

Diplomado Ciencia de datos

Práctica 2

Objetivo: El participante reunirá todas las nociones estadísticas adquiridas en el Módulo, para responder algunas preguntas formuladas sobre un conjunto de datos a analizar.

Instrucciones: La resolución será de manera escrita individualmente por cada participante, fuera de la sesión, pero con la guía del instructor. Los participantes se podrán auxiliar de todo el material proporcionado (presentaciones, libros y scripts de Jupyter Notebook).

- Esta actividad tendrá una duración estimada de 3hrs., programada como actividad de cierre, realizada fuera de la sesión. Se espera que los participantes utilicen la herramienta Jupyter Notebook para realizar su reporte escrito, ya que esta permite incluir el código utilizado. De otra forma, pueden reportar su análisis y respuestas a las preguntas a través de otra herramienta de su preferencia, pero se debe incluir el código empleado.
- Al finalizar la última sesión de este módulo, los participantes enviarán la práctica en un documento en formato PDF en la fecha establecida, deberán guardar el archivo como Apellidos_nombre_ejercicio1 ([Ejemplo: Torres_Sonia_ejercicio1](#))

Cada inciso tiene un valor de 2 puntos para sumar un total de 10 puntos.

Porcentaje de evaluación (30%).

El conjunto de datos 'Birthweight' contiene la información de 42 bebés al nacer. La pregunta de investigación es saber si existe una relación entre el peso al nacer y el tiempo de gestación. La variable dependiente es Peso al nacer (dada en libras) y la variable independiente para esta actividad es la edad gestacional del bebé al nacer (en semanas).

- a) Realiza una descripción gráfica y de medidas estadísticas (descriptivas) de los datos.

Valor 2 puntos

- b) Realiza un análisis de regresión lineal y proporcionar estimadores puntuales de los parámetros.

Valor 2 puntos

- c) Usando el error estándar, establece intervalos de confianza al 95% para los parámetros de la regresión.

Valor 2 puntos

- d) Realiza las pruebas de hipótesis para los parámetros y para determinar la significancia de la regresión.

Valor 2 puntos

- e) Con base en tu análisis, concluye sobre el contexto del problema y responde la pregunta de investigación

Valor 2 puntos

