

Capstone Project for Springboard Data Science Intensive Course  
Alona Varshal  
July 2016

#### Project Idea 1:

Wine Quality Data Set (UCI Machine Learning Repository)  
(<http://archive.ics.uci.edu/ml/datasets/Wine+Quality>)

The wine quality data set available at UCI Machine Learning Repository provides information on a number of physicochemical parameters of Portuguese “Vinho Verde” wines. Grapes are indigenous to the Vinho Verde area in Portugal. We look at how the physicochemical parameters can be used to classify the red and the white variants of the wine as well as tell whether these parameters can differentiate the qualities of wines. Are all the 11 parameters measured or just a few of them are necessary to classify better wines from poor ones?

#### Project Idea 2:

Twin Gas Sensor Arrays Data Set (UCI Machine Learning Repository)  
(<http://archive.ics.uci.edu/ml/datasets/Twin+gas+sensor+arrays>)

A gas sensor made up of an array of eight metal oxide (MOX) sensors was tested whether it can be calibrated to measure concentrations of various gases of environmental significance (ethanol, methane, ethylene and carbon monoxide). This project attempts to see whether the findings of the paper published using the data are reproducible. It was found that the sensor can be calibrated and that one unit's result can be mapped to the other unit.

#### Project Idea 3:

Austin (Texas, USA) Animal Center Intakes and Outcomes (separate data)

Austin is known to be an animal-friendly city. Does the open data on the city's animal center intakes and outcomes reflect this? There are numerous aspects that can be examined by looking at the data. Does the rise in intake of animals correlate with any town events? What breed, sex, age among dogs is the most popular to be lost? Which area in Austin has the most lost animal? Which animal gets adopted most? Which breed of dogs get adopted most? What is the average time spent by animals in the center before adoption? Does the state of the animal affect probability of being adopted?