

Estructura de Datos

Nombre: _____

1. Pensamiento de ingeniero

(50%)

La organización Netflix necesita saber más de sus usuarios y la aceptación de sus producciones. Por lo que ha decidido realizar un algoritmo que permita saber la percepción de aceptación de la producción tipo serie (película segmentada por episodios) entre sus usuarios.

Netflix tiene la siguiente información de interacción de los usuarios con su plataforma (independiente del medio, sea televisor, computador, celular y videoconsola):

- Series buscadas y a cuales les dio click para ver (no significa que la vio completa)
- Series seleccionadas de las opciones escogidas por los algoritmos de Netflix
- Frecuencia de reproducción de los episodios
- Orden de reproducción de los episodios
- Tiempo entre episodios
- Abandono o reproducción total de la serie
- Números de reproducciones de segmentos del episodio
- Fecha de las acciones realizadas por el cliente

Cada serie tiene:

- Género de la serie
- País de producción
- Año de edición
- Director

Necesitamos lo siguiente:

(para entregar)

1. Un diseño de la estructura de datos que se empleará para dar respuesta a la tendencia del cliente (gusto de consumo) (valor:30%)
2. Si las tendencias tienen influencia de similares, es decir, **personas con un comportamiento similar, sus gustos son similares**, ¿Cómo conformaría el proceso para saber cual es el próximo episodio que a un cliente X le gustará? Debe entregar los esquemas (el para qué) de técnicas propuestas de ordenamiento, búsqueda, etc. en la solución del problema (valor: 50%)
3. ¿Cómo define (con los datos) el adjetivo “similar”? No es una definición lingüística de la palabra similar, es una descripción por operaciones informáticas el concepto similar. (valor: 20%)

2. Codificación

(50%)

En MEData están los datos de la Proyección de Población Medellín 2017

(<https://drive.google.com/file/d/1g4Qg4NOM54ccyvY1Yx3UGByDWFvHeEEk/view?usp=sharing>)

Realice un programa en el lenguaje de programación de su preferencia para:

1. Por comuna, presentar los órdenes de:
 - a. Densidad de la diferencia entre hombres y mujeres por comuna, presentado la información de mayor diferencia a menor diferencia. Pueden pensar en el caso de la curva de ballena - acumulado.
 - b. Comuna según el número de infantes (personas menor a 10 años). De mayor a menor, para programas de primera infancia.
 - c. Comuna según mujeres jóvenes (mujeres entre 10 años a 24 años). De mayor a menor para programas de educación sexual.
 - d. Comuna según adultos mayores (personas mayores a 55 años). De mayor a menor para programas de tercera edad.

Es importante que presente un análisis de los datos observados.

Cada punto tiene un valor de 25%