CASE ANALYSIS:
FORECASTING FOOD AND BEVERAGE SALES

Suzanne Michelle Gager

QNT 531: Advanced Problems in Statistics and Research Methods

Terrance C. Feravich

March 20, 2006

Case Analysis: Forecasting Food and Beverage Sales

Problem Definition

The Vintage Restaurant is on Captiva Island, a resort community near Fort Myers, Florida. The restaurant, which is owned and operated by Karen Payne, has just completed its third year of operation. During that time, Karen has sought to establish a reputation for the restaurant as a high-quality dining establishment that specializes in fresh seafood. The efforts of Karen and her staff have proven successful, and her restaurant has become one of the best and fastest-growing restaurants on the island. Karen has concluded that to plan for the growth of the restaurant in the future, she needs to develop a system that will enable her to forecast food and beverage sales by month for up to one year in advance. Karen has the following data ($1000s) on total food and beverage sales for the three years of operation.

Lost Beverage and Food Sales Case
Vintage Restaurant Sales

MONTH	First Year	Second Year	Third Year
January	242	263	282
February	235	238	255
March	232	247	265
April	178	193	205
May	184	193	210
June	140	149	160
July	145	157	166
August	152	161	174
September	110	122	126
October	130	130	148
November	152	167	173
December	206	230	235

The statistical Summary of the data is shown below:

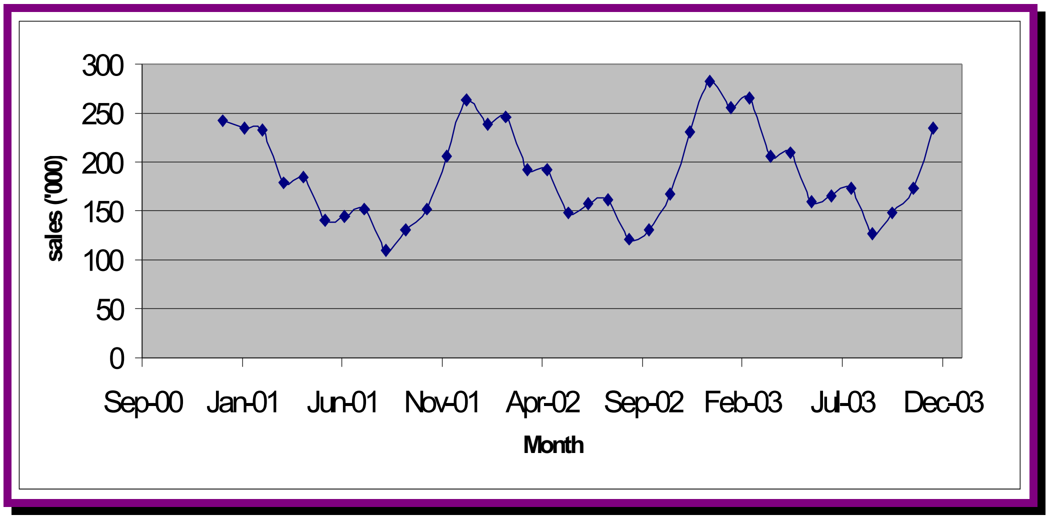
Year	January	February	March	April	May	June	July	August	September	October	November	December	Total Sales
1	242	235	232	178	184	140	145	152	110	130	152	206	2106
2	263	238	247	193	193	149	157	161	122	130	167	230	2250
3	282	255	265	205	210	160	166	174	126	148	173	235	2399
Total:	787	728	744	576	587	449	468	487	358	408	492	671	6755
Mean:	262.3333333	242.66667	248	192	195.6667	149.6667	156	162.3333	119.33333	136	164	223.6667	2251.666667
Variance:	400.3333333	116.33333	273	183	174.3333	100.3333	111	122.3333	69.333333	108	117	240.3333	21464.33333
StDev:	20.0083316	10.785793	16.52271	13.52775	13.20353	10.01665	10.53565	11.06044	8.326664	10.3923	10.81665	15.50269	146.5071102

Analyze the Data

There are two main goals of time series analysis: (a) identifying the nature of the phenomenon represented by the sequence of observations, and (b) forecasting (predicting future values of the time series variable). Both of these goals require that the pattern of observed time series data is identified and more or less formally described. Once the pattern is established, we can interpret and integrate it with other data (i.e., use it in our theory of the investigated phenomenon, e.g., seasonal commodity prices). Regardless of the depth of our understanding and the validity of our interpretation (theory) of the phenomenon, we can extrapolate the identified pattern to predict future events.

Time Series Graph

To discover the characteristic of the time series, the visual inspection of the graph is the first step in any time series analysis and forecasting. The graph of Vintage Restaurant food and beverage sales time series is plotted below:



Seasonal Index

Seasonal index represents the extent of seasonal influence for a particular segment of the year. The calculation involves a comparison of the expected values of that period to the grand mean.

A seasonal index is how much the average for that particular period tends to be above (or below) the grand average. Therefore, to get an accurate estimate for the seasonal index, we compute the average of the first period of the cycle, and the second period, etc, and divide each by the overall average. The formula for computing seasonal factors is $S_i = D_i/D$,

where:

S_i = the seasonal index for i^{th} period,
D_i = the average values of i^{th} period,
D = grand average,
i = the i^{th} seasonal period of the cycle.

A seasonal index of 1.00 for a particular month indicates that the expected value of that month is 1/12 of the overall average. A seasonal index of 1.25 indicates that the expected value for that month is 25% greater than 1/12 of the overall average. A seasonal index of 80 indicates that the expected value for that month is 20% less than 1/12 of the overall average. A summary of Vintage Restaurant food and beverage sales calculations are as shown below (See Appendix for detail of calculations):

Month	Seasonal index
January	144.36%
February	129.97%
March	134.41%
April	104.12%
May	104.94%
June	80.04%
July	82.83%
August	85.30%
September	62.80%
October	70.03%
November	85.28%
December	115.93%

As is clear, January, February and March have high seasonal indices, whereas June, September and October have low seasonal indices. Clearly, since the business is located in Florida, there is increased tourist traffic in the mild months of January, February and March. In the peak summer seasons in May, June, and the subsequent rains falling in June to October, the tourist traffic is declined. Hence the seasonal indices give quite an intuition into forecasting future predictions, as is evidenced by the following test of seasonality.

Test of Seasonality:

We conduct the test for the presence of any significant seasonal component in a given time series using its seasonal index vector.

$\mathbf{H_0}$: There is no significant seasonal component in the time series.

$\mathbf{H_A}$: The null hypothesis is false.

Chi Square Value: 69.72245

p-value = 0

Conclusion:

There exists very strong evidence against the null hypothesis; hence there is significant amount of seasonal component in the time series, as was evidenced by intuition.

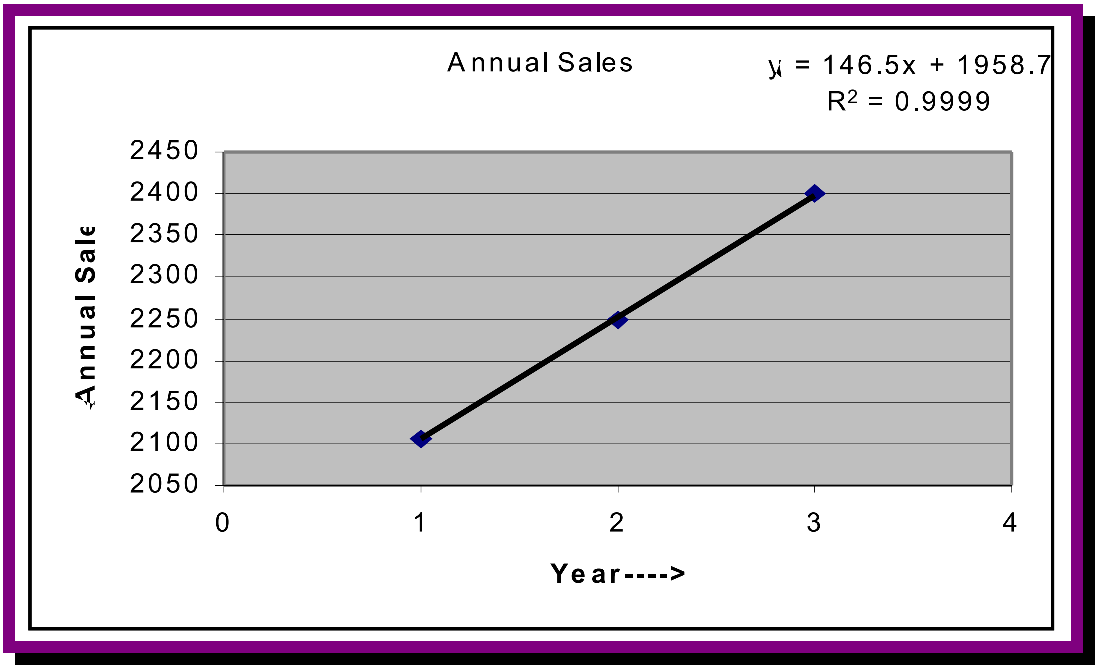
FORECASTING

Incorporating seasonality in a forecast is useful when the time series has both trend and seasonal components. The final step in the forecast is to use the seasonal index to adjust the trend projection. One simple way to forecast using a seasonal adjustment is to use a seasonal factor in combination with an appropriate underlying trend of total value of cycles. To forecast January to December sales for the fourth year, we proceed as follows:

1. We calculate the annual sales for each year:

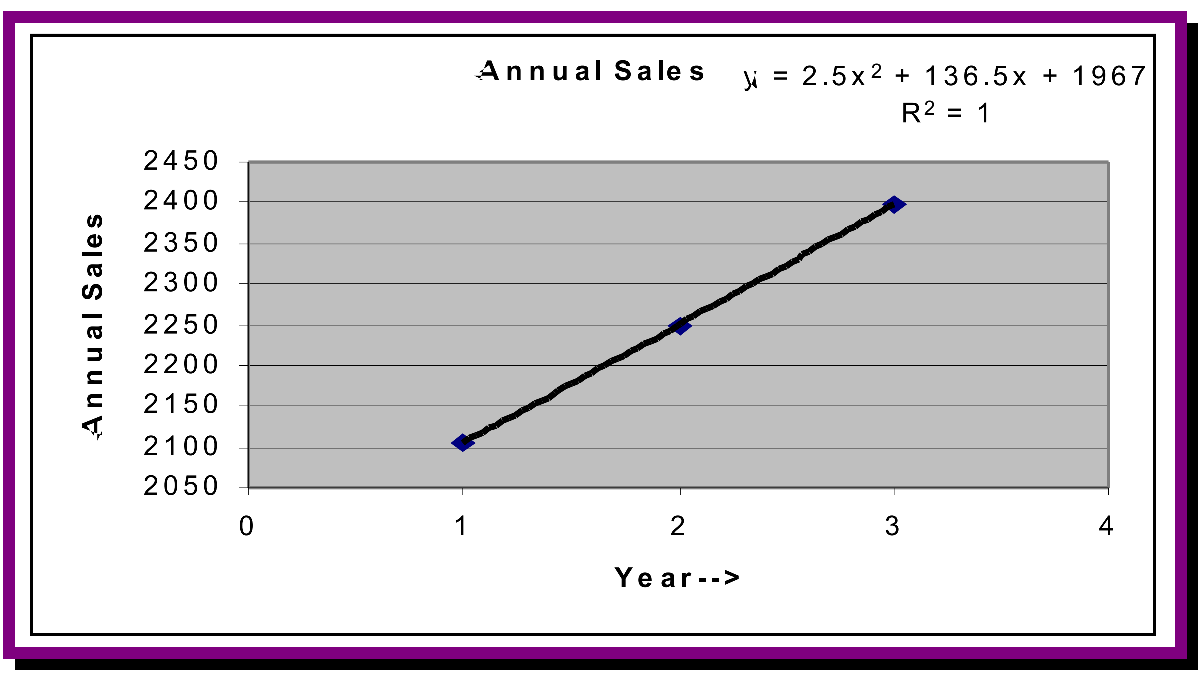
Lost Beverage and Food Sales Case
Vintage Restaurant Sales

MONTH	First Year	Second Year	Third Year
January	242	263	282
February	235	238	255
March	232	247	265
April	178	193	205
May	184	193	210
June	140	149	160
July	145	157	166
August	152	161	174
September	110	122	126
October	130	130	148
November	152	167	173
December	206	230	235
Annual Sales:	**2106**	**2250**	**2399**



Y = 146.5T + 1958.7

The main question is whether this equation represents the trend. Often fitting a straight line to the seasonal data is misleading. By constructing the scatter-diagram, we compare whether a straight line or a parabola would fit better. The estimated quadratic tend is:



$$Y = 2.5T^2 + 136.5T + 1967$$

By observation, a linear trend might be fair enough to give a suitable prediction.

For T = 5, we get Y = 146.5*4 + 1958.7 = 2544.7. The average monthly sale during next year, therefore, is 2544.7/12 = 212.06

Finally, the forecast for each month of the month is calculated by multiplying the average monthly sales forecast by the seasonal index. The table is as below:

Month	Seasonal Index	Forecast for 4th year
January	1.398	296.46
February	1.293	274.19
March	1.322	280.34
April	1.023	216.94
May	1.043	221.18
June	0.798	169.22
July	0.831	176.22
August	0.865	183.43
September	0.636	134.87
October	0.725	153.74
November	0.874	185.34
December	1.192	252.78

Now assuming that the January sales for the fourth year turn out to be $295,000, the forecast error would be: $295000 - 252780 = 42220$.

Since a forecast error of $42,220 is quite large, Karen might be puzzled about the difference between your forecast and the actual sales value. The cycles can be easily studied if the trend itself is removed. This is done by expressing each actual value in the time series as a percentage of the calculated trend for the same date. The resulting time series has no trend, but oscillates around a central value of 100.

A variety of factors are likely influencing data. It is very important in the study that these different influences or components be separated or decomposed out of the 'raw' data levels. Decomposition Analysis is the pattern generated by the time series and not necessarily the individual data values that offers to the manager who is an observer, a planner, or a controller of the system. Therefore, the Decomposition Analysis is used to identify several patterns that appear simultaneously in a time series. In general, there are four types of components in time series analysis: Seasonality, Trend, Cycling and Irregularity.

$$X_t = S_t \cdot T_t \cdot C_t \cdot I$$

The first three components are deterministic and are called "Signals", while the last component is a random variable, "Noise". To be able to make a proper forecast, we must know to what extent each component is present in the data. Hence, to understand and measure these components, the forecast procedure involves initially removing the component effects from the data (decomposition). After the effects are measured, making a forecast involves putting back the components on forecast estimates (recomposition). Pursuing this can help Karen take a proper judgment.

APPENDIX

Step 1

Calculate the 12 month centred moving average

1	2	3	4	5	6	7
Year	**Months**	**Sales**	**12 month moving total**	**12 month moving average= col (4) /12**	**12 month centered moving average**	**% of Actual to moving average values col (3)/(6)**
1	January	242				
	February	235				
	March	232				
	April	178				
	May	184				
	June	140	2106	175.5		
	July	145	2127	177.25	176.38	82.21%
	August	152	2130	177.5	177.38	85.69%
	September	110	2145	178.75	178.13	61.75%
	October	130	2160	180	179.38	72.47%
	November	152	2169	180.75	180.38	84.27%
	December	206	2178	181.5	181.13	113.73%
2	January	263	2190	182.5	182	144.51%
	February	238	2199	183.25	182.88	130.14%
	March	247	2211	184.25	183.75	134.42%
	April	193	2211	184.25	184.25	104.75%
	May	193	2226	185.5	184.88	104.39%
	June	149	2250	187.5	186.5	79.89%
	July	157	2269	189.08	188.29	83.38%
	August	161	2286	190.5	189.79	84.83%
	September	122	2304	192	191.25	63.79%
	October	130	2316	193	192.5	67.53%
	November	167	2169	180.75	193.71	86.21%
	December	230	2344	195.33	194.88	118.02%
3	January	282	2353	196.08	195.71	144.09%
	February	255	2366	197.17	196.63	129.69%
	March	265	2370	197.5	197.34	134.29%
	April	205	2388	199	198.25	103.40%
	May	210	2394	199.5	199.25	105.40%
	June	160	2399	199.92	199.71	80.12%
	July	166				
	August	174				
	September	126				
	October	148				
	November	173				
	December	235				

Ilustration

176.38	=(175.5+177.25)/2
82.21%	=145/176.38
2106	=242+235+232+178+184+140+145+152+110+130+152+206
2127	=235+232+178+184+140+145+152+110+130+152+206+263

Step 2

Calculate Seasonal index (Index picked up from above table)

Year	January	February	March	April	May	June	July	August	September	October	November	December	
1							82.21%	85.69%	61.75%	72.47%	84.27%	113.73%	
2	144.51%	130.14%	134.42%	104.75%	104.39%	79.89%	83.38%	84.83%	63.79%	67.53%	86.21%	118.02%	
3	144.09%	129.69%	134.29%	103.40%	105.40%	80.12%							
													Total
Mean=	144.30%	129.91%	134.35%	104.08%	104.89%	80.00%	82.80%	85.26%	62.77%	70.00%	85.24%	115.88%	**1199.48%**
Adjusted mean=	144.36%	129.97%	134.41%	104.12%	104.94%	80.04%	82.83%	85.30%	62.80%	70.03%	85.28%	115.93%	**1200.00%**

Illustration

144.36% =1.0004*144.3%

Total of means=	1199.4850%	
This total is to be adjusted to	1200.00%	(for 12 months)
Adjusting factor=	1.0004	=1200%/1199.485%

Step 3
Deseasonalize data

Deseasonalzed Attendance					
Year	Months	Sales	Seasonal index	Deseasonalized Attendance= Attendance/ Seasonal index	
1	January	242	144.36%	168	=242/144.36%
	February	235	129.97%	181	=235/129.97%
	March	232	134.41%	173	=232/134.41%
	April	178	104.12%	171	=178/104.12%
	May	184	104.94%	175	
	June	140	80.04%	175	
	July	145	82.83%	175	
	August	152	85.30%	178	
	September	110	62.80%	175	
	October	130	70.03%	186	
	November	152	85.28%	178	
	December	206	115.93%	178	
2	January	263	144.36%	182	
	February	238	129.97%	183	
	March	247	134.41%	184	
	April	193	104.12%	185	
	May	193	104.94%	184	
	June	149	80.04%	186	
	July	157	82.83%	190	
	August	161	85.30%	189	
	September	122	62.80%	194	
	October	130	70.03%	186	
	November	167	85.28%	196	
	December	230	115.93%	198	
3	January	282	144.36%	195	
	February	255	129.97%	196	
	March	265	134.41%	197	
	April	205	104.12%	197	
	May	210	104.94%	200	
	June	160	80.04%	200	
	July	166	82.83%	200	
	August	174	85.30%	204	
	September	126	62.80%	201	
	October	148	70.03%	211	
	November	173	85.28%	203	
	December	235	115.93%	203	

Step 4
Use Deseasonalize data to predict trend sales for the next 12 months by first calculating the regression equation and then using the equation to predict sales for the next 12 months

Year	Month	Sales
1	January	168
	February	181
	March	173
	April	171
	May	175
	June	175
	July	175
	August	178
	September	175
	October	186
	November	178
	December	178
2	January	182
	February	183
	March	184
	April	185
	May	184
	June	186
	July	190
	August	189
	September	194
	October	186
	November	196
	December	198
3	January	195
	February	196
	March	197
	April	197
	May	200
	June	200
	July	200
	August	204
	September	201
	October	211
	November	203
	December	203

1) Determine regression equation from previous data (months are coded as 1-36 , Jan of year 1= 1, Dec of year 3=36)

Determine regression equation

Regression equation is determined as shown in calculations below

Y=a+b1 X1= **Y= 169.4143 + 1.0181 X**

	Y	X	XY	X^2
	168	1	168	1
	181	2	362	4
	173	3	519	9
	171	4	684	16
	175	5	875	25
	175	6	1050	36
	175	7	1225	49
	178	8	1424	64
	175	9	1575	81
	186	10	1860	100
	178	11	1958	121
	178	12	2136	144
	182	13	2366	169
	183	14	2562	196
	184	15	2760	225
	185	16	2960	256
	184	17	3128	289
	186	18	3348	324
	190	19	3610	361
	189	20	3780	400
	194	21	4074	441
	186	22	4092	484
	196	23	4508	529
	198	24	4752	576
	195	25	4875	625
	196	26	5096	676
	197	27	5319	729
	197	28	5516	784
	200	29	5800	841
	200	30	6000	900
	200	31	6200	961
	204	32	6528	1024
	201	33	6633	1089
	211	34	7174	1156
	203	35	7105	1225
	203	36	7308	1296
Σ=	**6777**	**666**	**129330**	**16206**

n= 36

Calculation of regression coefficients using the equations

$\Sigma Y = na + b1\Sigma X$

$6777 = 36a + 666b1$

$\Sigma XY = a\Sigma X1 + b1\Sigma X^2$

$129330 = 666a + 16206b1$

36	666	a	=	6777
666	16206	b1	=	129330

Use matrix and inverse of matrix to calculate the coefficients

Inverse of the matrix

0.115873016	-0.0047619		6777		a
-0.004761905	0.0002574	X	129330	=	b1

Solving

a= 169.4143
b1= 1.0181

Hence regression equation= Y= **169.4143 + 1.0181 X**

2) Predict Trend Sales for the next 12 months using the regression equation

(x=37 to 48 corresponding to Jan of year 4 to Dec of year 4)

Year	Month	X	Predicted Trend Sales Y	
4	January	37	207.08	=169.4143+1.0181*37
	February	38	208.1	=169.4143+1.0181*38
	March	39	209.12	=169.4143+1.0181*39
	April	40	210.14	
	May	41	211.16	
	June	42	212.17	
	July	43	213.19	
	August	44	214.21	
	September	45	215.23	
	October	46	216.25	
	November	47	217.27	
	December	48	218.28	

Step 5
Multiply the predicted trend sales by the seasonal index to get the Predicted Sales value for the next 12 months

Year	Month	Seasonal index	Predicted Trend Sales	Predicted sales= Seasonal index * Trend sales	
4	January	144.36%	207.08	299	=207.08*144.36%
	February	129.97%	208.1	270	=208.1*129.97%
	March	134.41%	209.12	281	=209.12*134.41%
	April	104.12%	210.14	219	=210.14*104.12%
	May	104.94%	211.16	222	=211.16*104.94%
	June	80.04%	212.17	170	=212.17*80.04%
	July	82.83%	213.19	177	
	August	85.30%	214.21	183	
	September	62.80%	215.23	135	
	October	70.03%	216.25	151	
	November	85.28%	217.27	185	
	December	115.93%	218.28	253	

The monthly forecasts for the 12 months of the fourth year are as shown below:

Month (yr.4)	January	February	March	April	May	June	July	August	September	October	November	December
S. Index	1.398	1.293	1.322	1.023	1.043	0.798	0.831	0.865	0.636	0.725	0.874	1.192
Forecast	296.45755	274.19143	280.3411	216.9357	221.1768	169.2226	176.2205	183.4305	134.8691	153.7423	185.339	252.7735

Suppose the actual January sales for the fourth year turn out to be $295,000. The forecasted January sales are $296,458.
Error between actual and forecasted sales = $296,458 - $295,000 = $1458

Percentage Error = $\frac{Error}{Actual\,Sales} * 100 = \frac{1458}{295000} * 100 = 0.49\%$

This is an extremely small percentage error. Karen does not have to worry about this error and she can be assured that her forecast model is extremely good.

REFERENCES

Bowerman, B. L., & O'Connell, R. T. (2003). *Business Statistics in Practice* (3rd). : McGraw Hill.

Cooper, D. & Schindler, S. (2003). *Business Research Methods*. Boston: McGraw-Hill Irwin.