



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

TIPO DE EVALUACIÓN: EXAMEN DIAGNÓSTICO.

MATERIA/UNIDAD DE APRENDIZAJE: Minería de datos.

Isabel Alejandra Rangel Vallejo

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?
discretos
- b) Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) de los datos muestrales.
nominal
- c) ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) se utiliza?
aleatorio
- d) Si se calcula el número promedio (la media) de acciones. ¿El resultado es un estadístico o un parámetro?
Estadístico
- e) Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?
La fecha en que se adquirió las acciones
- 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?
Que los entrevistados deciden dejar para el rato contestar el cuestionario y después no lo contestan

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

Se utiliza para describir los datos

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

La media es el promedio de los datos, la mediana es el dato que más se repite y la desviación estándar es la varianza de los datos

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

Es un campo de estudio donde recaba datos y los analiza para descubrir algún patrón

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



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

FACULTAD DE CIENCIAS FISICO MATEMÁTICAS



R 30% y visual basic 70%

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

He trabajado en Jupyter y google collab

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

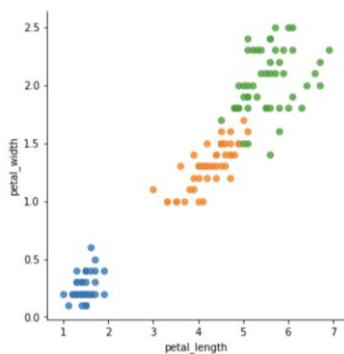
No

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

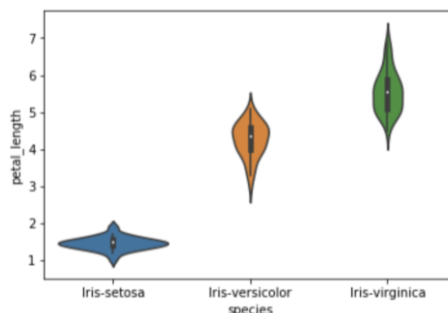
No

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

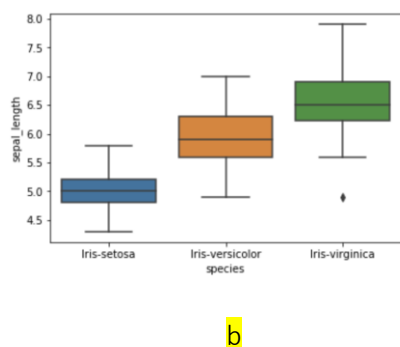
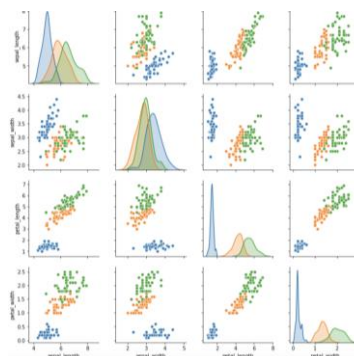
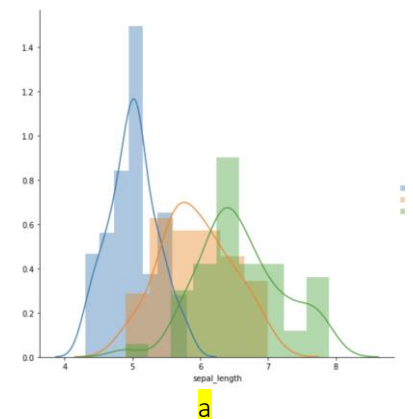


UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

FACULTAD DE CIENCIAS FISICO MATEMÁTICAS



- X. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos)
No
- XI. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno)
No
- XII. Que sistema operativo tienes: Windows 10
- XIII. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.
- Github. 10%
 - Jupyter Notebook 30%
 - Google Collab 30%
 - Python 3 40%
 - Terminal de tu computadora. 0%