

Instituto Tecnológico de Costa Rica

TEC

Tecnológico
de Costa Rica

Proyecto I

Profesor:
Armando Arce Orozco

Grupo 1
Integrantes:

Calero Rios Manuel De Jesus
Quirós Gómez Harlen Daniel

Temática de proyecto:
Análisis exploratorio de datos

Periodo:
I Semestre 2024

Análisis exploratorio de datos

Introducción

Para el presente proyecto se planea utilizar la herramienta de RStudio para la elaboración de gráficas, pero además de ello se hará uso de un archivo Markdown para poder administrar todas las gráficas y datos dentro de un sitio web. Para el desarrollo del sitio web, gracias a diversas investigaciones se pudo hallar la manera de vincular RStudio con GIT, siendo este último el comunicador con GitHub el cual es el sitio público dónde se hospedará el código fuente junto las imágenes y los html.

En nuestro caso, elegimos información acerca del COVID-19 del año 2020 al 2023, dicha información resultará útil para el desarrollo de gráficas con RStudio ya que constituye un grán grupo de información, incluso en diversos sectores, ya que se pueden realizar estudios generales a nivel de país o provincia, incluso casos más concretos como lo son localidades de cantones o provincias o grupos diversos, como hombres y mujeres. En este caso se usará en su gran mayoría la información relevante de las provincias y uno que otro caso de grupos como hombres y mujeres.

Descripción del problema

El propósito de este proyecto es realizar un análisis exploratorio (EDA) de un conjunto de datos de interés. Para ello se deberá publicar una página Web en donde se describa el EDA realizado, incorporando una serie de visualizaciones que describen el comportamiento de los datos y los patrones identificados en los mismos.

Definición de fuentes de datos

La fuente de datos a utilizar consistirá en información relacionada a los casos de COVID-19 que afectaron al país, dichos datos son contenidos dentro de un documento de excel el cual tiene información precisa proporcionada por una entidades públicas y recopilada en una sola.

Descripción detallada y explicación de las secciones principales del documento.

El documento consta de una única parte, la cual es la de las gráficas, aquí se mostrarán cada una de las diversas gráficas solicitadas. Usando herramientas de R en ggplot2 se implementaron algunas ayudas que permitirán al usuario navegar dentro de la información de manera más certeza ya que si presiona sitios puede observar datos en específico o incluso acercar o alejar la gráfica.

Conclusiones

Gracias al estudio se pudo aprender más acerca de las herramientas que proporciona R a las gráficas, llegando incluso a facilitar el diseño de páginas web en sitios como RStudio. Gracias a ellos se pudo observar los diversos comportamientos del COVID-19 en Costa Rica, ya que las gráficas proporcionaban distintos puntos de vista al tema en general, incluso facilitando su observabilidad ya que había herramientas que permitían que la navegación en los datos fuera mucho más sencilla.

Link

<https://manollito.github.io/index.html>

*Nota, por alguna razón el index dentro del github no redirecciona bien, pero si corre el HTML de index por aparte si aparece bien