



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? **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? **En qué estado hay más accionistas**
- f) ¿Qué es lo que está incorrecto al evaluar la opinión del accionista enviando un cuestionario por correo que los accionistas podrían llenar y regresar por el mismo medio? **La fiabilidad de la información**

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

Para analizar una única cualidad o característica de un individuo.

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

Media o promedio: valor medio de un conjunto de números.

Mediana: es el valor que se encuentra en el medio de un conjunto de datos, ordenados por tamaño.

Desviación estándar: valor numérico que representa el promedio de diferencia que hay entre los datos y la media.

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

Al igual que en la minería real, en la minería de datos se "mina" o busca entre los datos para encontrar: relaciones, patrones, características, etc cualquier cosa que sea de valor para después usar esa información en diversas áreas.

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



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



C/C++ - 80%

Python - 60% (no usado en actualidad)

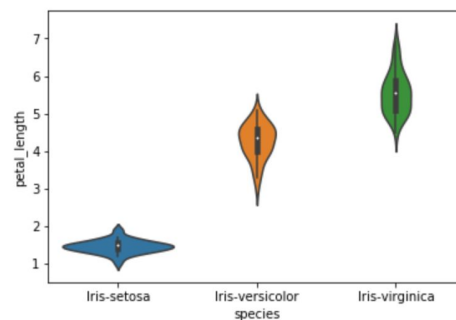
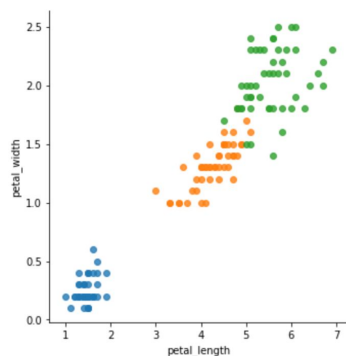
C# - 50%

Java - 85%

Javascript/Typescript - 90%

- VI. Sabes manejar notebooks digitales como Jupyter, Spyder o algún otro (Si la respuesta es otro, anota cuál): **No**
- VII. Has utilizado alguna herramienta de aprendizaje de máquina. Si la respuesta es afirmativa, anota cuál: **No**
- VIII. 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.





UANL

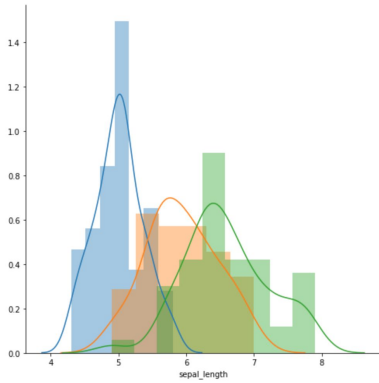
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

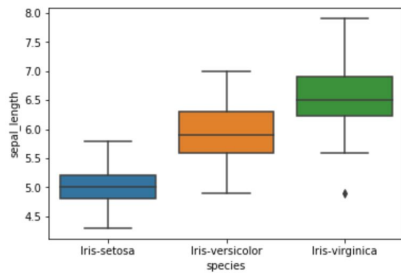
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



A

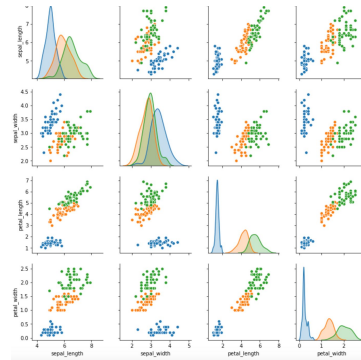


B



C

D



E

- IX. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos): **No**
- X. Que sistema operativo tienes: **Windows / Arch Linux**
- XI. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.
- Github. 90%
 - Jupyter Notebook 0%
 - Python 3 60%
 - Terminal de tu computadora. 100%



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

