Predictive sales analysis according to the effect of weather

Alessandro Massaro,
Donato Barbuzzi, Valeria Vitti,
Angelo Galiano
Dyrecta Lab srl
ResearchInstitute,
70014 Conversano (BA), Italy
alessandro.massaro@dyrecta.com

Melisa Aruci
Departmenti i Informatikës
Fakulteti i Shkencave të Natyrës,
Tiranë, Albania
melisa.aruci@gmail.com

Giuseppe Pirlo
Dept. of Informatics
Bari University,
70125 Bari, Italy
giuseppe.pirlo@uniba.it

Abstract

This work presents an empirical simulation to estimate the extent to which weather affects consumer spending. For the spending prediction purpose, the regression technique has been applied on the historical daily data of the following meteorological parameters: temperature, snow, sun, rain and humidity. The experimental results demonstrate the effectives of the model for predicting sales trends and for applying optimal business intelligence rules.

1Introduction

Weather has strong effects on human behavior and a lot of scientific research was devoted to the investigation on the direct link between weather and social activities. For instance, [Coh90a] and [Coh90b] reported that higher temperatures are correlated with increases in violent assaults and homicides. Also [Bar94] and [Sto99] studied the number of suicides related to the barometric pressure and demonstrated their decrease related to the wind.

Weather also influences human behavior in the sales. For instance, we buy warm clothing in winter and cool clothing in summer. Moreover, in the finance field, weather variables can affect human behavior and his mood [Goe05], [Hir03], [Sau93], [Tro97]. Coca-Cola company proposed a dynamic pricing strategy based on changes in the environment temperature, so the vending machines increase the price of a soda as the weather gets hotter [Kin00]. Nevertheless, the effect of weather on consumer spending has received only limited attention in the marketing literature [Par00] [Par01] [Ste51].

Our approach is consistent with the aforementioned studies, which analyzes the direct link between the weather variables and the customers' behavior in order to predict the sales of a particular product in the next future. This is important to establish optimal business intelligence rules, i.e. for warehouse management. The work begins with an analysis of daily sales in one shoe store, which have an effect onweather forecasts. The paper is organized as follows: Section 2 reports an overview on the effects that weather can have on consumer behavior. Section 3 describes the technique for sales prediction based on Regression; Section 4 presents the experimental results. The conclusion of the paper and some of the most interesting future perspectives are reported in Section 5.

2 Weather effects: An overview

In literature, the weather effects have been investigated for the consumer behavior analysis. Three different consumer categories, affected by: (1) bad weather; (2) seasons and (3) his moods, were considered.

In the first category (1), the rain, the snow and extreme temperature *keep people at home*. In this case,the weather negatively affects both sales and store traffic[Par01].An intelligent marketing solution would be to focus on online sales. The second set of consumers (2) influences both sales volume and store traffic in particular product categories [Mau95]. For example, when temperatures fall, ice cream sales decrease, while sales of oatmeal increase. Similarity, people tend to purchase more clothing and footwear in the winter and more food and drinks in the summer [Ros00].Finally, it has been suggested that weather can influence sales by affecting consumers' internal states (3). Although there is very little research forthis third category of effects, few studies have providedpreliminarysupport for this idea. In [Par00], the authors present a global climate-based model of the

effect of weather on consumer behavior, which predicts variation in consumption patterns in response to different temperatures and exposure to sunlight. They argue that consumers do adapt to changes in the environment by modifying their purchasing behavior to both maintain physiological homeostasis and to achieve optimal stimulation levels. In [Mur10], the authors report an empirical evidence of how weather can impact consumer spending. In particular, they claimed that:

- Weather variables and, sunlight in particular, affect consumer spending;
- Exposure to sunlight reduces negative affect;
- As negative affect decreases, consumer spending increases;
- Negative affect mediates the effect of sunlight on consumer spending.

Their research demonstrates a cause-and-effect relationship between exposure to sunlight and an increased willingness to pay for common product.

3Sales Predictive Analysis by Regression

In this section, we will provide some aspects concerning the approach used for sales prediction based on weather effect.
For the purpose, we consider the empirical model for the iterative calculation of the regression coefficients:

$$Sales_{ij}=a_{ij}+b_1Temp_{ij}+b_2Snow_{ij}+b_3Sun_{ij}+b_4Rain_{ij}+ +b_5Humid_{ij}+b_6Sun_{ij}Temp_{ij}+b_7Temp^2_{ij}+e_{ij}(1)$$

Table 1 shows the averagemeteorological values,in February 2010, related to:

- Temperature;
- Snow;
- Sun;
- Rain;
- Humidity.

Itreports thesimulated sales data in that day.

Table 1: Meteorological and sales data for the calculation of regression coefficients

Feb 2010	Sales	Temp	Snow	Sun	Rain	Humid	TempSun	$Temp^2$
1	7	8,7	0	66,75	14,79	69,54	573,92	77,21
3	12	7,1	0	72,5	0	57,38	493,58	56,17
4	15	9,13	0	88,46	0	74,75	796,54	85,36
8	3	8,37	0	35,67	39,38	79,33	312,33	70,71
10	1	13,08	0	29	55	83,38	431,71	176
11	2	10,25	0	48,67	42,08	76,83	492,17	105,5
12	10	9,71	0	72,42	8,13	69,79	687,21	96,96
13	6	7,29	0	59,21	1,88	68,46	433,17	54,04
15	3	9,08	0	46,04	22,29	81,75	429,21	83,17
16	5	11,83	0	21,13	42,29	82,79	209,42	143,25
17	3	13	0	37,04	51,25	87,75	502,42	169,75
19	14	14,96	0	79,21	1,25	78,71	1209,29	237,38
22	3	10,83	0	44	41,04	81,63	469,79	119,67
23	8	13,25	0	70,13	21,25	80,5	928,38	180,25
24	2	13,25	0	55,33	82,5	85,21	752,83	177,75
25	10	12	0	72,88	12,08	82,54	879,25	144,58
26	12	12,96	0	76,58	6,88	78,25	1002,38	176,38

Starting from the value of Table 1, the regression coefficients have been computed with Microsoft Excel tool. So, Table 2 shows thisspecific information. More specifically, it reports:

- **Coefficient**. This column shows the least squares estimate.
- **Standard Error**. This column shows the least squares estimate of the standard error.
- **T Statistic**. This column shows the T Statistic for the null hypothesis vs. the alternate hypothesis.
- **P Value**. This column shows the p-value for the hypothesis test.
- **Lower 95%**. This column shows the lower boundary for the confidence interval.
- **Upper 95%**. This column shows the upper boundary for the confidence interval.

Table 2: Calculation of regression coefficients

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercepts	-13,28	11,54	-1,15	0,28	-39,38	12,82
x_1	-1,68	2,26	-0,75	0,47	-6,79	3,42
x_2	0,00	0,00	65535	#NUM!	0,00	0,00
x_3	0,44	0,15	2,97	#NUM!	0,11	0,78
x_4	-0,11	0,03	-3,47	0,01	-0,18	-0,04
x_5	0,17	0,17	1,03	0,33	-0,21	0,55
x_6	-0,03	0,01	-2,38	0,04	-0,06	0,00

x_7	0,16	0,10	1,57	0,15	-0,07	0,40

The formula (1) becomes:

$Sales_{ij}$=-13,28-1,68x_1+0x_2+0.44x_3-0.11x_4+0.17x_5--0.03x_6+0.16x_7+1.67(2)

Table 3shows regression statistics. These measures show how well the calculated regression equation fits data. In particular:

- **Multiple R**. This is the correlation coefficient. It explains how strong the relationship is. A value of 1 means a perfect positive relationship and a value of zero means no relationship at all. It is the square root of r squared.
- **R squared**. This is *rxr*, the Coefficient of Determination. It explains how many points fall on the regression line.The best value for this measure is 80%. It means that 80% of the variation of y-values around the mean is explained by the x-values.
- **Adjusted R squared**. The adjusted R-square adjusts for the number of terms in a model. Users have to use this measure instead of R squared, if there are more than one x variable.
- **Standard Error of the regression**.This measure is an estimate of the standard deviation of the error μ. This is not the same as the standard error in descriptive statistics. The standard error of the regression is the precision that the regression coefficient is measured; if the coefficient is large compared to the standard error, then the coefficient is probably different from 0.
- **Observations**. Number of observations in the sample.

Table 3: Statistical Regression Performance

Regression Statistics	
Multiple R	0,962117902
R Squared	0,925670858
Adjusted R squared	0,867859304
Standard Error of regression	1,666985637
Observations	17

Table 4 describes the ANOVA parameters. An ANOVA test is a way to find out if survey or experiment results are significant. More specifically, these parameters are used to accept or to reject the null hypothesis for the alternate hypothesis. In particular:

- **df**. This measure explains degrees of freedom.
- **SS**. This measure explains the sum of squares.
- **Regression MS**. This measure is computed dividing Regression SS with Regression degrees of freedom.
- **Residual MS**. This measure is computed dividing Residual SS with Residual degrees of freedom and it explains the mean squared error.
- **F**. This measure explains overall F test for the null hypothesis.
- **Significance F**. This measure explains the significance associated P-Value.

Table 4: ANOVA Analysis

ANOVA Analysis					
	df	**SS**	**MS**	**F**	**Significance F**
Regression	7	311,4610	44,4944	16,0119	0,0002
Residual	9	25,0096	2,7788		
Total	16	336,4706			

4 Prediction Results

Figure 1 shows the sales prediction, in the next fifteen days, concerning a store (Store 1)located in Conversano (BA). Figure 2 and Figure 3 report weather forecast in February 2016 related to the city.

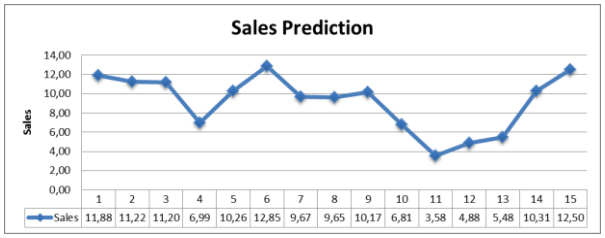

Figure 1: Sales prediction in Store 1

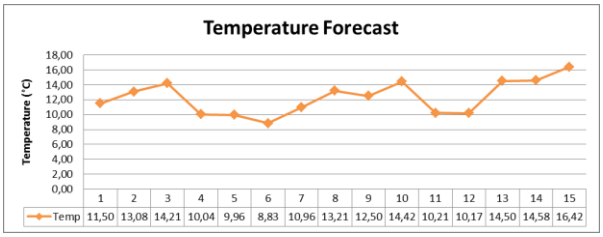

Figure 2: Temperature Forecast in Conversano (BA)

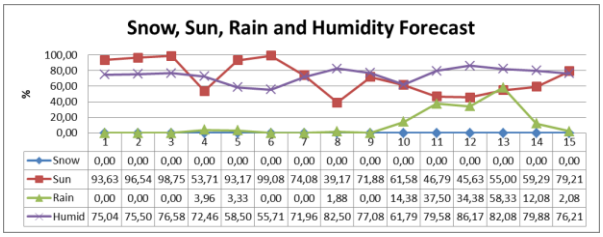

Figure 3: Sun, Snow, Rain and Humidity Forecast in Conversano (BA)

Figure 4 shows the sales prediction in Store 2, located in Milano. Figure 5 and Figure 6 report weather forecast in February 2016 related to the city.

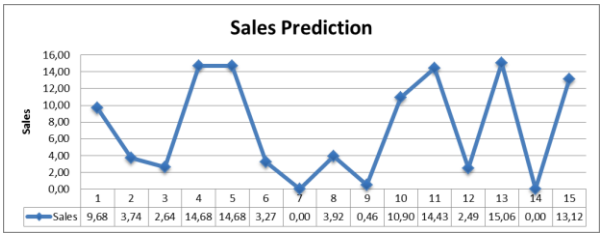

Figure 4: Sales prediction in Store 2

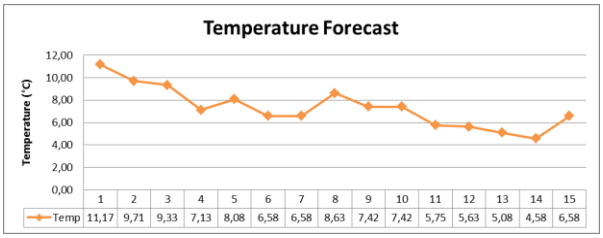

Figure 5: Temperature Forecast in Milano

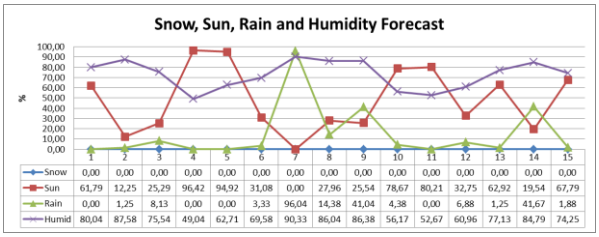

Figure 6: Sun, Snow, Rain and Humidity Forecast in Milano

Data of Figure 1 andFigure 4 have been calculated respect to the Equation (2). More specifically, the weather forecasts in Conversano (BA) and Milano have been used to draw the trend line, respectively.

5 Conclusion

A predicted model based on weather variables has been introduced in order to optimize business intelligence rules. The empirical simulation shows the direct link between the temperature, the snow, the sun, the rain and the humidity and the customers' behavior in order to predict the sales of a particular fashion product in the next future**.** Finally, the sales prediction for two weeks has been calculated respect to the historical data in the month. In the future other models such as the "Decision Tree" in the Data Mining field will be compared with the regression approach.

Acknowledgments

The work has been developed within the framework of the project titled: *Sistema di Business Intelligence di tipo "embedded" basatosu prediction real time in sistemi Big Data orientate al settore di calzature e accessorimoda "Predishoes"(Business Intelligence Embedded Systembased on real time prediction in BigData systems oriented on fashion"Predishoes").*

References

[Goe05] Goetzmann, W.N., Zhu, N., 2005. Rain or shine: where is the weather effect? European Financial Management 11 (November) 559–578

[Hir03] Hirshleifer, D., Shumway, T., 2003. Good day sunshine: stock returns and the weather. Journal of Finance 58 (June), 1009–1032.

[Sau93] Saunders, E.M., 1993. Stock prices and wall street weather. The American Economic Review 83 (December), 1337–1345.

[Tro97] Trombley, M.A., 1997. Stock price and wall street weather: additional evidence. Quarterly Journal of Business and Economics 36 (Summer), 11–21.

[Coh90a] Cohn, E.G., 1990a.Weather and crime. British Journal of Criminology 30 (1), 51–64.

[Coh90b] Cohn, E.G., 1990b.Weather and violent crime. Environment and Behavior 22 (March), 28–94

[Bar94] Barker, A., Hawton, K., Fagg, J., Jennison, C., 1994.Seasonal and weather factors in parasuicide. British Journal of Psychiatry 165 (3), 375–380.

[Sto99] Stoupel, E., Abramson, E., Sulkes, J., 1999. The effect of environmental physical influences on suicide. How long is the delay? Archives of Suicide Research 5 (September) 241–244.

[Kin00] King, C., Narayandas, D., 2000. Coca-cola's new vending machine (A): pricing to capture value, or not? Harvard Business School Case # 9-500-068.

[Par00] Parker, P.M., Tavassoli, N.T., 2000.Homeostasis and consumer behavior across cultures. International Journal of Research in Marketing 17 (March), 33–53.

[Par01] Parsons, A.G., 2001. The association between daily weather and daily shopping patterns. Australasian Marketing Journal 9 (2), 78–84

[Ste51] Steele, A.T., 1951. Weather's effect on the sales of a department store. Journal of Marketing 15 (April), 436–443.

[Mur10] Murray, K. B., Di Muro, F., Finn, A., &Leszczyc, P. P. (2010). The effect of weather on consumer spending.*Journal of Retailing and Consumer Services*,*17*(6), 512-520.

[Mau95] Agnew, Maureen D., and John E. Thornes."The weather sensitivity of the UK food retail and distribution industry."*Meteorological Applications* 2.2 (1995): 137-147.

[Ros00] Roslow, S., Li, T., & Nicholls, J. A. F. (2000).Impact of situational variables and demographic attributes in two seasons on purchase behaviour.*European Journal of Marketing*, *34*(9/10), 1167-1180.