



IIC2115 – Programación como Herramienta para la Ingeniería (I/2018)

Control 6

Tiempo: 10 mins

Nombre: _____

1. Para visualizar los valores extremos de una variable en un conjunto de datos, es preferible aplicar:
 - a) un *boxplot*
 - b) una transformación logarítmica, seguida de un *boxplot*.
 - c) una transformación logarítmica, seguida de un histograma.
 - d) ninguna de las anteriores.
2. ¿Cuál de las siguientes es la mejor alternativa para completar los valores faltantes de una variable categórica?
 - a) Utilizar el promedio.
 - b) Utilizar una regresión lineal.
 - c) Utilizar un árbol de decisión.
 - d) No es posible completar los valores faltantes de una variable categórica.
3. Con respecto a la función `fillna()`, es posible afirmar que:
 - a) Permite rellenar los puntos faltantes en un gráfico.
 - b) No puede llamarse sobre columnas con variables categóricas.
 - c) Puede recibir una función como argumento, sólo cuando actúa sobre los valores de una variable categórica.
 - d) Ninguna de las anteriores.
4. ¿Que principal ventaja trae el uso de la biblioteca Pandas?
 - a) Permite construir modelos predictivos.
 - b) Facilita la construcción de herramientas de cálculo científico, gracias al arreglo n-dimensional.
 - c) Permite manejar datos estructurados de manera sencilla.
 - d) Permite construir visualizaciones de los datos.

5. Al calibrar y evaluar un modelo predictivo usando validación cruzada, es posible afirmar que:
- a) El rendimiento siempre será superior al obtenido usando sólo el set de calibración para evaluar.
 - b) El rendimiento siempre será inferior al obtenido usando sólo el set de calibración para evaluar.
 - c) Se obtiene una estimación más cercana del rendimiento real del modelo predictivo, sobre datos no vistos.
 - d) Al agregar más variables, el rendimiento se mantendrá igual, o se reducirá.
6. Con respecto a la función `head()`, es posible afirmar que:
- a) Proporciona el conteo, media, desviación, mínimo, cuartiles y máximo de los datos.
 - b) Recibe como parámetro un entero, que indica la cantidad de filas a mostrar.
 - c) No es posible aplicarla a un `DataFrame`.
 - d) Los valores de las variables categóricas, son representadas con NaN.