

Análisis dinámico a la Situación Actual de Desminado Humanitario en Colombia haciendo uso de la programación literaria y el paquete Knitr

Wilmar Diaz Rodriguez
Ingeniero de Electrónico
Universidad Distrital Francisco José de Caldas
Bogotá D.C., Colombia
Email: wilmardiazr@gmail.com

Resumen—Hoy en día se muchas empresas pretenden hacer uso de los datos abiertos, día a día se encuentra mas información en línea confiable como por ejemplo el catálogo de datos abiertos del gobierno de Colombia. En esta fuente se pueden encontrar varios datasets recopilados de diferentes fuentes como por ejemplo Central Hidroeléctrica de Caldas, Colpensiones, Departamento administrativo de la Presidencia entre otros. En el caso particular el artículo pretende mostrar como haciendo uso de la programación literaria se pueden hacer análisis de datos en línea sobre dichos dataset. Para nuestro artículo los entregados por el Departamento administrativo de la Presidencia el tema que hace referencia a la Situación de Desminado Humanitario en Colombia. Así mismo se pretende mostrar la forma de realizar experimentación reproducible haciendo uso de los programas de latex, R, y el paquete de knitr

Keywords—Minas, Data Mining, Desminado, Dataset, Modelo predictivo.

I. INTRODUCCIÓN

El análisis de Desminado en Colombia es tema relevante para la actualidad ya que podría mostrar como al acercarse un acuerdo de Paz se ha aumentado o disminuido el Desminado en Colombia así como poder determinar cuales son los departamentos que mas se han visto afectados entre otros.

II. METODOLOGIA

Para el desarrollo de este documento se plantean los siguientes hitos [1].

II-A. Obtención de datos de Catálogo abierto para programación literaria y Variables del Catálogo

- **Como Acceder a datos Abiertos:** Lo primero que se realizo fue investigar como poder obtener los datos para esto se ingreso a la sección de desarrolladores de la pagina de datos abiertos de Colombia y allí se encuentra la información detallada que muestra el uso de dichos datos, ¿Qué son datos abiertos?, ¿Que características tienen los datos?, se puede ingresar al siguiente link [Datos Abiertos Colombia](#) , existe una guía de ejemplo don se muestra como se puede consultar de forma sencilla los datos en el siguiente link [Guía Desarrolladores](#) .

- **Como integrar los datos que están publicados con la programación Literaria:** Para poder hacer un análisis reproducible y que adicionalmente permita cambiar el conjunto de datos y observar el nuevo resultado se requiere poder leer los datos del catalogo en línea, luego procesarlos para esto y tal como se muestra en la guía para desarrolladores se puede consultar los datos usando la cadena del catalogo y añadiendo el formato requerido, existen dos opciones una es un formato json y la otra es un formato atom. Para nuestro analisis se utilizo el formato Json el cual implica colocar al final del catalogo el siguiente texto `¿$format=json`
- **Como usarlos con knitr :** Para poder integrar dichos datos con programación literaria se requiere usar un paquete que permita leer los datos en formato JSON y poderlos tratar con los chunks de código se uso el siguiente paquete `”jsonlite”` el cual se instala con el comando `install.packages(’jsonlite’, dependencies = TRUE)”` luego de esto se observaron problemas al cargar directo desde la web los datos por que le sobra un identificador por este motivo se necesito descargarlos en una carpeta local sin embargo la línea usada para la carga de estos datos es la siguiente
 - `library(jsonlite)`
 - `json_file <- ”http://servicedatosabiertoscolombia.cloudapp.net/v1/Departamento_Administrativo_de_la_Presidencia/eventosminasantipersonal?format=json”`
 - `fromJSON(json_file)`
 - `json_data1 <- fromJSON(json_file, flatten = TRUE)`

Finalmente se determino que se requería hacer referencia a los datos anteponiendo una `d` que viene en el formato que entrega la pagina de Datos Abiertos de Colombia y se podrían leer en línea

- **Datos e identificación de variables:** En particular el set de datos que se tomo es el que esta relacionado con el Desminado Humanitario en Colombia en el hipervínculo se puede acceder en el siguiente link [Datos Desminado Humanitario Colombia](#) de igual forma se consulto la metadata de dicho catalogo

la cual puede ser vista en el siguiente enlace [Metadatos Desminado Humanitario Colombia](#) y luego dando click en la opción Ficha Técnica de la pagina allí se muestra que

- **Objetivos:** Los objetivos están enfocados a tener datos que puedan ser obtenidos de forma dinamica y que muestren la situación actual de Desminado Humanitario ya que estos datos se van alimentando mes a mes por el gobierno Colombiano tal como se indica en la ficha técnica y se podría simplemente volver a compilar el análisis obteniendo los resultados más recientes para el momento de realizarlo.
 - Identificar cual departamento ha sido en el que se ha podido despejar mas área minada
 - Saber la media de área minada despejada en Colombia con los datos proporcionados.
 - Obtener por año cual de los departamentos ha sido el que mas área se ha podido despejar
 - Determinar si en los meses corridos del año 2016 con respecto al anterior existe aumento de área despejada con el posible acuerdo de Paz

III. PREGUNTAS DE INVESTIGACIÓN

Las preguntas de investigación serán enfocadas a resolver los objetivos planteados

III-A. Preguntas para análisis dinámico de Desminado Humanitario en Colombia

Para este artículo se pretende solucionar las preguntas descriptivas de forma dinámica:

- ¿Cuál es la Media de área despejada para los años 2014, 2015, 2016?
- ¿Qué departamento tiene la mayor área despejada de Minas en el año 2016?
- ¿Qué departamento tiene la mayor área despejada de Minas en el año 2015?
- ¿Qué departamento tiene la mayor área despejada de Minas en el año 2014?
- ¿Determinar si el acuerdo de paz ha permitido aumentar el área media despejada ?
- ¿Determinar si el acuerdo de paz ha permitido aumentar el área media despejada ?

IV. SOLUCIÓN A PREGUNTAS

El análisis dinámico de datos en la investigación reproducible debe permitir, reproducir el experimento propuesto para nuestro caso obtener el resultado de las preguntas de investigación de manera dinámica por tal razón de cambiar un dato en el dataset se va a ver reflejado en las respuestas del documento esto generalmente ocurrirá mes a mes.

En este momento hay que tener varias consideraciones, se requiere hacer referencia a los datos anteponiendo la letra d ya que así está construido el json que entrega

la pagina de datos abiertos colombianos como ejemplo `”json_data1dareadespejada[1]”` esto mostraria el valor de la `areadespejada` `—as.numeric(json_data1dareadespejada)”` lo que realiza esta instrucción

Respuesta dinámica de las preguntas planteadas:

En el siguiente apartado se muestran los datos estadísticos del área despejada así como el total de artefactos destruidos.

Tabla I: Total de Artefactos y Área Despejada

Statistic	N	Mean	St. Dev.	Min	Max
totalartefactos	530	9.872	44.424	0	476
totalarea	530	4,868.119	6,899.598	0	54,370

- ¿Cuál es la Media de área despejada para los años 2014, 2015, 2016?

Para poder manejar estos datos de forma dinámica lo que se debe realizar es una selección específica de datos para este caso se realiza de la siguiente manera. vamos a verlo con el año 2016 datos 2016;- `subset(json_data1$d, json_data1$d$ano == 2016)` de esta manera se seleccionan los datos que cumplan con el

Tabla II: Total de Artefactos y Área Despejada para el año 2016

Statistic	N	Mean	St. Dev.	Min	Max
totalartefactos	10	0.300	0.675	0	2
totalarea	10	978.600	1,359.855	0	4,179

Tabla III: Total de Artefactos y Área Despejada para el año 2015

Statistic	N	Mean	St. Dev.	Min	Max
totalartefactos	97	2.330	5.554	0	37
totalarea	97	3,970.443	5,994.439	0	28,150

Tabla IV: Total de Artefactos y Área Despejada para el año 2014

Statistic	N	Mean	St. Dev.	Min	Max
totalartefactos	98	2.306	4.368	0	20
totalarea	98	5,612.071	9,010.254	0	43,194

Para solucionar las preguntas que siguen a continuación se usaron los subconjuntos de datos para poder obtener los resultados requeridos de forma dinámica así como algunas funciones propias de R

¿Qué departamentos tiene la mayor y menor área despejada de Minas en el año 2016?

Los departamentos que tiene la mayor área despejada para 2016 son:

[1]. ^NTIOQUIA”

Los departamentos que tiene la menor área despejada para 2016 son:

[1] .^ANTIOQUIAANTIOQUIASANTANDERSANTANDER TOLIMA”

- ¿Qué departamentos tiene la mayor y menor área despejada de Minas en el año 2015?

Los departamentos que tiene la mayor área despejada para 2015 son:

[1] .^ANTIOQUIA”

Los departamentos que tiene la menor área despejada para 2015 son:

[1] .^ANTIOQUIAANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIA”[6] .^ANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIAANTIOQUIA”[11] .^ANTIOQUIAANTIOQUIA ANTIOQUIAANTIOQUIAANTIOQUIA”[16] .^ANTIOQUIA ANTIOQUIAANTIOQUIAANTIOQUIAANTIOQUIA”[21] .^ANTIOQUIAANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIA”[26] .^ANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIABOLIVAR”[31] ”BOLIVARCALDAS SANTANDERSANTANDERSANTANDER”[36] ”SANTANDERSANTANDERSANTANDER”

- ¿Qué departamentos tiene la mayor y menor área despejada de Minas en el año 2014?

Los departamentos que tiene la mayor área despejada para 2014 son:

[1] ”BOLIVAR”

Los departamentos que tiene la menor área despejada para 2014 son:

[1] .^ANTIOQUIAANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIA”[6] .^ANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIAANTIOQUIA”[11] .^ANTIOQUIAANTIOQUIA ANTIOQUIAANTIOQUIAANTIOQUIA”[16] .^ANTIOQUIA ANTIOQUIAANTIOQUIAANTIOQUIAANTIOQUIA”[21] .^ANTIOQUIAANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIA”[26] .^ANTIOQUIAANTIOQUIAANTIOQUIA ANTIOQUIAANTIOQUIA”[31] ”BOLIVARBOLIVAR BOLIVARCALDASCALDAS”[36] ”METASANTANDER SANTANDERSANTANDERSANTANDER”[41] ”SANTANDERSANTANDER”

- ¿Determinar si el acuerdo de paz ha permitido aumentar el área media despejada ? en conclusión no se ha visto aumento de desminado en colombia

- ¿Determinar si el acuerdo de paz ha permitido aumentar el área media despejada ? No

REFERENCIAS

- [1] Cornel University library The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects [URL1](#)
- [2] Situación de Desminado Humanitario en Colombia [URL2](#)
- [3] New package: jsonlite. A smart(er) JSON encoder/decoder. [URL3](#)
- [4] LINEAMIENTOS PARA LA IMPLEMENTACIÓN DE DATOS ABIERTOS EN COLOMBIA [URL4](#)