



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? **Sistemático**
- d) Si se calcula el número promedio (la media) de acciones. ¿El resultado es un estadístico o un parámetro? **Parámetro**
- e) Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante? **Promedio de acciones por 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? **Podría suponer un poco mas de dificultad a la hora de tomar datos, o deja de forma libre las respuestas esto podría generar ciertos problemas de no ser muy especifico el cuestionario**

II. Describe con tus palabras para que se utiliza el análisis uni-variante de datos: Para saber el comportamiento o características de una sola variable

III. Describe qué es la media, la mediana y la desviación estándar:
La media es el promedio (La suma de los datos dividida entre el total de datos), la mediana es el dato que esta en medio si se ordena de forma ascendente o descendente, y la desviación es el margen de error entre un dato y lo que se espera que valga

IV. Con tus palabras explica qué es minería de datos y en donde puede ser aplicada:
Entiendo que es como la obtención y análisis de ciertos datos, separándolos por algún parámetro (Personas en ciertas áreas, ventas, gustos, etc.), se puede aplicar para dar pronósticos , estrategias o separar información.

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



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



C – 60%

VB – 40%

R – 30%

Son los lenguajes que hemos visto durante la carrera, pero tampoco soy muy bueno programando.

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

La verdad no.

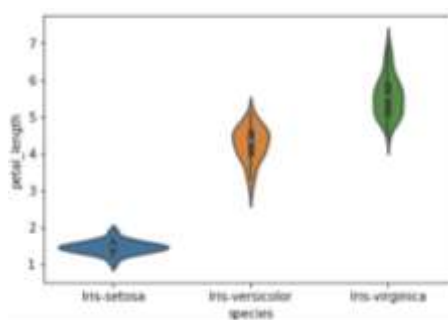
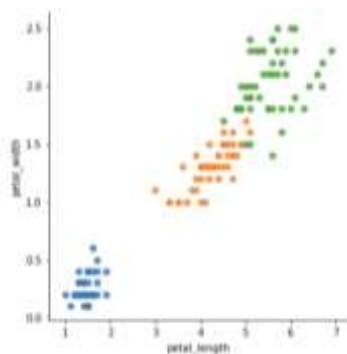
VII. ¿Trabajas o has trabajado alguna vez en alguna empresa que maneje técnicas de ciencia de datos? No he trabajado aun en nada relacionado con eso.

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

No por ahora

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

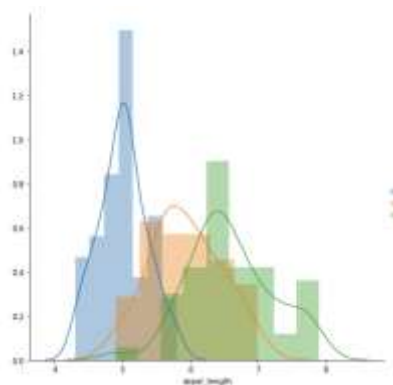
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

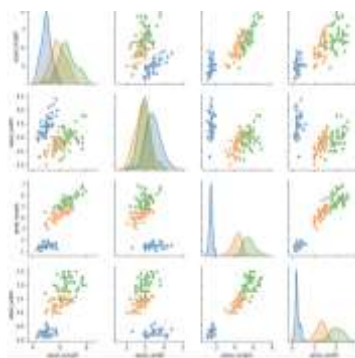


A)

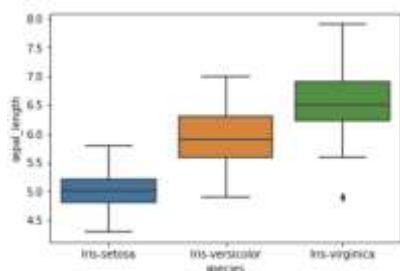


D)

C)



E)



B)

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



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS