

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO**  
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, IIMAS  
*Ciencia de Datos*  
**Visualización de la Información**

**EPIDEMIA DE ÉBOLA**  
**TAREA-4**

## 1. Introducción

La enfermedad por el virus del Ébola es grave y a menudo mortal causada por un virus. Los síntomas incluyen fiebre, diarrea, vómitos, sangrado y, con frecuencia, causa la muerte. Esta afección se descubrió en 1976 cerca del río Ébola en la República Democrática del Congo. Desde entonces, se han presentado varios brotes pequeños en África. El brote de la enfermedad del Ébola en África occidental, que inició en marzo de 2014, fue la epidemia viral hemorrágica más grande en la historia.

## 2. Actividades

Importe el conjunto de datos de casos de una epidemia de ébola simulada utilizando el archivo [`linelist\_cleaned.csv`](#)

Como puede observar, entre algunas características de la información mostrada podemos encontrar lo siguiente:

<b>case_id</b>	Número de identificación del caso
<b>date_onset</b>	Inicio de los síntomas
<b>date_hospitalisation</b>	Fecha de hospitalización
<b>date_outcome</b>	Fecha de salida del paciente
<b>outcome</b>	Condición de salida del paciente: recuperado, deceso, no se sabe
<b>gender</b>	Género: masculino, femenino, se desconoce
<b>age</b>	Edad
<b>age_cat</b>	Rango de categoría de edad
<b>hospital</b>	Lugar donde fue internado el paciente
<b>wt_kg</b>	Peso corporal del paciente
<b>ht_cm</b>	Altura del paciente en centímetros
<b>fever</b>	Síntomas de fiebre
<b>chills</b>	Síntomas de escalofríos
<b>cough</b>	Síntomas de tos
<b>aches</b>	Síntomas de dolores
<b>vomit</b>	Síntomas de vomito
<b>temp</b>	Temperatura corporal
<b>time_admission</b>	Hora de ingreso al hospital
<b>bmi</b>	Índice de masa corporal

De acuerdo con la información anterior y con los datos de la importación del archivo, realice las siguientes actividades:

- Elabore una gráfica (la que considere conveniente) que muestre una distribución sobre la cantidad de casos de Ébola en relación con la edad de los pacientes (**age**).
- Elabore una **gráfica de dispersión** que indique el peso de los pacientes (**wt\_kg**) en relación con la edad de los pacientes (**age**) y el género (**gender**). Esto es, toda la información solicitada debe observarse en una sola gráfica.

Una **curva epidémica** (también conocida como "epicurva") es un gráfico epidemiológico central que se usa típicamente para visualizar el patrón temporal de aparición de la enfermedad entre un grupo o una epidemia de casos. El análisis de la epicurva puede revelar tendencias temporales, valores atípicos, la magnitud del brote, el período de tiempo más probable de exposición, intervalos de tiempo entre generaciones de casos e incluso puede ayudar a identificar el modo de transmisión de una enfermedad no identificada (por ejemplo, fuente puntual, transmisión continua, fuente común, propagación de persona a persona).

- Elabore una gráfica que indique el número de casos agregados semanales por hospital.
- Elabore una gráfica de barras donde se muestre el total de hospitalizaciones por grupo de edad (**age\_cat**). Todos los grupos de edad deben observarse en una sola grafica.
- Elabore una gráfica de barras que indique el número de casos semanales por salida del paciente (**outcome**): recuperado, deceso, no se sabe. Toda la información se requiere en una sola grafica.

### 3. Conclusiones

De acuerdo con todas sus gráficas anteriores y resultados obtenidos, describa las conclusiones que se pueden obtener a partir del conjunto de datos de casos de una epidemia de ébola simulada.