

## Informe de Participación en Conecta R

### Valor 5% correspondiente a la Tarea 5

**Objetivo:** conocer acerca de los beneficios que tiene el lenguaje R para el análisis de datos y poder conocer los logros que se realizan en Costa Rica sobre este tema.

1. **Nombre del estudiante:** Sofia Elena Arias Juárez
2. **Periodo de tiempo en que participó en el evento:** 9:00 a 10:35
3. **¿Qué fue lo que más le gustó del evento:**

Las charlas tenían diferentes enfoques en las cuales R es la principal herramienta para el análisis de datos, en sí que dicho lenguaje se puede usar para datos de diferentes temas, lo cual considero enriquecedor. Aun así, se mostraron también otras herramientas como terraform y dCurl

4. **¿Por qué considera que esta temática es importante para su formación como estudiante universitario?**

Como estudiante, considero que las charlas son importantes ya que nos presentan varias herramientas, paquetes, funciones, en los cuales facilitan a las personas analizar los datos, además de que dicho lenguaje es demandado en el mercado laboral y tiene una adaptabilidad en diferentes disciplinas.

5. **¿Considera usted que el lenguaje R es importante para el análisis?, ¿Por qué? Explique ampliamente.**

El lenguaje R su principal propósito por el cual fue creado fue para el análisis de datos, muchas personas que están en esta área, utilizan el lenguaje para simplificar su trabajo, en este caso es usado por su gran diversidad de paquetes y librerías por su gran cantidad de funciones, además de que permite visualizar los datos, incluso por medio de datos, también R permite escribir scripts y documentos de R Markdown.

6. **Complete el siguiente cuadro para cada una de las charlas que participó:**

#### Charla #1

Nombre de la Charla	Proceso de Estimación de Población para el Censo 2023
Nombre del Expositor	Jesús Molina
Temática de la exposición: ¿Paquete nuevo, aplicación sobre el paquete,	El uso de R en análisis de datos censales en el INEC es el tema de la charla. Se usan diferentes funciones en R: read_csv, save, load, readxl, st_read, haven, write.csv, writexl, st_write. Por ejemplo: read_csv, podemos leer un csv para mostrarlo en graficos y columnas.

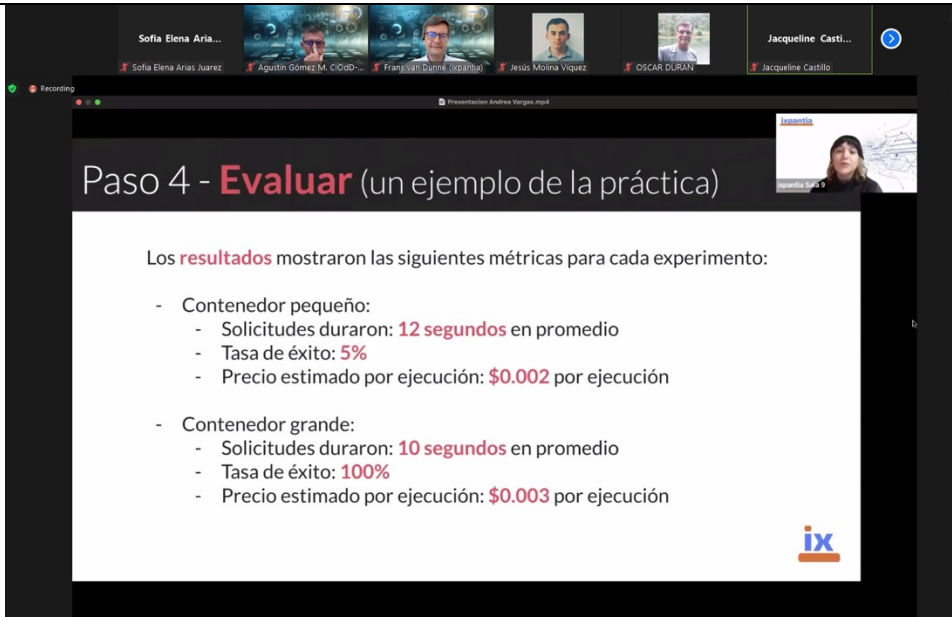
resultados interesante sobre un paquete? Todas las anteriores	
¿Cómo se llama el paquete que la persona presenta? Sino es un paquete detalle que fue lo que explicó	El expositor menciona RODBC que es una librería para la conexión de base de datos , además del uso de pandas para la generación de dataframes.
¿Dónde puedo encontrar el paquete expuesto?	La mayoría de los de paquetes pueden ser importados , por ejemplo aunque pandas es de Python, puede ser utilizado para el análisis de datos, por otro lado RODBC si es parte de R.
¿Por qué esta charla considera que fue valiosa con respecto al curso de análisis de datos?	Considerando que la entidad INEC es nacional ,el expositor menciona un poco lo que todo el proceso de recolección, además de ciertos ejemplos de código para realizar los cálculos, incluso los paquetes utilizados y las razones del uso de estos paquetes, son posibles herramientas para usar el curso de análisis de datos.
Beneficios del tema que el expositor investigó[puede ser en el tema de salud, educación o datos de cualquier índole]	Con las estimaciones poblacionales, hay un leve margen de error por lo que al crear funciones eliminan ese margen y generan más precisión en los resultados. Finalmente , a través de todo esta recolección de datos , como beneficio se puede ver el promedio de ocupantes por vivienda según operación estadística.

Capturas de pantalla sobre esta charla en particular [Debe tener una foto del expositor, de sus resultados y de la lista de asistencia de Zoom donde venga su nombre completo o bien el chat]



## Charla #2

Nombre de la Charla	Métodos cuantitativos para estimar costos de infraestructura en nube
Nombre del Expositor	Andrea Vargas & Kimberley Orozco
Temática de la exposición: ¿Paquete nuevo, aplicación sobre el paquete, resultados interesante sobre un paquete? Todas las anteriores	El tema es como optimizar y rendimiento los costos de infraestructura en la nube. Diseñar la arquitectura para soluciones data, se utilizan herramientas, entre ellas terraform, , lleva registro de cuándo y cómo se crea cada recurso, tiene replicabilidad, automatizar la creación de infraestructura, estandariza la creación de recursos en la organización, etc. Algunas funciones : \$terraform plan \$terraform apply \$terraform delete.
¿Cómo se llama el	Terraform, la expositora menciona que R tiene limitaciones en análisis de datos, por lo que puede optimizar el uso de recursos. Es una herramienta para

paquete que la persona presenta? Sino es un paquete detalle que fue lo que explico	crear infraestructura como código y nos permite crear recursos en la nube en formas de archivos.
¿Dónde puedo encontrar el paquete expuesto?	Terraform tiene que estar instalado en el sistema y una vez hecho en un IDE de R lo utilizamos.
¿Por qué esta charla considera que fue valiosa con respecto al curso de análisis de datos?	Hay una serie de pasos los cuales la expositora menciona para generar una solución, toma por ejemplo los : recursos necesarios, experimentar como código para configurar los experimentos, que tantos datos vamos a analizar, evaluar y documentar el rendimiento y los costos de los experimentos realizados, identificar áreas para mejorar y optimizar , lo cual es importante a tomar en cuenta en el análisis de datos
Beneficios del tema que el expositor investigó[puede ser en el tema de salud, educación o datos de cualquier índole]	Se espera que cada experimento sea de bajo costo para la organización, además de que sean replicables , documentar cada fase y sea fácil de entender, representativo y conciso.
Capturas de pantalla sobre esta charla en particular [Debe tener una foto del expositor, de sus resultados y de la lista de asistencia de Zoom donde venga su nombre completo o bien el chat]	 <p>The screenshot shows a Zoom meeting in progress. At the top, there is a gallery view of participants: Sofia Elena Arias, Agustín Gómez M. Córdova, Francisco Durán, Jesús Molina Viquez, OSCAR DURAN, and Jacqueline Castillo. The main window displays a presentation slide titled "Paso 4 - Evaluar (un ejemplo de la práctica)". The slide content is as follows:</p> <p>Los <b>resultados</b> mostraron las siguientes métricas para cada experimento:</p> <ul style="list-style-type: none"> <li>- Contenedor pequeño: <ul style="list-style-type: none"> <li>- Solicitudes duraron: <b>12 segundos</b> en promedio</li> <li>- Tasa de éxito: <b>5%</b></li> <li>- Precio estimado por ejecución: <b>\$0.002</b> por ejecución</li> </ul> </li> <li>- Contenedor grande: <ul style="list-style-type: none"> <li>- Solicitudes duraron: <b>10 segundos</b> en promedio</li> <li>- Tasa de éxito: <b>100%</b></li> <li>- Precio estimado por ejecución: <b>\$0.003</b> por ejecución</li> </ul> </li> </ul> <p>The slide also features the "ix" logo in the bottom right corner.</p>

### Charla #3

Nombre de la Charla	Inteligencia y Análisis en la toma de decisiones: Dashboards en R y variables significativas con Dinamic CUR
Nombre del Expositor	Greivin Villegas
Temática de la exposición: ¿Paquete nuevo, aplicación sobre el paquete, resultados interesante sobre un paquete? Todas las anteriores	Uso del paquete Dinamic CUR en R y C# con paquetes en R. Es un paquete desarrollado por la escuela de estadística de la UCR, Dinamic CUR es para la reducción de la dimensionalidad de matrices de datos. Se utilizaron diferentes paquetes para la creación de este
¿Cómo se llama el paquete que la persona presenta? Sino es un paquete detalle que fue lo que explicó	El paquete se llama Dinamic CUR o dCur.
¿Dónde puedo encontrar el paquete expuesto?	Se puede descargar en la página Web o RStudio.
¿Por qué esta charla considera que fue valiosa con respecto al curso de análisis de datos?	Simplifica así las hojas con una gran cantidad de datos, lo cual en nuestro curso podría ayudarnos a entender más fácil los datos de la investigación y utilizar funciones sobre esta.
Beneficios del tema que el expositor investigó[puede ser en el tema de salud, educación o datos de cualquier índole]	En el análisis de datos, este paquete beneficia en: Encuentra las mejores columnas que minimiza el error, obtiene las variables más relevantes de forma rápida, se puede visualizar la progresión del error de manera dinámica y la visualización de grandes cantidades de datos es más simple.

Capturas de pantalla sobre esta charla en particular [Debe tener una foto del expositor, de sus resultados y de la lista de asistencia de Zoom donde venga su nombre completo o bien el chat]

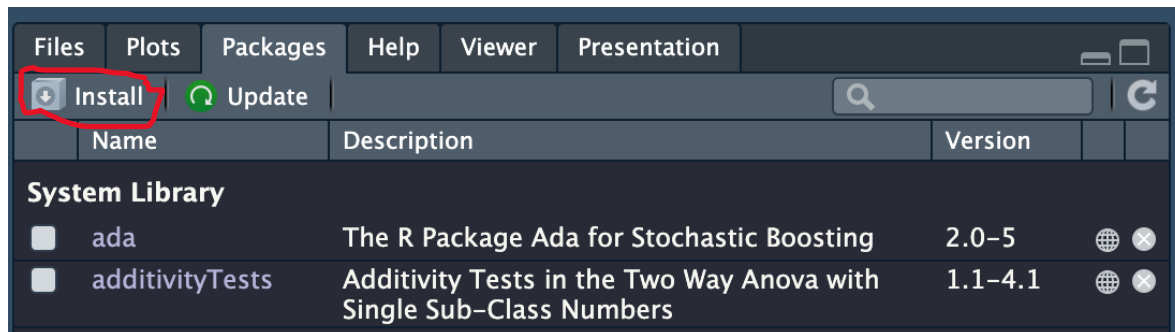


Nota: 25 puntos extra (sobre la tarea 5) si usted instala uno de los paquetes y documenta su instalación, haciendo eso en este mismo documento.

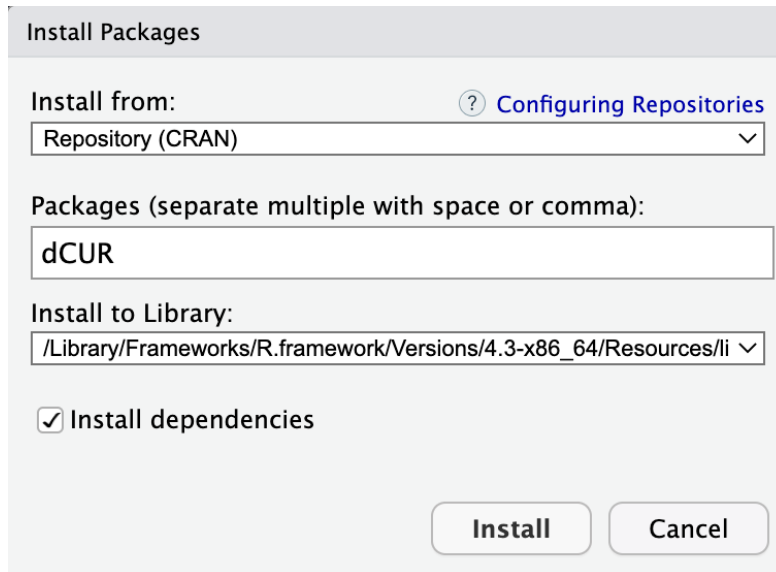
#### Instalación de dCurl en RStudio:

Para instalar el paquete dCurl, en el entorno de desarrollo RStudio, simplemente se siguen los siguientes pasos:

1. Vamos al apartado de Packages y damos click en el logo de descarga que tiene al lado la etiqueta "Install".



2. Seleccionamos de donde queremos descargar la librería, en este caso se puede instalar del Repository CRAN. La ventana además, nos indica los paquetes a instalar y donde se instalaran.



Install Packages

Install from: [? Configuring Repositories](#)

Repository (CRAN) ▼

Packages (separate multiple with space or comma):

dCUR

Install to Library:

/Library/Frameworks/R.framework/Versions/4.3-x86\_64/Resources/li ▼

☒ Install dependencies

Install Cancel

3. Rstudio por sí solo empezara a descargar el paquete y todas sus dependencias, ya con esto podemos hacer uso del paquete.

```
error: unable to access object (object is from previous session and is now archived)
> install.packages("dCUR")
also installing the dependencies 'rbibutils', 'ppcor', 'Rdpack'

trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-x86_64/contrib/4.3/rbibutils_2.2.16.tgz'
Content type 'application/x-gzip' length 1055419 bytes (1.0 MB)
=====
downloaded 1.0 MB

trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-x86_64/contrib/4.3/ppcor_1.1.tgz'
Content type 'application/x-gzip' length 28304 bytes (27 KB)
=====
downloaded 27 KB

trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-x86_64/contrib/4.3/Rdpack_2.6.tgz'
Content type 'application/x-gzip' length 746805 bytes (729 KB)
=====
downloaded 729 KB

trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-x86_64/contrib/4.3/dCUR_1.0.1.tgz'
Content type 'application/x-gzip' length 582017 bytes (568 KB)
=====
downloaded 568 KB

The downloaded binary packages are in
  /var/folders/gx/4y1_cgwd1ll9ym15w79m14sm0000gn/T//RtmpvD9p39/downloaded_packages
> |
```

