



Visualización de datos y automatización de reportes con ggplot2 y R Markdown



Visualización de datos y automatización de reportes con ggplot2 y R Markdown

Cursos de perfeccionamiento profesional – receso académico 2024

Organizado por: Dirección de investigación - Universidad Tecnológica de Panamá

Autor: Danny Murillo González

Centro de Investigación, Desarrollo e Innovación en Tecnologías de la Información y las Comunicaciones - CIDITIC

<https://orcid.org/0000-0003-0297-7213>

<https://scholar.google.com/citations?user=YNx08I0AAAAJ&hl=es>

Marzo 2024



Esta obra está bajo licencia internacional Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0.

<https://ridda2.utp.ac.pa/handle/123456789/18250>

Acerca del curso

El material presentado está relacionado con curso de **Visualización de datos y automatización de reportes con ggplot2 y R Markdown** con una duración de 40 horas en formato virtual sincrónico organizado por la dirección de Investigación de la universidad Tecnológica de Panamá.

Objetivo: Hacer uso efectivo de la reglas, principios y recomendaciones en la visualización de datos al generar gráficos con la herramienta Rstudio utilizando lenguaje R y el paquete ggplot2 para crear visualización de datos e informes reproducibles y automatizados con Rmarkdown.

Descripción: La visualización de datos es la técnica utilizada para brindar información sobre los datos utilizando señales visuales como gráficos, tablas, mapas y muchos otros. Esto es útil ya que ayuda a comprender fácilmente y de forma intuitiva grandes cantidades de datos y, por lo tanto, a tomar mejores decisiones al respecto.

La visualización de datos tiene dos propósitos principales:

- Análisis de datos exploratorios e identificación de problemas de datos.
- Comunicar conocimientos y resultados

Las herramientas populares de visualización de datos disponibles son Tableau, Plotly, R, Phytion, Google Charts, DataWrapper, Flourish , entre otros. Las distintas plataformas de visualización de datos tienen diferentes capacidades, funcionalidades, sin embargo, una de las mayores diferencias de R, es que es lenguaje diseñado para análisis de datos, estadística, visualización e investigación científica. Si bien como lenguaje de programación, su curva de aprendizaje suele resultar compleja, el componente de reproducibilidad y flexibilidad en las funciones que contienen son una de las mayores ventajas de quienes lo utilizan.

Contenido del Curso:

Tema 1: Conceptos de EDA y visualización de datos

Tema 2: Paquete Ggplot2

Tema 3: Tipos de gráficos con variable numéricas

Tema 4: Tipos de gráficos con variables categóricas

Tema 5: Principios y recomendaciones en la visualización de datos

Tema 6: Rmarkdwon y automatización de Reportes

Requisitos

Crear proyecto

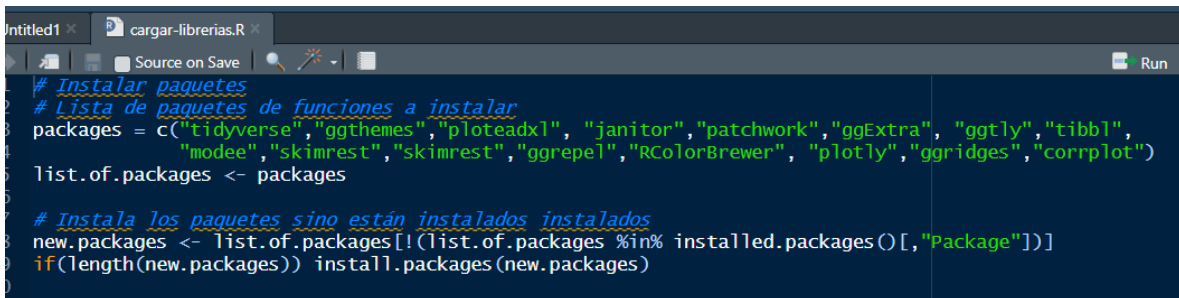
Para seguir este curso debe descargar el documento **curso_VIZ-DI.zip** que se encuentra en el [Repositorio Ridda2](#).

Al descargar el documento, coloque el archivo en cualquier ubicación de su computador y descomprímalo.

Dentro del documento aparecerá un archivo llamado **curso_VIZ-DI.Rproj** con un icono con un cubo celeste y una R haga doble click sobre él. Se abrirá el Rstudio con la configuración del proyecto a utilizar. esta es la carpeta con los códigos y datos iniciales el curso.

code	03/01/2024 10:53 a. m.	Carpeta de archivos	
data	03/01/2024 9:26 a. m.	Carpeta de archivos	
html	03/01/2024 1:53 p. m.	Carpeta de archivos	
img	03/01/2024 9:26 a. m.	Carpeta de archivos	
.Rhistory	03/01/2024 11:53 a. m.	Archivo RHISTORY	1 KB
curso_VIZ-DI	03/01/2024 10:59 a. m.	R Project	1 KB

En la carpeta **code** del proyecto creado, hay un archivo con el nombre **cargar-librerias.R**, puede ejecutar el código que se muestra en la figura para instalar los paquetes que necesita para el curso.



```

# Instalar paquetes
# Lista de paquetes de funciones a instalar
packages = c("tidyverse", "ggthemes", "ploteadxl", "janitor", "patchwork", "ggExtra", "ggTly", "tibble",
            "modee", "skimrest", "skimrest", "ggrepel", "RColorBrewer", "plotly", "ggridges", "corrplot")
list.of.packages <- packages

# Instala los paquetes sino están instalados instalados
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[, "Package"])]
if(length(new.packages)) install.packages(new.packages)

```

En esa misma carpeta **code** se encuentra el archivo **tema-ggplot-histograma.R** que contiene todos los códigos a utilizar en la primera clase.

En la carpeta **html** se encuentra el contenido del material de la primera clase **tema-ggplot-histograma.html** en formato de página web con la visualización de los gráficos a practicar.