

Ciencia de Datos

Tarea 1

Para entregar el 12 de febrero de 2020

1. Concluye el ejercicio de análisis en los datos `olive` según los objetivos que discutimos en clase, es decir, tratar de identificar las regiones de los aceites y también las áreas dentro de cada región.
2. Considera los datos que se encuentran en el archivo `suicide_data.csv`, que compila información relacionada con suicidios en 101 países en diferentes años e incluye también algunos indicadores de desarrollo, como el Producto Interno Bruto (GPD) y per capita, Índice de Desarrollo Humano (HDI), entre otra información.
 - a) ¿Qué preguntas de interés puedes formular para analizar éste fenómeno en base a estos datos?
 - b) Realiza un análisis exploratorio y descriptivo de los datos con las herramientas que consideres apropiadas. ¿Qué estructuras o correlaciones encuentras? Comenta tus hallazgos.
 - c) Trata de responder las preguntas que formulaste en el primer inciso. ¿Puedes obtener conclusiones interesantes?
 - d) Considera solamente los datos de México y Estados Unidos. Realiza los incisos 2b y 2c por separado y compara los resultados. ¿Qué puedes concluir?
 - e) ¿Qué otra información te gustaría tener disponible para analizar éste fenómeno¹?

El resultado, para ambos ejercicios, debe ser un reporte corto donde describas los pasos que seguiste y las conclusiones a las que llegas, incluyendo las gráficas más ilustrativas o informativas que tu consideres.

Nota: La base de datos de suicidios contiene datos faltantes. Indica en el reporte cómo los manejaste.

¹Quizá pueda servirte el documento disponible en https://www.who.int/mental_health/suicide-prevention/national_strategies_2019/en/ para generar ideas.