



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?
Discretos
- b) Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) de los datos muestrales.
De razón
- c) ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) se utiliza?
Muestreo estratificado
- d) Si se calcula el número promedio (la media) de acciones. ¿El resultado es un estadístico o un parámetro?
Un estadístico ya que está en función de los datos de una muestra
- e) Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?
La distribución de la cantidad de acciones/accionistas por cada estado para saber si el reparto de acciones es equilibrado y si existe alguna amenaza a mí puesto como dueño de la empresa
- 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?
Las opiniones a veces son subjetivas por lo que los valores obtenidos tienen la posibilidad de estar sesgados

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

Se utiliza para analizar los datos en base a una sola variable

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

La media es el promedio de un conjunto de datos o el resultado más probable de algún evento, la mediana es el valor central de dicho conjunto de datos y la desviación estándar representa la dispersión de los datos respecto a la media.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

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

La minería de datos es una parte de la estadística que sirve para analizar el comportamiento de grandes cantidades de datos con la finalidad de encontrar tendencias y realizar predicciones. Se puede aplicar en sondeos de preferencias (rating de audiencia, procesos electorales, encuestas) y en predicciones de eventos futuros (tendencias en el precio de acciones, productos más vendidos, efectividad de tratamientos médicos, pronóstico del clima, etc)

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

C, C++, R, Python (son los que más uso), Visual Basic (no mucho) y SQL (muy poco)

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

No

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.



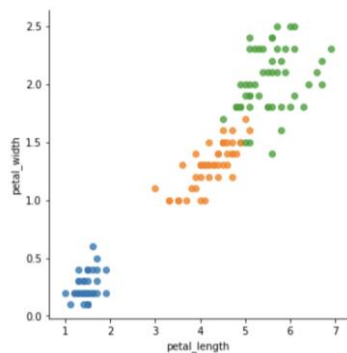
UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

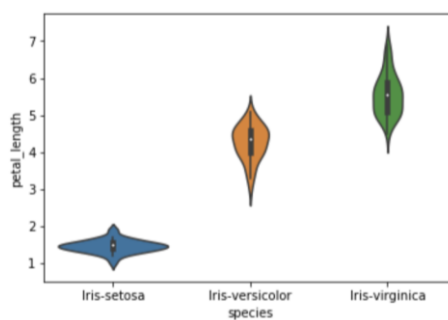
FCFM



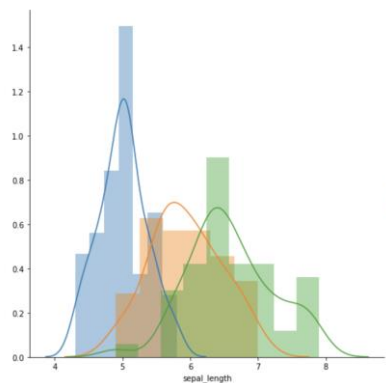
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



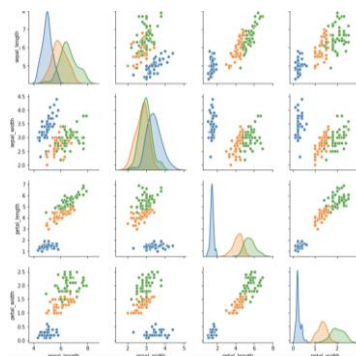
D



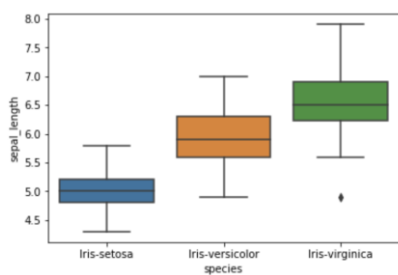
C



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)
Usualmente uso audífonos y ya no me funciona el micrófono con ellos por lo que para hablar tengo que desconectarlos.



UANL

FCFM



UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

- XI. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno)
Recientemente no he tenido problemas pero sé que ciertos en días no carga la página ya sea por saturación o por mantenimiento
- XII. Que sistema operativo tienes: Microsoft Windows 10 Home
- XIII. De las siguientes herramientas menciona si tienes conocimiento de alguna y en qué porcentaje.
- Github.
 - Jupyter Notebook
 - Google Collab
 - Python 3
 - Terminal de tu computadora.
- Solo tengo conocimiento de Python 3, diré que más del 50% pues conozco lo necesario y sé que me faltan cosas por aprender pero no cuáles son esas cosas