



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.  
Ordinal
- 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?  
Un parámetro
- e) Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?  
Que su varianza no sea tan alta.
- 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 lo podría contestar cualquier otra persona.

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

*Significa que se evalúa o se analiza cada variable por separado*

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

*La media es el promedio de los datos, la suma de todos los valores dividida entre el número de datos. La mediana es el valor que se encuentra en el centro, tras haber ordenado los datos de menor a mayor. La desviación estándar es una medida de cuán alejados están los datos de la media.*

### 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

*Es un proceso mediante el cual se seleccionan aquellos datos que son útiles de un conjunto muy grande de los mismos. Puede ser utilizada en numerosos ámbitos para ayudar a resolver problemas, realizar predicciones, calcular probabilidades, entre otras cosas.*

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

C (70-80%), Visual Basic (Excel, 70-80%), R (50%), SQL (50%)

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

OneNote de Microsoft

**VII. ¿Trabajas o has trabajado alguna vez en alguna empresa que maneje técnicas de ciencia de datos?**

Actualmente trabajo en una empresa que las utiliza.

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

El programa Darwin, en la empresa en que trabajo.

**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.



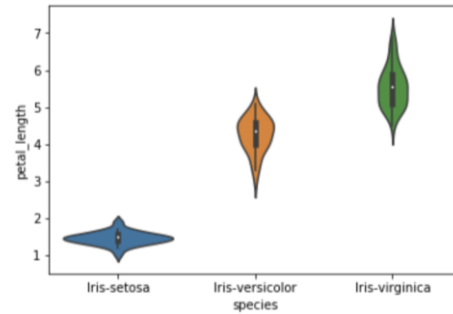
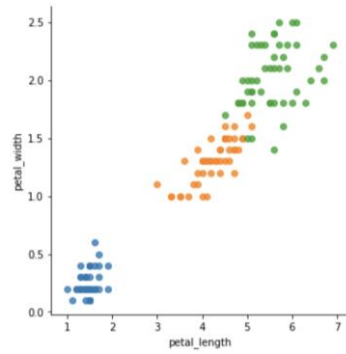
UANL

FCFM



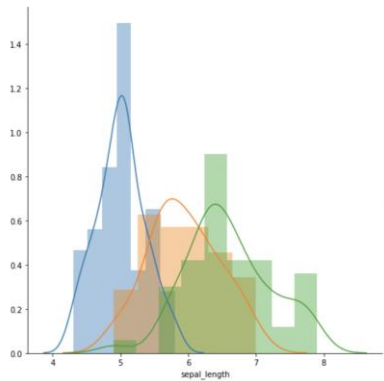
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

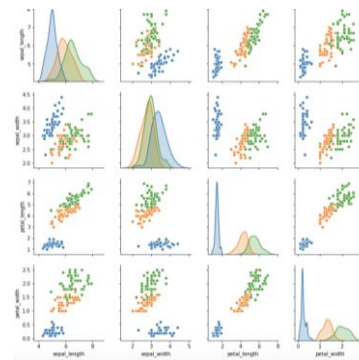


c)

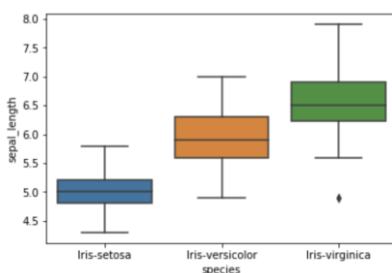
d)



a)



e)



b)

- X. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos): NO
- XI. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno): Sí, no estoy agregado al equipo de Minería de datos.



# UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

# FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

XII. Que sistema operativo tienes:

Windows 10

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

- Github. 20%
- Jupyter Notebook. 0%
- Google Collab: 0%
- Python 3: 20%
- Terminal de tu computadora: 50%