**MANEJO DE DATOS (2026-1)**

**MINERÍA DE DATOS (GUÍA PARA LA CREACION DEL REPORTE)**

|  |  |
| --- | --- |
| Equipo No. |  |

|  |
| --- |
| Nombres: |
|  |
|  |
|  |
|  |

**Indicaciones:**

* Programe los nodos en Python para realizar las indicaciones que se mencionan a continuación.
* Desarrolle el reporte de esta práctica en BIRT en formato PDF.
* Integre un anexo con la imagen del flujo de trabajo, la configuración de los nodos.

**Proceso:**

1. Cargue el archivo de Excel llamado: visita23.xlxs.
2. Considerando el diccionario de datos presentado en el documento llamado descripcion\_bd\_museos\_2024 en la página 113 del pdf, seleccione (si aún no están) los datos con número: 2 (Ent\_regis), 3 (Mes\_entrev), 5 (Sexo), 6 (Edad), 12 (Escolarida), 14 (Ocupacion), 18 (Estim\_fam), 37 (Medio\_tran), 38 (Tiempo\_tra), 61 (Opin\_expos), 63 (Dur\_vis\_h), 64 (Dur\_vis\_m) y 67 (Eval\_gral).
3. Realice el diccionario de datos de las variables indicadas.
4. Agregue las etiquetas de datos a las clases de TODAS las variables que las necesiten. El significado de los datos se encuentra en el mismo documento (descripcion\_bd\_museos\_2024). Cuidar la importación adecuada de las etiquetas de datos.
5. Crear la siguiente variable, la llama *tiempoTotal*, la cual será el tiempo que duró la visita de las personas en el museo, generada a partir de: Dur\_vis\_h y Dur\_vis\_m.
6. Asigne la región a cada una de las Entidades Federativas (2) en donde se realizó el registro de la encuesta (*Región*).
7. Responda las siguientes consultas. Agregar el enunciado de cada una de las consultas.
8. La edad (6) promedio de cada nivel de escolaridad (12).
9. Crear una tabla cruzada que indique el estímulo familiar (18) que recibieron hombres y mujeres (5) para asistir a los museos.
10. La calificación promedio (67) que asignó cada grupo según la opinión de la exposición (61).
11. Calcular el tiempo promedio en minutos (tiempoTotal) que duró la visita de los hombres y las mujeres en el museo.
12. La cantidad de hombres y de mujeres (5) que asistieron al museo cada mes (3).
13. La edad promedio (6) de los hombres y las mujeres (5) por nivel de escolaridad (12) en cada entidad federativa en donde se aplicó la encuesta (2).
14. Mostrar una tabla cruzada del medio de transporte (37) que utilizaron los hombres y las mujeres (5) para llegar al museo.
15. Mostrar la cantidad de personas por ocupación (14) que asistieron al museo cada mes (3) en cada entidad federativa (2).
16. Calcular la edad promedio (6) de los hombres y las mujeres en cada región (Región).
17. Genere dos clústeres considerando los atributos que sean de su interés. Describa los grupos (clústeres) generados y el despliegue de los mismos.
18. Genere dos ciclos con los siguientes enunciados:
    1. Crear un archivo de Excel con cada mes en una hoja. En cada hoja se debe indicar la cantidad de visitantes por entidad federativa de registro.
    2. Crear un archivo de Excel donde cada hoja índique la escolaridad de visitantes y en cada pestaña agregar la edad promedio y la calificación promedio que asignaron los hombres y mujeres.