Instituto

Politécnico

Nacional

Escuela Superior de Cómputo

MUSEOS

PRACTICA 1

Materia:

Data Mining

Grupo:

3CV19

Profesor:

Ocampo Botello Fabiola

Alumnos:

Castro Cruces Jorge Eduardo

Medina Granados Alan Alejandro

Guzmán Gutiérrez Manuel

Fecha:

Jueves, 23 de septiembre de 2021

**MINERÍA DE DATOS**

**PRÁCTICA DE TRATAMIENTO DE DATOS DE MUSEOS**

*Jueves 23 de septiembre de 2021*

|  |  |
| --- | --- |
| Grupo: | 3CV19 |

|  |
| --- |
| **Nombres:** |
| Castro Cruces Jorge Eduardo |
| Medina Granados Alan Alejandro |
| Guzmán Gutiérrez Manuel |

En esta práctica de tratamiento de datos se considera para fines educativos el estudio realizado por el INEGI.

Enlace de acceso:

<https://www.inegi.org.mx/programas/museos/#Datos_abiertos>

Descripción del proyecto:

Se trata de un estudio para producir y difundir información sobre las acciones museales en relación con las comunidades que los visitan a través de los registros de información que se generan en los establecimientos destinados para esos espacios culturales, lo que coadyuva a la prestación del servicio público de información, a la formulación de políticas culturales, a la toma de decisiones respecto a los museos y la gestión de los mismos.

Última actualización: 30 de junio de 2021

La descripción de los datos de los visitantes a los museos se encuentra en la página 11 del libro llamado Estadística de museos. Síntesis metodológica.

El conjunto de datos está compuesto por 70 columnas y 13158 registros.

Realizar las siguientes acciones:

1. **ELECCIÓN DE DATOS**

Elegir 20 atributos (los atributos compuestos cuentan como 1), debe elegir por lo menos un atributo compuesto. Crear un diccionario de datos con el siguiente formato

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nombre de variable** | **Tipo de dato** | **Dominio** | **Significado** | **Valores** |
| Sexo | Nominal  (String) | 1, 2 | Condición biológica que distingue a las mujeres y hombres | 1-Hombre  2-Mujer |
| Edad | Numérico | 12..99 | Años transcurridos entre la fecha de nacimiento de la persona y la fecha en la que se le encuesta. | 12...97- años  98- 98 y más años  99- no especifica |
| Lugar de registro | String | 1,2,3 | Referencia del registro. | 1-Entidad federativa  2-Municipio o Alcaldía  3-Localidad |
| Año de creación de la institución | date | DD-MM-YYYY | Cuando se creó la institución |  |
| Superficie del terreno | Float | >0 | Metros cuadrados del terreno |  |
| Señalización | String | 1,2,3,4 | Tipos de señales para los visitantes | 1-Orientadoras  2-Informativas  3-Identificativas  4-Reguladoras |
| Accesibilidad | String | 1,2,3,4 | Tipos de accesibilidad para los museos | 1-Auditiva  2-Motriz  3-Cognitiva  4-Visual  5-Otra |
| Página Web | Boolean | True/False | Si cuentan con una página web | 1.- True  2.- False |
| Tipo de establecimiento | String | 1,2,3,4,  5,6,7,8 | Se refiere a la categoría de museo | 1-Museo  2-Galería de arte  3-Jardín botánico  4-Jardín zoológico  5-Acuario  6-Planetario  7-Museo en zona arqueológica  8-Otro tipo |
| Tema principal | String | 1,2,3,4,  5,6,7,8 | Temática del museo. | 1-Arqueología  2-Arte  3-Paleontología  4-Historia  5-Industria  6-Ciencia  7-Tecnología  8-Ambiental/Ecología |
| Grupos de apoyo | String | 1,2,3,4,5 | Grupos que sirven de apoyo a los museos | 1-Sociedad de amigos  2-Fundación  3-Patronato  4-Fideicomiso  5-Otro |
| Entrada | String | 1,2,3 | Si es gratuita o tiene un costo | 1-Siempre es gratis  2-Sólo algunos días es gratis  3-Siempre tiene costo |
| Cobro a extranjeros | Boolean | True/False | Si se cobra a extranjeros | True/False |
| Cuota de ingreso | String | 1,2,3,4 | Tipos de cuota del museo | 1-Precio de entrada para un adulto  2-Precio de entrada para un extranjero  3-Visita nocturna  4-Visita con servicios integrados |
| Nacionalidad | String | 1,2 | Nacionalidad del visitante | 1-Mexicana  2-Extranjera |
| Frecuencia de visita | Compuesto(  Id:int  Numero:int  ) | 1,2,3 | Cuantas visitas lleva | 1-Primera visita  2-Visitas en el año  3-Años trascurridos de la última visita |
| Visita a otros lugares de exhibición | Compuesto(  visito:boolean  cuantos:int  ) |  | Sí se visito a otros lugares de la exhibición | Visito:True/False  cuantos>=0 |
| Planeación de la visita | String | 1,2 | Sí se planeo la visita o no | 1-Planeada  2-Espontanea |
| Tiempo de traslado | time | HH:mm:ss | Cuanto tardo en trasladarse |  |
| Tamaño del grupo | Int | >0 | Cuantos visitantes venían en su grupo | >0 |

1. **DESCRIPCIÓN DE SALIDAS**

Generar salidas a archivos mediante ciclos.

1. **GRÁFICAS Y TABLAS**

Gráfico, Gráfico circular

Descripción generada automáticamente

Figura 1.1 .- Visitantes por sexo.

Figura 1.2.- Grafica de dispersión de edades

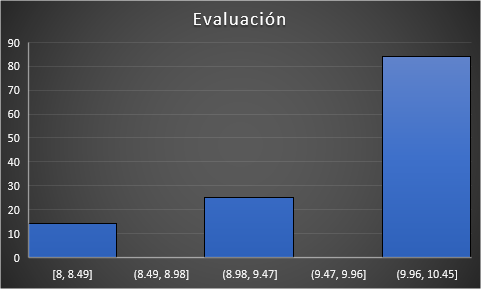


Figura 1.3.- Histograma de evaluación.

Figura 1.4.- Grafica de dispersión de mes entrevistado.

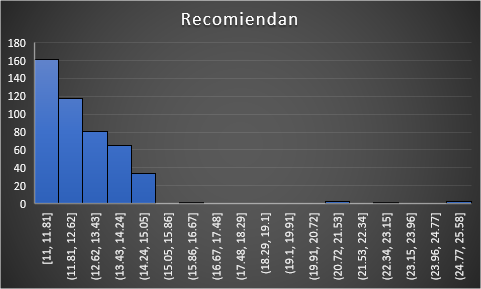


Figura 1.5.- Histograma de recomendaciones.

1. **TRATAMIENTO DE DATOS**

Elegir atributos para realizar tres de cada una de las siguientes acciones:

* 1. Imputación
  2. Normalización
  3. Transformación
  4. Creación de banderas
  5. Contenedores (*Binner*) de diverso tipo
  6. Reclasificación

1. **ANEXOS**

Agregar el flujo de trabajo y las ventanas de configuración.