



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.

razón

- c) ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) se utiliza?

estratificado

- 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?

El porcentaje de acciones que pose respecto al total , y respecto al estado

- 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?

El sesgo que se podría tener dado que son dueños de la empresa y siempre verán por su beneficio.

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

Sirve para poder obtener información estadística de la variable en cuestión, como su comportamiento , distribución, sus medidas descriptivas etc.

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

La media es el promedio de la muestra

la mediana es el numero que ocupa el valor central de todos los datos

la desviación estándar es una medida de dispersión y se utiliza para saber que tan alejados están los datos de su promedio y es la raíz cuadrada de la varianza

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



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

La minería de datos es un área de la programación relacionada con la estadística, que se basa en reunir una gran base de datos, y a partir de estos obtener información significativa según el tipo de datos, como por ejemplo con la base de datos que genera cada persona en internet se puede usar la minería de datos para dar recomendaciones de anuncios, productos entre otras cosas

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

R 30%

C 40%

Visual basic 60%

Python 30%

VI. Sabes manejar notebooks digitales como Jupyter, Spyder, Google Collab, etc. (Si la respuesta es otro, anota cuál): e utilizado jupyter pero no demasiado

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.

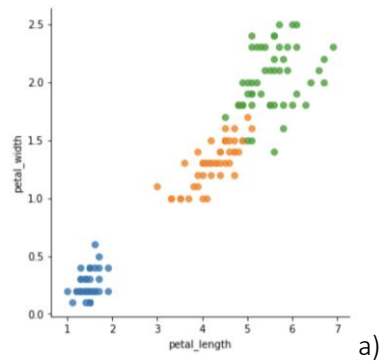


UANL

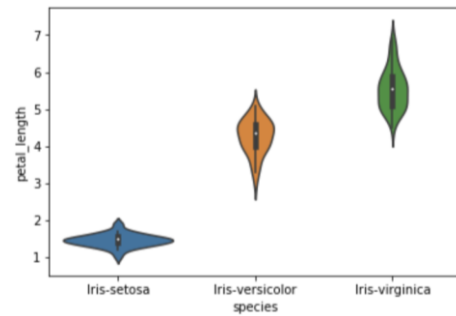
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

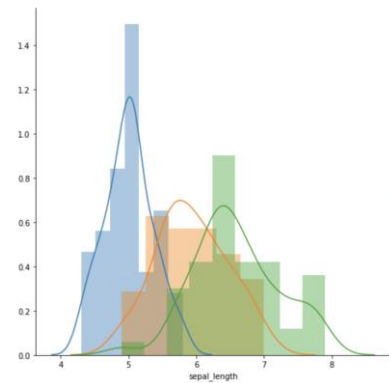
FACULTAD DE CIENCIAS FISICO MATEMÁTICAS



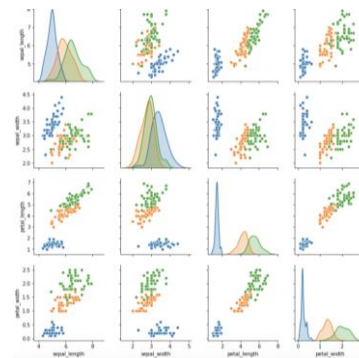
a)



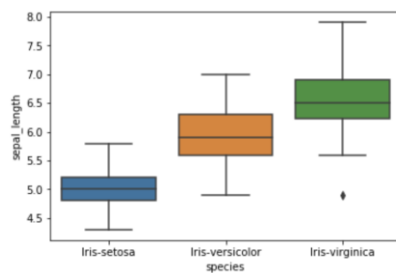
b)



d)



e)



b)

- X. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos): mi lap de repente no me detecta el micrófono , pero solo es en contadas ocasiones



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) No me aparece la materia en nexus
- XII. Que sistema operativo tienes:
Windows 10
- XIII. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.
- Github. 0%
 - Jupyter Notebook 15%
 - Google Collab 0%
 - Python 3 30%
 - Terminal de tu computadora. 25%