

Reto sobre aplicación de R para respuesta a preguntas de analítica

Maestría en Ciencia de Datos

Profesores:

Christian Urcuqui (ccurcuqui@icesi.edu.co)

Universidad Icesi

Objetivos de esta sesión

1. Usar R para la exploración de un conjunto de datos.
2. Responder a un conjunto de preguntas a partir de los datos
3. Aplicar R para el desarrollo de funciones, ciclos, asignación de variables y explorar la información de un dataframe.

Taller práctico

Contexto

Análisis de datos

PetFinder.my Adoption Prediction **How cute is that doggy in the shelter?**

Millones de animales sufren en la calle o son sacrificados en refugios todos los días en todo el mundo. Si se pueden encontrar hogares para ellos, se pueden salvar muchas vidas preciosas – y se pueden crear más familias felices.

Esta competencia busca que se los participantes puedan proponer algoritmos que permitan predecir la capacidad de adopción de las mascotas, ¿con qué rapidez se adopta una mascota?

<https://www.kaggle.com/c/petfinder-adoption-prediction>

PREGUNTAS DE ANÁLISIS

Vamos a utilizar solo el archivo train.csv

1. ¿Cómo están los valores de la velocidad de adopción?
2. ¿Hay más datos de perros o de gatos?
3. ¿Cómo es la velocidad de adopción para perros y gatos?
4. Elimine los registros con NA en los nombres