



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: Estratificado
- d) Si se calcula el número promedio (la media) de acciones. ¿El resultado es un estadístico o un parámetro?
R: parámetro
- e) Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?
R: Que el muestreo fuera de manera aleatoria, para así evitar sesgos y trayendo consigo un menor margen de error en los resultados
- 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: Que la información recabada pudieran ser datos no estructurados

II. Describe con tus palabras para que se utiliza el análisis univariante de datos:

R: Principalmente para encontrar patrones entre los datos, y puede ser con ayuda de medidas como media, moda, mediana, entre otras.

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

Media: Es el promedio de un conjunto de datos

Mediana: Es el valor central del conjunto ordenado de datos

Desviación estándar: Es la raíz cuadrada de la varianza y nos da información respecto a la dispersión de los datos



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

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

R: Es un proceso en el cual se trata de encontrar patrones dentro de un conjunto de datos haciendo uso del análisis de cada una de las variables, y esta puede ser aplicada en el sector laboral, ya sea en empresas que desean optimizar procesos, y desarrollar un trato eficiente hacia los clientes.

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

- Visual basic: 75%
- C: 70%

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

R: Solo se manejar Jupyter

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

R: No

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

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.

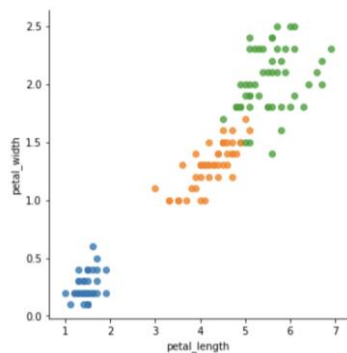


UANL

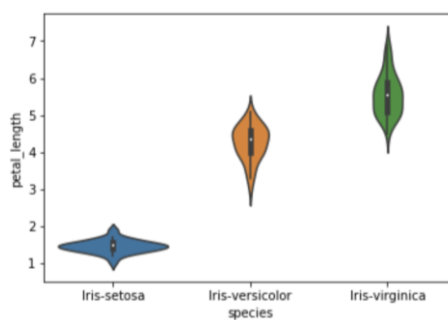
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

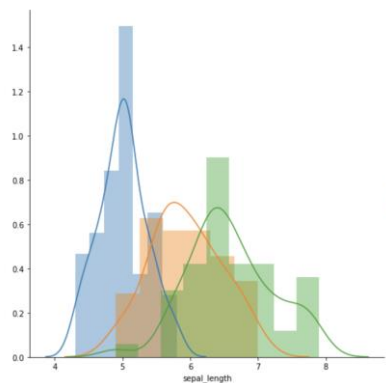
FACULTAD DE CIENCIAS FISICO MATEMÁTICAS



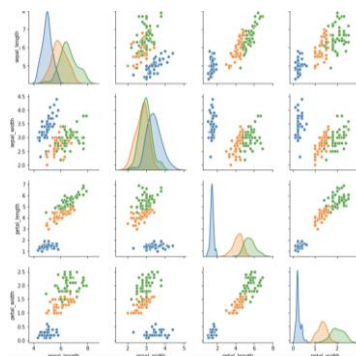
D



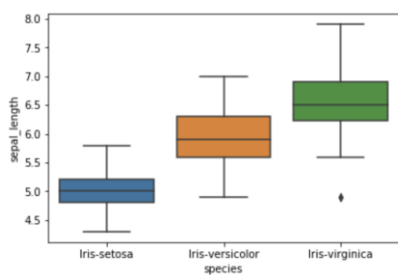
C



A



E



B

- X. Tienes problemas de compartir pantalla, o hablar en micrófono en el Temas: (Menciona cuales son los problemas en caso de tenerlos)
R: Al compartir pantalla, mi equipo se pone lento
- 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

R: No tengo problemas para ingresar a Nexus, pero la materia no me aparece en Nexus

XII. Qué sistema operativo tienes:

R: Windows 10

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

- Github.
- Jupyter Notebook: 25%
- Google Collab
- Python 3:
- Terminal de tu computadora.