

Proyecto Integrador

Armando Ocampo

Preguntas Proyecto

A partir del dataset *covid_mexico.RDS*, realiza o responderas lo siguiente. Asimismo, compartirás el código utilizado para responder cada pregunta. Antes de iniciar corre la siguiente semilla aleatoria *set.seed(456)*.

1. Obten la media aritmética, la moda y la mediana de la variable *total_cases*.
2. Obten la varianza y desviación estándar de la variable *new_tests*.
3. Genera los 3 cuartiles de la variable *total_tests*
4. ¿En qué cuartil se encontraría el valor 9,458,964 para la variable *total_tests*
5. ¿Cuándo finalizó la captura de la información del dataset *covid_mexico*?
6. ¿Qué medida de tendencia central utilizarías para recodificar los valores faltantes de la variable *total_tests*? Explica y comparte el flujo de trabajo realizado para explicar tu respuesta.
7. Realiza una recodificación de los valores faltantes en la variable *total_test* con la medida de tendencia central elegida en la pregunta previa. A continuación toma tres muestreos sin reemplazo del dataset *covid_mexico*. El primero con una *n* de 50, el segundo con 200 muestras y el tercero con 500. Posteriormente obten las tres medidas de tendencia central de cada uno de los datasets y comparalas con el conjunto de datos original. A partir de este proceso, ¿que conjunto de datos se asemeja al original?
8. Realiza la limpieza y transformación de los datos que creas conveniente para las variables *total_cases* y *total_deaths*. Genera un modelo de regresión y explica la relación entre estas variables. Asimismo, agrega los estadísticos necesarios que avalen tu modelo.
9. Grafica el modelo previo utilizando la función *ggplot()*. Asimismo, agrega los cambios estéticos que permitan entender el gráfico.
10. ¿Qué variables del dataset *covid_mexico*, presentan mayor influencia sobre la variable *total_tests*? (Realiza el modelo que creas pertinente).
11. Identifica las fechas en las cuales se presentaron aumento en los casos de COVID-19 durante la pandemia.

Nota: debes de llamar el dataset al ambiente de trabajo y utilizar las librerías vistas en clase