

## **Tarea 05**

### **Nombres y código:**

Ruiz Ortiz Cesar Adrián 208020817

Padilla Martin Juan Pablo 217294261

Álvarez Gutiérrez David Alejandro 217294016

**Sección:** D01

**Materia:** MINERIA DE DATOS

**Profesor:** ISRAEL ROMAN GODINEZ

**CARRERA:** INGENIERIA EN INFORMATICA

**CICLO ESCOLAR:** 2020 A



## Descripción del problema

Buscar un problema de análisis y descubrimiento de datos, que le parezca interesante, en alguno de los repositorios vistos en clase, por ejemplo, Kaggle. Elaborar un documento con la información que se pide en la lista. Quizás no toda la información necesaria se podrá obtener del repositorio, si es necesario deberá consultar otras fuentes para completar la información de esta etapa.

- **Objetivos del negocio (coronavirus)**
  - **Antecedentes**

el 31 de diciembre de 2019, la OMS recibió una alerta sobre varios casos de neumonía en la ciudad de Wuhan, provincia de Hubei en China. El virus no coincidía con ningún otro virus conocido. Esto generó preocupación porque cuando un virus es nuevo, no sabemos cómo afecta a las personas.

- **Objetivo comercial o de investigación del proyecto**

Buscar información importante con la información que se conoce de los casos de contagio, muertes y personas recuperadas para ver las tendencias de la epidemia.

- **Criterios de rendimiento**

N/A

- **Valoración de la situación**

- **Inventario de recursos**

Conjunto de datos en archivo CSD

### Columnas:

- Sno - Número de serie
- Fecha - Fecha y hora de la observación en MM / DD / AAAA  
HH: MM: SS
- Provincia / Estado - Provincia o estado de la observación (podría estar vacío cuando falta)
- País - País de observación
- Última actualización: hora en UTC en la que se actualiza la fila para la provincia o país en cuestión. (No estandarizado actualmente. Así que límpiélos antes de usarlo)
- Confirmado - Número de casos confirmados
- Muertes - Número de muertes
- Recuperado - Número de casos recuperados

- **Requisitos, supuestos y restricciones:**

- supuestos

- 1.- casos de infección no reportados.

- 2.-actualización de los datos en tiempo real

- Requisitos

- N/A

- Restricciones

- N/A

- **Riesgos y contingencias**

Riesgos: difundir información falsa con el estudio que pueda malinterpretar la situación de la epidemia

Contingencias: usar el resultado para fines académicos

- **Terminología**

N/A

- **Análisis de costo/beneficio**

N/A

- **Objetivos de la minería de datos**

- **Metas de la minería de datos**

- 1.- Determinar Cambio de casos a lo largo del tiempo a nivel de país para determinar la gravedad de la epidemia

- 2.- Determinar Cambios en el número de casos afectados a lo largo del tiempo ver la evolución de la epidemia

- **Criterio de rendimiento**

- N/A

- **Plan de proyecto**

- **Plan de proyecto**

- N/A

- **Valoración de herramientas técnicas**

- N/A

**Link del sitio**

<https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset>

[CSD coronavirus:](#)

**Link del sitio 2**

<https://www.kaggle.com/hamdinizar/bitcoin>

[CSD bitcoin.](#)