

Reto sobre aplicación de Python para respuesta a preguntas de analítica

Profesores:

Christian Urcuqui (ccurcuqui@icesi.edu.co)

Universidad Icesi

Objetivos de esta sesión

1. Usar Python para la exploración de un conjunto de datos.
2. Responder a un conjunto de preguntas a partir de los datos
3. Aplicar Python 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

Complejidad 1

1. ¿Cómo están los valores de la velocidad de adopción?
 1. Vea su distribución
 2. Cuales son los estadísticos descriptivos para esta variable.
2. ¿Hay más datos de perros o de gatos?
3. ¿Cómo es la velocidad de adopción para perros y gatos?
4. Encuentre los nombres sin registros y asígneles un valor que corresponda al contexto del problema.

PREGUNTAS DE ANÁLISIS

Vamos a utilizar solo el archivo train.csv

Complejidad 2

1. ¿Cuales son los cinco nombres más populares y su relación con la velocidad de adopción?
2. ¿Cuál es la relación de los registros sin nombre y la velocidad de adopción?
3. Como es la distribución de la variable edad.
4. Revise la relación entre la variable de edad y la velocidad de adopción.