

# Usos de los recursos hídricos en el Uruguay

Facundo Morini, Mariana Ceresa

2023-07-30

## 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 exploratorio de los datos disponibles utilizando las diversas herramientas que proporciona R para el análisis de datos. Crearemos una aplicación utilizando la biblioteca Shiny, que permitirá visualizar de forma interactiva los resultados obtenidos en el informe, principalmente enfocados en responder la tres preguntas que se realizan en el informe. Por último, vamos a ajustar un modelo predictivo con el objetivo de predecir el uso del agua utilizando las diversas variables de respuesta presentes en los datos.

## Datos

Como se mencionó, los datos utilizados pertenecen a la Dirección Nacional del Agua, un ente perteneciente al Ministerio de Ambiente. Estos datos comprenden el registro de todas las empresas y entidades públicas que realizan extracciones de agua para diversos usos productivos o abastecimiento a la población. Los datos abarcan un total de 5167 registros que se encuentran actualmente habilitados para la explotación de recursos hídricos, con 19 variables que van desde el tipo de forma en que se extrae el agua hasta el departamento donde se encuentra la extracción.

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

Nombre Variable	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
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.
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
codcuenca	Permite identificar la cuenca hidrográfica de nivel 2 a la cual se está realizando la extracción.
codcuenca_nv1	Permite identificar la cuenca hidrográfica de nivel 1 a la cual se está realizando la extracción.
días_permiso	Diferencia en días entre la fecha inicio y fin

La introducción y las preguntas se explican muy bien y está muy enfocado al problema. La descripción de los datos también es bastante acertada (como comentario general, las tablas también se enumeran) me gustó que consideraran el resumen de las variables a nivel de metadatos en la introducción al leerlo el documento se vuelve aún más autocontenido.

Al final incluyeron los paquetes utilizados (que está muy bien mencionarlos) pero hubiera estado mejor que lo incluyeran como una referencia cruzada a las referencias del documento, de esta manera es más fácil acceder a la documentación.

## Análisis exploratorio de los datos

Con el objetivo de responder las preguntas planteadas, se centrará en el volumen anual permitido de extracción de agua, así como en el uso y destino asignado a la misma. También se tomará en cuenta la cuenca hidrográfica de origen de la extracción. Para llevar a cabo este análisis, utilizaremos la librería GeoUY, la cual nos facilita el uso de datos geográficos de Uruguay de manera cómoda y sencilla en R. El siguiente gráfico muestra la cantidad total de solicitudes de extracción agrupadas por su uso. En Uruguay, un país cuya producción está estrechamente vinculada al sector agrícola, se observa una marcada mayoría de empresas que realizan extracciones para uso agropecuario, tanto para riego como para otros fines, que incluyen el uso para abrevaderos de ganado, tambos, entre otros.

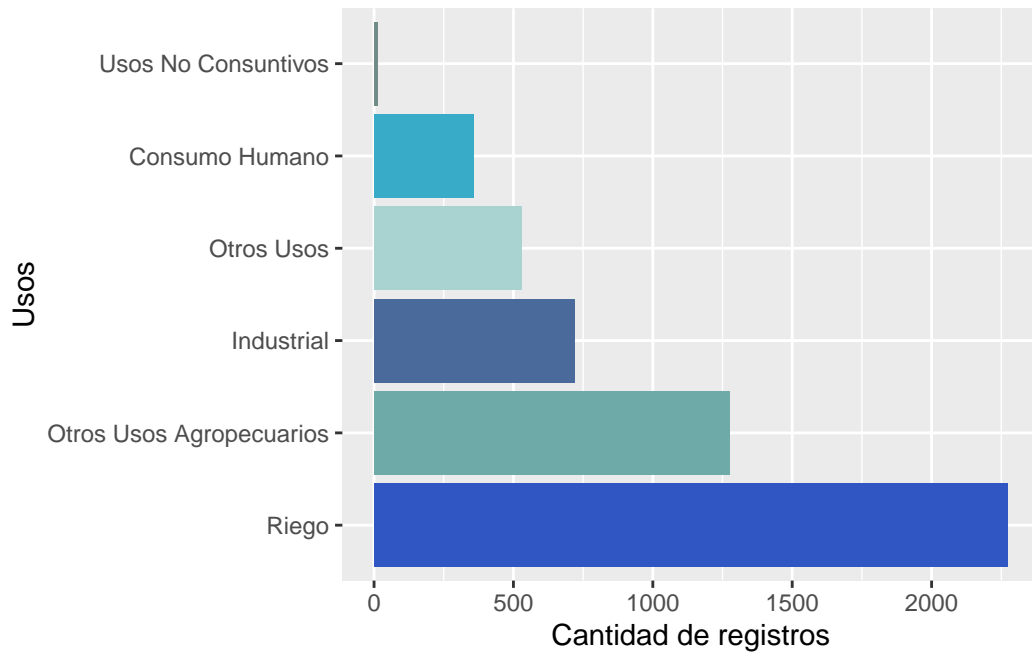


Figura 1: Cantidad de registros por uso

Deberían usar referencias cruzadas para mencionar las figuras en el texto. En el título de la figura debe contener el tipo de gráfico y las variables involucradas algunas veces también se puede incluir en el caption un breve comentario y extenderlo en el texto. Por ejemplo en la Figura 1 el título podría ser, Gráfico de barras de la cantidad de registros habilitados según usos del agua, se podrá incluir la fecha a la que se tiene los datos. Este es un comentario general para todos los gráficos, el título debe dar la información de lo que estamos dibujando, deben estar autocontenidos.

En principio, parece que los usos agropecuarios representarán la mayoría del consumo de agua. Ahora, analicemos el volumen anual permitido de extracción según los diferentes usos.

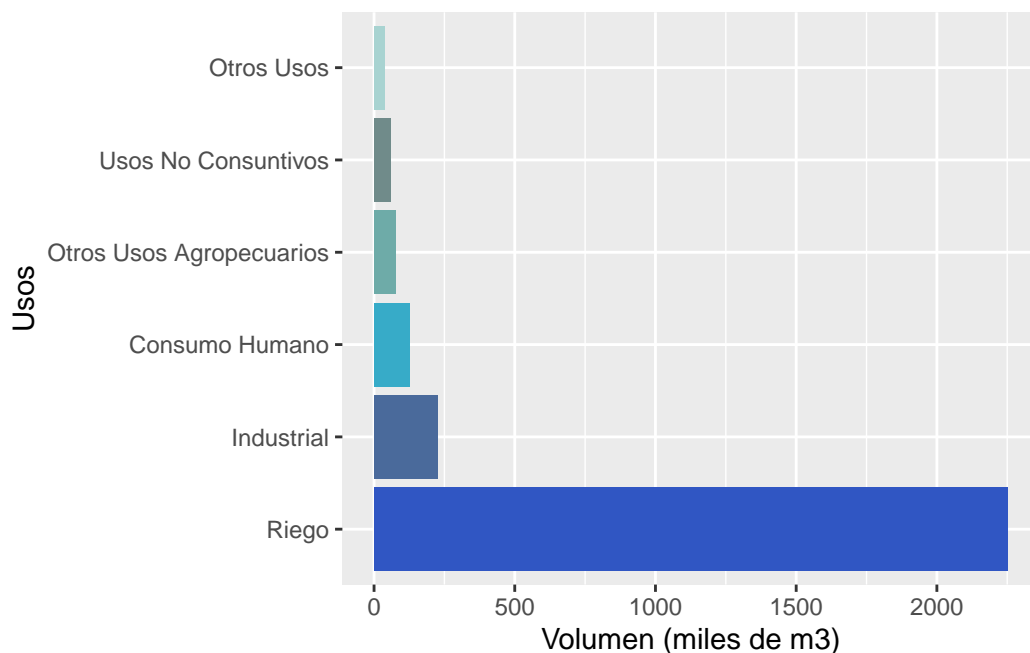


Figura 2: Volumen anual disponible por uso

Observamos una marcada diferencia en el volumen destinado al riego en comparación con los demás usos, lo cual es significativo. Aproximadamente el 80 % del volumen anual permitido de extracción en el país se destina al riego.

Dado que el riego pareciera ser el uso mayoritario del agua, vamos a examinar cómo se distribuye en todo Uruguay el uso del agua. El siguiente gráfico muestra, para cada cuenca hidrográfica, cuál es el uso con el mayor volumen autorizado en cada cuenca.

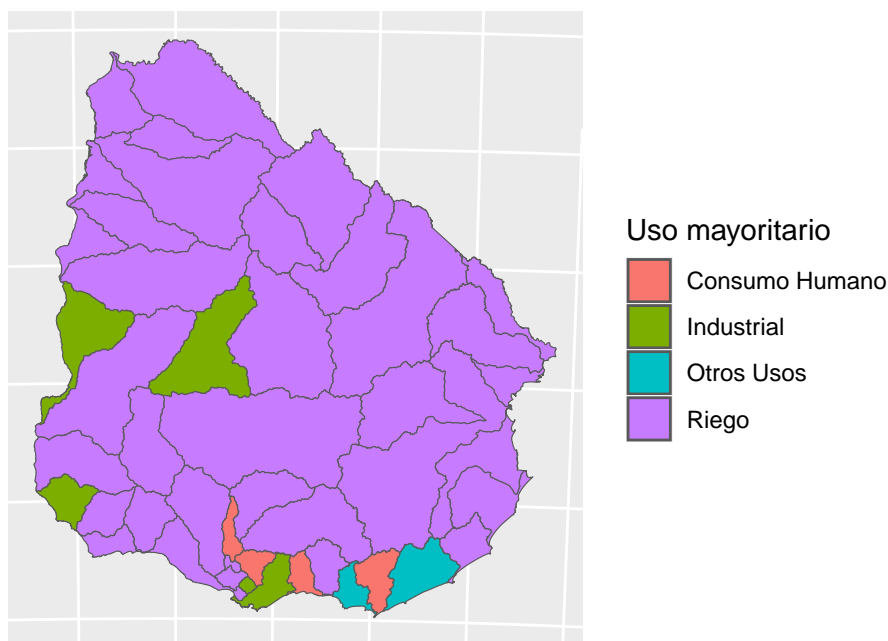


Figura 3: Mapa uso mayoritario en cada cuenca

Podemos observar cómo lo que hemos visto en los gráficos anteriores se refleja en el mapa, donde casi todas las cuencas del país presentan un uso mayoritario hacia el riego. Además, es interesante notar que en las cuencas del sur del país, donde se concentra una mayor población, el uso mayoritario pasa a ser el consumo humano.

Ahora, examinaremos un gráfico más detallado que diferencie por destinos que se le asignan al agua y el volumen anual permitido para su extracción en cada cuenca. Obteniéndose el siguiente gráfico:

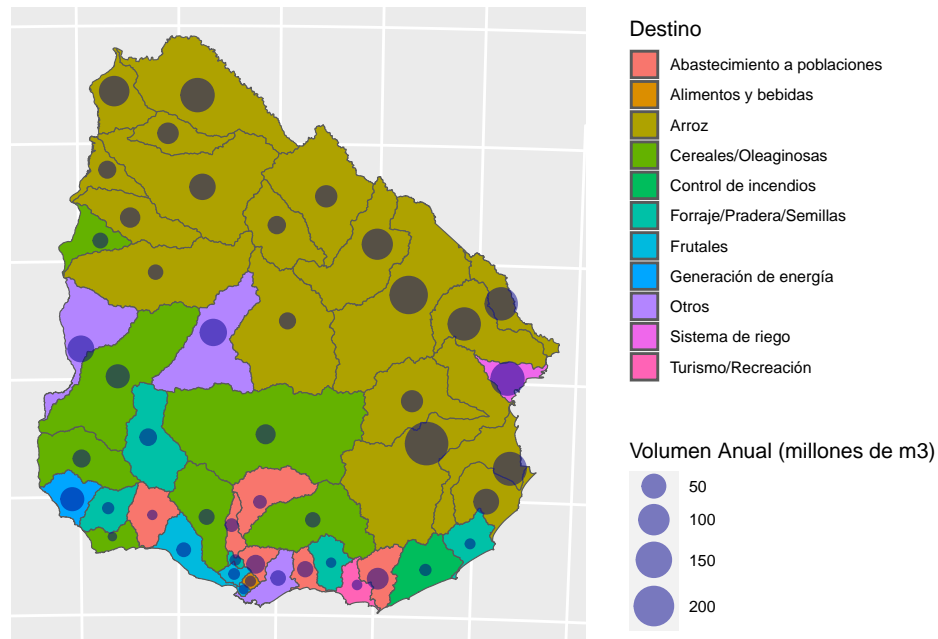


Figura 4: Mapa destino mayoritario en cada cuenca y volumen anual permitido a extraer

En el gráfico se puede observar que en la mayoría de las cuencas, el destino con los mayores volúmenes de agua es el cultivo de Arroz, seguido por el uso de “Cereales/Oleaginosas”. Ahora vamos a examinar de manera diferenciada cómo se distribuye el volumen anual permitido para el uso “Riego”.

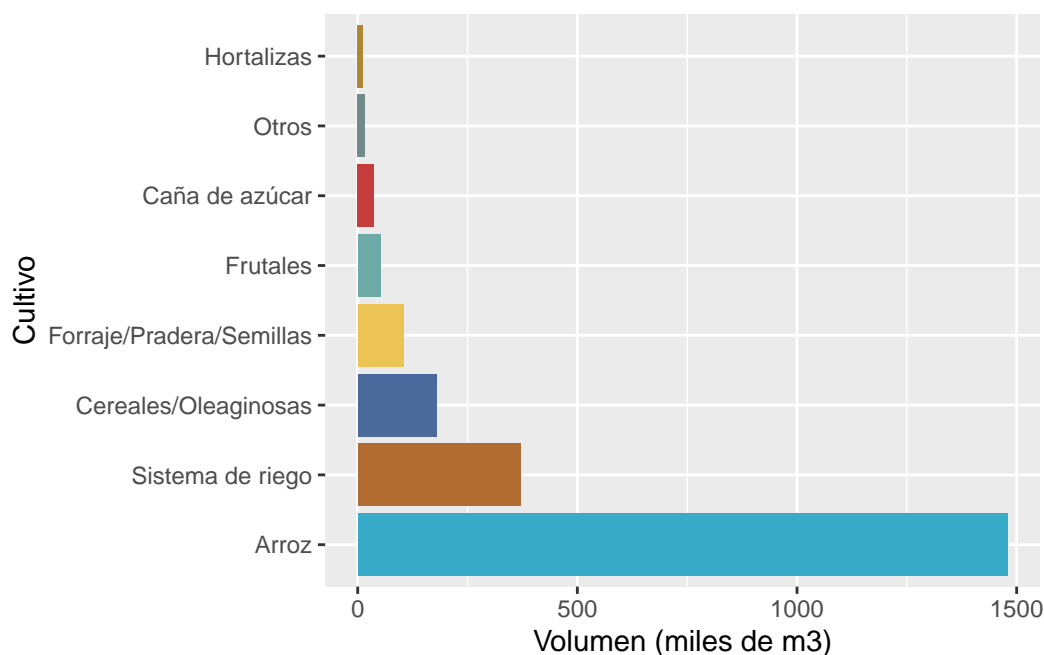


Figura 5: Grafico volumen anual disponible por cultivo

Visto los datos, queda evidente que el cultivo de arroz se posiciona como el cultivo que precisa mas extracción de agua, por encima de cualquier otro cultivo. Podemos preguntarnos también quienes tienen el máximo volumen permitido para uso viendo los 10 registros con mayor volumen tenemos los siguientes:

Codigo de registro	Nombre de registro	Departamento	Uso	Volumen anual
2021/14000/003303	PLANTA	DURAZNO	Industrial	66.0
2021/14000/005002	UPM S.A.	RÍO NEGRO	Industrial	60.0
2021/14000/004315	PLANTA	RÍO NEGRO	Usos No Consuntivos	59.0
2021/14000/005235	*SISTEMA ZAPATAROTACIÓN CULTIVO ZAFRA 2021_2022	TREINTA Y TRES	Riego	58.5
2019/14000/012327	CASARONE AGROINDUSTRIAL S.A.	CERRO LARGO	Riego	51.0
2019/14000/012284	POLÍGONO N° 17 SISTEMA CEBOLLATÍ	ROCHA	Riego	49.5
2022/36001/005342	CEL.Y ENER. P.PEREIRA. SA Y Z.FRANCA P.PEREIRA SA	COLONIA	Industrial	44.5
2022/36001/009660	CHACRA (ROTACIÓN PUSA)	TREINTA Y TRES	Riego	42.0
2019/14000/008554	CHACRA	TREINTA Y TRES	Riego	40.5
SGRH-2017-10-004-0510-1-401	CHACRAS SISTEMA CORRALES	LAVALLEJA	Riego	40.4

Al examinar las principales empresas registradas, se observa que las dos primeras corresponden a las plantas de celulosa de UPM. Estas plantas tienen los mayores volúmenes anuales permitidos de extracción, lo cual tiene sentido considerando su producción y dimensiones. Además, se destaca que una gran proporción de los registros en los primeros puestos están relacionados con empresas dedicadas al cultivo del arroz. .

Como conclusión del análisis de los datos, vemos que el enfoque al agro de la producción uruguaya repercute en los usos que se le da al agua donde es predominantemente utilizada para el riego, se destaca claramente cómo en las zonas del este del país, donde el arroz es el cultivo predominante, el volumen de agua utilizado es considerablemente mayor. Por otro lado, en las zonas del sur se percibe que el consumo humano se presenta como el uso mayoritario, especialmente en las cuencas cercanas a Montevideo. En la zona oeste del país, en cambio, prevalece el uso del agua para la producción de cereales y oleaginosas. Es importante tener en cuenta un detalle de los datos: al solo estar contabilizándose las extracciones, cultivos como el arroz, que requieren un gran traslado de agua para su producción, van a quedar en evidencia. No se podrían sacar conclusiones con los datos disponibles sobre qué cultivos tienen un mejor rendimiento de producción/volumen de agua consumidos, ya que no se estaría considerando, por ejemplo, el efecto que tienen los cultivos de soja en las aguas subterráneas, que no requieren de una extracción y no aparecen en estos datos. Este informe es solo un punto de partida para lo que podrían ser estudios más detallados sobre el tema, haciéndose preguntas como si deberíamos revisar nuestra matriz productiva en pos de asegurar los recursos hídricos en el futuro, o qué

efectos tienen otras industrias, como la forestal, que no están contemplados en estos datos en la disponibilidad del agua.

A lo que refiere al análisis exploratorio abordaron bien sus dos variables principales tanto a nivel individual como las georeferencias. En la Figura 3 deberían Poner algo como Mapa de Uruguay agrupado por cuenca hidrográfica y coloreado por uso mayoritario. En la Figura 4 se puede mejorar el título para que quede mas claro. En la Figura 4 bien en intentar agregar una capa adicional con información, un gráfico interactivo podría aportar para ello incluyendo información adicional cuando selecciono una cuenca. O estático incluyendo intensidad de color en vez de tamaño para la variable volumen.

## Aplicacion de Shiny

Para visualizar de manera interactiva los diferentes usos del agua, optamos por utilizar un mapa que muestre los destinos en las distintas cuencas y permita filtrarlos según los usos. Para lograr esto, utilizamos nuevamente los mapas disponibles en la biblioteca **GeoUY** y la librería **GGIRAPH** para crear un mapa interactivo. La idea es crear un mapa que resuma en gran medida lo que hemos observado en el análisis exploratorio.

Después de analizar varias opciones de visualización, los datos se presentan a realizar un diagrama Sankey. Este tipo de diagrama nos permite visualizar de manera bivariada los datos de volúmenes y los diferentes usos del agua, proporcionando una representación clara de como se distribuye el volumen de agua. Utilizaremos **NetworkD3** para la creación de los diagramas una librería que se adapta perfectamente para utilizar en una aplicación Shiny.

La aplicación se encuentra publicada en la siguiente ruta: [https://vdq66d-facundo-morini.shinyapps.io/App\\_Agua/](https://vdq66d-facundo-morini.shinyapps.io/App_Agua/)

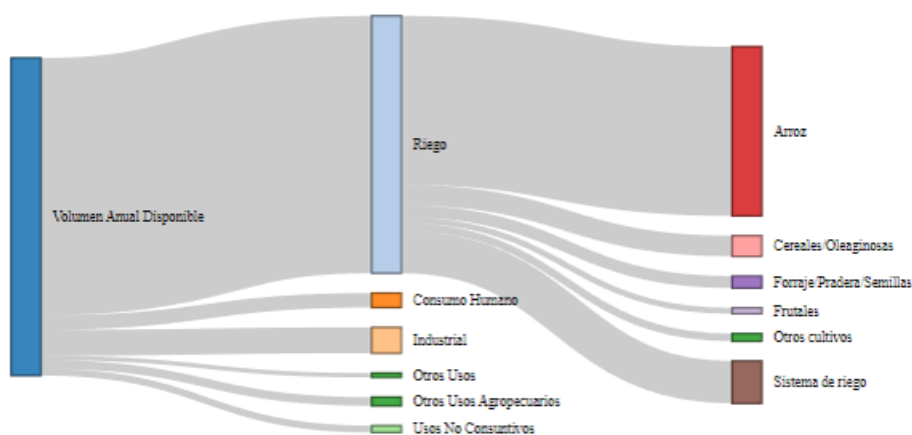


Figura 6: Diagrama Sankey para el volumen anual total

La shiny quedó muy bien, interactiva y bastante minimalista, es decir no tiene cosas de más, todo lo que se encuentra es útil. Incluir visualizaciones interactivas que incluyan un link entre plots es la primera vez que lo hacen en un proyecto en estos 5 años así que felicitaciones! Nos gustó la idea sobre el tooltip de información para darle contexto a un usuario que ingresa por primera vez, yo lo hubiera puesto al ingresar pero son gustos.



## Modelo predictivo

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

Vamos a analizar la relación entre nuestras variables estrictamente numéricas; .

- Volumen anual
- Días transcurridos entre el inicio y el fin de la solicitud

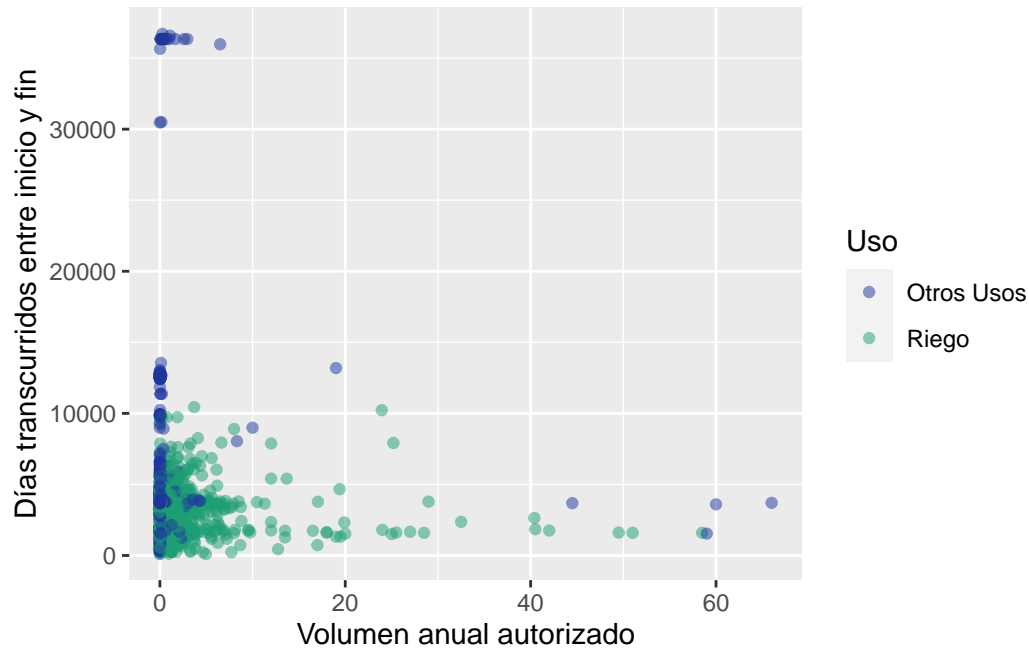


Figura 7: Relación de las variables Volumen Anual y Días entre Inicio y Fin de Solicitud

Del gráfico se desprende no solo que los menores volúmenes tienen en general destino “Otros Usos”, sino que aquellas solicitudes con mayor cantidad de días entre inicio y fin, y poco volumen autorizado, son destinados exclusivamente a “Otros Usos” también.

Tabla de proporción de Usos según la cantidad de registros:

uso	proporcion
Otros Usos	0.5585
Riego	0.4415

## Construcción de Modelos

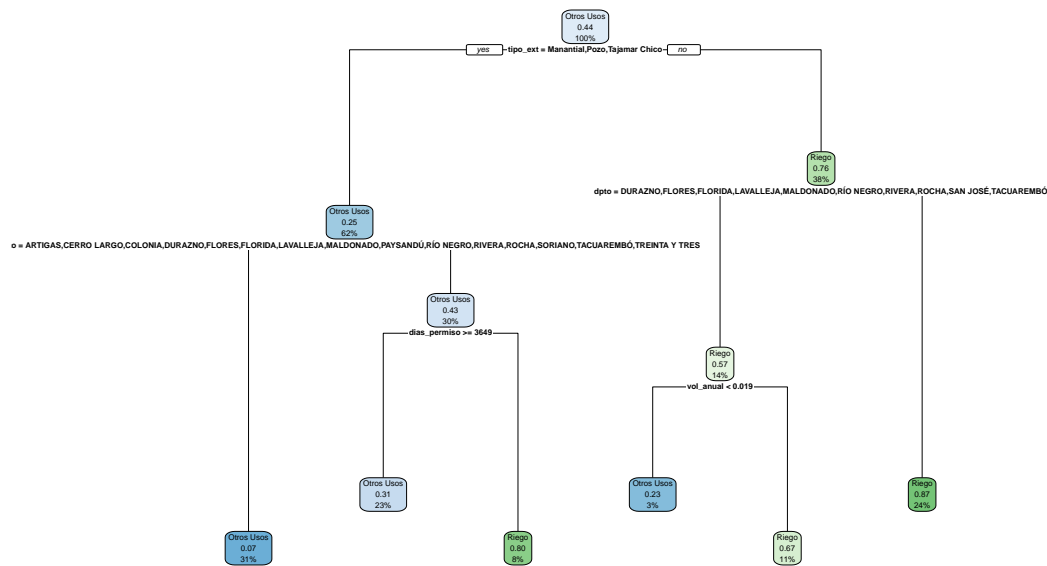
Realizamos tres modelos con los siguientes datos;

- codcuenca
- vol\_anual
- tipo\_ext
- dias\_permiso
- codcuenca\_nv1

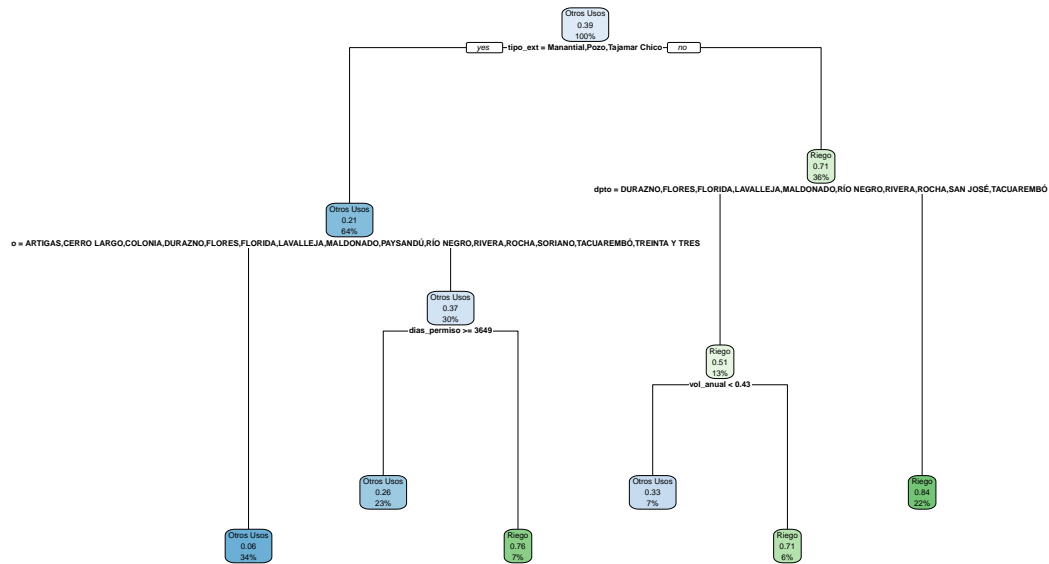
Dos árboles, y un Random Forest. La diferencia entre el primer Árbol construido y el segundo, fue que con el segundo intentamos mejorar los resultados, tratando el desbalance con el parámetro “weight”.

Los diseños de los Árboles son los siguientes:

Árbol 1



## Árbol 2



## Evaluación de modelos

Para poder evaluar nuestros modelos, presentamos en dos tablas los resultados obtenidos:

	Árbol1	Árbol2	RandomForest
Accuracy	0.82	0.81	0.86
Kappa	0.62	0.61	0.72
AccuracyLower	0.79	0.79	0.84
AccuracyUpper	0.83	0.83	0.88
AccuracyNull	0.56	0.56	0.56
AccuracyPValue	0.00	0.00	0.00
McnemarPValue	0.01	0.00	0.68

	Árbol1	Árbol2	RandomForest
Sensitivity	0.86	0.90	0.88
Specificity	0.75	0.70	0.84
Pos Pred Value	0.82	0.79	0.87
Neg Pred Value	0.81	0.85	0.85
Precision	0.82	0.79	0.87
Recall	0.86	0.90	0.88
F1	0.84	0.84	0.88
Prevalence	0.56	0.56	0.56
Detection Rate	0.49	0.51	0.50
Detection Prevalence	0.59	0.64	0.57
Balanced Accuracy	0.81	0.80	0.86

Para la selección del modelo elegimos centrarnos en las siguientes métricas; - Accuracy: Fracción de predicciones que el modelo realizó correctamente. - Sensitivity: Indica la proporción de ejemplos positivos que están identificados correctamente por el modelo entre todos los positivos reales. - Specificity: Indica la proporción de negativos clasificados correctamente.

Miremos estos valores para los tres modelos ejecutados:

	Árbol1	Árbol2	RandomForest
Accuracy	0.82	0.81	0.86
Sensitivity	0.86	0.90	0.88
Specificity	0.75	0.70	0.84

Únicamente para la Sensibilidad el modelo de Árbol 2 supera al Random Forest (0.02 “más sensible”), entendemos que al ser pequeña la diferencia, y ser mejor el Random Forest en el resto de los indicadores, es éste último el modelo que seleccionamos para poder predecir el uso final que tendrá una solicitud de extracción entre “Riego” y “Otros Usos”

Las ideas que consideraron estuvieron bien respecto al modelado. No queda claro en el texto ni en las tablas que presentan si son con datos training o test, mirando el código lo hicieron pero deberían aclarar esto en el texto. Les faltó interpretar los resultados del los árboles las conclusiones pueden ser interesantes. A lo que refiere a la evaluación del modelo la interpretación de los indicadores son correctas y podrían utilizar el F1 score que mezcla especificidad y sensibilidad. Podrían haber incluido una medida de importancia de variables en RF para interpretar los resultados.

## Paquetes utilizados y fuente de los datos

**Datos:** Visualizador de recursos hídricos. Dirección Nacional de Aguas (DINAGUA). Ministerio de ambiente. [https://www.ambiente.gub.uy/informacion\\_hidrica/](https://www.ambiente.gub.uy/informacion_hidrica/)

**Paquetes:** Wickham (2023)

HICIERON UN MUY BUEN TRABAJO Y SE NOTA QUE APRENDIERON MUCHO. SIGAN ESTUDIANDO EN LA LICENCIATURA Y VAN A PODER PROFUNDIZAR MAS SOBRE ALGUNAS COSAS QUE NO LLEGAMOS A PROFUNDIZAR EN EL CURSO Y QUE MEJORARÍA MUCHO EL MODELO DE ESTOS DATOS ENTRE OTROS. PUNTOS 95/100

Allaire, J. J., Christopher Gandrud, Kenton Russell, and CJ Yetman. 2017. *networkD3: D3 JavaScript Network Graphs from r*. <https://CRAN.R-project.org/package=networkD3>.

Breiman, Leo, Adele Cutler, Andy Liaw, and Matthew Wiener. 2022. *randomForest: Breiman and Cutler's Random Forests for Classification and Regression*. <https://www.stat.berkeley.edu/~breiman/RandomForests/>.

Detomasi, Richard. 2023. *Geouy: Geographic Information of Uruguay*. <https://CRAN.R-project.org/package=geouy>.

Gohel, David, and Panagiotis Skintzos. 2023. *Ggiraph: Make Ggplot2 Graphics Interactive*. <https://davidgohel.github.io/ggiraph/>.

Therneau, Terry, and Beth Atkinson. 2022. *Rpart: Recursive Partitioning and Regression Trees*. <https://CRAN.R-project.org/package=rpart>.

Wickham, Hadley. 2023. *Tidyverse: Easily Install and Load the Tidyverse*. <https://CRAN.R-project.org/package=tidyverse>.