

In-class Brainstorming: Bike Rental Forecast

Ling-Chieh Kung*

September 14, 2016

1 Introduction

Many business decisions depend on accurate forecasts. Two obvious examples are that inventory planning depends on demand forecasts and personnel scheduling depends on the forecasts of consumer arrivals. In this in-class brainstorming challenge, we will consider bike rental forecasts.

Consider the MS Excel file “Bike.xlsx” provided to you. The sheet “Day” contains the daily rental records of a public bike rental system in a city from 2011/1/1 to 2012/12/31.¹ There are 731 rows representing the 731 days in the time horizon. The 16 columns are explained below:

- (Serial number) instant: A serial number from 1 to 731.
- (Date information) date, year, season, month: the labels of that date.
- (Working information) holiday, weekday, workingday: “holiday” is 1 if that day is a national holiday not in a weekend and 0 otherwise; “weekday” labels whether it is Sunday (labeled as 0), Monday (labeled as 1), ..., or Saturday (labeled as 6); “workingday” is 1 if that day is a working day (neither a weekend nor a holiday) and 0 otherwise.

*Department of Information Management, National Taiwan University; lckung@ntu.edu.tw.

¹This set of data is real and publicly available online. For instructional purpose, the name of the city is hided. The instructor is glad to reveal the city name at the end of the semester.

- (Weather information) Five attributes are recorded in this category:
 - weathersit (weather situation): 1 for sunny or partly cloudy, 2 for misty and cloudy, 3 for light snow or light rain, and 4 for heavy snow or thunderstorm.
 - temp (temperature) and atemp (apparent temperature): the daily average of temperature and apparent temperature (in Celsius), respectively.
 - humidity: the daily average of the humidity (in %).
 - windspeed (wind speed): the daily average of the wind speed (in knot; 1 knot is around 1.852 km/h).
- (Rental data) casual, registered, cnt: “casual” is the number of rentals made by unregistered users, “registered” is the number of rentals made by registered members, and “cnt” is the sum of the two numbers.

The sheet “Month” aggregates the daily records into monthly records. You may see that there are 24 rows, one for each month in 2011 and 2012. Among the 11 columns, “workingday” is the number of working days in that month, “weathersit” is the number of days whose weather situation is labeled as 1, “temp”, “atemp”, “humidity”, and “windspeed” are the monthly average of the corresponding attributes, and “casual”, “registered”, and “cnt” are the monthly sum of the corresponding attributes.

2 Tasks

Here are two tasks for you:

1. Using the monthly data in 2011 and 2012, forecast the total number of bike rentals in January, 2013. Give us one number and tell us how you obtain it.
2. Then use the historical monthly data to forecast the numbers of monthly bike rentals in the first six months in 2013. Give us six numbers and tell us how you obtain it.

Statisticians and practitioners have developed several widely recognized ways to do forecast. Some methods will be covered in this semester. Nevertheless, different methods always have different pros and cons; no one says that you cannot find a good way for this specific problem. Just try it!