

TIPO DE EVALUACIÓN: EXAMEN DIAGNÓSTICO.

MATERIA/UNIDAD DE APRENDIZAJE: Minería de datos.

I. 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.

- a) ¿Los valores obtenidos son discretos o continuos?

R: Discretos

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

R: Razón

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

R: Aleatorio

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

R: Estadístico

- e) Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?

R: El total de acciones que tienen los accionistas encuestados

- f) ¿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?

R: Podrían mentir, lo que provocaría que la información recolectada no sea confiable, al igual que el análisis de la misma.

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

R: Para ver el comportamiento de la variable analizada y tomar acciones que la pongan a "nuestro favor"

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

R: La "media" es el promedio de la muestra o población. La "mediana" es el dato medio al ordenar los datos de menor a mayor. La "desviación estándar" es qué tanto cambian las observaciones de la muestra o población

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

R: Es recopilar los datos necesarios, de una base de datos proporcionada, para analizarlos y encontrar patrones para hacer predicciones o simplemente hacer un resumen consistente de lo que se ha hecho hasta el momento del análisis.

Se puede usar en bancos, empresas de seguros, aplicaciones, plataformas de streaming, etc.

V. Que lenguajes de programación dominas o en qué porcentaje lo manejas:

R: Manejo la programación en Visual Basic, R (como al 20%) y llegué a manejar C++, pero ya no me acuerdo de cómo programar ahí.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

FACULTAD DE CIENCIAS FÍSICO-MATEMÁTICAS



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

R: No. Ni idea de cómo se usan

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

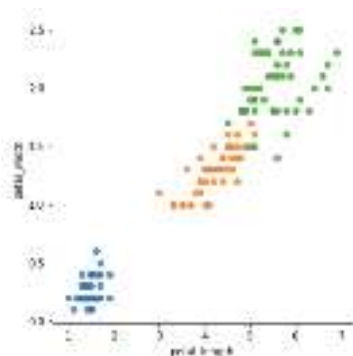
R: No

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

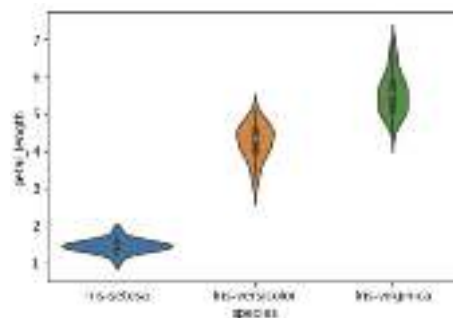
R: No

IX. Relaciona los conceptos con sus gráficas correspondientes:

- a) Se utilizan para evaluar visualmente cómo se distribuyen los puntos de datos con respecto a su frecuencia.
- b) 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.
- c) 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.
- d) 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.
- e) 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.



d)



c)

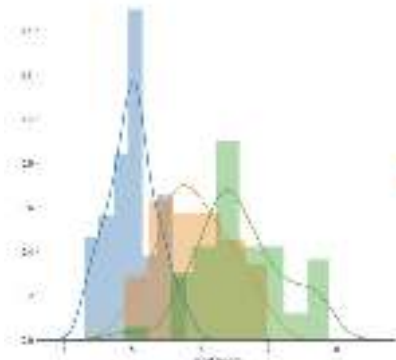


UANL

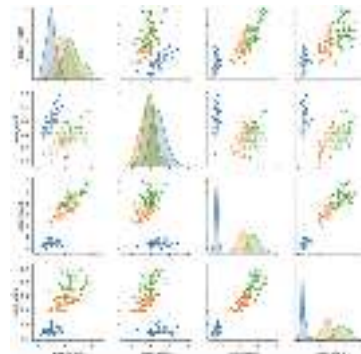
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

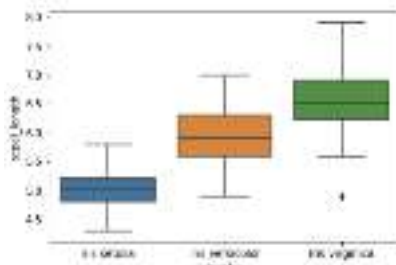
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



a)



e)



b)

- X. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos)
R: Para compartir pantalla es el mayor problema, ya que, a pesar de que puedo compartirla y los demás la pueden ver sin problema, yo creo que mi pantalla parpadea hasta llegar al punto en el que se pone totalmente negra, por lo que no puedo ver nada de lo que estoy haciendo.
Otro problema es que a veces se pone lenta mi computadora, por lo que tardo en activar el micrófono, pero el micrófono funciona bien una vez puedo activarlo
- XI. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno)
R: No. Funciona a la perfección
- XII. Que sistema operativo tienes:
R: Windows 10
- XIII. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.
- Github. **R:** Sí. Solo lo que vimos en clase, pero se ve muy intuitiva
 - Jupyter Notebook **R:** No
 - Google Collab **R:** No
 - Python 3 **R:** Sé que existe, pero no tengo idea de cómo se usa
 - Terminal de tu computadora. **R:** Lo llegué a usar un poco para unas cosas, pero no sé cómo se usa. Solo seguí instrucciones de internet.