CE 88 Final: Bike Sharing

In this take-home exam you will explore the data and build a forecast for the total number of hourly bike rentals in the Capital Bikeshare, a bike sharing network of Washington, D.C.

The data. The dataset (**bikeshare_2011.csv**) provided contains hourly bike rental data of 2011. The following variables are available:

- index: record index, from the beginning of the year (1 to 8760)
- date: date
- season: season (winter, spring, summer, fall)
- year: year (0: 2011, 1: 2012)
- month: month (1 to 12)
- hour: hour (0 to 23)
- holiday: if a day is a holiday or not (if a day is a regular weekend, it is still 0)
- weekday: day of the week (0 - Sunday to 6 - Saturday)
- workingday: if day is neither weekend nor holiday is 1, otherwise is 0.
- weather: general weather conditions
  1: Clear, Few clouds, Partly cloudy, Partly cloudy
  2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist
  3: Light Snow, Light Rain + Thunderstorm, Light Rain + Scattered clouds
  4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog, Apocalypse
- temp: normalized temperature, from values in Celsius. The values are derived via (t-t_min)/(t_max-t_min), t_min=-8, t_max=+39
- feels_like: normalized chill-factor adjusted temperature, from original values in Celsius. The values are derived via (t-t_min)/(t_max-t_min), t_min=-16, t_max=+50
- hum: normalized humidity. The values are divided to 100 (max).
- windspeed: normalized wind speed. The values are divided to 67 (max).

Finally, the dependent variables:

- casual: count of casual users
- registered: count of registered users
- count: total count of bikes rented, casual + registered

The Problem. You will be asked to provide a forecast of the number of hourly bike rentals for 30 days in 2012. You will predict the "count" for all 30 days. The submission will have 2 parts:

- A Kaggle competition entry where you upload a csv with your predictions.
- A submission on bcourses where you describe your methods and include your code – this can either be one document containing both code and methodology or two separate documents.

Good luck!