**Data Science Python Technical Test**

**Background:**

Use the following datasets to predict the number of future bike rentals. We would like you to produce code, in Python, using a Jupyter notebook format, that demonstrates:

* your knowledge of Python
* your approach to problem solving
* your data science skillset

**Instructions:**

1. We need you to submit one Python Jupyter notebook and a presentation (with a couple of slides) to [Griselda.dias@saga.co.uk](mailto:Griselda.dias@saga.co.uk) and [Kerry.austin@saga.co.uk](mailto:Kerry.austin@saga.co.uk). If you need to compress the files before sharing, only use the zip format
2. If you are successful in the tech test, Kerry will be in touch to book you in for a technical interview where you’ll take us through your work, explaining your thought process and methodology, and answer any questions we may have about your work

**Tips for success:**

1. Please read this document carefully and get back to us with any questions
2. There’s no “right” answer. We’re looking for how you combine problem solving, coding, and presentation skills.

**Files:**

* hour.csv : bike sharing counts aggregated on hourly basis. Records: 17,379
* day.csv - bike sharing counts aggregated on daily basis. Records: 731

**Dataset characteristics:**

Both hour.csv and day.csv have the following fields, except hr which is not available in day.csv

* instant: record index
* dteday : date
* season : season (1:springer, 2:summer, 3:fall, 4:winter)
* yr : year (0: 2011, 1:2012)
* mnth : month ( 1 to 12)
* hr : hour (0 to 23)
* holiday : weather day is holiday or not
* weekday : day of the week
* workingday : if day is neither weekend nor holiday is 1, otherwise is 0.
* weathersit :
  + 1: Clear, Few clouds, Partly cloudy, Partly cloudy
  + 2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist
  + 3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds
  + 4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog
* temp : Normalized temperature in Celsius. The values are divided to 41 (max)
* atemp: Normalized feeling temperature in Celsius. The values are divided to 50 (max)
* hum: Normalized humidity. The values are divided to 100 (max)
* windspeed: Normalized wind speed. The values are divided to 67 (max)
* casual: count of casual users
* registered: count of registered users
* cnt: count of total rental bikes including both casual and registered