

2020

Gestión de Datos Proyecto Final

Programa Profesional

Analista Científico de Datos.

DAT002 – Gestión de Datos

para la Toma de Decisiones.

Prof.: José A. Cabezas Jaikel, MTBD.

Autor: Ing. Diego Araya Chaves

Versión: 1.0

Tabla de contenidos

[Tabla de Ilustraciones 1](#_Toc40906890)

[Introducción 2](#_Toc40906891)

[Descripción de las fuentes de datos 3](#_Toc40906892)

[Descripción destino de datos 5](#_Toc40906893)

[Bibliografía. 6](#_Toc40906894)

# Tabla de Ilustraciones

[Ilustración 1:Diccionario de datos de CRI\_CSV\_GENERAL 3](#_Toc40906815)

[Ilustración 2: Diccionario de datos de WHO-COVID-19-global-data 4](#_Toc40906816)

[Ilustración 3: Diccionario de datos de covid-19-sitrep-114 4](#_Toc40906817)

[Ilustración 4: Diccionario de datos de Capacidad América Latina covid-19 - BBC News Mundo 5](#_Toc40906818)

# Introducción

El siguiente documento presenta el proyecto final del curso de Gestión de Datos para la Toma de Decisiones, el cual consiste en la elección de un tema en el que se quiera profundizar y realizar un análisis con las herramientas y metodologías que se aprendieron el curso, para posterior a eso presentar un documento formal de los resultados obtenidos.

Para la elaboración de este proyecto se procede a la obtención de datos de varias fuentes con un mismo tema relacionado, estos datos fueron obtenidos de diferentes fuentes públicas de internet y luego condensados en un solo repositorio para el su respectivo análisis, transformación, limpieza y presentación con aplicaciones de visualización de datos que se han visto en el curso.

# Descripción de las fuentes de datos

**Nombre de la fuente de datos:** CRI\_CSV\_GENERAL.csv

**Descripción del dataset:** Se trata de un archivo csv que contiene los datos generales del covid-19 en Costa Rica, elaborado por geovisión de las Universidad Estatal a Distancia (UNED), esta fuente de datos se actualiza cada día, por lo que para la presentación del proyecto es probable que los datos estén desactualizados, pero se pueden actualizar desde el link que se presenta.

**Link de la fuente:** <http://geovision.uned.ac.cr/oges/archivos_covid/05_19/05_19_CSV_GENERAL.csv>

**Diccionario de datos:**

[](https://github.com/diegoaraya/ProyectoGentionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx)

Ilustración 1:Diccionario de datos de CRI\_CSV\_GENERAL

Para poder tener acceso al diccionario de datos completo lo puede encontrar en:[**https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx**](https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx)

**Nombre de la fuente de datos:** WHO-COVID-19-global-data.csv

**Descripción del dataset:** Es un archivo csv que contiene los datos mundiales sobre los casos generales de contagio y defunción por corona virus, en el cual se presentan agrupaciones por región según la Organización mundial de la salud (WHO), los datos además se presentan por país. Este archivo es tomado de la WHO.

**Link de la fuente:** <https://covid19.who.int/>

**Diccionario de datos:**



Ilustración 2: Diccionario de datos de WHO-COVID-19-global-data

Para poder tener acceso al diccionario de datos completo lo puede encontrar en:[**https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx**](https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx)

**Nombre de la fuente de datos:** covid-19-sitrep-114.pdf

**Descripción del dataset:** Es un archivo pdf que contiene los datos mundiales actuales de casos de contagio de corona virus, así como de las defunciones. Se presentan también los datos totalizados acumulados de estos, pero lo que más destaca en el archivo, es que se prestan por clasificación de contagio, la fuente de este archivo es la organización mundial de la salud.

**Link de la fuente:** <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>

**Diccionario de datos:**



Ilustración 3: Diccionario de datos de covid-19-sitrep-114

Para poder tener acceso al diccionario de datos completo lo puede encontrar en:[**https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx**](https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx)

**Nombre de la fuente de datos:** Capacidad América Latina covid-19 - BBC News Mundo.pdf

**Descripción del dataset:** Es un archivo pdf el cual se trata de un reportaje hecho por la cadena de noticias BBC, el cual expone datos de la capacidad de Latinoamérica para combatir al covid-19 de acuerdo a las cifras que se tienen de inversión en salud, equidad en el acceso a la salud y capacidad instalada en los diferentes países de la región

**Link de la fuente:** <https://www.bbc.com/mundo/noticias-america-latina-51916767>

**Diccionario de datos:**



Ilustración 4: Diccionario de datos de Capacidad América Latina covid-19 - BBC News Mundo

Para poder tener acceso al diccionario de datos completo lo puede encontrar en:[**https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx**](https://github.com/diegoaraya/ProyectoGestionDatos/blob/master/Meta%20datos%20proyecto%20final.xlsx)

# Descripción destino de datos

El repositorio de datos para este proyecto se encuentra en GitHub, la cual es una compañía sin fines de lucro que ofrece un servicio de hosting de repositorios almacenados en la nube que ayuda a los desarrolladores a almacenar y administrar su código, al igual que llevar un registro y control de cualquier cambio sobre este código.

El para acceder al repositorio es el siguiente:

<https://github.com/diegoaraya/ProyectoGestionDatos>

# Bibliografía.

Organización Mundial de la Salud (WHO). (20 de Mayo de 2020). Obtenido de WHO Coronavirus Disease (COVID-19) Dashboard: https://covid19.who.int/

Organización Mundial de la Salud (WHO). (Mayo de 2020). *Coronavirus disease (COVID-2019) situation reports.* Recuperado el 20 de Mayo de 2020, de https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

Universidad Estatal a Distancia. (2015). *Situación Nacional COVID-19*. Obtenido de Geovisión: http://geovision.uned.ac.cr/oges/

Wallace, A. (23 de Marzo de 2020). *BBC News Mundo*. Obtenido de Coronavirus: qué capacidad tienen realmente los países de América Latina para hacer frente a la epidemia de covid-19.: https://www.bbc.com/mundo/noticias-america-latina-51916767

Wikipedia. (14 de Diciembre de 2018). *GitHub*. Obtenido de https://es.wikipedia.org/wiki/GitHub