



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, ya que la población es finita.**
- 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?  
**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?  
**Que los datos obtenidos sean verídicos y fiables**
- 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?  
**Que este sea divulgado en alguna parte y que los datos generados por la encuesta no sean certeros.**

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

**Es el análisis de datos mediante una sola variable, encontrando tendencias en ese tipo de datos, así como ver la media, varianza, mínimo, máximo, etc. Por ejemplo, el análisis de la edad de una población.**

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

**Media: Es el valor que se espera obtener de la muestra de datos.**

**Mediana: Es el valor que se encuentra en el medio de los datos cuando estos tienen un orden de mayor a menor o viceversa.**

**Desviación estándar: que tanto se pueden alejar los datos del promedio**

### 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

Obtener modelos o predicciones de lo que puede pasar en un futuro tomando en cuenta lo que ya sucedió en el pasado esto a través del análisis de los datos. Yo imagino que puede ser aplicada en cualquier parte donde se tenga un buen registro de información de antes, por ejemplo una predicción de cuánto y si pagará a tiempo un cliente de acuerdo a su historial de pagos.

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

Visual Basic: 70%

R: 60%

C++: 70%

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

no jamás los he utilizado

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 aun 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.



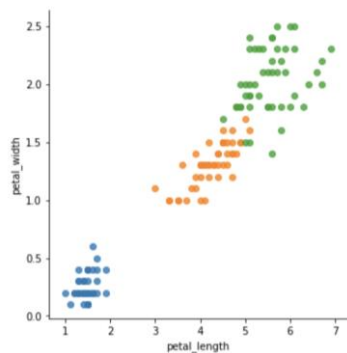
# UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

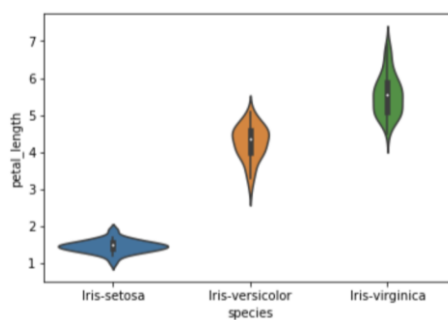
# FCFM



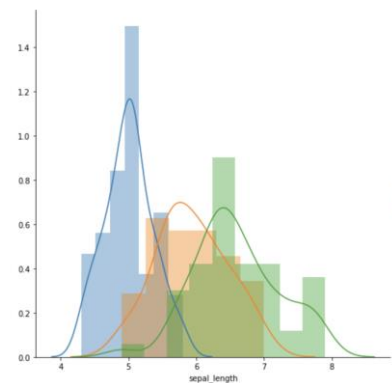
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



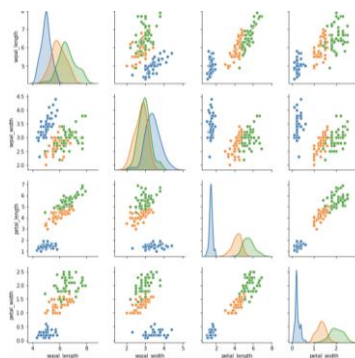
d



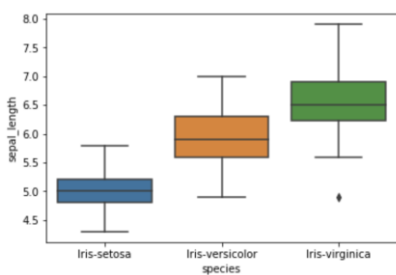
c



e



a



b

- X. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos)  
[ninguno](#)
- XI. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno)



# UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

# FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

No estoy dentro del grupo de nexus, fuera de eso nada.

- XII. Que sistema operativo tienes:  
[Windows 10](#)
- XIII. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.
- Github. [no](#)
  - Jupyter Notebook [no](#)
  - Google Collab [no](#)
  - Python 3 [no](#)
  - Terminal de tu computadora. [no](#)