas catalogadas por volumen anual disponible

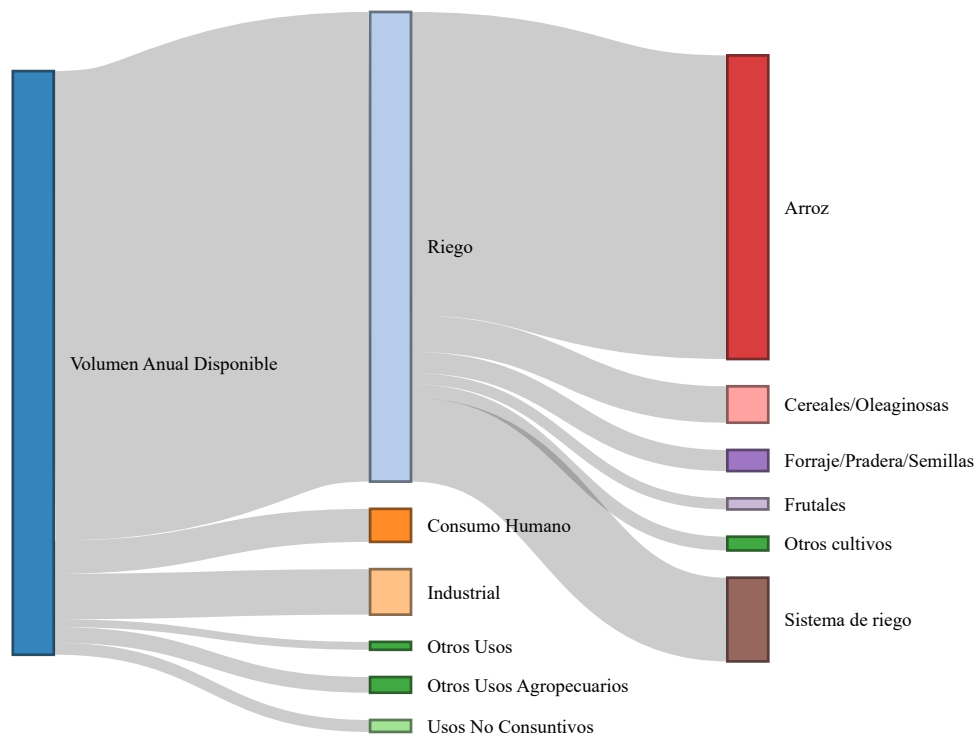
El siguiente mapa muestra el volumen anual total autorizado para cada una de las cuencas:



## **Volumen anual permitido distribuido por uso**

El siguiente esquema muestra como se desagrega el volumen total autorizado según su uso, y según cada destino dentro de cada Uso.





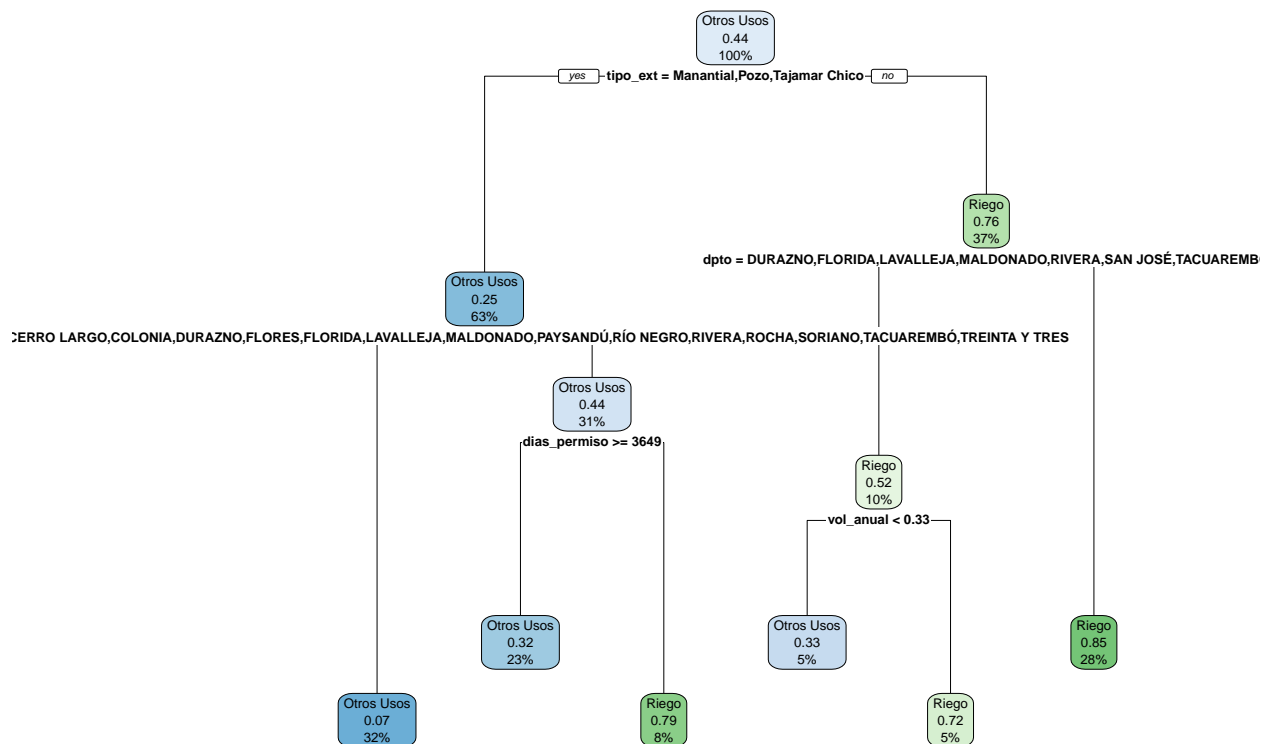
# MODELOS

Se construye un modelo para determinar si el Uso que tendrá la solicitud de extracción va a ser para Riego, o para otros usos distintos de Riego. La clasificación sería de la siguiente forma:

- Riego
- Otros Usos

## Arbol

Resultados primer Árbol construido:



```

## Confusion Matrix and Statistics
##
##               Reference
## Prediction  Otros Usos Riego
##  Otros Usos      741   177
##   Riego         120   503
##
##               Accuracy : 0.8073
##               95% CI : (0.7867, 0.8267)
##   No Information Rate : 0.5587
##   P-Value [Acc > NIR] : < 2.2e-16
##
##               Kappa : 0.6057
  
```

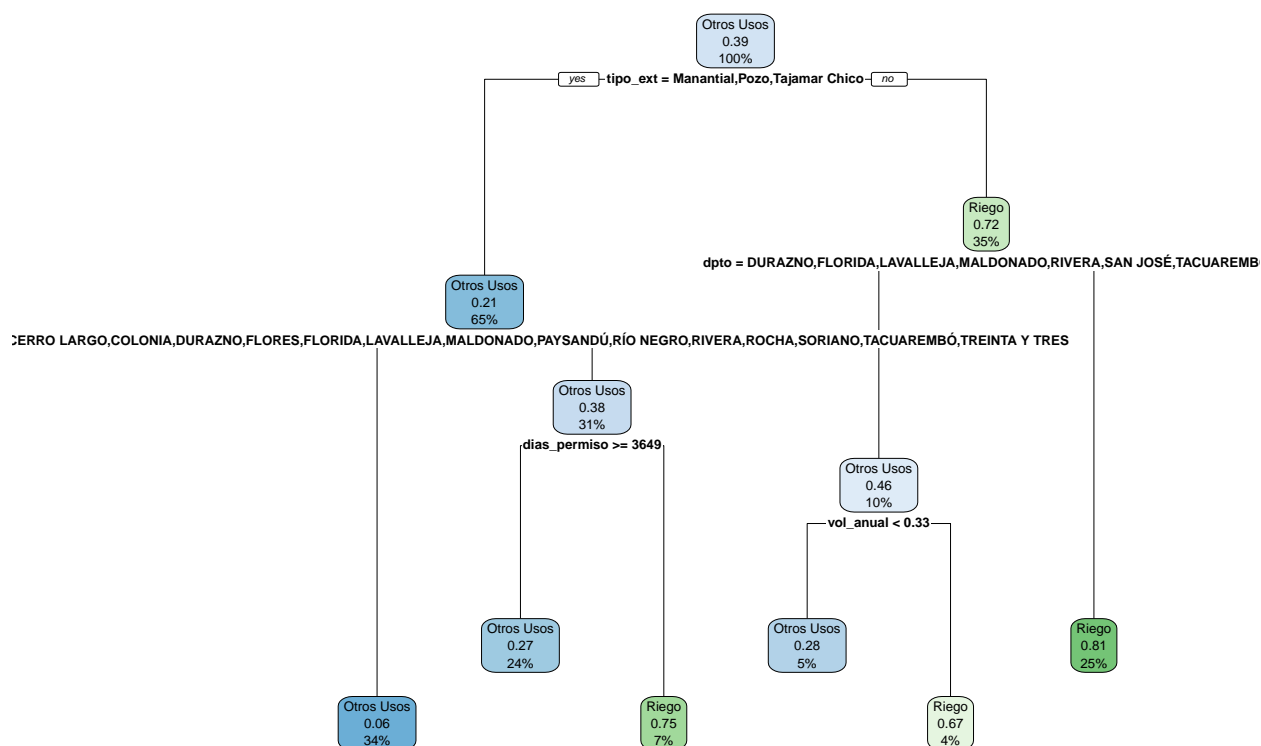
```
##
## McNemar's Test P-Value : 0.001156
##
##      Sensitivity : 0.8606
##      Specificity : 0.7397
##      Pos Pred Value : 0.8072
##      Neg Pred Value : 0.8074
##      Prevalence : 0.5587
##      Detection Rate : 0.4809
##      Detection Prevalence : 0.5957
##      Balanced Accuracy : 0.8002
##
##      'Positive' Class : Otros Usos
##
```

## Arbol 2 - tratando el desbalance

Debido a que nuestros datos se encuentran desbalanceados, realizamos un segundo árbol en el cual intentamos trabajar dicho desbalance

Tabla de proporción de Usos según la cantidad de registros:

```
## # A tibble: 2 x 2
##   uso      proporcion
##   <fct>      <dbl>
## 1 Otros Usos  0.557
## 2 Riego      0.443
```



## ## Confusion Matrix and Statistics

##

## Reference

## Prediction Otros Usos Riego

## Otros Usos 741 177

## Riego 120 503

##

## Accuracy : 0.8073

## 95% CI : (0.7867, 0.8267)

## No Information Rate : 0.5587

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

##

## Kappa : 0.6057

##

## McNemar's Test P-Value : 0.001156

##

## Sensitivity : 0.8606

## Specificity : 0.7397

## Pos Pred Value : 0.8072

## Neg Pred Value : 0.8074

## Prevalence : 0.5587

## Detection Rate : 0.4809

## Detection Prevalence : 0.5957

## Balanced Accuracy : 0.8002

##

## 'Positive' Class : Otros Usos

##

## Random Forest

Probamos mejorar nuestro Accuracy con un Random Forest, obteniendo los siguientes resultados:

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  Otros Usos Riego
##   Otros Usos          770   104
##   Riego              91   576
##
##           Accuracy : 0.8735
##           95% CI : (0.8558, 0.8897)
##   No Information Rate : 0.5587
##   P-Value [Acc > NIR] : <2e-16
##
##           Kappa : 0.7429
##
##   McNemar's Test P-Value : 0.3902
##
##           Sensitivity : 0.8943
##           Specificity : 0.8471
##           Pos Pred Value : 0.8810
##           Neg Pred Value : 0.8636
##           Prevalence : 0.5587
##           Detection Rate : 0.4997
##   Detection Prevalence : 0.5672
##   Balanced Accuracy : 0.8707
##
##   'Positive' Class : Otros Usos
##
```

\*El modelo Random Forest con un Accuracy de 0.87 es el que mejor performa\*\*