



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

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 = Valores Discretos.
- b) Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) de los datos muestrales.
R = De razón.
- c) ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) se utiliza?
R = Muestreo 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 = La cantidad de acciones de cada accionista.
- 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 = No se puede saber cuando el accionista regresara el cuestionario o si dejara algunas respuestas en blanco. No podemos verificar si en realidad el accionista recibió el cuestionario.

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

R = Como solo se analiza una variable solo podíamos describir lo que pasa con ella o si existen patrones en los datos.

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

R = Son medidas que te hacen resumir la información que obtienes y por ende es más fácil entenderla.

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

R = Es un proceso que te ayuda a encontrar relaciones o patrones en los datos. Se puede aplicar en el sector de comunicaciones para hacer predicciones sobre sus consumidores, en el sector asegurador para prever algún riesgo.

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

R = C++ - (60%) Visual Basic - (30%) R - (20%)



UANL

FCFM



UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

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

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

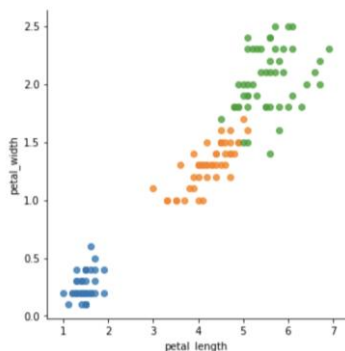
R = No

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

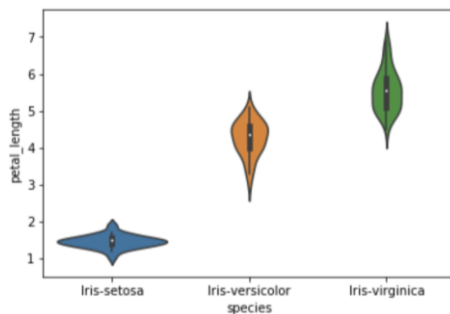
R = No

IX. Relaciona los conceptos con sus gráficas correspondientes:

- Se utilizan para evaluar visualmente cómo se distribuyen los puntos de datos con respecto a su frecuencia.
- 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.
- 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.
- 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.
- 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

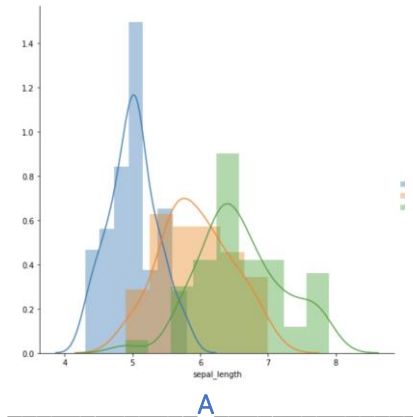


UANL

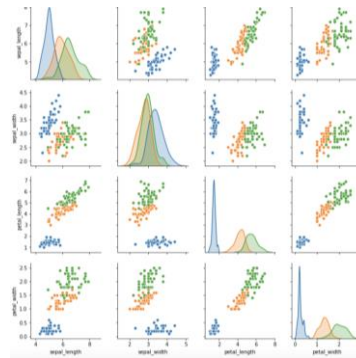
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

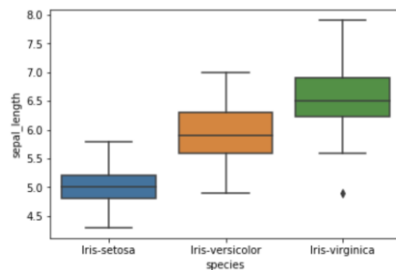
FACULTAD DE CIENCIAS FISICO MATEMÁTICAS



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 = No
- XI. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno)
R = Si, en este momento no me aparece la materia en mis unidades de aprendizaje y en las otras materias no he tenido ningún problema.
- XII. Qué sistema operativo tienes:
R = Windows 10 sistema operativo de 64 bits
- XIII. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.
- Github. 20%
 - Jupyter Notebook. 0%
 - Google Collab. 0%
 - Python 3. 0%
 - Terminal de tu computadora.