

datasciencemidterm

December 12, 2023

1 BSc Computer Science CM3005 (Data Science) Midterm Coursework on Linear Regression

1.1 Task 1: Domain-Specific Area and Objectives of the Project

“Weather has had a greater effect on sales than economic numbers, we’ve known that forever.” [Andy Street, John Lewis Managing Director, 2014*](#)

1.1.1 (a) Introducing the Domain: *Quantifying the Influence of Weather on Retail Sales in the UK*

This is a question aimed at anyone living in typically cool and rainy climate, like the UK: *Have you ever experienced the frustration and annoyance of going clothes shopping during a particularly blustery and wet February or March and finding that all the fashion collections are full of spring-themed floral, summery dresses and floaty cardigans, when all you want is snugly quilted parka or a wooly jumper?* It’s hard to believe that spring could be just around the corner when it’s sleeting outside - which continue for a couple more weeks if not months! As a Pole, my personal experience of shopping in “springtime” in Poland has involved seeing Spanish-owned international apparel retail companies such as Zara and Bershka selling their springtime collections in March in a country which frequently sees snow at the beginning of May. Personally, I have always wondered how clothing companies have managed to make money during these transition periods! As such, the purpose of this data analysis is to explore and quantify the impact of the weather on retail sales in the UK by implementing several statistical techniques as well as a linear regression model (in Python). Although it has been widely acknowledged in both economic research and the media that the weather has an important effect on consumer spending patterns and daily business activity, so far it has proven very difficult to determine precisely what the nature of this relationship is ([WeatherAds.io](#)). For instance, this *Guardian* article suggested that wet and cold weather during a specific month led to an unexpected fall in retail sales ([The Guardian: Wet weather dampens retail sales in Great Britain](#)), while this one on *Euronews* blamed the **above-average** temperatures for a decline in consumer spending ([Euronews: Cost of living and hot weather cause retail sales to fall in UK](#)). One could argue that warm weather would make consumers less likely to go shopping as they would engage more in sports and other outdoor activities, while on the other hand, one article posits that > “It has been proven that after exposure to sunlight consumers are more willing to self-reward (i.e. make purchases), and are also willing to pay more for goods and services. It has also been shown that when consumers are in a warm environment, their purchase intent increases, and they judge the value of certain items higher. In addition the same study found that warmth leads consumers to be more easily swayed by marketing messages.” ([WeatherAds.io](#))

Despite its complexity, it is undoubtable that examining the nature of this relationship more closely

be of value for a range of different stakeholders, including retail company managers, shareholders, governmental institutions and potential investors. As [Badorf and Hoberg](#) outline in their 2020 study on the impact of daily weather on sales in 673 German retail stores, weather “has an immediate effect on daily businesses and sales”, and precipitation and temperature “significantly influence people’s behavior”, although it is “unclear how”.

What makes the investigation of this relationship particularly topical and timely is that in recent years, the UK, together with other areas of north-west Europe, has experienced increasingly warm and dry summer temperatures due to the effects of climate change, with temperatures of over 40C being reached during summer 2022 in the UK for the first time ([UNDRR PreventionWeb](#)). Establishing the effect of hot/dry weather in the summer on retail sales is especially important in this context where this trend is expect to continue due to global warming. For example, shop owners might decide to invest in air-conditioning or different building materials if it the data suggests that temperatures over a certain value discourage people from browsing in their stores.

1.1.2 (b) Project Objectives

- To collect, pre-process, and statistically summarize datasets containing historic UK retail prices and weather data.
- To use visualization libraries in Python to get an overview of the important trends, patterns and relationships within the data.
- To determine the effect of weather on UK retail sales using statistical analysis and linear regression.
- To determine which weather parameters (e.g temperature, wind, precipitation) have the greatest effect on sales.
- To determine if a polynomial regression model should be used in case the relationship is not linear.

1.1.3 (c) Proposed Impact of Study

Retail sales are often considered as a “barometer” for the overall health of an economy - for instance, consumer spending accounts for 60% of the UK’s GDP per capita ([Admirals.com](#)), and the sector accounts for about 3 million jobs. Therefore, sales figures inform economists about consumer demand for finished products, whether the economy is expanding or contracting, and likely trends in the unemployment rate. Importantly, they are considered to be a **reliable** economic indicator of growth ([Finance.Yahoo.com](#)), as well as having a profound impact on foreign investment ([DailyFx.com](#)). Consequently, an examination of how weather impacts sales can potentially be of use to the following stakeholders:

1.1.4 (d) Key Stakeholders

- Governmental institutions have a vested interest in trends in retail sales in order to gauge how well businesses are doing and to assess consumers’ levels of disposable income ([Dailyfx.com](#))
- Retail companies can use these kinds of investigations for the basis of their data-driven decision making. To provide just one example, being able to sync advertisements to match appropriate weather conditions and decreasing ad-spend during periods of less-favourable weather can lead to the decrease of costs for the company ([WeatherAds.io](#))
- Central banking and monetary regulation institutions adjust interest rates based on consumer demand and sales to control inflation, thus building models to predict trends in spending can

lead to a better-informed adaptation of rates ([Admirals.com](#)).

- Shareholders, traders and potential investors in retail companies are interested in patterns in sales because retail sales exhibit a positive correlation with stock prices ([Dailyfx.com](#))
- Bondholders are negatively affected by an increase in retail sales, as when inflation rises, bond prices drop ([Finance.Yahoo.com](#))
- Everyday consumers and small businesses can be benefitted by knowing how demand fluctuates based on weather can be of interest in relation of when to make purchases ([WeatherAds.io](#))

1.1.5 (e) Possible Limitations and Challenges of this Project

- **Linear or polynomial model?:** Some academic research has suggested that the relationship between weather and sales is probably non-linear. This suggests that a possible solution would be a model which uses a higher-order polynomial. However, finding the correct degree of complexity can be challenging ([Badorf and Hoberg, 2020](#)).
- **Which parameters/features are important?:** “Weather” that might have an impact on sales is not constricted to just temperature, but also involves parameters such as wind, precipitation and humidity([Rose and Dolega, 2021](#)). As such, it can be difficult to select which features have the greatest influence on the outputs (sales figures). Therefore, using domain expertise to determine which parameters have been found the most significant could be a potential solution here.
- **Which retail sectors are more influenced by the weather?:** Retail sales in certain sectors (e.g. apparel/clothing) may be more influenced by weather fluctuations than other sectors such as electronics, therefore it might be useful to examine the impact of weather on the different sectors as well as generally

[]: