Correlation Between Weather & Sales for Businesses in the United Kingdom

Swaraag Sistla

July 2024

1 Abstract

There exists a well-documented correlation between ice cream sales and temperature: sales rise with increasing temperatures until reaching a peak, after which they decline as consumers shift to cold beverages. However, it is unclear if this trend applies to ice lollies (popsicles), a hybrid offering between ice cream and cold beverages. We undertake a study of United Kingdom ice lolly enterprise Ice Kitchen, and evaluate sales data from 97 businesses of various types (desserts, groceries, frozen goods, etc.) across the United Kingdom to examine how sales correlate with temperatures in London. By applying linear regression to compute normalized slopes and R^2 scores for each business's sales data, and applying thresholds of 0.03 for slopes and 0.4 for R^2 scores, we identified robust correlations in 18 businesses including Ice Kitchen. Our analysis reveals a nonlinear relationship for these primarily frozen-food businesses—the correlation between sales and temperature strengthens as temperatures rise. Unlike ice cream sales, which decline at certain high temperatures, ice lolly sales continue to increase with progressively higher temperatures. In particular, Ice Kitchen's data show a strong correlation between sales and temperatures above 17°C, while the data across the 18 correlated businesses reveal a mean threshold of 16.39°C.

2 Introduction

On a hot day, ice cream is a popular purchase, which is reflected in how sales rise alongside temperatures. This relationship eventually weakens when the temperature becomes 'too hot' for ice cream, as customers prefer cold drinks instead (Ziady, 2023). To determine whether this phenomenon applies to other hot-day choices, we analyzed sales data from 97 United Kingdom-based businesses, focusing on a young and independent ice lolly business called Ice Kitchen. Ice Kitchen (2014-present), has sold ice lollies for numerous years and has recently pivoted by introducing frozen sauces and cooking pastes. Ice lollies, being a fusion of cold drinks (ice cream's high-temperature alternative) and ice cream, raise a captivating question: how are their sales impacted as temperatures continue to rise?

This research aims to benefit Ice Kitchen, the focus of our analysis, by allowing them to predict demand using weather forecasts, allowing for precise supply preparation. The analysis investigates the varying levels of correlation between temperature and sales for different temperatures, highlighting the non-linear nature of this relationship. For instance, sales may share a low correlation with temperatures between 0°C and 15°C but substantially increase between 15°C and 30°C. Understanding this relationship will help Ice Kitchen predict and prepare for demand more accurately.

Additionally, optimizing marketing strategies based on temperature ranges that will entice more customers can improve efficiency and reduce unnecessary marketing costs.

Therefore, in doing this research, there were four goals in mind:

1. to determine the impact of temperature on sales
2. to identify the temperature at which sales reliably correlate with temperature
3. to assess if Ice Kitchen's sales decline at very high temperatures like ice cream
4. to provide insights to Ice Kitchen about its and its competitors' businesses.

To achieve these objectives, we preprocessed sales data from Ice Kitchen and weather data from London. After identifying a clear correlation, we expanded the analysis to 96 other businesses and investigated the relationship between their sales and the temperature.

3 Methods

This research began with the collection of data from 97 businesses and their weekly sales from July 5, 2021, to July 17, 2023, and Weather Underground's dataset of Charing Cross, London weather from November 8, 2021, to July 17, 2023. Charing Cross was chosen as a representative sub-locale of London, as it is widely regarded as the city's center. Weather Underground's data includes values for temperature, dew point, humidity, wind, speed, gust, pressure, precipitation rate, precipitation accumulation, UV, and solar data of Charing Cross in five-minute intervals.

The first leg of the analysis consisted only of Ice Kitchen data. Sales data were normalized by dividing each sales data point by the dataset's maximum, regularizing all values such that they were between 0 and 1. Weather data were pre-processed, but not normalized.

To account for differences in both date ranges and increments between sales and temperature data, the datasets were re-indexed and re-produced for further analysis and plotting. To quantify the correlation beyond visual inspection, a separate plot was created where each data point consisted of a temperature x-coordinate and a slope y-coordinate. These slopes were computed by a scikit-learn linear regression model using a 10°C rolling window. The R^2 scores of each data partition were also plotted on a twinned x-axis to monitor the strength of the correlation.

This process was repeated with the sales data of 96 of Ice Kitchen's competitors: plotting their sales against Charing Cross temperatures and examining the changes in the correlation across various temperature ranges.

In order to identify reliable and strong sales-temperature correlations, thresholds for R^2 scores and slopes—above which the correlation would be deemed sufficiently strong and reliable—were selected, according to the following procedure:

The R^2 score threshold was set after observing that, in many of the slope and R^2 score graphs, the increase in R^2 scores over temperature eventually plateaued. The chosen R^2 score threshold was 0.4, which signifies that 40% of the variability in sales could be explained by the temperature. This threshold was considered a significant benchmark for this analysis.

The slope threshold was set after noting a trend: slope over temperature initially increased at a slow rate, before transitioning to a much steeper rate of growth. This transition was consistently observed around a slope of 0.03 (this is a unitless value because the slope computation involved normalized sales data), so this was selected as the threshold. This pattern reflects the non-linearity of the correlation, as the correlation itself increased along with the temperature.

4 Results

Of the 97 businesses analyzed, 18 possessed slopes and R^2 scores that passed and stayed above the determined thresholds.

The graph of Ice Kitchen's sales versus time is shown in Figure 1.

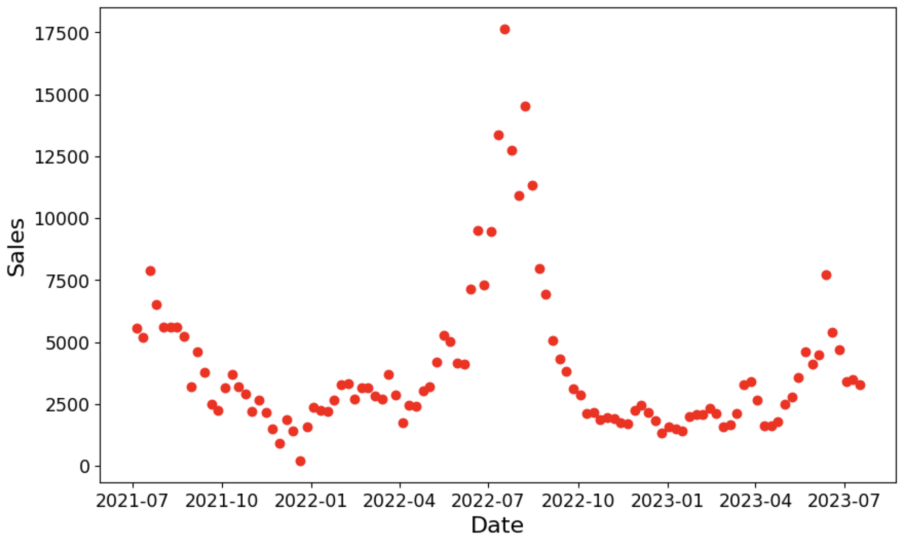

Figure 1: A sum of weekly sales of Ice Kitchen as a function of time.

There is a sizable increase in sales during the summer months of all three years of collected data, with an especially large peak in the summer of 2022. Sales are at their lowest in the first 3–5 months of the year before beginning a gradual assent in May.

Figure 2 shows normalized data of the cumulative sales in a week graphed against the average temperature of that week.

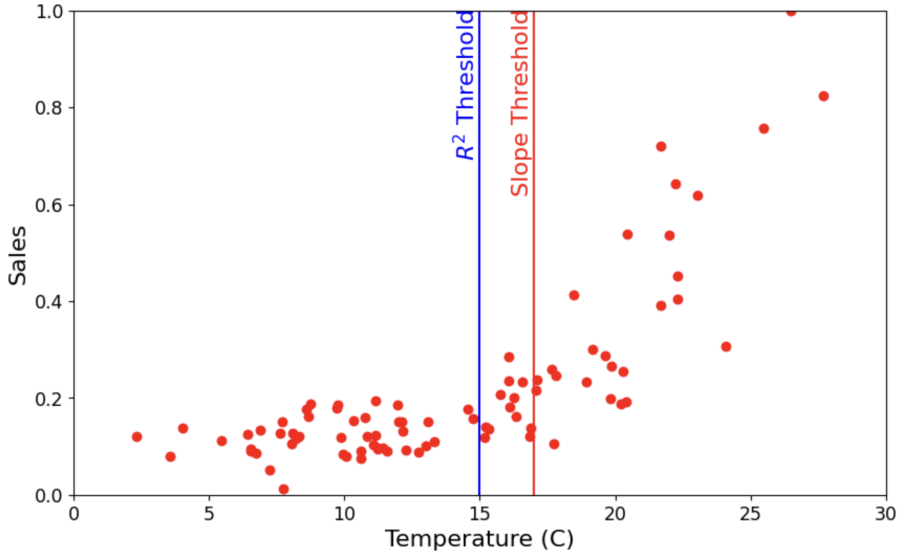

Figure 2: Normalized sales of Ice Kitchen as a function of temperature.

At roughly 15°C and above, sales begin to rise more rapidly. The thresholds for R^2 scores and slopes are depicted as vertical lines on this graph: they show the temperature values at which the slopes and R^2 scores in Figure 3 cross the thresholds.

Figure 3 represents a running derivative twinned with an R^2 score graph. Each red data point represents the slope of a linear regression in a 10°C interval (5°C in both directions of the x value of the data point), and each blue data point is the slope's associated R^2 score. The red data points use the left y-axis, and the blue data points follow the right.

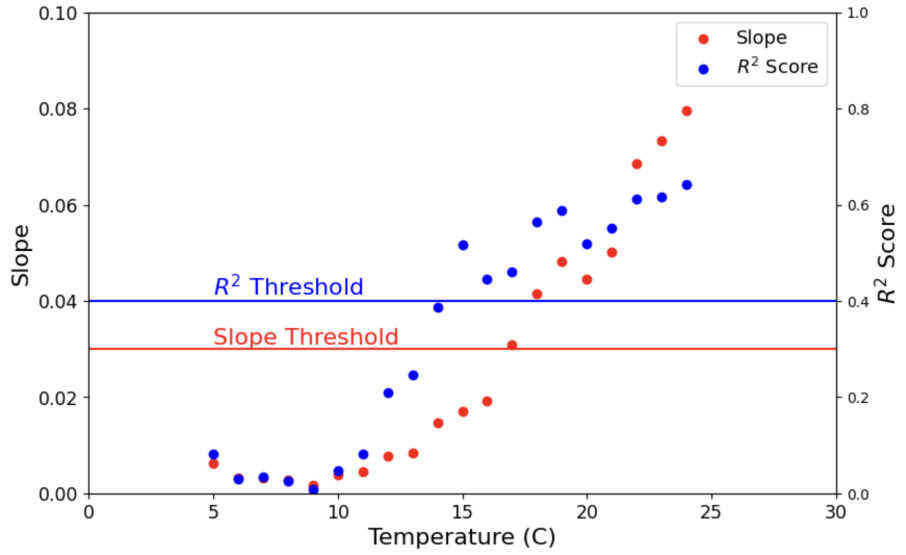

Figure 3: Ice Kitchen slope of sales and R^2 scores as a function of temperature.

The slope and R^2 score thresholds are placed as horizontal lines on this graph at $y = 0.03$ and $y = 0.4$ respectively. The slopes are unitless.

After repeating this process over the 18 businesses, we had 36 graphs (a sales versus temperature graph and its slope/R^2 score graph for each). The values at which the thresholds are crossed for each business are summarized in Table 1.

Table 1 shows the results of repeating this process on all 18 businesses. 21°C is the highest temperature value at which the slope threshold is crossed, and 12°C is the lowest. The mean of slope thresholds is 15.6°C, the median is 15°C, the mode is 14°C, and the range is 9°C. The minimum and maximum temperature values at which the R^2 score threshold is crossed are the same as for the slope threshold: 12°C and 21°C. The mean is 15.1°C, the median is 14.5°C, the mode is 12°C, and the range is 9°C.

Taking the highest of the two temperature values for each business results in Table 2.

5 Discussion

Ice Kitchen's sales begin correlating highly with the temperature at 17°C, and for the other 17 businesses with slopes and R^2 scores consistently above their respective thresholds, their sales began to correlate highly with the temperature at a mean of 16.35°C.

Table 1: The temperature values at which the slopes and R^2 scores of a business's sales cross their respective thresholds (0.03 and 0.4)

Business	**Slope Threshold Crossed**	**R^2 Threshold Crossed**
Pip Organic	14°C	12°C
Magnum	14°C	12°C
The Coconut Collab	12°C	14°C
Nestlé	14°C	14°C
R. White's	14°C	14°C
Yoomoo	14°C	14°C
Rowntree's	15°C	12°C
Cornetto	15°C	12°C
Remeo	16°C	15°C
Nuii	16°C	16°C
Wall's	15°C	16°C
Jude's	15°C	16°C
Ice Kitchen	17°C	15°C
LICKALIX	17°C	15°C
The Ice Co	16°C	20°C
Smooze	19°C	21°C
Northern Bloc	17°C	21°C
Del Monte	21°C	12°C

Table 2: The temperature value at which both the slopes and R^2 scores of a business's sales cross their respective thresholds (0.03 and 0.4)

Business	**Slope/R^2 Threshold Crossed**
Pip Organic	14°C
Magnum	14°C
The Coconut Collab	14°C
Nestlé	14°C
R. White's	14°C
Yoomoo	14°C
Rowntree's	15°C
Cornetto	15°C
Remeo	16°C
Nuii	16°C
Wall's	16°C
Jude's	16°C
Ice Kitchen	17°C
LICKALIX	17°C
The Ice Co	20°C
Smooze	21°C
Northern Bloc	21°C
Del Monte	21°C

The sales of frozen goods in general have been shown to possess a strong positive correlation with temperature. The British Retail Consortium showed that weather is incredibly influential on the behavior of consumers, second only to the region's economic state (Angell, 2023). Our results for these 18 businesses corroborate those claims.

Our results also indicate that the slopes and R^2 scores of these businesses' sales stay above the two thresholds even at high temperatures. This suggests that sales continue to increase along with the temperature, even as upper-temperature bounds are reached. This starkly contrasts ice cream sales, which plateau after the temperature reaches a certain point.

For the other 79 businesses, we did not find a correlation between their sales and the temperature. Other factors influence these correlations, causing noticeable imperfections in the data. For instance, if a business is located elsewhere in the United Kingdom, its local temperature may not match the temperature used for this analysis (Charing Cross). Additionally, for numerous businesses, there was not enough data available to judge the correlation between their sales and the temperature. Furthermore, some businesses may have multiple locations or sell retail in various climates, also contributing to potential inaccuracy. Finally, some of the businesses that we analyzed sell more than just food (such as clothing), which tallies into their total sales.

Figure 1 allows for visualization of Ice Kitchen's sales over time. The summer months witness the highest sales, which is consistent with intuition, as a hotter climate associated with the summer months may encourage higher ice lolly sales.

Figure 2 displays a clear relationship between the temperature and sales of Ice Kitchen. Sales stay low before roughly 15°C, where they begin consistently increasing. As the temperature continues to climb, sales follow suit.

Figure 3 adds more detail to the analysis of sales and the temperature, as the slope of the sales versus temperature begins to rise above the threshold at 17°C. This represents the temperature at which sales begin to show a consistent correlation with temperature. With lower values, small changes in temperature have too little of an effect on sales, whereas changes in temperature values above this threshold strongly affect sales. After the R^2 score threshold is crossed at 15°C, the slope values can be considered meaningful, as the linear regression can predict the relationship between sales and the temperature with strong accuracy.

6 Conclusion

In order to determine at what temperature a business begins to witness sales that have a strong correlation with temperature, we analyzed sales data from Ice Kitchen and temperature data from London. After finding a noticeable correlation between the business's sales and London's temperature, we analyzed the sales of other businesses and found strong correlations for 18 of the 97 using slope and R^2 score thresholds. Inaccuracy prevented other potential businesses from consistently crossing our chosen thresholds.

For each business, two temperature values were selected at the points at which the thresholds were crossed. As our analysis shows, sales data with a threshold-crossing slope but a low R^2 score—or threshold-crossing R^2 score and a low slope—both do not suggest a reliable slope of sales versus temperature. Thus, we can conclude that the temperature value at which a business begins having reliable slopes is only when both the slope and R^2 score thresholds are crossed. Therefore, it is 17°C at which Ice Kitchen begins seeing a reliably high slope of sales versus temperature. For the rest of the 18 businesses, the temperature value at which their sales have a reliably strong correlation with temperature is displayed in Table 2.

7 References

Angell, A. (2023). *How weather impacts consumer behaviour... beyond increased ice-cream sales on a Hot Day.* SHIFT. https://www.shiftconsultancy.co.uk/how-weather-impacts-consumer-behaviour-beyond-increased-ice-cream-sales-on-a-hot-day/

Ziady, H. (2023, July 26). The world's biggest ice cream company says heat waves may not help sales — CNN business. CNN. https://www.cnn.com/2023/07/25/business-food/unilever-heat-wave-ice- cream-sales/index.html