**TIPO DE EVALUACIÓN: EXAMEN DIAGNÓSTICO.**

MATERIA/UNIDAD DE APRENDIZAJE: Minería de datos.

1. **Lee el caso de ejemplo y responde las preguntas correspondientes:**

La Coca Cola Company tiene 366,000 accionistas y efectúa una encuesta por medio de la selección aleatoria de 30 accionistas de cada uno de los 50 estados de Estados Unidos. Se registra el número de acciones de cada accionista de la muestra.

1. ¿Los valores obtenidos son discretos o continuos?

Discretos

1. Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) de los datos muestrales.

nominal

1. ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) se utiliza?

aleatorio

1. Si se calcula el número promedio (la media) de acciones. ¿El resultado es un estadístico o un parámetro?

estadístico

1. Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?
2. ¿Qué es lo que está incorrecto al evaluar la opinión del accionista enviado un cuestionario por correo que los accionistas podrían llenar y regresar por el mismo medio?

Datos falsos

1. **Describe con tus palabras para que se utiliza el análisis uni-variante de datos:**

**Para evaluar las características del conjunto de datos**

1. **Describe qué es la media, la mediana y la desviación estándar:**

**La media es un estimador de lo que mas se repite, la mediana es el dato que se encuentra a la mitad de todos los datos ordenados, la sd es la cantidad promedio de dispersión de los datos con respecto a la media**

1. **Con tus palabras explica qué es minería de datos y en donde puede ser aplicada:**

**Es extraer información de un conjunto de datos puede ser con el propósito de pronosticar cierta característica, sacar provecho de la información**

1. **Que lenguajes de programación dominas o en que porcentaje lo manejas:**

**SQL 60%**

**R 70%**

**Python 60%, C lo dominaba bien pero no lo he usado en años**

1. **Sabes manejar notebooks digitales como Jupyter, Spyder, Google Collab, etc. (Si la respuesta es otro, anota cuál):**

**No**

1. **¿Trabajas o has trabajado alguna vez en alguna empresa que maneje técnicas de ciencia de datos?**

**No☹**

1. **Has utilizado alguna herramienta de aprendizaje de máquina. Si la respuesta es afirmativa, anota cuál:**

No

1. **Relaciona los conceptos con sus gráficas correspondientes:**
2. Se utilizan para evaluar visualmente cómo se distribuyen los puntos de datos con respecto a su frecuencia.

Grafica de barras

1. Forma estandarizada de mostrar la distribución de datos basada en un resumen de cinco números ("mínimo", primer cuartil (Q1), mediana, tercer cuartil (Q3) y "máximo"). Puede informarle si sus datos son simétricos, qué tan estrechamente están agrupados y si sus datos están sesgados.

boxplot

1. Método para visualizar la distribución de datos numéricos de diferentes variables. Es similar al diagrama de caja pero con un diagrama rotado en cada lado, que brinda más información sobre la estimación de densidad en el eje *y.*
2. Visualización de datos bidimensionales que utiliza puntos para representar los valores obtenidos para dos variables diferentes: una trazada a lo largo del eje “xy” otra trazada a lo largo del eje y.

Diagrama de dispersión

1. Produce una matriz de relaciones entre cada variable en sus datos para un examen instantáneo de nuestros datos. También puede ser un gran punto de partida para determinar los tipos de análisis de regresión a utilizar.

|  |  |
| --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |

1. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos):

Solo me trabo un poco

1. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno)

No aparece la materia

1. Que sistema operativo tienes:

windows

1. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.

* Github. NO
* Jupyter Notebook NO
* Google Collab NO
* Python 3 SI hasta manejar csv
* Terminal de tu computadora. NO