In-class Brainstorming: Bike Rental Forecast

Ling-Chieh Kung[*]

September 14, 2015

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);

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

[1]This 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.

"workingday" is 1 if that day is a working day (neither a weekend nor a holiday) and 0 otherwise.

- (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.

To somehow help you, 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.

Given the historical data in 2011 and 2012, how would you forecast bike rentals in 2013? 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!

2 Tasks

1. Maybe it is good to start from the following task: 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 whole 2013. Give us 12 numbers and tell us how you obtain it (or your plan).

3. How would you use the historical daily data to make a more accurate monthly forecast for 2013? How would you make a daily forecast?

3 Historical monthly data

year	month	workingday	weathersit	temp	atemp	humidity	windspeed	casual	registered	cnt
2011	1	20	17	8.11	10.22	58.44	7.23	3073	35116	38189
2011	2	19	19	11.58	14.22	56.01	8.46	6242	41973	48215
2011	3	23	17	13.6	16.26	56.94	8.6	12826	51219	64045
2011	4	20	14	19.32	22.84	66.83	9.04	22346	72524	94870
2011	5	21	19	23.67	27.54	71.34	6.71	31050	104771	135821
2011	6	22	21	28.42	32.33	59.32	6.59	30612	112900	143512
2011	7	20	27	31.1	35.53	58.97	6.35	36452	104889	141341
2011	8	23	23	28.92	32.77	62.68	7.06	28842	107849	136691
2011	9	21	11	25.13	28.43	78.41	5.66	26545	100873	127418
2011	10	20	18	19.27	22.94	70.68	6.5	25222	98289	123511
2011	11	20	20	16.5	19.97	67.53	6.69	15594	86573	102167
2011	12	21	20	13.33	16.49	64.92	6.22	8448	78875	87323
2012	1	20	20	11.28	13.76	58.73	8.04	8969	87775	96744
2012	2	20	18	12.93	15.73	57.46	7.52	8721	94416	103137
2012	3	22	19	18.43	21.93	60.75	7.88	31618	133257	164875
2012	4	20	23	19.22	22.87	50.78	8.32	38456	135768	174224
2012	5	22	18	25.11	29.03	66.46	6.83	44235	151630	195865
2012	6	21	25	27.68	31.5	55.84	7.13	43294	159536	202830
2012	7	21	24	30.85	34.88	60.61	5.94	41705	161902	203607
2012	8	23	20	29.18	32.38	64.86	5.74	43197	171306	214503
2012	9	19	22	25.42	29.5	64.54	6.62	43778	174795	218573
2012	10	22	14	20.5	24.26	68.07	6.46	34538	164303	198841
2012	11	20	22	13.78	16.77	57.44	6.91	21009	131655	152664
2012	12	20	12	13.24	16.12	68.29	6.85	13245	110468	123713

4 Historical daily data (some rows and columns are removed)

date	holiday	weekday	weathersit	temp	atemp	humidity	windspeed	casual	registered	cnt
2011/1/1	0	6	2	14.11	18.18	80.58	5.94	331	654	985
2011/1/2	0	0	2	14.9	17.69	69.61	9.20	131	670	801
2011/1/3	0	1	1	8.05	9.47	43.73	9.19	120	1229	1349
2011/1/4	0	2	1	8.20	10.61	59.04	5.93	108	1454	1562
2011/1/5	0	3	1	9.31	11.46	43.70	6.92	82	1518	1600
2011/1/6	0	4	1	8.38	11.66	51.83	3.31	88	1518	1606
2011/1/7	0	5	2	8.06	10.44	49.87	6.24	148	1362	1510
2011/1/8	0	6	2	6.77	8.11	53.58	9.87	68	891	959
				$\vdots$						
2012/12/24	0	1	2	9.48	12.95	79.13	2.86	174	746	920
2012/12/25	1	2	2	11.94	14.72	73.48	6.24	440	573	1013
2012/12/26	0	3	3	9.98	11.02	82.33	11.71	9	432	441
2012/12/27	0	4	2	10.42	11.33	65.29	12.95	247	1867	2114
2012/12/28	0	5	2	10.39	12.75	59.00	5.75	644	2451	3095
2012/12/29	0	6	2	10.39	12.12	75.29	4.60	159	1182	1341
2012/12/30	0	0	1	10.49	11.59	48.33	12.98	364	1432	1796
2012/12/31	0	1	2	8.85	11.17	57.75	5.73	439	2290	2729