SDAIA Data Science Bootcamp Project
Deliverable # 1: Proposal

Name: Tasneem Yousef Halawani **Cohort:** Nov 21 – Dec 16

THE DESIGN

Predicting used cars price

THE DATASET

Dataset

Saudi Arabia Used Cars Dataset

Content

The dataset contains 8,248 records of used cars collected from syarah.com. Each row represents a used car with a link to its webpage. Other information regarding each car is the brand name, model, manufacturing year, origin, the color of the car, options, capacity of the engine, type of fuel, transmission type, the mileage that the car covered, region price, and negotiable.

Number of records (rows): 8,248 cars
Number of attributes: 15 attributes

THE ALGORITHMS

Regression to be used to predict the **price** as target attribute

THE TOOLS

EDA: Pandas, Numpy
Modeling: Keras, SciKit-Learn
Visualization: matplotlib, seaborn