

We are provided hourly

rental data spanning two

years. You must predict the

total count of bikes

rented during each hour,

using only information

available prior to the rental

period.

Data Description:

**datetime** - hourly date + timestamp

**season** - 1 = spring, 2 = summer, 3 = fall, 4 = winter

**holiday** - whether the day is considered a holiday

**workingday** - whether the day is neither a weekend nor holiday

**weather** - 1: Clear, Few clouds, Partly cloudy

2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist

3: Light Snow, Light Rain + Scattered clouds

4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog

**temp** - temperature in Celsius

**atemp** - "feels like" temperature in Celsius

**humidity** - relative humidity

**windspeed** - wind speed

**casual** - number of non-registered user rentals initiated

**registered** - number of registered user rentals initiated

**count** - number of total rentals

**Questions:**

1. Identify the nature of each variable (categorical/numerical).
2. Which variables are useful for prediction?
3. Build model with useful features so that error in prediction should be less.

Answer: Temperature

Please do stepwise regression both forward selection and backward elimination method