

1.

¿La estadística descriptiva puede ser subjetiva en la definición de sus métricas?

Verdadero

2.

¿En qué etapa del flujo de trabajo de un proyecto de Data Science se usa la estadística descriptiva?

En la ingesta de datos y validación

REPASAR CLASE

3.

Un ejemplo de dato categórico es:

Género

4.

Un ejemplo de dato numérico es:

Precio

5.

Un diagrama de frecuencia nos sirve para:

Visualizar la probabilidad de ocurrencia de un dato.

REPASAR CLASE

6.

La media es una excelente medida de tendencia central cuando tenemos outliers en nuestro conjunto de datos. Este enunciado es:

Verdadero

REPASAR CLASE

7.

Para calcular la mediana de un dato en nuestro dataset lo hacemos con:

```
df['name'].median()
```

8.

¿Cuál de estas visualizaciones es muy útil para presentar los intercuatiles?

Diagrama de caja

9.

La desviación estándar se puede dividir por:

Población y muestra

10.

El rango de un dato que describe la altura en un conjunto de datos se puede calcular como:

```
df['height'].max() - df['height'].min()
```

11.

Un diagrama de dispersión nos sirve para:

Visualizar la correlación de todas las variables dentro de nuestro conjunto de datos.

12.

Una de las razones por las que realizamos escalamientos es:

Para que los optimizadores de nuestros modelos funcionen mejor.

13.

Las transformaciones no lineales nos sirven para tipos de datos:

Fuertemente sesgados y no simétricos

14.

Un tipo de mapeo numérico que causa una representación compacta es el:

One-hot

REPASAR CLASE

15.

El preprocesamiento por mapeo se puede aplicar tanto a variables categóricas y numéricas:

Verdadero

16.

Cuando el coeficiente de correlación es alto:

Los datos en nuestro diagrama de dispersión se agrupan con una pendiente positiva.

17.

Una matriz de covarianza nos sirve para:

Medir la varianza de una sola variable y comparando con el resto.

REPASAR CLASE

18.

El análisis de componentes principales es un proceso de:

Transformación lineal

REPASAR CLASE