



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: 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: Sistemático
- 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: Cantidad de acciones destinados a cada uno de los productos
- 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: El comparar los resultados de los de un estado con otro estado

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

R: Descripción detallada de una sola variable de los datos

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

R: Media: es el punto medio de los datos, es decir, un promedio.

Desviación Estándar: es la variabilidad que puede tener la media

Mediana: es el valor medio de los datos ordenados de mayor a menor

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

R: es el análisis de la información de los datos que un analista obtiene para poder mejorar la calidad del servicio y/o producto para asegurar al cliente con una excelente satisfacción.

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

R: Python-80%(Practicando lo que he aprendido:98%),R-15%,C-30%(Practicando lo que he aprendido:80%)



UANL

FCFM



UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

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

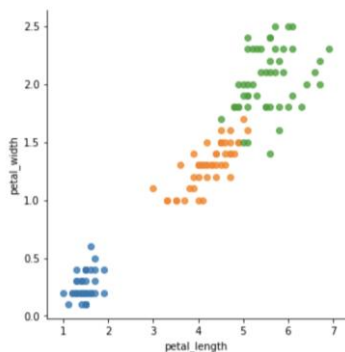
R: No

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

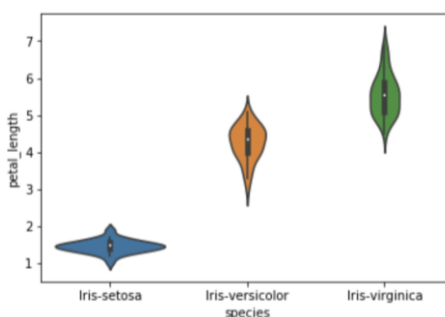
R: No

VIII. 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.



e



c

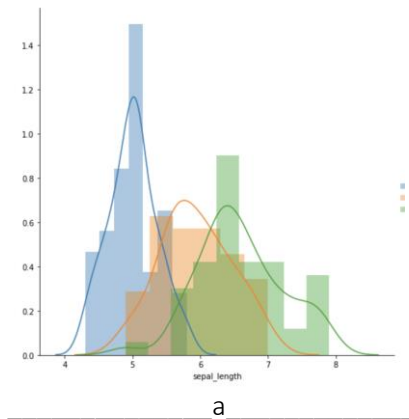


UANL

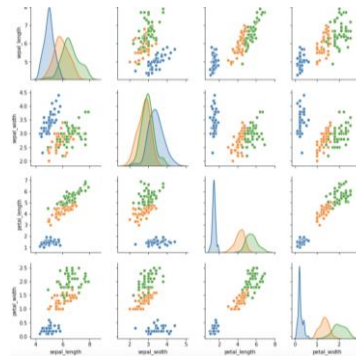
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

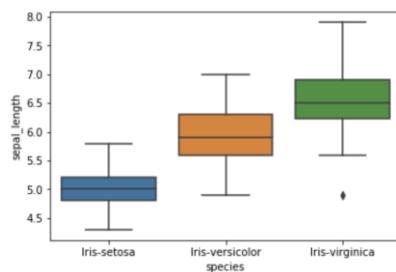
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



a



d



b

IX. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos)

R: No

X. Que sistema operativo tienes:

R: Windows 10

XI. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.

- Github.
- Jupyter Notebook
- Python 3: 80%
- Terminal de tu computadora : 15% (correr programas de Python y su simulador)