

**Data Science**

Modulo: Procesamiento de datos con Python Santander

**Equipo 17**

Bolaños Pardo Diego

Carrillo Rivera Rafael Sebastián

Espinosa Ávila Ana Yessica

Pérez Sánchez Diana Alejandra

Silva Reyes Luis Roberto

Velázquez Moreno Salvador

**Proyecto final**

COVID-19 Progreso de vacunación mundial

Profesor: Rafael Arias González

Viernes 13 de agosto de 2020

**COVID-19 Progreso de vacunación mundial**

1. **INTRODUCCIÓN**

Hace más de un año que se declaró la pandemia por coronavirus SARS-CoV-2 la cual ha provocado numerosos casos y muertes en todo el mundo. La pandemia está provocando en el mundo una grave crisis sanitaria, social y económica. Ha mostrado la debilidad de los sistemas sanitarios y de los organismos de salud pública para poder dar una respuesta rápida y adecuada a una situación inesperada (Casas & Mena, 2021).

Al día 12 de agosto de 2021, la Universidad Johns Hopkins contabiliza 204,688,069 casos, 4,323,993 muertes y 4.532.619.735 vacunas aplicadas. Las cifras anteriores indican que la tasa de letalidad de la enfermedad es de 2.11%

Nuestra vida ha cambiado tal y como la conocíamos hasta entonces. Una de las esperanzas que tenemos actualmente es poder disponer de vacunas seguras y efectivas para administrar a la población. Con estas vacunas se pretende conseguir la inmunidad colectiva que permita romper la cadena de transmisión. Desde la emergencia de esta enfermedad más de 250 grupos en el mundo están trabajando en el desarrollo de estas vacunas (Casas & Mena, 2021).

1. **PLANTEAMIENTO DEL PROBLEMA**

La rapidez con la que esta enfermedad se ha transmitido ha traído consigo la imperiosa necesidad de encontrar vacunas eficientes y esparcirlas en todo el mundo con la mayor rapidez posible. Es ahí donde el estudio de los datos puede ser de gran ayuda, ya que nos provee de información de gran utilidad para conocer el número de vacunas aplicadas a nivel mundial, además de conocer también qué vacunas son las que se han aplicado en cada país. El análisis de estos datos es primordial para la toma de decisiones y tener una idea más clara y precisa del avance que puede tenerse en cuanto al progreso de vacunación mundial se refiere. Además de lo anterior, el avance en tiempo de respuesta en el procesamiento y análisis de estos datos nos permite generar los datos requeridos con la premura que se necesitan.

**2.1 Justificación**

Se tomó la decisión de tomar este tema ya que comprender el análisis sobre temas referentes a salud y su impacto a nivel mundial puede ser enriquecedor tanto en nuestro conocimiento como científicos de datos tanto como para poder comprender y poder plasmar lo que se puede obtener mediante el estudio y análisis del conjunto de datos obtenido.

De igual manera, creemos que poder obtener algunos indicadores sobre el tema de vacunación contra la COVID-19 pueden proveernos una idea acerca de qué clase de información de interés podríamos obtener al elaborar un análisis con mayor profundidad de la información de la base de datos que se utilizará, con el objetivo de comprender como las campañas de vacunación han impactado el avance de la pandemia.

En este caso, el conjunto de datos se obtuvo del sitio web de Kaggle, esto debido a que el conjunto de datos brindado cuenta con información de interés y cumple con los requisitos que la entrega del proyecto requiere, como lo son algunos registros con valores NaN y necesidad de limpieza del conjunto de datos. Se tomó el conjunto de datos en formato csv, y contaba con una breve descripción de cada uno de los campos contenidos, estos son:

• País: este es el país para el que se proporciona la información de vacunación;

• Código ISO del país: código ISO del país;

• Fecha: fecha para la entrada de datos; para algunas de las fechas solo tenemos las vacunas diarias, para otras, solo el total (acumulativo);

• Número total de vacunaciones: este es el número absoluto de inmunizaciones totales en el país;

• Número total de personas vacunadas: una persona, según el esquema de inmunización, recibirá una o más (normalmente 2) vacunas; en un momento determinado, el número de vacunaciones puede ser mayor que el número de personas;

• Número total de personas completamente vacunadas: este es el número de personas que recibieron el conjunto completo de inmunizaciones de acuerdo con el esquema de inmunización (normalmente 2); en un momento determinado, puede haber un cierto número de personas que recibieron una vacuna y otro número (menor) de personas que recibieron todas las vacunas del esquema;

• Vacunas diarias (crudas): para una determinada entrada de datos, el número de vacunaciones para esa fecha / país;

• Vacunas diarias: para una determinada entrada de datos, el número de vacunaciones para esa fecha / país;

• Total de vacunaciones por cien - relación (en porcentaje) entre el número de vacunaciones y la población total hasta la fecha en el país;

• Número total de personas vacunadas por cien - relación (en porcentaje) entre la población inmunizada y la población total hasta la fecha en el país;

• Número total de personas totalmente vacunadas por cien - relación (en porcentaje) entre la población totalmente inmunizada y la población total hasta la fecha en el país;

• Número de vacunaciones por día: número de vacunaciones diarias para ese día y país;

• Vacunas diarias por millón: relación (en ppm) entre el número de vacunaciones y la población total para la fecha actual en el país;

• Vacunas utilizadas en el país: número total de vacunas utilizadas en el país (hasta la fecha);

• Nombre de la fuente: fuente de la información (autoridad nacional, organización internacional, organización local, etc.);

• Sitio web de la fuente: sitio web de la fuente de información;

* 1. **Objetivos**

Obtener conjunto de datos del progreso de vacunación mundial, limpieza y transformación de datos.

Plantear preguntas referentes al tema de investigación, su implantación y su análisis, así como algunos indicadores que ayuden a visualizar qué clase de información se puede utilizar para posteriores análisis con mayor profundidad.

1. **INVESTIGACIÓN**

Los coronavirus son una extensa familia de virus que pueden causar enfermedades tanto en animales como en humanos. En los humanos, se sabe que varios coronavirus causan infecciones respiratorias que pueden ir desde el resfriado común hasta enfermedades más graves como el síndrome respiratorio de Oriente Medio (MERS) y el síndrome respiratorio agudo severo (SARS). La COVID-19 también conocida como enfermedad por nuevo coronavirus es causada por el coronavirus 2 del síndrome respiratorio agudo severo (SARS-CoV2), su forma es redonda u ovalada y a menudo polimórfica, tiene un diámetro de 60 a 140 nm, la proteína espiga que se encuentra en la superficie del virus y forma una estructura en forma de barra, es la estructura principal utilizada para la tipificación, la proteína de la nucleocápside encapsula el genoma viral y puede usarse como antígeno de diagnóstico. Tanto el nuevo virus como la enfermedad eran desconocidos antes de que estallara el brote en Wuhan (Pérez, Gómez & Dieguez, 2020).

Produce síntomas similares a los de la gripe, entre los que se incluyen fiebre, tos, disnea, mialgia y fatiga. También se ha observado la pérdida súbita del olfato y el gusto (sin que la mucosidad fuese la causa). En casos graves se caracteriza por producir neumonía, síndrome de dificultad respiratoria aguda, sepsis y choque séptico que conduce a la muerte. La Covid-19 se identificó por primera vez el 1 de diciembre de 2019 en la ciudad de Wuhan, capital de la provincia de Hubei, en la China central, cuando se reportó a un grupo de personas con neumonía de causa desconocida, vinculada principalmente a trabajadores del mercado mayorista de mariscos del sur de China de Wuhan. El número de casos aumentó rápidamente en el resto de Hubei y se propagó a otros territorios. La rápida expansión de la enfermedad hizo que la Organización Mundial de la Salud, el 30 de enero de 2020, la declarara una emergencia sanitaria de preocupación internacional, basándose en el impacto que el virus podría tener en países subdesarrollados con menos infraestructuras sanitarias y la reconociera como una pandemia el 11 de marzo. (Pérez, Gómez & Dieguez, 2020).

**3.1 Planteamiento de preguntas**

Las preguntas planteadas para obtener del conjunto de datos son las siguientes:

1. ¿En qué país el programa de vacunación está más avanzado?
2. ¿Dónde hay más personas vacunadas por día?
3. ¿Cuál es el rating de vacunas más difundidas a nivel mundial?
4. ¿Qué países reportan sus datos con mayor frecuencia?
5. ¿Cuáles son los sitios web que brindan mayor información?
6. **BUSQUEDA DE SOLUCIONES ANTERIORES**

Uno de los países que más ha llamado la atención en los últimos meses ha sido Reino Unido, debido a que enfrentó una tercera ola durante la primavera y al ser uno de los países con mayor avance de vacunación, se podría observar si dicha campaña de vacunación tendría un efecto en cuanto a las hospitalizaciones y los fallecimientos a causa de la COVID-19.

El *Public Health England* publica un reporte semanalmente sobre el avance de la vacunación en el Reino Unido y el impacto que ha tenido en la pandemia a nivel nacional. Dichos reportes detallan información acerca de la efectividad de las vacunas en prevenir enfermedad sintomática, hospitalizaciones, defunciones y transmisibilidad, así como un desglosamiento del avance de vacunación por diferentes grupos etarios.

Para los propósitos de nuestro problema, la parte más importante del reporte corresponde a las estimaciones sobre el número de infecciones y de fallecidos que se han prevenido como resultado de la campaña de vacunación en el Reino Unido, los cuales según el reporte son 22,057,000 infecciones y 60,000 fallecimientos prevenidos, a fecha del 23 de julio de 2021.

1. **CONCLUSIONES**

Sin lugar a dudas el avance tecnológico en conjunto con el uso del análisis y ciencia de datos nos permite obtener, procesar, limpiar, ordenar, analizar y presentar datos que pueden ser de gran utilidad en la toma de numerosas decisiones, en el caso de tema elegido, COVID-19, al ser una emergencia sanitaria, podemos comprobar que el alcance de estos datos puede ser incluso a nivel mundial con un gran nivel de desglose. Dado lo anterior, podemos observar el alcance que se tiene el análisis de datos de la mano con diversas áreas temáticas para brindar información con mayor precisión.

**BIBLIOGRAFÍA**

Casas I., Mena G. (2021). *La vacunación de la COVID-19*. Recuperado el 11 de agosto de 2021 de: <https://www.sciencedirect.com/sdfe/pdf/download/eid/1-s2.0-S0025775321001469/first-page-pdf>

Pérez M., Gómez J., Dieguez R. (2020). *Características clínico-epidemiológicas de la COVID-1.*  Recuperado de: <https://www.medigraphic.com/pdfs/revhabciemed/hcm-2020/hcm202e.pdf>

Public Health England. (2021). COVID-19 vaccine surveillance report - Week 31. <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1008919/Vaccine_surveillance_report_-_week_31.pdf>