Technische Universität Berlin
Faculty IV – Electrical Engineering and Computer Science
Institute of Computer Engineering and Microelectronics
Dept. Computational Psychology

Weather's Cinematic Influence: Analyzing the Genre Preferences

Bachelor's Thesis

submitted by

Azer Mahjoub

in fulfilment of the requirements for the degree

Bachelor of Science (B.Sc.) in Computer Science

July 17, 2025

First examiner: **Prof. Dr. Marianne Maertens**

Second examiner: **Prof. Dr. Guillermo Gallego**

Declaration of authorship

I hereby declare that the thesis submitted is my own, unaided work, completed without any external help. Only the sources and resources listed were used. All passages taken from the sources and aids used, either unchanged or paraphrased, have been marked as such.

Where generative AI tools were used, I have indicated the product name, manufacturer, the software version used, as well as the respective purpose (e.g. checking and improving language in the texts, systematic research). I am fully responsible for the selection, adoption, and all results of the AI-generated output I use.

I have taken note of the Principles for Ensuring Good Research Practice at TU Berlin dated 15 February 2023. https://www.static.tu.berlin/fileadmin/www/10002457/K3-AMBl/Amtsblatt_2023/Amtliches_Mitteilungsblatt_Nr._16_vom_30.05.2023.pdf

I further declare that I have not submitted the thesis in the same or similar form to any other examination authority.

Berlin, July 17, 2025

Azer Mahjoub

Abstract

This thesis investigates how daily weather conditions influence film genre preferences, based on the assumption that weather affects mood, which, in turn, shapes viewing behavior. Mood-related genres were expected to respond more strongly to weather changes than non-mood-related genres. The study used daily viewing time per genre from the streaming platform Filmfriend, based on data from the Berlin postal area 10178, combined with weather data from Berlin between 2020 and 2023. Four genres were analyzed: comedy, drama, and adventure as mood-related and knowledge as non-mood-related. Weather variables were classified into subjective levels to reflect perceived states. Using exploratory data analysis, correlation analysis, and one-way ANOVA, the results confirm that colder, cloudier, and windier weather is associated with higher viewing of mood-related genres, particularly drama and comedy. All genres followed the same trend across all weather parameters except precipitation height, where the effects were small and mostly not statistically significant.

Keywords: Film genres, Weather, Weather perception, Mood, Media consumption, ANOVA

Zusammenfassung

Diese Arbeit untersucht, inwiefern tägliche Wetterbedingungen auf die Vorlieben bestimmter Filmgenres auswirken. Als Ausgangspunkt dient die Annahme, dass Wetter die Stimmung verändert und dadurch das Sehverhalten beeinflusst. Es wurde erwartet, dass stimmungsbezogene Genres sensibler auf Wetterveränderungen reagieren als nicht stimmungsbezogene Genres. Grundlage der Analyse sind tägliche Streaming-Zeiten pro Genre von der Streamingplattform Filmfriend aus dem Berliner Postleitzahlengebiet 10178, kombiniert mit Wetterdaten aus Berlin im Zeitraum von 2020 bis 2023. Betrachtet wurden vier Genres: Comedy, Drama und Adventure als stimmungsbezogene sowie Knowledge als nicht stimmungsbezogenes Genre. Die Wetterdaten wurden in subjektive Kategorien eingeteilt, um wahrgenommene Wetterlagen besser abzubilden. Mithilfe explorativer Datenanalyse, Korrelationsanalyse und einer einfaktoriellen Varianzanalyse konnte gezeigt werden, dass kühles, bewölktes und windiges Wetter mit einer erhöhten Nutzung stimmungsbezogener Genres einhergeht, insbesondere bei Drama und Comedy. Alle Genres zeigten unter allen Wetterbedingungen ein ähnliches Muster. Bei der Niederschlagshöhe hingegen fielen die Effekte gering aus und waren überwiegend nicht statistisch signifikant.

Schlüsselwörter: Filmgenres, Wetter, Wetterwahrnehmung, Stimmung, Mediennutzung, Varianzanalyse

Acknowledgements

I would like to express my sincere gratitude to Prof. Dr. Marianne Maertens for her valuable support, constructive feedback, and for the opportunity to write my thesis in her department. Her guidance was essential throughout the project.
I also thank Prof. Dr. Guillermo Gallego for serving as the second examiner and for evaluating this work.
I gratefully acknowledge Filmfriend for providing access to the streaming data used in this study, and the Deutscher Wetterdienst (DWD) for making their weather data openly available.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Visual media consumption is a regular part of everyday life and has evolved significantly in recent years. One of the most notable developments is the rise of streaming platforms, which have transformed viewing habits by offering a broader selection of content than traditional television. These platforms also provide greater flexibility, allowing users to choose what to watch, when to watch it, and how to watch it. In Germany, the average person watched television for 182 minutes per day in 2023, marking the lowest level since 1997. Meanwhile, video-on-demand usage reached a record 36 minutes per day in the fourth quarter of the same year [2, 3]. This shift shows a growing preference for on-demand and personalized content. As a result, selecting what to watch has become a routine decision for many. One factor known to influence everyday decisions is weather. Weather affects behaviours such as transportation, clothing choices, and social activities [4, 5]. Since people often adjust their daily habits based on weather changes, it is reasonable to assume that these changes may also affect media consumption patterns. Therefore, this thesis investigates whether daily weather conditions influence film genre selections.

Expressions like “under the weather” reflect a common belief that weather affects emotional well-being. This emotional impact may in turn lead individuals to engage with media content, as media consumption is often used to regulate or improve mood [6]. The popular saying “A rainy day calls for a cozy movie” supports this idea by suggesting that certain weather conditions are associated with specific viewing preferences. The term “cozy” typically refers to comfort and relaxation, often associated with genres such as romance or family films, rather than action or horror. This points to an indirect connection between weather and genre preference by affecting emotions.

Mood is clinically defined as “a pervasive and sustained feeling tone that is endured internally and which impacts nearly all aspects of a person’s behaviour in the external world” [7]. Mood influences decision-making; for example, people in a positive mood tend to make different choices than those experiencing a negative mood [8]. This influence extends to media consumption, where emotional state can influence film preferences. For instance, individuals feeling sad often prefer social dramas or dark comedies, while happy viewers are more likely to choose action films or slapstick comedies [9]. In an attempt to investigate whether mood plays a role in genre preferences, a mood-aware recommendation model was developed by integrating mood into a traditional collaborative filtering approach. The results demonstrated that this mood-based model performed better than conventional systems that do not account for users’ emotional states [10].

Weather, on the other hand, is known to influence mood. However, research on the impact of weather on mood has yielded mixed results. Some studies did not find consistent effects of weather on mood, suggesting that daily weather variations may not significantly alter emotional states [11, 12]. Newer research built on previous studies identified different personality-based responses to weather [13]. The study classified individuals into four distinct weather reactivity types: “Summer Lovers”, who feel better in warm and sunny weather; “Unaffected”, who show little mood change with weather; “Summer Haters”, who experience worse moods in warm and sunny conditions; and “Rain Haters”, who report negative moods on rainy days. These findings suggest that weather does not affect all individuals uniformly but instead interacts with personal characteristics. Nevertheless, certain patterns may still emerge at the population level.

If weather influences mood, and mood shapes genre preferences, then it is reasonable to consider that weather could indirectly affect what people choose to watch. In this thesis, I investigate whether such a connection exists by analyzing viewing behaviour across different weather conditions. I expect to see that the weather has stronger effects on mood-related genres than on non-mood-related genres.

1.2 Prior Work

During the course “Programming Project: Data Science in Python and R” at the Technical University of Berlin, supervised by Prof. Dr. Maertens, I conducted an initial analysis to explore how viewing time for different film genres varies by season. Seasons were defined meteorologically as winter (December–February), spring (March–May), summer (June–August), and autumn (September–November). I used data from Filmfriend [14], which provided daily viewing time per genre from 2020 to 2023, recorded by ZIP code. The dataset included both films and episodes. Figure 1.1 shows the total viewing time for selected genres across seasons. The percentages indicate the share of each season in the total viewing time of a genre from 2020 to 2023. For example, 30.4% of comedy viewing occurs in winter, while only 20.3% happens in summer. The results show a seasonal pattern for genres like comedy and romance, while no such trend appears for crime and thriller. This supports the idea that weather may influence genre preferences, and it raises the question of whether some genres are more sensitive to weather-related factors than others. While the findings confirm seasonal differences, this thesis aims to investigate which specific weather conditions are responsible for these patterns rather than treating the seasons as a whole.

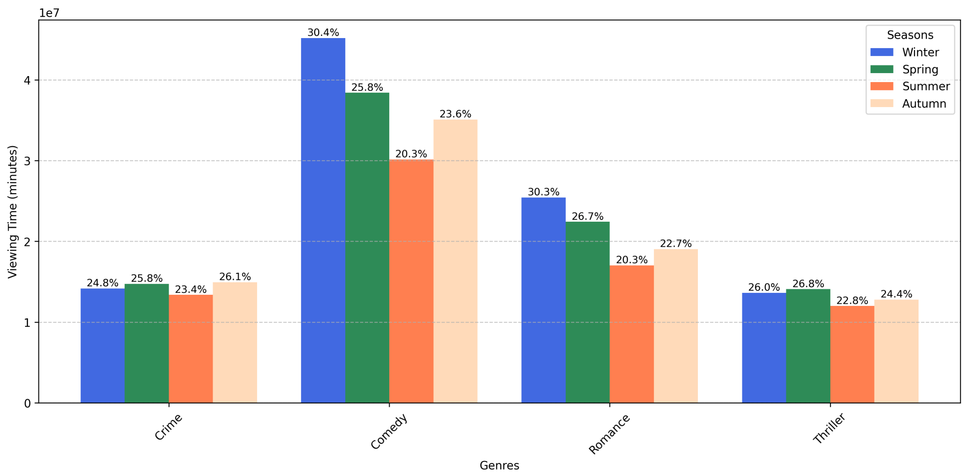


Figure 1.1. Total viewing time per genre per season by Filmfriend users from 2020 to 2023. Adapted from an earlier analysis conducted during the course “Programming Project: Data Science in Python and R”.

1.3 Data Selection

While the aim of this thesis may appear straightforward, it involves complex methodological considerations that can influence the results and complicate the direct attribution of viewing patterns to weather. Addressing this question requires careful decisions, including how to define weather conditions in objective or subjective terms, how to interpret genre preferences based on the available data, and how to choose appropriate temporal and geographic resolutions. The following sections outline how these challenges were addressed and the reasons behind the methodological choices made in this study.

1.3.1 Weather Data

Weather is defined by the World Meteorological Organization (WMO) as "the state of the atmosphere at a particular time, as defined by the various meteorological elements, including temperature, precipitation, atmospheric pressure, wind, and humidity" [15]. These meteorological factors are recorded through physical instruments, providing objective, precise, and consistent measurements. However, people usually do not describe the weather in technical terms. In everyday life, we tend to use subjective expressions such as "it is warm", "it is dark", or "it is sunny" rather than stating exact values like "12 degrees Celsius with 5 millimeters of rain". This difference between objective measurement and personal subjective experience of weather raises the question of what is a good quantitative indicator of weather in this context.

Perceptions of weather are shaped by several influences. Factors such as age, health status, and acclimatization affect how people experience temperature, wind, and humidity. For example, older adults may feel cold even under mild conditions, and someone in a hot climate may find moderate temperatures uncomfortably cold. Social and cultural contexts also influence the perception of the weather. In some cultures, overcast skies may be associated with sadness or low energy, while in others, they may evoke calmness or focus. Individuals living in urban areas often see rain as inconvenient, especially for commuting, so they stay indoors, whereas individuals in rural or agricultural regions may see overcast skies as beneficial, so they become active. Capturing subjective weather at this level of nuance would require extensive surveys, which are resource-intensive and sensitive to issues like memory bias and social desirability.

Because of these challenges, this study relies on objective meteorological data as the main source. The focus is on temperature, wind velocity, precipitation height, and cloud cover. I created categories corresponding to weather states that reflect common ways people describe weather in everyday life, assuming that the variables and the derived states are both perceptible and behaviourally meaningful.

1.3.2 Genre Data

Genre is a term in film studies and theory that groups films together based on shared characteristics, which can include visual style as well as social, cultural, institutional, and psychological factors [16]. Ideally, genre preferences would be studied using detailed viewing diaries kept by individual users. Such diaries would include information about why a film was chosen, such as interest in the genre, the actors, the release date, or other factors. In cases where genre was the main reason, the diary would note which one was most relevant, recognizing that most films fall under multiple genres. It would also record how long the viewer stayed engaged and describe their overall viewing experience. This kind of data would offer deep insights into how people select and experience films. However, collecting this type of detailed personal data is highly challenging. It would require consistent and active participation from users, and managing such a dataset on a large scale would be time-consuming and complex. Since Filmfriend, the data provider for this study, does not offer such data, alternative indicators must be used to infer genre preferences.

Given these limitations, this thesis uses viewing time per genre as the primary indicator of genre preference. While this measure does not capture the viewer's motivations or level of engagement, it offers a practical way to estimate preferences within the constraints of the available data.

At Filmfriend, genres are assigned based on the subjective judgments of individual curators. This thesis focuses on comedy, drama, and adventure, considered mood-related, and knowledge, considered a non-mood-related genre. Mood-related genres are those that evoke strong emotional responses and are often associated with mood-influenced entertainment choices [17, 6]. These genres differ in the specific emotions they evoke and relate to, which is why I examine multiple mood-related genres separately to capture this emotional diversity. In contrast, non-mood-related genres involve content driven primarily by factual information and learning, rather than emotional engagement [17]. This categorization enables the analysis to determine whether any observed effect of weather on viewing behaviour is likely mediated by mood.

In contrast to the prior work, genre data include only viewing time for films. Episodes were excluded because their continuous viewing patterns could bias the results, as the decision to watch a particular genre might be driven by narrative continuity rather than an active genre preference.

1.3.3 Temporal and Spatial Scope

In this thesis, I examine day-to-day fluctuations in viewing behaviour, recognizing that some weather effects may not be immediate and could persist beyond the specific weather event.

Film consumption data are recorded by ZIP code, while weather measurements come from fixed stations, so I must select a region that is large enough to provide sufficient viewing records yet small enough that labels such as “cold” or “windy” remain consistent across the area. Unlike areas near the Alps, which are affected by foehn winds, or coastal regions subject to maritime winds and variable precipitation, Berlin’s flat terrain within the Brandenburg region experiences relatively uniform weather conditions. The city’s abundant cultural and leisure offerings further reduce the likelihood that high viewing times are driven by a lack of alternative activities.

Chapter 2

Methodology

The methodological approach of this study was shaped by both the opportunities and constraints of the available data. The following sections outline how the dataset was collected and prepared for analysis, as well as the analytical strategy used to address the research question.

2.1 Data Collection and Preparation

This section outlines the sources of both weather and genre data, the definition of subjective weather states to better reflect perceived conditions, and the temporal and geographic alignment of the two datasets to create a unified basis for analysis.

2.1.1 Weather Dataset

The weather data were obtained from the Deutscher Wetterdienst (DWD), Germany's national meteorological service [18]. A full description of the original dataset can be found in the appendix. The weather data were restricted to the same time frame as the genre data, covering 4 years from 2020 to 2023. While the original dataset includes a broader range of meteorological variables, this study focuses on daily measurements of the following parameters:

- Daily mean of temperature (°C)
- Daily mean of wind velocity (m/s)
- Daily precipitation height (mm)
- Daily mean of cloud cover (okta)

The primary weather station used for this analysis was located in Berlin-Dahlem (52.4537°N, 13.3017°E). Because wind velocity measurements were missing at this station, wind data were supplemented from the Berlin-Brandenburg station (52.3807°N, 13.5306°E). Four additional missing wind velocity values were filled using linear interpolation.

2.1.2 Weather States

Based on the daily measurements, each weather variable was discretized into 3 levels, ranging from 0 to 2, with each level corresponding to a distinct weather state. To ensure that the thresholds presented in Table 2.1 are both intuitive and scientifically grounded, I defined them using internationally recognized sources. The temperature classification adopts the "no thermal stress" range from the Universal Thermal Climate Index, so I categorized temperatures between 10 and 20°C as moderate [19]. Wind levels follow the Beaufort scale, distinguishing between calm conditions (< 5 mph), a breeze (5–12 mph), and windy conditions (> 12 mph) [20]. For precipitation, I focus on the daily liquid-water-equivalent total. Based on WMO hourly intensity guidelines and assuming a typical 3-hour precipitation event per day, the thresholds are set as: < 1 mm for dry to drizzly days, 1–10 mm for light to moderate precipitation, and > 10 mm for moderate to heavy precipitation [21]. Cloud cover categories follow the WMO oktas system: 0–3 oktas (clear to partly cloudy), 4–6 (partly to mostly cloudy), and 7–8 (mostly to fully overcast) [22].

Table 2.1. Weather levels and states

Category	Level	State
Temperature (°C)	0	Cold: < 10 °C
	1	Moderate: 10 °C – 20 °C
	2	Warm: > 20 °C
Wind Velocity (mph)	0	Calm: < 2.24 m/s (≈ 5 mph)
	1	Breezy: 2.24 m/s – 5.36 m/s
	2	Windy: > 5.36 m/s (≈ 12 mph)
Precipitation (mm per day)	0	Dry: < 1 mm
	1	Wet: 1 mm – 10 mm
	2	Very Wet: > 10 mm
Cloud Cover (okta)	0	Sunny: < 4
	1	Cloudy: 4 – 6
	2	Overcast: > 6

2.1.3 Genre Dataset

The film data for this study were obtained from Filmfriend [14], a digital streaming platform associated with public and academic libraries across several European countries. The dataset captures daily viewing time from January 1, 2020, to December 31, 2023. Each record includes the date, the associated ZIP code, the total minutes viewed, and the minutes viewed per genre. This study focuses on the daily viewing time of the following genres:

- **Mood-related genres:**
  - Comedy
  - Drama
  - Adventure
- **Non-mood-related genre:**
  - Knowledge

The dataset was filtered using the 194 ZIP codes listed for Berlin by the Open Data Informationsstelle Berlin (ODIS) [1]. The resulting dataset contained viewing time records from only 5 ZIP codes. Among them, ZIP code 10178 accounted for approximately 97% of the total viewing time recorded in the Berlin region during the study period (see appendix, Figure A.1). Thus, summing or averaging across all ZIP codes would distort the analysis. To ensure a consistent and interpretable investigation of weather effects on viewing behaviour, this thesis focuses exclusively on ZIP code 10178.

2.1.4 Data Integration

After preprocessing, the weather and genre datasets were merged using the date as a common key. Figure 2.1 visualizes the data sources and the merging process. The chosen weather stations, Berlin-Dahlem and Berlin-Brandenburg, are located within or near the Berlin metropolitan area. Since the viewing data were limited to ZIP code 10178, which lies approximately 10 km northeast of Berlin-Dahlem and about 20 km northwest of Berlin-Brandenburg Airport, the weather observations can be considered representative of local conditions for the target population. The resulting dataset, which was used for the analysis and is shown in Figure 2.2, contains 1461 rows, covering the period from January 1, 2020, to December 31, 2023. Each row includes 13 columns, comprising the date, daily measures of weather parameters, their corresponding classified levels, and daily viewing time for the selected genres.

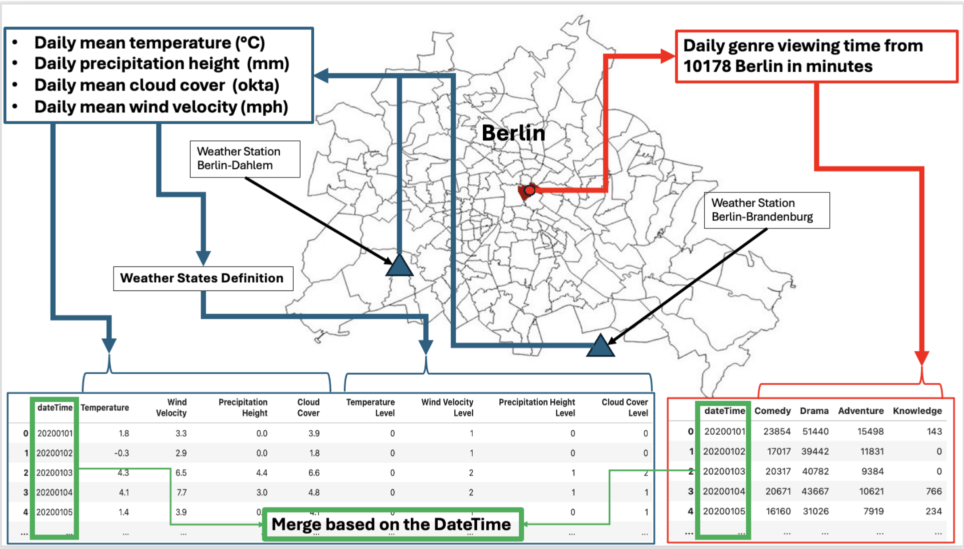


Figure 2.1. Overview of the data integration process. Berlin map adapted from ODIS [1]

	dateTime	Temperature	Wind Velocity	Precipitation Height	Cloud Cover	Temperature Level	Wind Velocity Level	Precipitation Height Level	Cloud Cover Level	Comedy	Drama	Adventure	Knowledge
0	20200101	1.8	3.3	0.0	3.9	0	1	0	0	23854	51440	15498	143
1	20200102	-0.3	2.9	0.0	1.8	0	1	0	0	17017	39442	11831	0
2	20200103	4.3	6.5	4.4	6.6	0	2	1	2	20317	40782	9384	0
3	20200104	4.1	7.7	3.0	4.8	0	2	1	1	20671	43667	10621	766
4	20200105	1.4	3.9	0.0	4.1	0	1	0	1	16160	31026	7919	234
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1456	20231227	4.0	4.2	0.1	5.3	0	1	0	1	26804	39684	10376	605
1457	20231228	8.4	7.1	0.0	6.6	0	2	0	2	25617	41318	9868	1434
1458	20231229	9.5	8.9	4.1	6.5	0	2	1	2	20475	38425	8002	1223
1459	20231230	6.1	6.0	0.0	5.3	0	2	0	1	29773	50291	6368	961
1460	20231231	5.9	4.1	0.4	6.8	0	1	0	2	21165	33227	6932	590

Figure 2.2. Sample view of the final dataset after merging weather and viewing records by date.

Figure 2.3 displays the overall viewing time per genre across the study period. Drama clearly dominates with a total of approximately 5.14×10^7 minutes (around 856 000 hours or 586 hours per day). Comedy follows with 2.57×10^7 minutes (around 428 000 hours or 293 hours per day), and adventure accounts for 1.41×10^7 minutes (approximately 235 000 hours or 161 hours per day). In contrast, knowledge shows a substantially lower total with 9.68×10^5 minutes (roughly 16 000 hours or 11 hours per day). Due to this small volume, findings related to knowledge should be interpreted with caution, as the low number of observations may limit the reliability of trend detection and statistical comparisons.

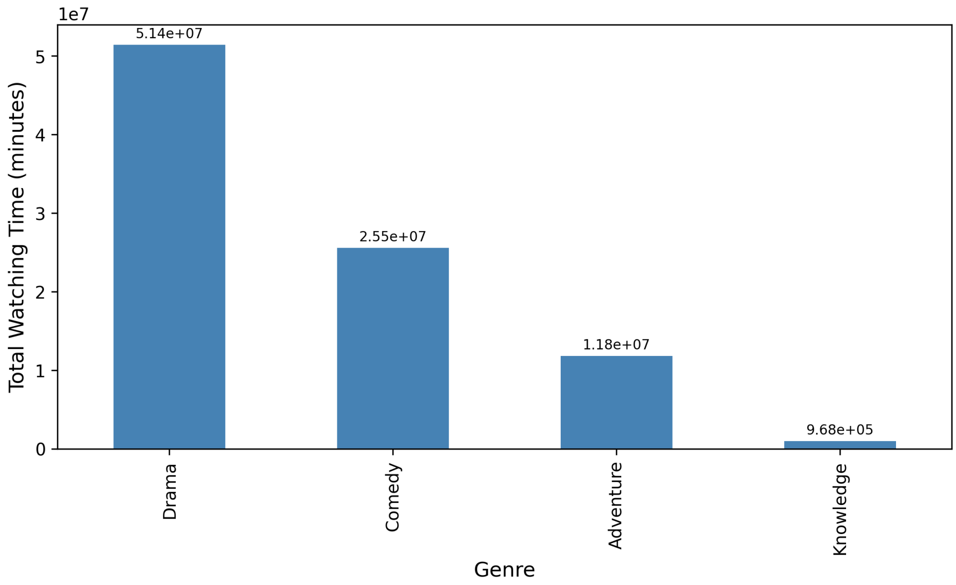


Figure 2.3. Total viewing time per genre in postal area 10178 Berlin from 2020 to 2023.

2.2 Data Analysis

The analysis consisted of two main steps. First, I conducted an exploratory data analysis to summarize the distribution, central tendencies, and seasonal trends of both weather variables and genre-specific viewing times. This step provided descriptive context for the subsequent statistical analysis. Next, Pearson correlation coefficients were calculated to assess the linear relationship between each weather parameter and daily viewing time per genre.

The second part of the analysis aimed to test whether mean daily viewing times vary across different levels of each weather variable and whether mood-related genres (comedy, drama, adventure) respond differently than the non-mood-related genre (knowledge). Ideally, this could be evaluated using a two-way Analysis of Variance (ANOVA) with weather state and genre type (mood-related vs. non-mood-related) as factors. However, this approach was not suitable because it did not meet important statistical requirements. First, the data could not be treated as independent. This is because many films have multiple genres, with the same amount of viewing time being counted for all the genres, and because viewing sessions can continue across different days, making daily values dependent on each other. These issues go against the ANOVA assumption that each observation must be independent. Second, the large differences in average viewing time between genres make it difficult to interpret interaction effects between independent variables.

To address these constraints, I conducted separate one-way Welch ANOVAs for each genre across the 3 categorical levels of each weather parameter. Welch's ANOVA was chosen because Levene's test showed violations of homogeneity of variances in most cases, and normality checks via the Shapiro–Wilk test also failed (see appendix, Table A.1). Welch's test adjusts for unequal group variances and sample sizes [23] and is considered robust to non-normality when sample sizes per group are large ($N > 100$) [24]. When a significant effect was detected ($p < 0.05$), I followed up with Games–Howell post hoc tests to determine which pairwise comparisons of levels were statistically different.

To compare the responsiveness of genres to weather while controlling for differences in baseline viewing time, I also computed relative percentage changes between weather levels (0 to 1, 1 to 2, and 0 to 2) based on the mean values. This was used as an indirect approximation of the interaction between weather and genre type, especially regarding the expectation that mood-related genres are more sensitive to environmental changes. This approach allowed me to avoid the issue of genre interdependence by treating each genre separately. However, it still relies on the assumption that viewing sessions spanning midnight are infrequent and do not introduce substantial bias.

Chapter 3

Results

This chapter presents the findings of the study in two parts. The first part explores descriptive patterns in the data. The second part reports the results of statistical tests assessing whether viewing behaviour varies across different weather states.

3.1 Exploratory Data Analysis

In this section, the weather and genre data are analyzed separately to explore their individual characteristics. For both datasets, temporal patterns, descriptive statistics, and distribution plots are provided. Lastly, a correlation analysis with visualizations offers a preliminary view of the potential interaction between weather and genre.

3.1.1 Weather

Figures 3.1 to 3.4 show how the weather parameters change over the months. Temperature follows a regular seasonal cycle. It rises gradually in spring, reaches its peak in summer, and then slowly drops during autumn and winter. Cloud cover also changes with the seasons, with more cloudy days in the colder months, though the pattern is less clear than for temperature. Wind velocity does not follow a simple yearly pattern. Interestingly, similar fluctuations appear in alternating years, like 2020 and 2022 compared to 2021 and 2023, which might point to a rhythm that spans two years instead of one. Precipitation stands out as the most irregular. There is no clear seasonal trend, and the amount of rainfall seems to change from year to year.

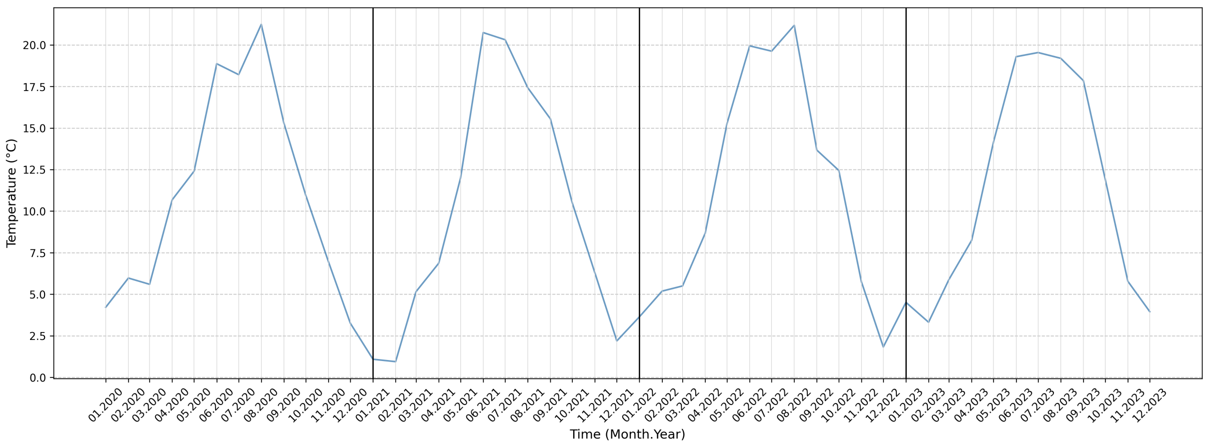


Figure 3.1. Monthly average temperature in Berlin-Dahlem from January 2020 to December 2023.

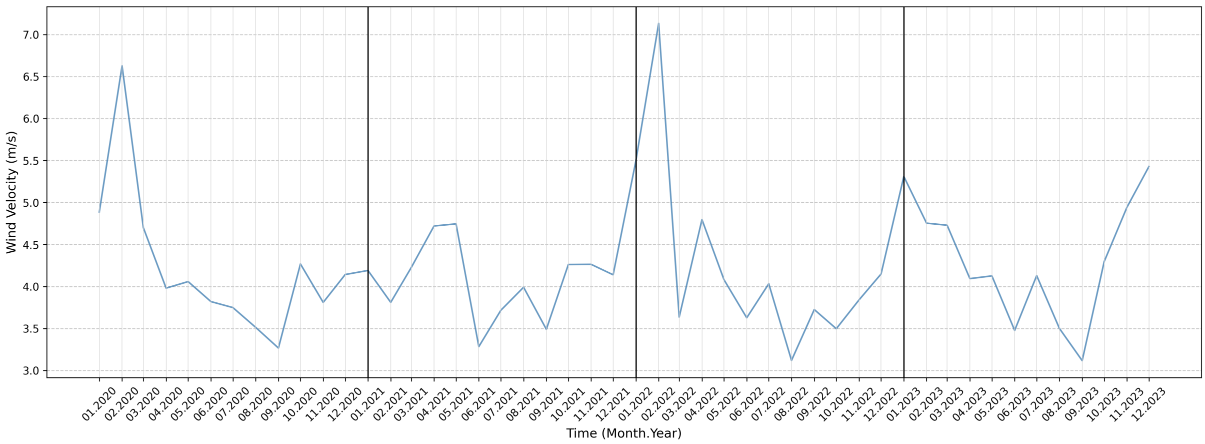


Figure 3.2. Monthly average wind velocity in Berlin-Brandenburg from January 2020 to December 2023.

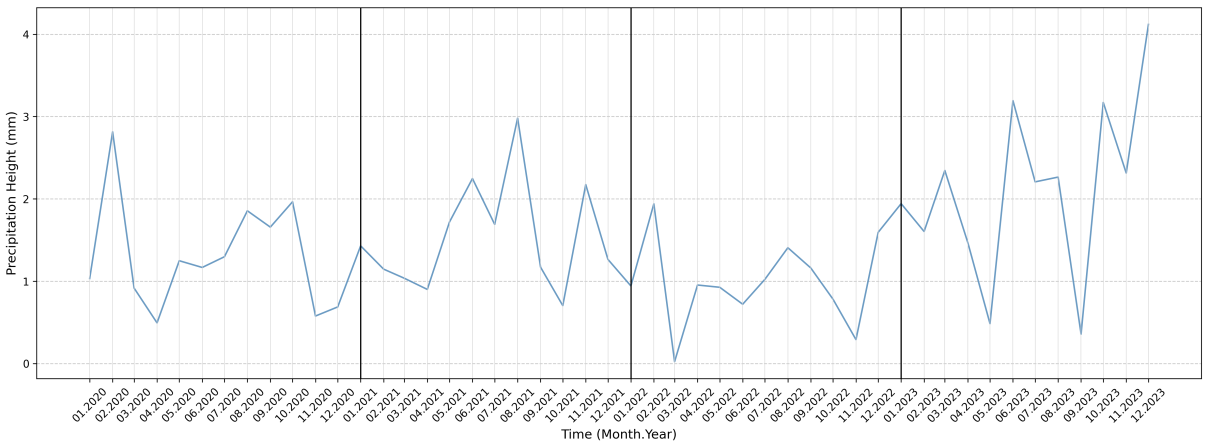


Figure 3.3. Monthly average precipitation height in Berlin-Dahlem from January 2020 to December 2023.

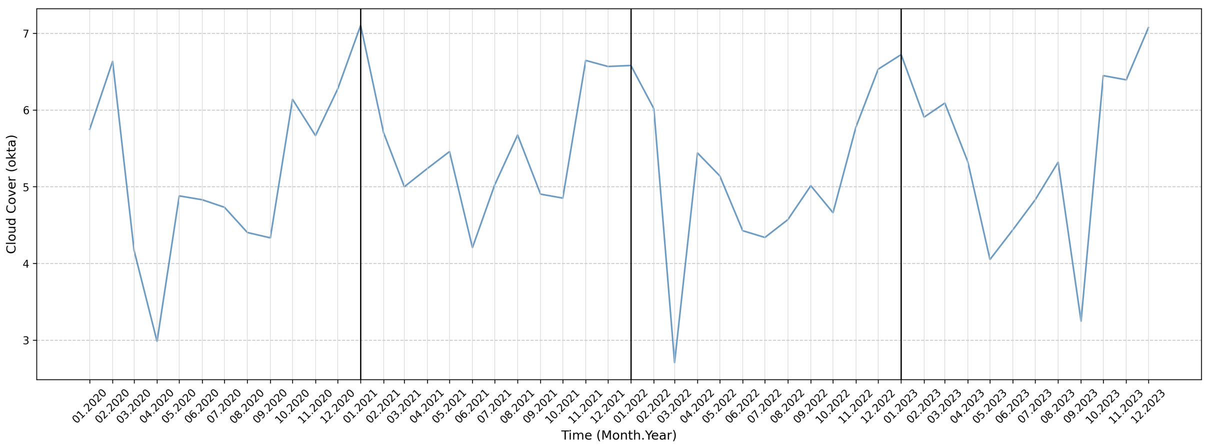


Figure 3.4. Monthly average cloud cover in oktas in Berlin-Dahlem from January 2020 to December 2023.

The daily time series in Figures 3.5 to 3.8 reveal more detailed day-to-day variations. Temperature shifts gradually and predictably, with colder days in winter and warmer days mostly in summer. Some unusually cold or warm days also appear in spring and autumn. Wind conditions are usually breezy, but strong winds occur more often in winter and early spring, while calm days show up mainly in summer. Precipitation is highly uneven. Most days are dry, but occasional spikes in rainfall appear suddenly, and they are spread throughout the year without a clear pattern. Cloud cover is more balanced, with values ranging across the full spectrum from clear to overcast. However, fully cloudy days are a bit more common in the winter months.

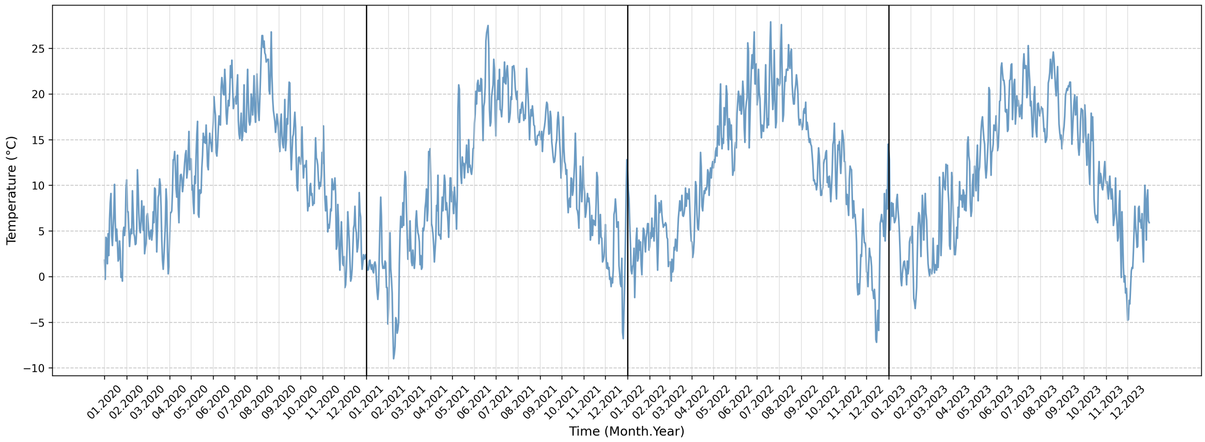


Figure 3.5. Daily mean of temperature in Berlin-Dahlem from January 2020 to December 2023.

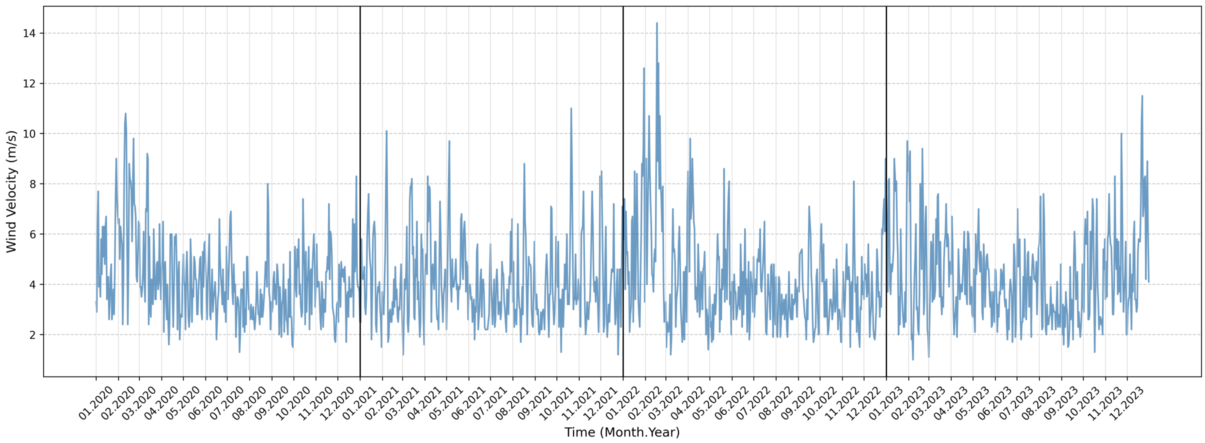


Figure 3.6. Daily mean of wind velocity in Berlin-Brandenburg from January 2020 to December 2023.

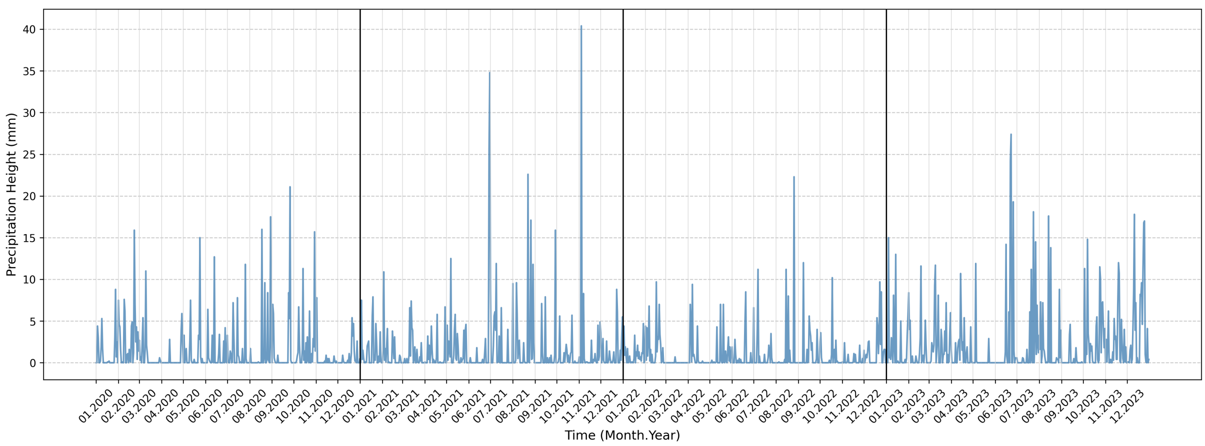


Figure 3.7. Daily precipitation height in Berlin-Dahlem from January 2020 to December 2023.

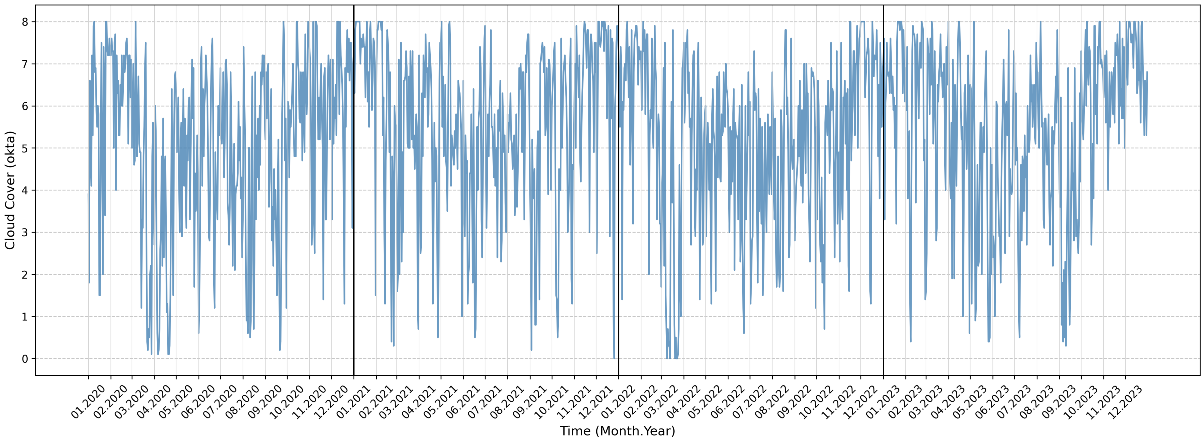
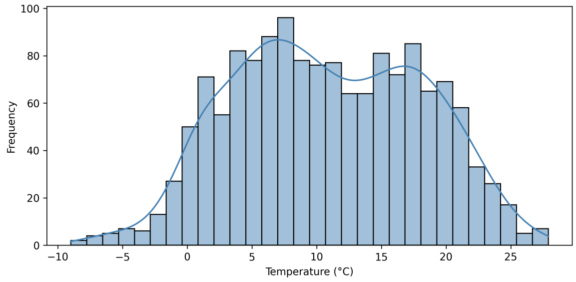


Figure 3.8. Daily mean of cloud cover in Berlin-Dahlem from January 2020 to December 2023.

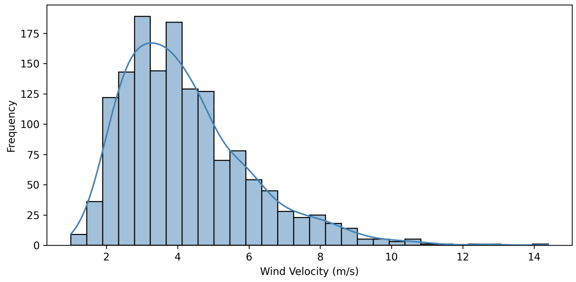
Table 3.1 and Figure 3.9 provide an overview of the distribution and variability of the weather parameters. Temperature is symmetrically distributed around a mean of approximately 11 °C, with values spanning from –9 °C to nearly 28 °C. Wind velocity, by contrast, shows a mild right skew, with occasional peaks up to 14.4 m/s, though most days cluster around the average of 4.2 m/s. Precipitation is the most unevenly distributed variable: its mean is relatively low (1.49 mm), but rare high-intensity days drive the maximum up to 40.4 mm. This right-skewed pattern is also reflected in the strong imbalance of precipitation levels shown in Table 3.3, where over 90% of days are classified as dry. Cloud cover spans the full range from clear to fully overcast (0 to 8 oktas) and appears more evenly distributed, with a slight concentration in the higher levels. These observations suggest that temperature and cloud cover offer relatively balanced input for further analysis, while precipitation and, to a lesser extent, wind may present challenges due to class imbalance and skewed distributions.

Table 3.1. Summary statistics for weather parameters from 2020 to 2023, including mean, standard deviation, minimum, and maximum values.

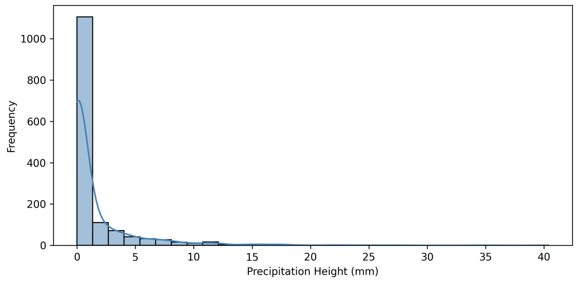
Variable	Mean	Std. Dev.	Min	Max
Temperature (°C)	10.85	7.21	-9.00	27.90
Wind Velocity (m/s)	4.22	1.80	1.00	14.40
Precipitation Height (mm)	1.49	3.47	0.00	40.40
Cloud Cover (Okta)	5.30	2.03	0.00	8.00



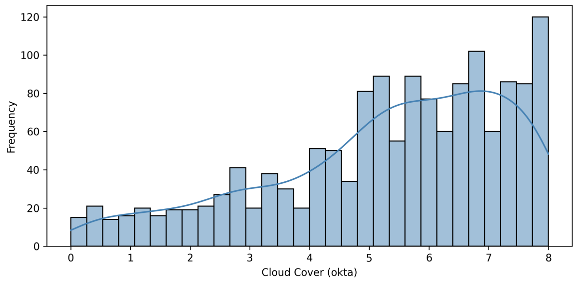
(a) Distribution of daily mean of temperature.



(b) Distribution of daily mean of wind velocity.



(c) Distribution of daily precipitation height.



(d) Distribution of daily mean of cloud cover.

Figure 3.9. Distribution of daily weather measurements from 2020 to 2023.

3.1.2 Genre

Figures 3.10 to 3.13 illustrate the monthly genre viewing time series. They reveal shared seasonal dynamics across all genres. Viewing time tends to rise during the colder months and decline in summer, suggesting a seasonal component that may be linked to indoor leisure preferences. Even the knowledge genre generally follows this trend, although with less consistency. The early years show a different pattern: while 2021 to 2023 consistently exhibit peaks around the turn of the year, 2020 features a more modest winter peak followed by a distinct increase in April. This spike may be related to the first lockdown measures in response to the COVID-19 pandemic [25, 26]. All genres show a remarkably similar structure throughout 2022 and 2023, which could indicate a return to more stable, post-pandemic viewing habits.

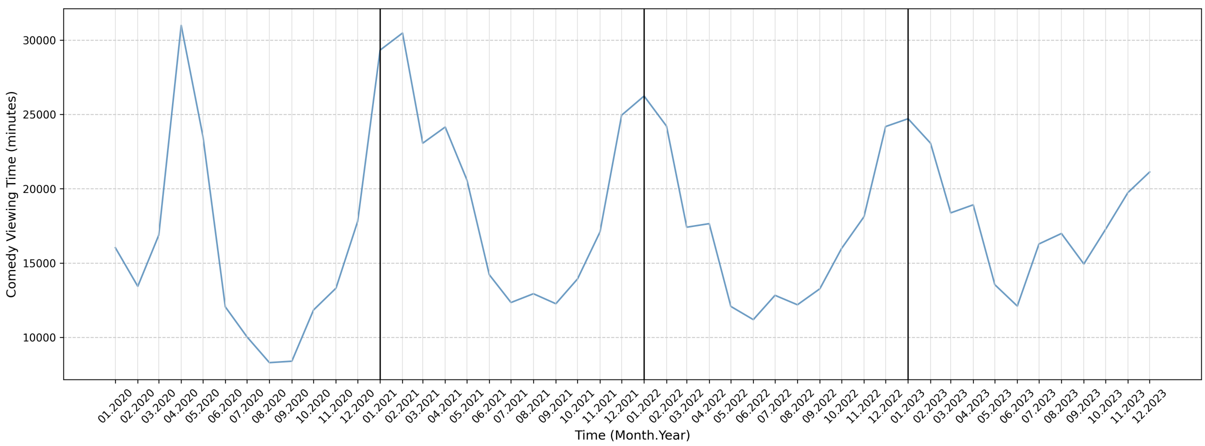


Figure 3.10. Monthly average viewing time for comedy in 10178 Berlin from 2020 to 2023.

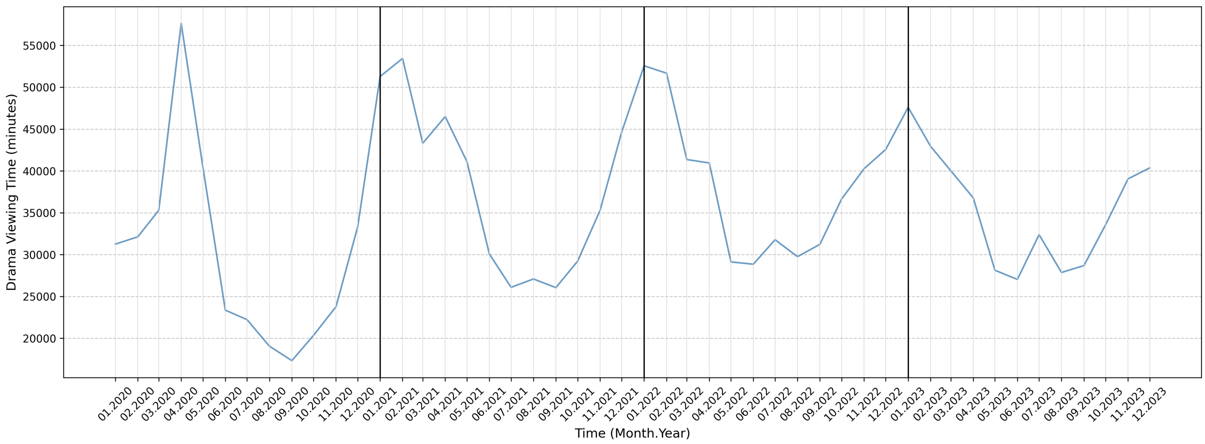


Figure 3.11. Monthly average viewing time for drama in 10178 Berlin from 2020 to 2023.

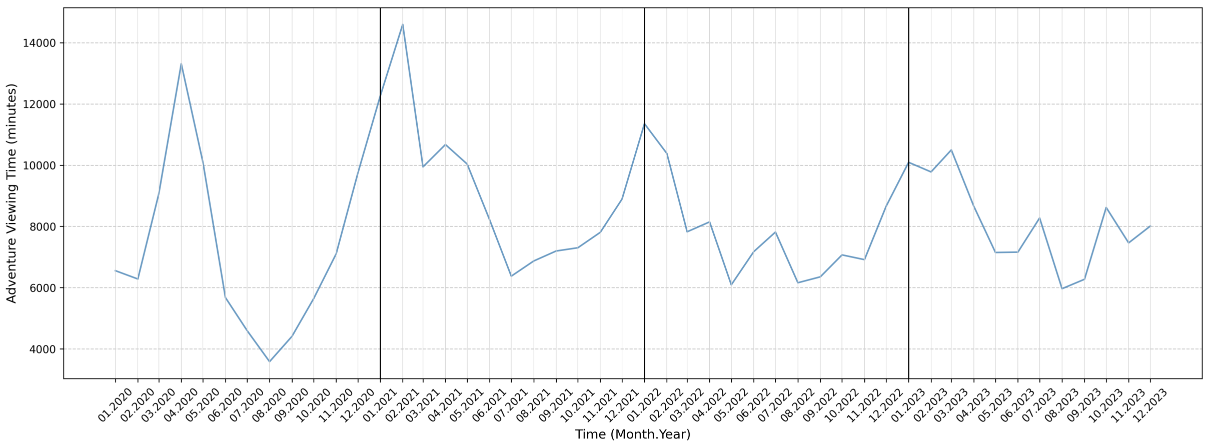


Figure 3.12. Monthly average viewing time for adventure in 10178 Berlin from 2020 to 2023.

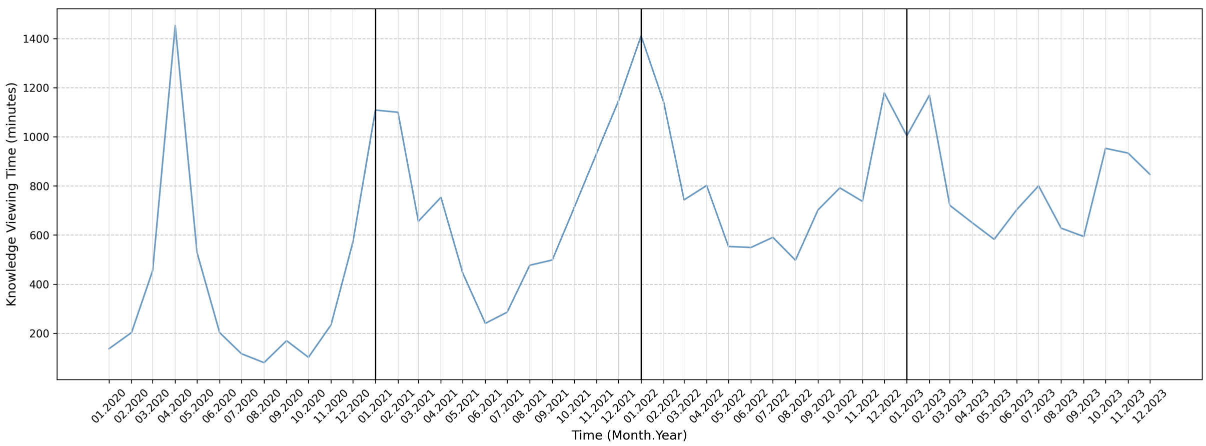


Figure 3.13. Monthly average viewing time for knowledge in 10178 Berlin from 2020 to 2023.

Figures 3.14 3.17 present the daily genre viewing time series. All genres exhibit noticeable fluctuations in daily viewing time, often following an oscillating pattern that likely reflects regular shifts between weekdays and weekends. Among the mood-related genres, comedy shows relatively smooth and gradual oscillations, which points to a more stable rhythm of engagement over time. Drama and adventure, in contrast, experience sharper short-term spikes and drops. These variations tend to become more pronounced during colder months, when overall viewing is elevated. The knowledge genre stands out due to its irregular pattern. It includes several distinct spikes. These abrupt increases may be driven by targeted bursts of interest that occur in response to educational campaigns, new content releases, or current events.

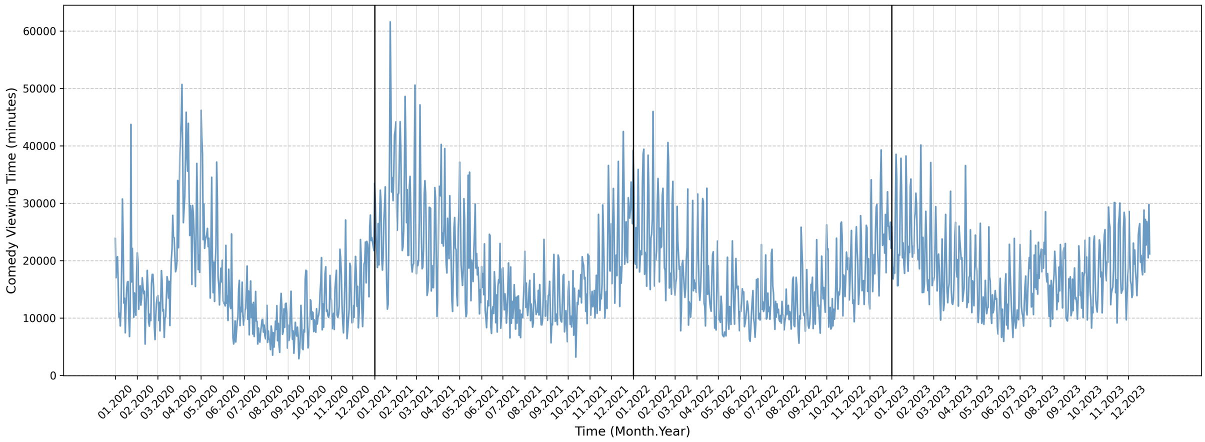


Figure 3.14. Daily viewing time for comedy in 10178 Berlin from 2020 to 2023.

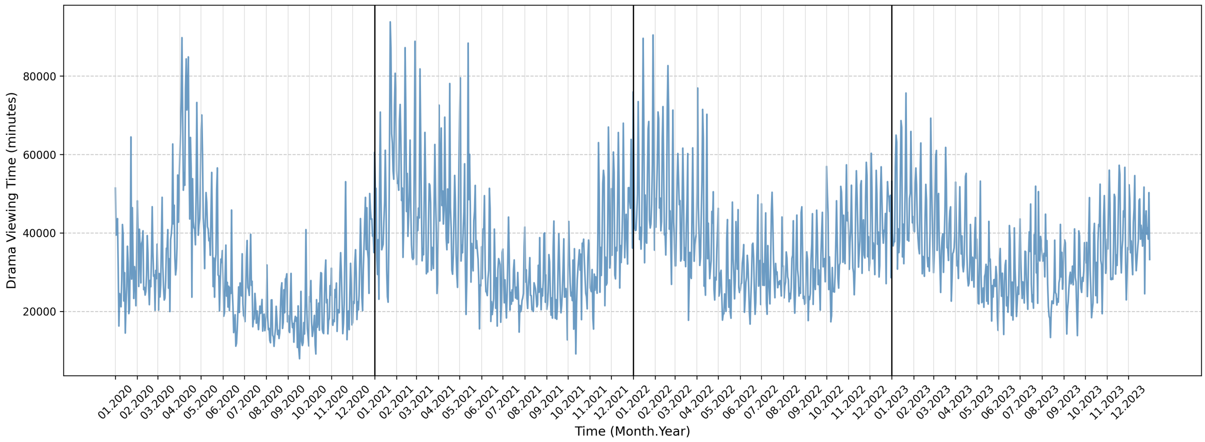


Figure 3.15. Daily viewing time for drama in 10178 Berlin from 2020 to 2023.

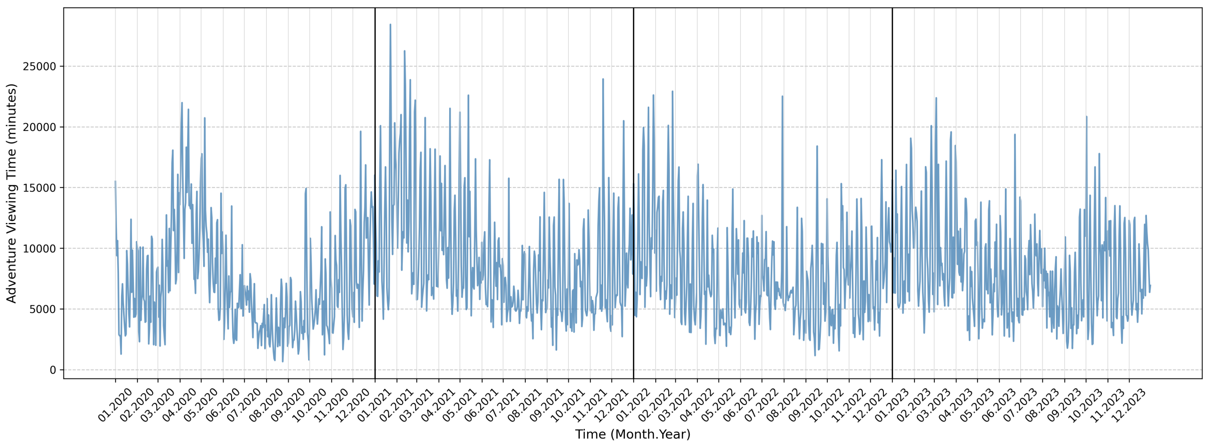


Figure 3.16. Daily viewing time for adventure in 10178 Berlin from 2020 to 2023.

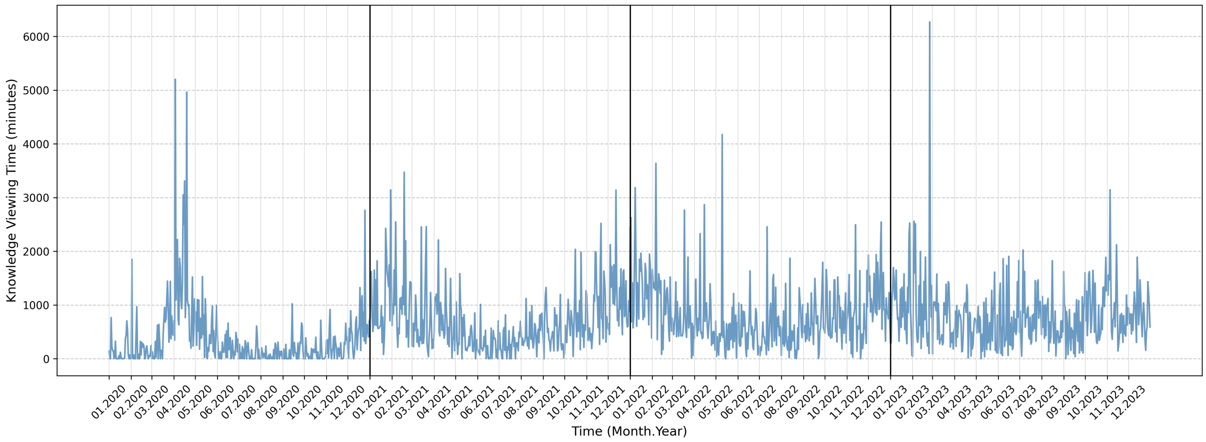
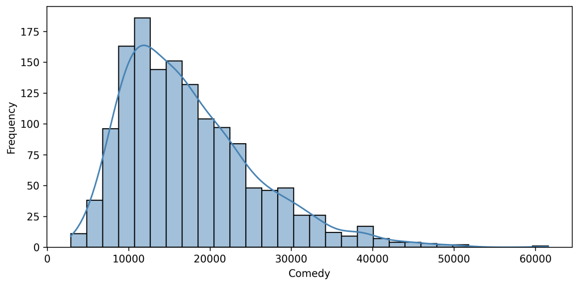


Figure 3.17. Daily viewing time for knowledge in 10178 Berlin from 2020 to 2023.

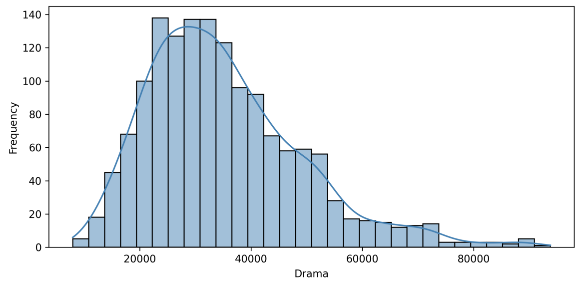
Table 3.2 provides summary statistics for daily viewing time across the genres. Drama registers the highest average viewing time at over 35 000 minutes per day (around 583 hours per day), along with the largest standard deviation, indicating strong engagement and substantial variability in consumption. Comedy follows with a lower mean but still considerable variation. Adventure shows more moderate viewing levels, while knowledge records the lowest average and the smallest overall volume. However, the standard deviation for knowledge is relatively high compared to its mean, suggesting that although most values are low, a very high deviation is driven by occasional spikes. These distributional characteristics are visually supported by Figure 3.18. All genres exhibit right-skewed distributions, where most values cluster at lower levels and fewer days experience unusually high viewing. The skew is most pronounced in the knowledge genre, which shows a steep concentration near zero and a long tail extending toward rare, high-usage days. This suggests that mood-related genres maintain more consistent engagement patterns, while knowledge consumption is sporadic and likely event-driven.

Table 3.2. Summary statistics for genre viewing time from 2020 to 2023 in minutes, including mean, standard deviation, minimum, and maximum values.

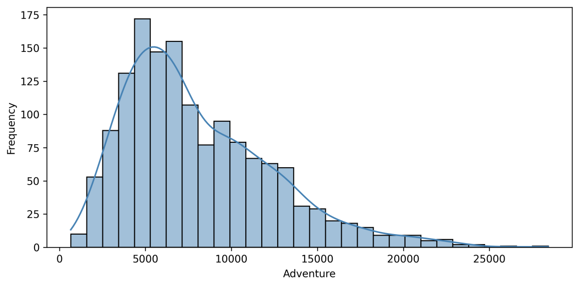
Genre	Mean	Std. Dev.	Min	Max
Comedy	17473.74	8175.59	2889.00	61593.00
Drama	35177.17	13974.79	7937.00	93816.00
Adventure	8071.64	4367.07	656.00	28432.00
Knowledge	662.41	613.06	0.00	6271.00



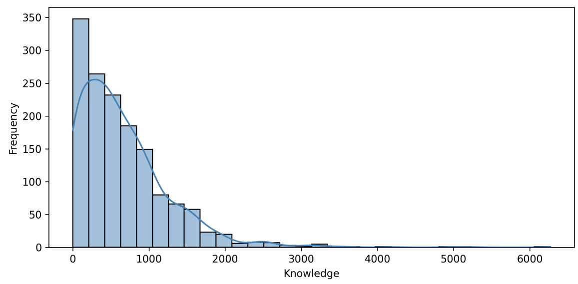
(a) Distribution of comedy viewing time.



(b) Distribution of drama viewing time.



(c) Distribution of adventure viewing time.



(d) Distribution of knowledge viewing time.

Figure 3.18. Distribution plots of daily genre viewing times.

3.1.3 Correlation Between Weather and Genre Viewing Time

Figures 3.20 to 3.23 display scatter plots illustrating how daily weather conditions relate to genre-specific viewing time. Across all weather variables, the direction of influence is consistent for all genres. Temperature stands out with the most pronounced effect: viewing time tends to decline as temperature increases. This relationship is shown by Figure 3.19, where comedy and drama were strongly negatively correlated with $r = -0.49$ and $r = -0.48$, respectively, both significant with $p < 0.001$ (see appendix, Table A.2). Cloud cover and wind velocity also exhibit statistically significant positive associations, but with a more modest strength. For example, cloud cover correlates with comedy at $r = 0.24$ and wind with drama at $r = 0.19$, both with $p < 0.001$.

Precipitation has the least consistent impact. While correlations with comedy, drama, and adventure were statistically significant, they are small and visually subtle due to the predominance of dry days. More generally, knowledge, the only non-mood-related genre, differs from the other genres in its scatter pattern across all weather parameters. Its plots are more dispersed and correlations consistently weaker, indicating that its consumption is less sensitive to short-term weather changes.

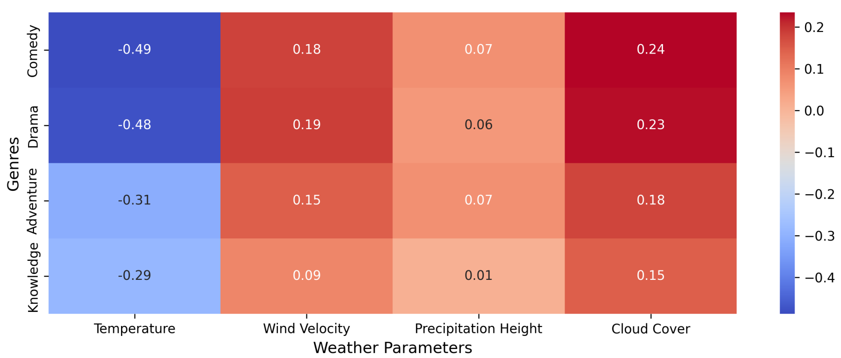


Figure 3.19. Correlation heatmap between weather parameters and genre viewing time.

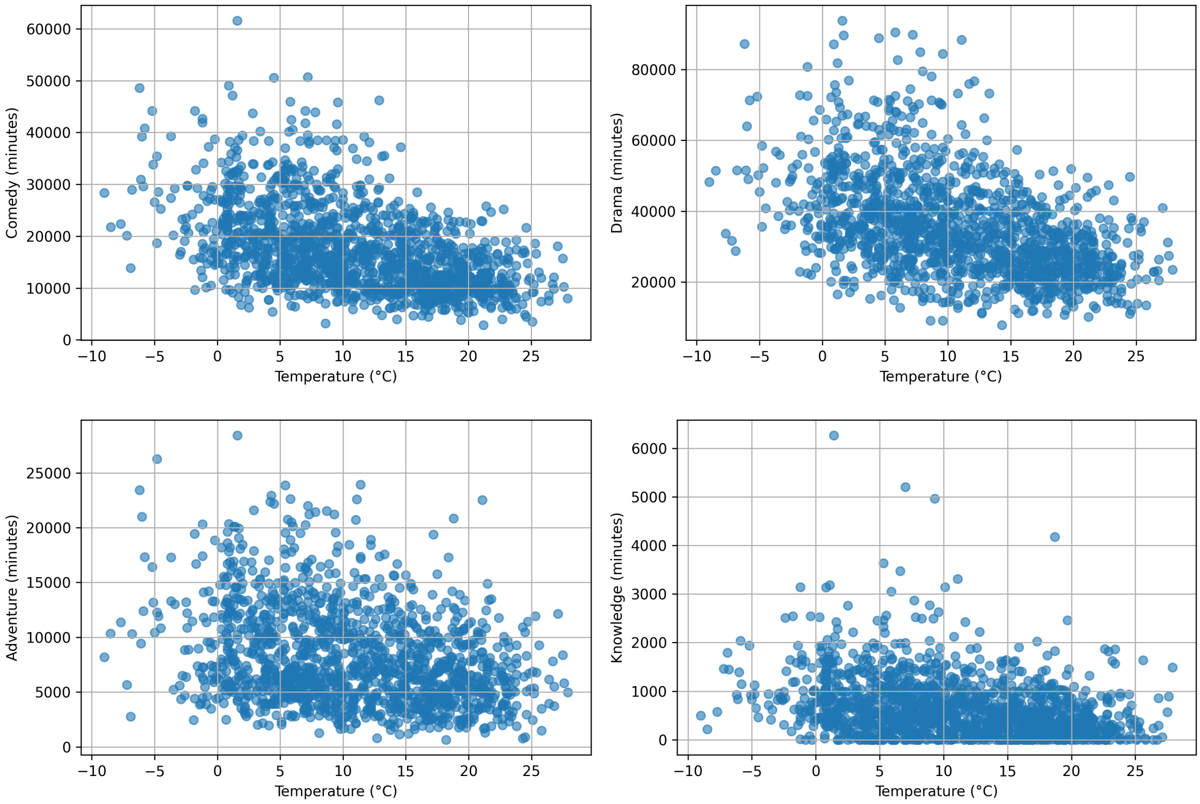


Figure 3.20. Genre viewing time by temperature.

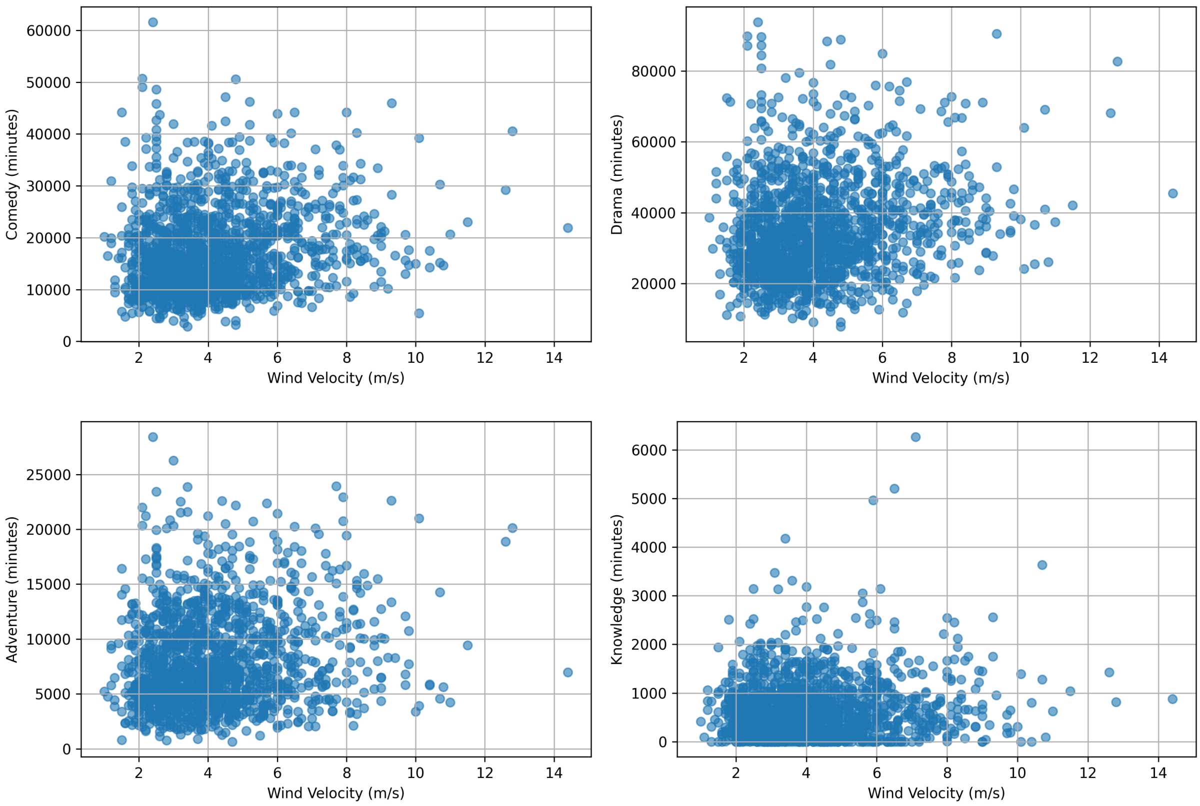


Figure 3.21. Genre viewing time by wind velocity.

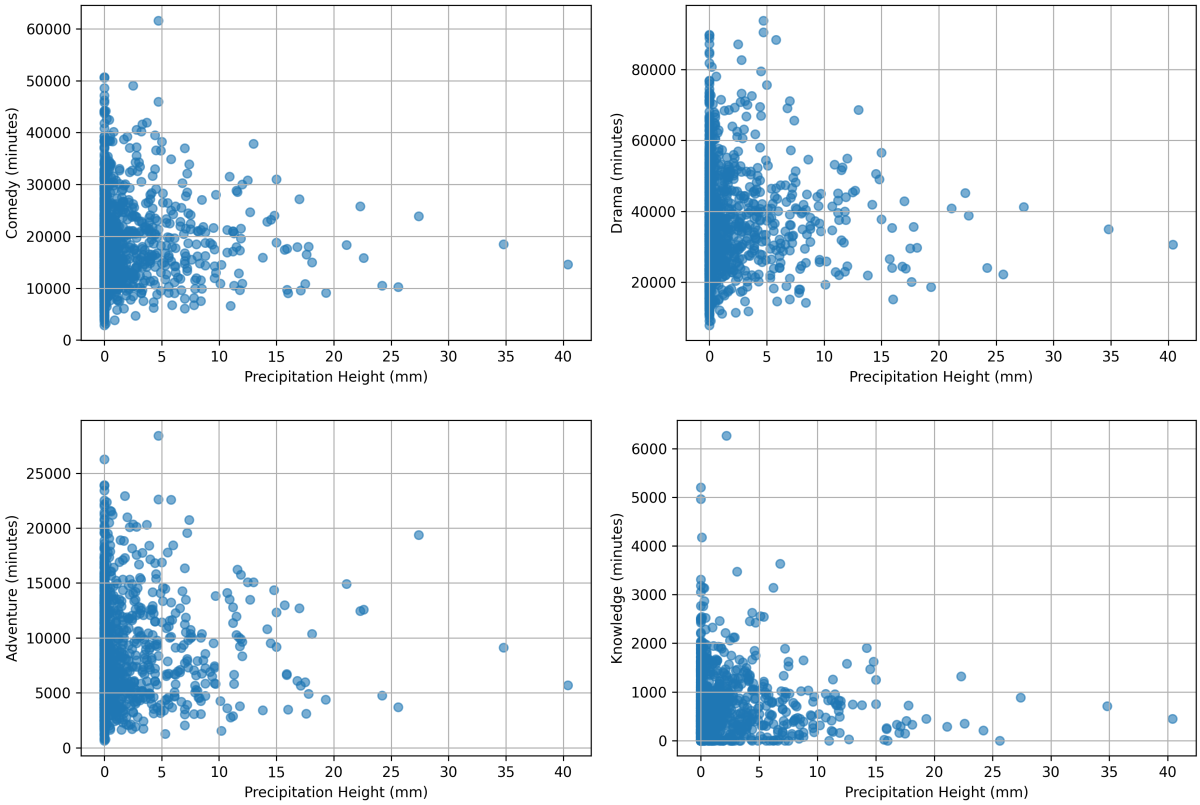


Figure 3.22. Genre viewing time by precipitation height.

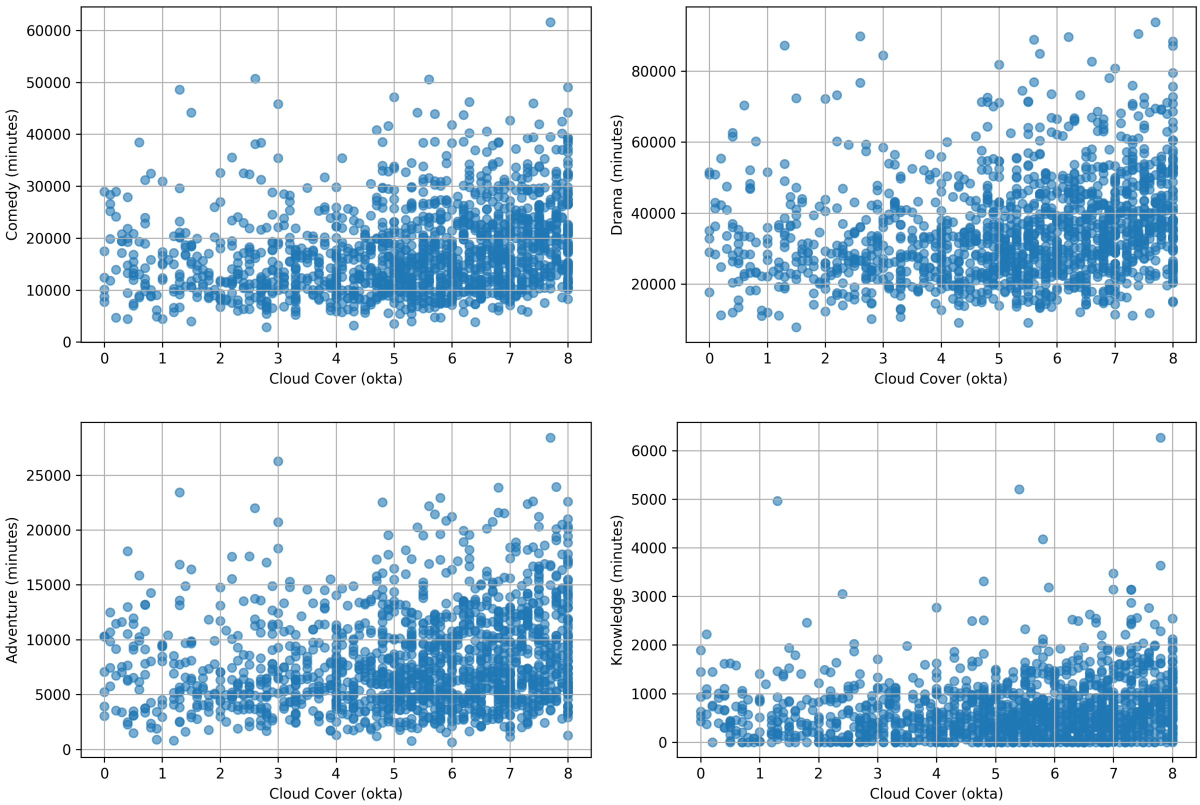


Figure 3.23. Genre viewing time by cloud cover.

3.2 Genre Viewing Time Across Weather States (ANOVA)

Table 3.3 reports the mean and standard deviation of viewing time (in minutes per day) across all weather levels. Figures 3.24 to 3.27 show the distribution via box plots. With the exception of precipitation, all genres show consistent directional changes across weather levels. Viewing times either increase or decrease continuously across levels, and the difference in viewing time between levels 0 and 2 is larger than between 0 and 1 or 1 and 2. Welch ANOVAs revealed significant differences in mean viewing time for all weather–genre combinations, and pairwise Games–Howell tests were performed for all cases (See appendix, Table A.3).

Outliers are visible across all conditions. Drama and comedy, with larger viewing volumes, show broad distributions with fewer extreme points. Adventure reveals several upward spikes, possibly due to events or special releases. Knowledge exhibits large relative deviations for such a small base, with a standard deviation nearly matching and in some cases surpassing its mean, suggesting volatility and low baseline stability.

Table 3.3. Number of days (N) in each weather level and mean viewing time (minutes)± standard deviation.

Parameter	Level	N	Comedy	Drama	Adventure	Knowledge
Temperature	Cold (0)	704	21035 ± 8621	41250 ± 14553	9249 ± 4734	823 ± 695
	Mild (1)	585	14817 ± 6375	30561 ± 11146	7203 ± 3809	552 ± 497
	Warm (2)	172	11935 ± 4373	26024 ± 7966	6205 ± 3063	380 ± 391
Wind Velocity	Calm (0)	138	16416 ± 8399	33395 ± 13630	7218 ± 3965	613 ± 452
	Breezy (1)	1001	16793 ± 8068	33957 ± 13785	7898 ± 4255	632 ± 564
	Windy (2)	322	20044 ± 7911	39735 ± 13788	8978 ± 4731	779 ± 784
Precipitation	Dry (0)	1052	16869 ± 8146	34107 ± 13866	7824 ± 4266	630 ± 593
	Wet (1)	359	19080 ± 8178	38201 ± 14143	8658 ± 4578	765 ± 677
	Very Wet (2)	50	18673 ± 7167	35974 ± 11890	9069 ± 4452	609 ± 458
Cloud Cover	Sunny (0)	337	15223 ± 7594	31499 ± 13168	7190 ± 4042	562 ± 547
	Cloudy (1)	473	16095 ± 7689	32950 ± 13327	7518 ± 3964	593 ± 573
	Overcast (2)	651	19641 ± 8288	38699 ± 14036	8930 ± 4649	765 ± 657

Table 3.4. Relative changes between weather levels with * for pairwise significance ($p < .05$) and ns for non significance.

Parameter	**Comparison**	**Comedy**	**Drama**	**Adventure**	**Knowledge**
Temperature	Cold to Mild	–29.56%*	–25.91%*	–22.12%*	–32.93%*
	Mild to Warm	–19.45%*	–14.85%*	–13.85%*	–31.16%*
	Cold to Warm	–43.26%*	–36.91%*	–32.91%*	–53.83%*
Wind Velocity	Calm to Breezy	+2.29%ns	+1.68%ns	+9.42%ns	+3.08%ns
	Breezy to Windy	+19.36%*	+17.01%*	+13.69%*	+23.20%*
	Calm to Windy	+22.10%*	+18.98%*	+24.39%*	+27.00%*
Precipitation	Dry to Wet	+13.11%*	+12.00%*	+10.66%*	+21.48%*
	Wet to Very Wet	–2.14%ns	–5.83%ns	+4.75%ns	–20.43%ns
	Dry to Very Wet	+10.70%ns	+5.47%ns	+15.92%*	–3.34%ns
Cloud Cover	Sunny to Cloudy	+5.72%ns	+4.61%ns	+4.55%ns	+5.63%ns
	Cloudy to Overcast	+22.03%*	+17.45%*	+18.79%*	+28.97%*
	Sunny to Overcast	+29.02%*	+22.86%*	+24.20%*	+36.23%*

Temperature had the strongest and most consistent impact on viewing behaviour. Across all genres, the most pronounced changes occurred between cold and mild conditions, where the average daily viewing time dropped significantly. From a cold to a warm day, drama decreased by 10 689 minutes (approximately 178 hours, +36.91%), and comedy by 6 218 minutes (about 104 hours). Adventure showed a drop of just over 34 hours, while knowledge declined by more than 4.5 hours per day (-53.83%). All pairwise Games–Howell tests were statistically significant.

Wind Velocity showed moderate but clear effects, with most of the change occurring between breezy and windy conditions. From a calm to a windy day, viewing time increased by 6 340 minutes (around 105 hours) for drama, 3 628 minutes (around 60 hours) for comedy, and 1760 minutes (29 hours) for adventure. Knowledge increased by about 2.7 hours (+27%). The transition from calm to breezy was not significant, while breezy to windy changes were statistically significant for most genres. These findings suggest that stronger wind conditions may encourage indoor media consumption.

Precipitation Height exhibited the weakest and most irregular effects. Viewing time increased from dry to light rain for all genres, but often slightly declined under very wet conditions. For example, drama rose by 4 094 minutes (about 68 hours) from a dry to a wet day, then dropped by 2 227 (37 hours) minutes when moving to very wet. Knowledge also showed a small rise followed by a big decline of 20.43%. Interestingly, adventure continued to increase even under heavy precipitation, a pattern that contrasts with the trend in other genres. Statistically significant effects were found only for the transition from dry to wet, while other comparisons were non-significant, likely due to the low number of very wet days.

Cloud Cover showed a similar pattern to Wind Velocity, with most of the change occurring between cloudy and overcast conditions. On an overcast day, the drama viewing time increased by 7 200 minutes (120 hours), comedy by 4 418 minutes (about 73 hours), drama by 5,200 minutes (87 hours), and adventure by 1 749 minutes (29 hours). Knowledge also saw a rise of roughly 3.4 hours (+36.23%). The shifts from sunny to cloudy were not statistically significant.

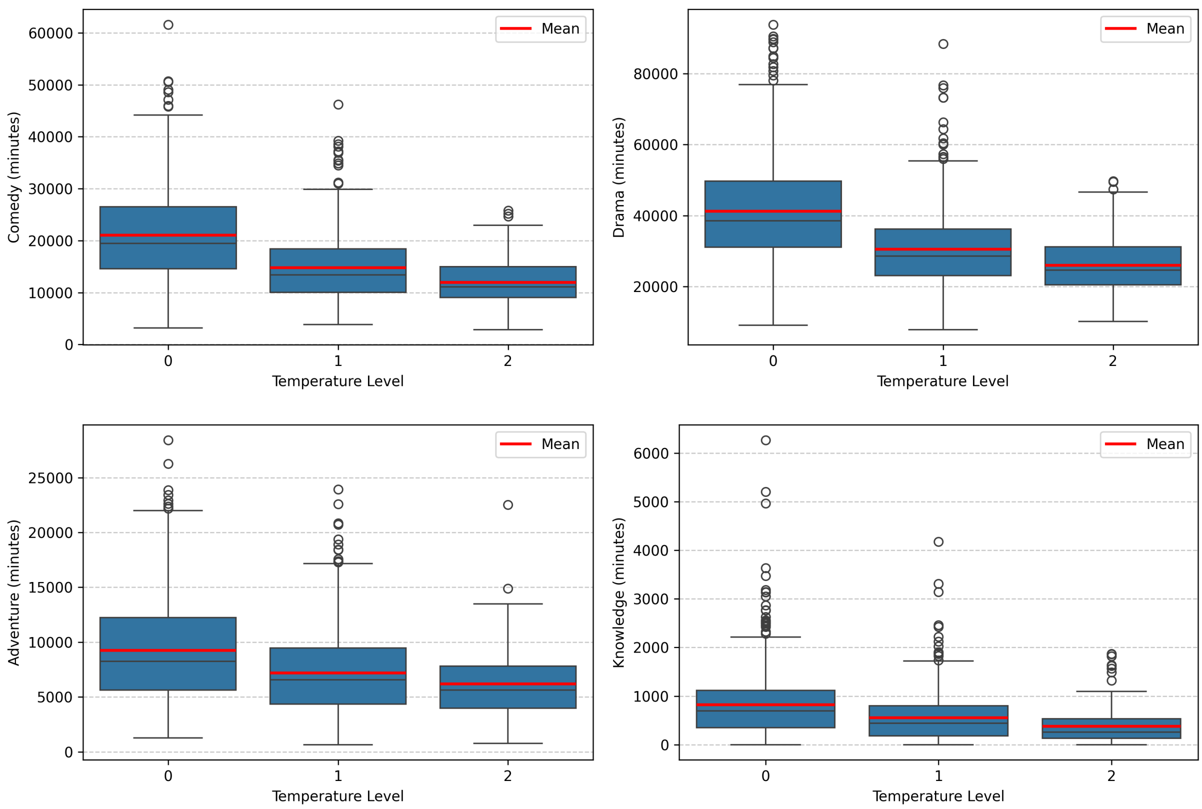


Figure 3.24. Box-plots of genre viewing time by temperature level.

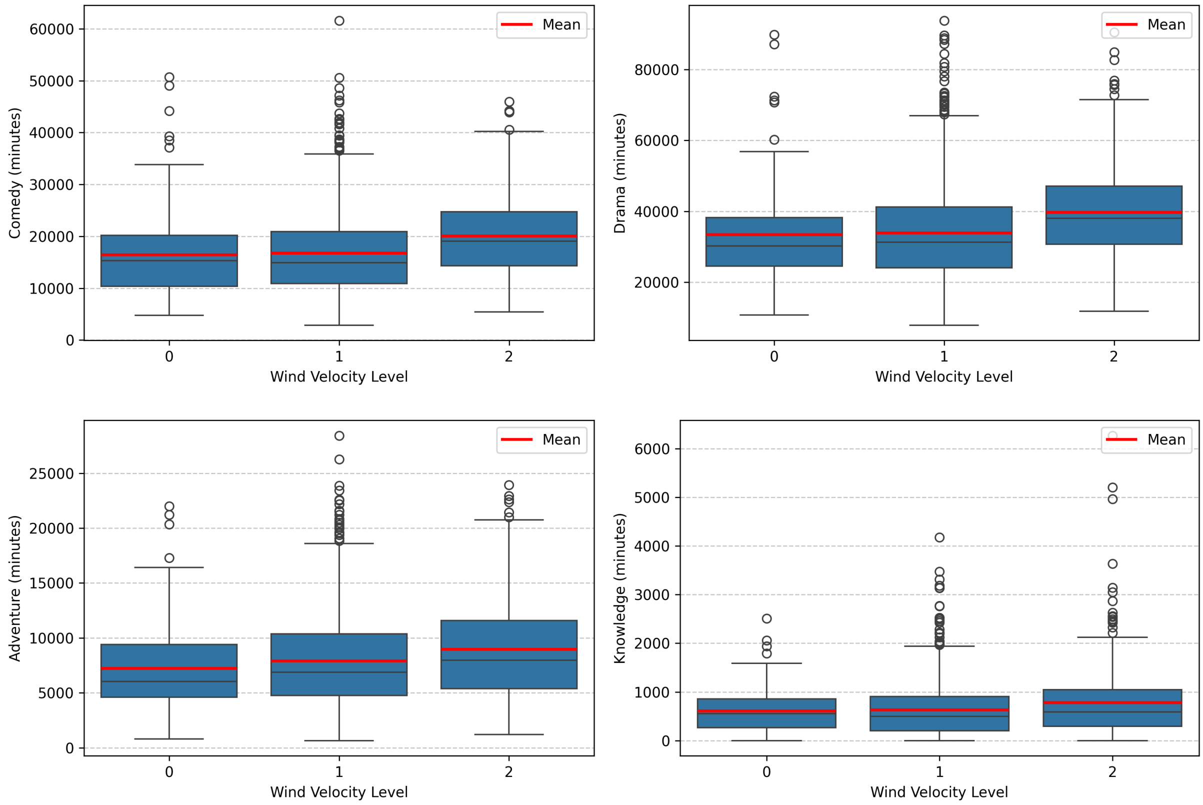


Figure 3.25. Box-plots of genre viewing time by wind velocity level.

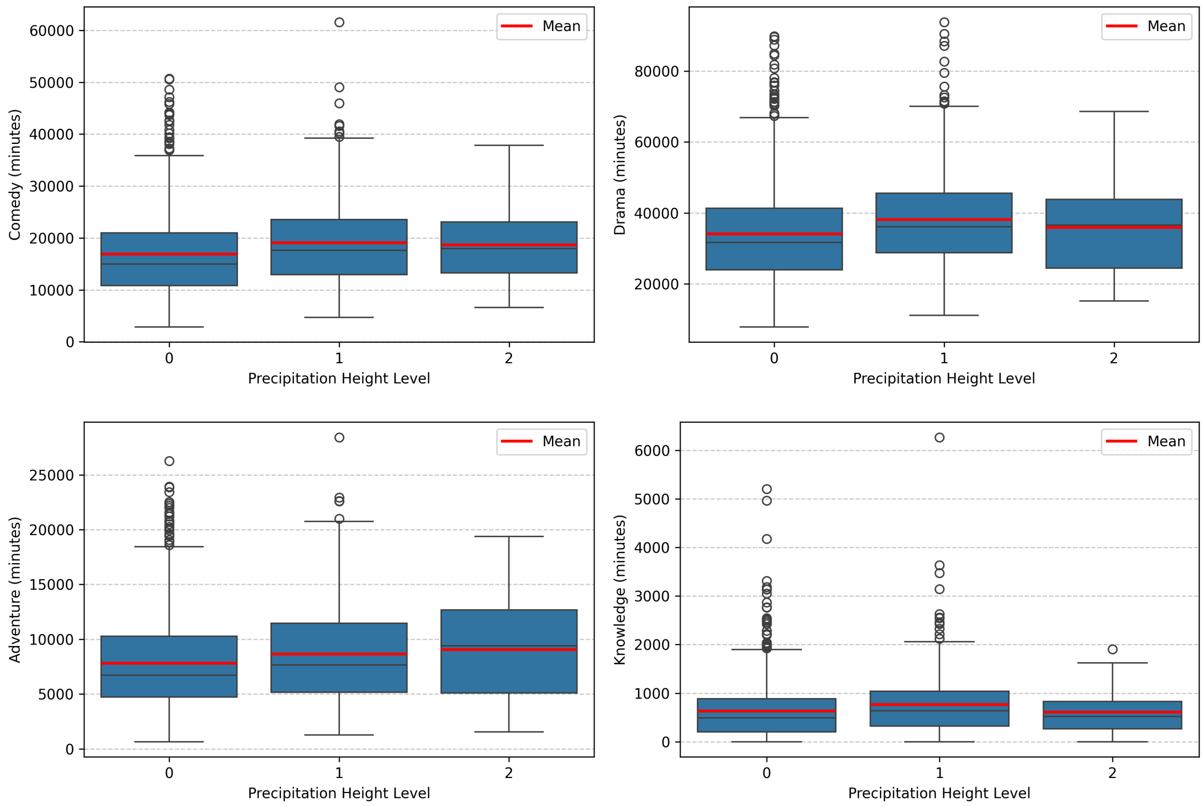


Figure 3.26. Box-plots of genre viewing time by precipitation height level.

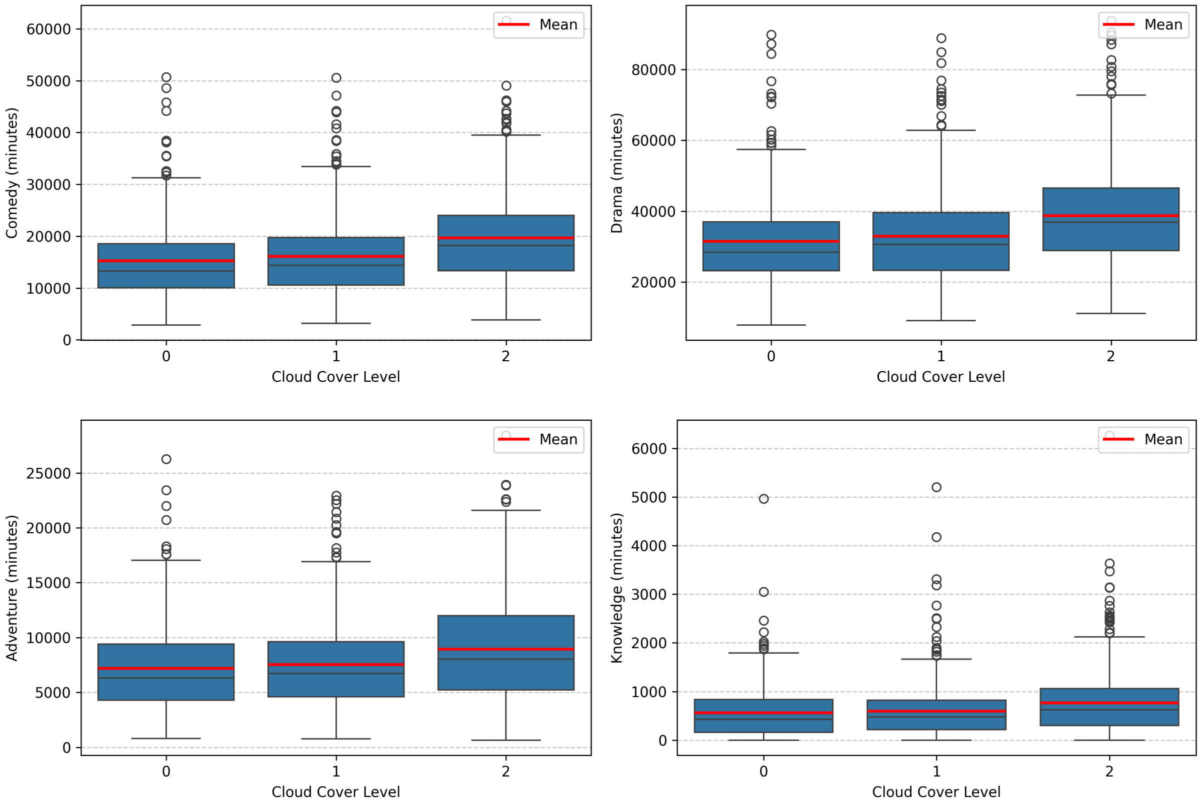


Figure 3.27. Box-plots of genre viewing time by cloud cover level.

Chapter 4

Discussion

This study set out to explore whether daily weather conditions influence film genre preferences, based on the idea that changes in weather may affect mood, which in turn could influence viewing behaviour. The findings provide partial support for this assumption. Mood-related genres showed consistent and interpretable changes in response to temperature, wind velocity, and cloud cover. These patterns were statistically significant in most cases and consistent both in total viewing time and relative terms. This consistency suggests a potential connection between environmental discomfort and the tendency to choose emotionally engaging content, possibly as a way to regulate mood or cope with external conditions.

Among the mood-related genres, drama and comedy were most clearly affected by weather changes, followed by adventure. This trend was also supported by the correlation analysis. The differences in how genres responded may reflect the emotional engagement each one offers. Adventure can be emotionally engaging, but possibly not to the same extent as drama or comedy. However, other factors could also play a major role in shaping genre choices. These include the design of the streaming platform, the availability of certain content, and general popularity patterns among viewers.

In contrast, viewing patterns for the genre knowledge were harder to interpret. Although relative changes were sometimes larger than those seen in the mood-related genres, such as a decrease of over 50 percent from cold to warm days, these changes were based on very low overall viewing levels. In absolute terms, the daily difference represented just a few additional hours of viewing across the entire ZIP code area. Without more detailed information about users, such as who is watching, for example, their age, profession, or gender, and how many individuals are involved, it is difficult to tell whether these changes reflect actual behaviour or random variation. Even small fluctuations in the number of people watching knowledge content could create large percentage changes. Additionally, the correlation analysis showed that knowledge had the weakest links to weather parameters, suggesting that these relative changes may not be reliable or meaningful.

Across all genres, viewing time increased under colder, windier, and darker circumstances. In contrast, for precipitation, the presence of precipitation itself seemed to have some influence on behaviour more than its intensity. At the same time, the results point to a curious contradiction. If unpleasant weather conditions, such as cold, wind, and cloudiness, lead to higher viewing levels across genres, one might assume this results from people staying indoors more frequently on such days. Yet precipitation, which is usually considered the weather factor most likely to keep people indoors, showed the weakest connection to changes in viewing time. This raises further questions about what truly drives shifts in media consumption. One possibility is that staying indoors doesn't automatically increase film viewing. Another possibility is that precipitation may not influence indoor behaviour as much as is often assumed. To understand this better, further analysis using more balanced and varied precipitation data would be beneficial.

While this study provides primary evidence for weather-related variation in genre preferences, several methodological and conceptual limitations should be noted. These challenges relate to the weather data, the genre data, and the analytical approach.

- **Weather-Related Limitations:**
  - **Subjective weather states and their definition:** The categorization of weather variables into perceived levels (e.g., cold, mild, warm) was based on physical thresholds rather than established psychological frameworks. While the classification aimed to reflect typical human perception, more research is needed to validate which thresholds correspond to actual behavioural effects. A standardized model for mapping meteorological measures to subjective experiences would strengthen interpretability.
  - **Mismatch between weather measurement and viewing location:** The weather data used in the study were recorded at the city level and may not perfectly represent conditions in the postal area 10178. Local microclimates can vary within a city, particularly in urban settings. The assumption of uniformity across Berlin was necessary due to data limitations, but the degree of deviation between recorded and experienced weather was not quantified.
  - **Correlated weather parameters:** Weather variables such as temperature, wind, cloud cover, and precipitation are often interdependent. For example, colder days tend to be windier and cloudier. This co-occurrence complicates the isolation of effects, as it is difficult to determine whether a behavioural change is due to one variable or the combination of several.
  - **Delayed or lasting weather effects:** The analysis assumes immediate effects of daily weather, yet some impacts on mood or behaviour may be delayed or cumulative. For instance, several dark days in a row might influence viewing differently than a single overcast afternoon. Capturing such patterns would require a temporal model beyond daily aggregation.
- **Genre-Related Limitations:**
  - **Viewing time does not capture engagement:** The measure used in this study was total daily viewing time, which does not necessarily reflect attentiveness or emotional engagement. For example, users might leave content playing passively, fall asleep, or multitask. This introduces noise into the relationship between weather and actual content preference.
  - **Multiple genre assignments:** Many films on the platform were labeled with multiple genres, which poses analytical challenges. A single session may simultaneously contribute to comedy, drama, and adventure, potentially inflating correlations and complicating causal interpretation.
  - **Non-independence of daily data:** The assumption of independent observations is violated in some cases, such as when viewing sessions begin on one day and continue into the next. This introduces autocorrelation across days that standard ANOVA procedures cannot account for.
  - **Restricted geographic sample:** The analysis focused exclusively on ZIP code 10178, which, limits generalizability. No information was available about the population size, stability, or user turnover in this area. Demographics could substantially influence genre preferences and weather sensitivity.
  - **Curated genre definitions:** Genre labels were assigned by platform curators and may reflect organizational or editorial decisions rather than audience perception. Without a standardized or user-informed tagging system, it is uncertain whether genre boundaries align with user expectations or emotional content.

- **Methodological considerations:**
  - **Data quality and resolution:** The analysis would benefit from a richer dataset that includes user-level information, session duration, device type, and more localized weather data. Improved outlier detection and treatment will also enhance robustness.
  - **Simplified analytical approach:** The study relied on separate correlations, one-way Welch ANOVAs, and relative changes, which allowed for a straightforward analysis. However, this approach does not account for the interdependence, potential correlations among weather variables, or temporal autocorrelation in daily viewing. More sophisticated statistical methods could better handle dependent observations and the simultaneous influence of multiple factors.
  - **Lack of distinction between snowfall and rainfall:** The study treated all precipitation as a single variable, without separating snow from rain. These forms of precipitation might have distinct effects on mood, and future analyses should differentiate them if data permit.

Chapter 5

Conclusion

This thesis offers an initial exploration of how daily weather conditions relate to changes in film genre viewing patterns, particularly for emotionally engaging genres. The findings suggest that colder, cloudier, and windier days are associated with increased viewing of genres such as comedy and drama, pointing to a potential connection between environmental discomfort and a preference for emotionally engaging media. However, these results should not be interpreted as definitive evidence. Mood was not directly measured, and other factors may also influence viewing behaviour.

These patterns should therefore be viewed as a preliminary step, prompting further investigation. Future research could build on this work by using higher-resolution datasets, incorporating user-specific metrics, or directly capturing emotional states to better identify the underlying mechanisms. Comparative studies across multiple locations and an extended study period would help clarify whether these patterns are consistent. Moreover, weather may shape other aspects of digital behaviour, such as music listening, reading, or social media use. Understanding these broader effects could inform more adaptive recommendation systems or strategically timed content releases and production planning.

Bibliography

[1] Open Data Informationsstelle Berlin, “PLZ - Postleitzahlgebiete Berlins,” accessed: April 20, 2025. [Online]. Available: https://daten.odis-berlin.de/de/dataset/plz/

[2] AGF, “Average Daily Television Viewing Time in Germany from 1997 to 2023 (in Minutes),” 2024, accessed: May 14, 2025. [Online]. Available: https://www.statista.com/statistics/380182/tv-consumption-viewing-time-germany/

[3] Seven.One Media GmbH, “Average Daily Viewing Time of Pay Video on Demand (VOD) in Germany from 2015 to 2023 (in Minutes),” 2024, accessed: May 14, 2025. [Online]. Available: https://www.statista.com/statistics/1300249/vod-use-average-daily-duration-germany/

[4] S. Gössling, C. Neger, R. Steiger *et al.*, “Weather, Climate Change, and Transport: A Review,” *Natural Hazards*, vol. 118, pp. 1341–1360, 2023. [Online]. Available: https://doi.org/10.1007/s11069-023-06054-2

[5] K. Minor, E. Moro, and N. Obradovich, “Worse Weather Amplifies Social Media Activity,” *Psychological Science*, vol. 36, no. 1, pp. 35–54, 2025. [Online]. Available: https://doi.org/10.1177/09567976241306099

[6] D. Zillmann, “Mood Management through Communication Choices,” *American Behavioral Scientist*, vol. 31, no. 3, pp. 327–340, 1988. [Online]. Available: https://doi.org/10.1177/000276488031003005

[7] S. Sekhon and V. Gupta, “Mood Disorder,” accessed: Jun. 12, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK558911/

[8] S. Zulfiqar and A. Islam, “Exploring the Role of Emotions and Moods in Decision Making: Study on the Use of Structured Decision Approach and Intuition,” *International Journal of Engineering and Management Sciences*, vol. 2, pp. 140–149, 2017. [Online]. Available: https://doi.org/10.21791/IJEMS.2017.3.14.

[9] D. Greenwood, “Of Sad Men and Dark Comedies: Mood and Gender Effects on Entertainment Media Preferences,” *Mass Communication and Society*, vol. 13, no. 3, pp. 232–249, 2010. [Online]. Available: https://doi.org/10.1080/15205430903186526

[10] P. Winoto and T. Y. Tang, “The role of user mood in movie recommendations,” *Expert Systems with Applications*, vol. 37, no. 8, pp. 6086–6092, 2010. [Online]. Available: https://doi.org/10.1016/j.eswa.2010.02.117.

[11] J. J. A. Denissen, L. Butalid, L. Penke, and M. A. G. van Aken, “The Effects of Weather on Daily Mood: A Multilevel Approach,” *Emotion*, vol. 8, no. 5, pp. 662–667, 2008. [Online]. Available: https://doi.org/10.1037/a0013497

[12] M. C. Keller, B. L. Fredrickson, O. Ybarra, S. Côté, K. Johnson, J. A. Mikels, A. Conway, and T. Wager, "A Warm Heart and a Clear Head: The Contingent Effects of Weather on Mood and Cognition," *Psychological Science*, vol. 16, no. 9, pp. 724–731, 2005. [Online]. Available: https://doi.org/10.1111/j.1467-9280.2005.01602.x

[13] T. A. Klimstra, T. Frijns, L. Keijsers, J. J. A. Denissen, Q. A. Raaijmakers, M. A. G. van Aken, H. M. Koot, P. A. C. van Lier, and W. H. J. Meeus, "Come Rain or Come Shine: Individual Differences in How Weather Affects Mood," *Emotion*, vol. 11, no. 6, pp. 1495–1499, 2011. [Online]. Available: https://doi.org/10.1037/a0024649

[14] Filmfriend, "About Us - Filmfriend," accessed: Mar. 18, 2025. [Online]. Available: https://www.filmfriend.de/de/about-us

[15] World Meteorological Organization, "Weather," 2025, accessed: Mar. 18, 2025. [Online]. Available: https://wmo.int/topics/weather

[16] I. Bondebjerg, "Film: Genres and Genre Theory," in *International Encyclopedia of the Social and Behavioral Sciences*, J. D. Wright, Ed. Elsevier, 2015, pp. 160–164. [Online]. Available: https://doi.org/10.1016/B978-0-08-097086-8.95052-9

[17] E. Zwiky, P. König, R. M. Herrmann, A. Küttner, J. Selle, L. E. Ptasczynski, K. Schöniger, M. Rutenkröger, V. Enneking, T. Borgers, M. Klug, K. Dohm, E. J. Leehr, J. Bauer, U. Dannlowski, and R. Redlich, "How Movies Move Us – Movie Preferences Are Linked to Differences in Neuronal Emotion Processing of Fear and Anger: An fMRI Study," *Frontiers in Behavioral Neuroscience*, vol. 18, 2024. [Online]. Available: https://doi.org/10.3389/fnbeh.2024.1396811

[18] Deutscher Wetterdienst, "German Weather Service (DWD) – About us," accessed: March 20, 2025. [Online]. Available: https://www.dwd.de/EN/aboutus/aboutus_node.html

[19] European Environment Agency, "Thermal Comfort Indices — Universal Thermal Climate Index, 1979–2020," 2024, accessed: June 27, 2025. [Online]. Available: https://climate-adapt.eea.europa.eu/en/metadata/indicators/thermal-comfort-indices-universal-thermal-climate-index-1979-2019

[20] National Weather Service, "Beaufort Wind Scale," 2025, accessed: June 27, 2025. [Online]. Available: https://www.weather.gov/mfl/beaufort

[21] World Meteorological Organization, "Guide to Meteorological Instruments and Methods of Observation (WMO-No. 8), Annex: Criteria for Light, Moderate and Heavy Precipitation Intensity," 2018, accessed: June 27, 2025. [Online]. Available: https://www.weather.gov/media/epz/mesonet/CWOP-WMO8.pdf

[22] ——, "Code Table D-8: Cloud Amount Reported at Aerodrome," 2019, accessed: June 27, 2025. [Online]. Available: https://codes.wmo.int/49-2/_CloudAmountReportedAtAerodrome

[23] D. C. Howell, *Statistical Methods for Psychology*, 7th ed., Belmont, Calif., 2010.

[24] M. J. Blanca, R. Alarcón, J. Arnau, R. Bono, and R. Bendayan, "Non-normal data: Is ANOVA still a valid option?" *Psicothema*, vol. 29, no. 4, pp. 552–557, 2017. [Online]. Available: https://doi.org/10.7334/psicothema2016.383

[25] World Health Organization, "Coronavirus disease (COVID-19) pandemic," 2025, accessed June 29, 2025. [Online]. Available: https://www.who.int/europe/emergencies/situations/covid-19

[26] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. Vallina-Rodriguez, O. Hohlfeld, and G. Smaragdakis, “The Lockdown Effect: Implications of the COVID-19 Pandemic on Internet Traffic,” in *Proceedings of the Internet Measurement Conference (IMC ’20)*. Virtual Event, USA: ACM, Oct. 2020, pp. 1–18. [Online]. Available: https://doi.org/10.1145/3419394.3423658

[27] Deutscher Wetterdienst, “DATASET DESCRIPTION,” accessed: March 20, 2025. [Online]. Available: https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl//DESCRIPTION_obsgermany-climate-daily-kl_en.pdf

Appendix

Supporting Materials for the Methodology

Official Weather Dataset Documentation

The following pages contain the official description of the weather dataset provided by Deutscher Wetterdienst [27].

DATASET DESCRIPTION

Daily station observations (temperature, pressure, precipitation, sunshine duration, etc.) for Germany

Version: v24.3

Publication date: 2024

Cite data set as:	Daily station observations (temperature, pressure, precipitation, sunshine duration, etc.) for Germany, Version v24.3
Dataset-ID:	urn:wmo:md:de-dwd-cdc:obsgermany-climate-daily-kl
Dataset-URL:	https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/recent/
Dataset-URL:	https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/historical/
Dataset-URL:	https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl //timeseries_overview//

ABSTRACT

These data originate from the stations of the DWD and legally as well as qualitatively equal partner network stations. Extensive station metadata, such as station relocations, instrument changes, reference time changes, algorithm changes or operator information are included.

The dataset is divided into a versioned part with completed quality check, in the directory ./historical/.
And a part for which the quality check has not yet been completed, in the directory ./recent/.

The folder ./timeseries_overview/ contains information about long time series.

POINT OF CONTACT

Deutscher Wetterdienst
CDC - Vertrieb Klima und Umwelt
Frankfurter Strasse 135
63067 Offenbach
Tel:+ 49 (0) 69 8062-4400
Fax:+ 49 (0) 69 8062-4499
E-Mail:klima.vertrieb@dwd.de

DATASET DESCRIPTION

Parameter	precipitation parameters, form of precipitation, relative humidity, vapor pressure, air pressure at station level, sunshine duration, air temperature at 2 m, wind velocity, wind gust, air temperature near ground, snow depth, cloud coverage, precipitation height
Unit(s)	eighth, beaufort, °C, %, mm, m/s, cm, hPa
Statistical processing	daily mean, daily sum, time series, daily min, daily max
Temporal coverage	1781-01-01 -- ...
Spatial coverage	stations in Germany
Projection	WGS 84 (EPSG:4326)
Format description	Recent daily station observations (temperature, pressure, precipitation, sunshine duration, etc.) for Germany, quality control not completed : In the folder recent/ for each station a zip-archive is provided. The zip-archive contains the data and meta information about the station, instruments and algorithms. The naming schema of the zip-archives is: *_{product_code}_{station_id}_akt.zip

Format description [Historical daily station observations (temperature, pressure, precipitation, sunshine duration, etc.) for Germa](#) :
In the folder recent/ for each station a zip-archive is provided.
The zip-archive contains the data and meta information about the station, instruments and algorithms.

The naming schema of the zip-archives is: *_{product_code}_{station_id}_{begin_date}_{end_date}_hist.zip

Format description [timeseries overview](#) :
In the folder ./timeseries_overview, information on long time series is available. The files provided (TimeSeries_[DataType]_[Interval]_GE_[XXXYears]_[Parameter].html or ***.txt) contain a sorted overview of stations for which time series of >=100, >=50 and >=30 years are available. Information on the proportion of missing values is also provided.

Content description
Stations_id := Identifier of the station;
Start := Start date of the time series;
End := End date of the time series;
Number_years := Number of years of measurement operation;
Missing_Years := Number of missing years of measurement operation;
Missing_values := number of missing values ;
max(Missing_period)>=25 := More than 25 years missing in the time series: Indication of start date and end date;
Station name := Station name of the current location ;
Federal state := Name of the federal state

application schema csv dialect description

delimiter	line terminator	header	quote char
;	\\r\\n	true	\"

csv content description

column name	description	uom	type	format
STATIONS_ID	Station ID		VARCHAR2	
MESS_DATUM	reference date		NUMBER	YYYYMMDD
QN_3	quality level of the following columns		NUMBER	numerical code
FX	daily maximum of windgust	m/s	NUMBER	9990.0
FM	daily mean of wind velocity	m/s	NUMBER	9990.0
QN_4	quality level of the following columns		NUMBER	numerical code
RSK	daily precipitation height	mm	NUMBER	9990.0
RSKF	precipitation form	numerical code	NUMBER	
SDK	daily sunshine duration	h	NUMBER	9990.0
SHK_TAG	daily snow depth	cm	NUMBER	9990.0
NM	daily mean of cloud cover	1/8	NUMBER	9990.0
VPM	daily mean of vapor pressure	hPa	NUMBER	9990.0
PM	daily mean of pressure	hPa	NUMBER	9990.0
TMK	daily mean of temperature	°C	NUMBER	9990.0
UPM	daily mean of relative humidity	%	NUMBER	9990.0
TXK	daily maximum of temperature at 2 m height	°C	NUMBER	9990.0
TNK	daily minimum of temperature at 2m height	°C	NUMBER	9990.0
TGK	daily minimum of air temperature at 5 cm above ground	°C	NUMBER	9990.0

Quality Information The QUALITAETS_NIVEAU (QN) shows the quality control procedure applied for a data report (of several parameters) for a certain reporting time.

Data before and including 1980 can reach as best quality check level QN=5. Data after 1980 can reach QN=10 as best quality check level.

QN = 1 : only formal control;
QN = 2 : controlled with individually defined criteria;
QN = 3 : automatic control and correction;
QN = 5 : historic, subjective procedures;
QN = 7 : second control done, before correction;
QN = 8 : quality control outside ROUTINE;
QN = 9 : not all parameters corrected;
QN = 10 : quality control finished, all corrections finished.

The QUALITAETS_BYTE (QB) denotes whether the value was objected to and/or corrected.

QB = 0 : denotes not flagged,
QB = 1 : had no objections (either checked and not objected, or not checked and not objected, this can be interpreted only when considering QN);
QB = 2 : corrected;
QB = 3 : confirmed with objection rejected;
QB = 4 : added or calculated;
QB = 5 : objected;
QB = 6 : only formally checked;
QB = 7 : formal objection;
QB = -999 : quality flag does not exist.

DATA ORIGIN

The data are taken from the station measuring networks of Deutscher Wetterdienst as well as its predecessor organisations.
The dataset is regularly updated with recent as well as with recovered historical data.
From 1997 onwards, the data have been imported operationally into the central specialist database and archived, see Behrendt et al., 2011, and Kaspar et al., 2013.
Note that when going back to historical times, guidelines on observation procedure, instruments and observation times were issued by the authority in charge (see, e.g., Freydank, 2014), and might be incompletely recorded in the metadata.
As explained in Kaspar et al., 2013 in the early years numerous meteorological agencies were active in the area of todays Germany. After establishment of the der International Meteorological Organization (IMO) in 1873, the various standards were gradually harmonized, resulting in a single standard 1936.
After 1945, the standards in East and West Germany developed differently, and were harmonized again after re-unification in 1990.
Between the end of the nineties and 2009 many stations were changed from manual to automated.

RESOURCE MAINTENANCE

[Data maintenance]: In the directory recent/ the data files are updated daily. On a rolling basis, the data of the last 500 days - up to yesterday - are exchanged.
Quality control has not yet been completed for these data, so there may always be changes in the values.

In the directory historical/ the data files are updated annually.
Quality control has been completed for this data, so that the values for the version are constant.
During the annual version change, both corrections and historical additions are incorporated.

VALIDATION AND UNCERTAINTY ESTIMATE

The quality control (see Spengler, 2002) of this data is not completed yet. Various levels of quality control (see Kaspar et al., 2013) are in progress.

UNCERTAINTIES

The stations are nowadays selected and operated according to WMO guidelines. Though these guidelines aim at minimizing possible local effects, still some applications of certain parameters may require the consideration of local and regional effects.

CONSIDERATIONS FOR APPLICATIONS

For any data analysis, the metadata available in the *.zip files should be taken into account.

ADDITIONAL INFORMATION

For extending the time series into the past, see subdirectories ../historical/. When data from both directories "historical" and "recent" are used together, the difference in the quality control procedure should be considered. For the long term stability consider the uncertainties explained in the data set descriptions within subdirectories /historical/.

LITERATURE

Behrendt, J., et al.: Beschreibung der Datenbasis des NKDZ. Version 3.5, Offenbach, 15.02.2011.

DWD Vorschriften und Betriebsunterlagen Nr. 2 (VuB 2), Wetterschlüsselhandbuch Band D, Nov 2013.

DWD Vorschriften und Betriebsunterlagen Nr. 3 (VuB 3), Beobachterhandbuch (BHB) für Wettermeldestellen des synoptisch-klimatologischen Mess- und Beobachtungsnetzes, März 2014a.

DWD Vorschriften und Betriebsunterlagen Nr. 3 (VuB 3), Technikerhandbuch (THB) für Wettermeldestellen des synoptisch-klimatologischen Mess- und Beobachtungsnetzes, März 2014b.

Kaspar, F., et al.: Monitoring of climate change in Germany – data, products and services of Germany`s National Climate Data Centre. Adv. Sci. Res., 10, doi:10.5194/asr-10-99-2013, 99–106, 2013.

Spengler, R.: The new Quality Control- and Monitoring System of the Deutscher Wetterdienst. Proceedings of the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation, Bratislava, 2002.

COPYRIGHT

REVISION HISTORY

This document is maintained by Deutscher Wetterdienst, Climate Data Center (CDC) - Betrieb, last edited at 2024-06-05.

Viewing Time from Berlin

zipCode	totalMetric	Comedy	Drama	Adventure	Knowledge
10178	85293533	25529132	51393849	11792665	967783
10623	1230793	367160	780934	110141	16005
10115	363264	117989	234830	30290	3012
15566	315948	105459	186653	47644	4831
15537	151149	46820	86284	41528	949

Figure A.1. Total minutes of viewing time per ZIP code within Berlin from 2020 to 2023, highlighting the dominant contribution of ZIP code 10178.

ANOVA Assumptions Tests

The Shapiro–Wilk test was conducted to assess normality. A p under 0.05 is considered significant, which means that the null hypothesis of normality is rejected and the data for that group deviate from a normal distribution. The Levene test was conducted to assess homogeneity of variances. A p under 0.05 is considered significant, which means that the null hypothesis of equal variances is rejected and the variances across groups are unequal.

Table A.1. Normality (Shapiro–Wilk) and homogeneity (Levene) tests for each weather–genre combination

Weather	Genre	Shapiro p level 0	Shapiro p level 1	Shapiro p level 2	Sig. Shapiro	Levene F	Levene p	Sig. Levene
Temperature	Comedy	3.51×10^{-14}	6.69×10^{-16}	9.11×10^{-05}	Yes	43.86	3.18×10^{-19}	Yes
Temperature	Drama	5.19×10^{-12}	5.70×10^{-14}	3.34×10^{-04}	Yes	33.92	3.98×10^{-15}	Yes
Temperature	Adventure	9.44×10^{-16}	1.05×10^{-14}	2.57×10^{-07}	Yes	30.35	1.22×10^{-13}	Yes
Temperature	Knowledge	4.86×10^{-25}	1.33×10^{-22}	3.47×10^{-13}	Yes	25.15	1.82×10^{-11}	Yes
Wind Velocity	Comedy	1.96×10^{-09}	4.80×10^{-23}	2.67×10^{-07}	Yes	0.16	0.849	No
Wind Velocity	Drama	1.62×10^{-08}	5.33×10^{-20}	4.07×10^{-07}	Yes	0.49	0.613	No
Wind Velocity	Adventure	1.27×10^{-07}	4.12×10^{-21}	1.36×10^{-09}	Yes	5.30	0.005	Yes
Wind Velocity	Knowledge	1.65×10^{-06}	3.54×10^{-27}	9.16×10^{-21}	Yes	7.85	0.0004	Yes
Precipitation Height	Comedy	3.18×10^{-23}	2.68×10^{-11}	9.93×10^{-02}	Yes	0.35	0.708	No
Precipitation Height	Drama	2.42×10^{-20}	2.15×10^{-10}	1.88×10^{-01}	Yes	0.14	0.866	No
Precipitation Height	Adventure	2.26×10^{-22}	5.05×10^{-11}	6.65×10^{-02}	Yes	2.27	0.103	No
Precipitation Height	Knowledge	1.81×10^{-31}	5.51×10^{-19}	4.63×10^{-03}	Yes	2.82	0.060	No
Cloud Cover	Comedy	3.41×10^{-15}	9.34×10^{-16}	2.57×10^{-14}	Yes	4.72	0.009	Yes
Cloud Cover	Drama	2.41×10^{-13}	1.64×10^{-14}	6.91×10^{-12}	Yes	2.50	0.083	No
Cloud Cover	Adventure	3.52×10^{-13}	3.19×10^{-14}	1.57×10^{-14}	Yes	11.95	7.15×10^{-06}	Yes
Cloud Cover	Knowledge	4.25×10^{-19}	9.13×10^{-24}	4.56×10^{-23}	Yes	7.23	0.0008	Yes

Supporting Materials for the Results

Significance of Correlations

Table A.2. Correlation Significance between Weather Parameters and Genre Viewing Time ($p < 0.05$)

Weather Parameter	Genre	p	Significant
Temperature	Comedy	6.53×10^{-88}	Yes
Temperature	Drama	1.55×10^{-83}	Yes
Temperature	Adventure	4.47×10^{-34}	Yes
Temperature	Knowledge	1.85×10^{-29}	Yes
Wind Velocity	Comedy	2.03×10^{-12}	Yes
Wind Velocity	Drama	2.80×10^{-13}	Yes
Wind Velocity	Adventure	2.21×10^{-8}	Yes
Wind Velocity	Knowledge	7.02×10^{-4}	Yes
Precipitation Height	Comedy	7.61×10^{-3}	Yes
Precipitation Height	Drama	2.88×10^{-2}	Yes
Precipitation Height	Adventure	6.77×10^{-3}	Yes
Precipitation Height	Knowledge	7.02×10^{-1}	No
Cloud Cover	Comedy	7.60×10^{-20}	Yes
Cloud Cover	Drama	1.32×10^{-18}	Yes
Cloud Cover	Adventure	4.69×10^{-12}	Yes
Cloud Cover	Knowledge	2.45×10^{-8}	Yes

Welch ANOVA and Games–Howell Post-hoc p-values

Table A.3. Welch ANOVA tests and Games–Howell pairwise p-values for mean viewing time by weather level and genre.

Weather	Genre	Welch F	Welch p	Games–Howell p: 0 - 1	Games–Howell p: 0 - 2	Games–Howell p: 1 - 2
Temperature	Comedy	203.687	≈ 0.000	6.82×10^{-14}	≈ 0.000	1.27×10^{-10}
	Drama	192.448	≈ 0.000	≈ 0.000	≈ 0.000	1.79×10^{-8}
	Adventure	63.060	≈ 0.000	≈ 0.000	≈ 0.000	1.31×10^{-3}
	Knowledge	66.065	≈ 0.000	≈ 0.000	≈ 0.000	≈ 0.000
Wind Velocity	Comedy	21.597	≈ 0.000	0.865	6.76×10^{-5}	1.10×10^{-9}
	Drama	22.646	≈ 0.000	0.895	2.38×10^{-5}	4.25×10^{-10}
	Adventure	9.825	≈ 0.000	0.196	1.50×10^{-4}	8.40×10^{-4}
	Knowledge	5.268	0.006	0.938	0.0130	0.0057
Precipitation	Comedy	10.476	≈ 0.000	3.30×10^{-5}	0.204	0.927
	Drama	11.389	≈ 0.000	7.00×10^{-6}	0.533	0.451
	Adventure	5.923	0.003	7.16×10^{-3}	0.138	0.815
	Knowledge	5.925	0.003	2.30×10^{-3}	0.948	0.0932
Cloud Cover	Comedy	44.337	≈ 0.000	0.246	≈ 0.000	8.46×10^{-13}
	Drama	40.089	≈ 0.000	0.293	≈ 0.000	1.56×10^{-11}
	Adventure	23.210	≈ 0.000	0.487	5.42×10^{-9}	1.58×10^{-7}
	Knowledge	16.684	≈ 0.000	0.745	9.03×10^{-7}	1.04×10^{-5}

Further Notes

- I used OpenAI's ChatGPT (GPT-4, versions in the period February-July 2025) to check and improve the language.
- The full project files, including original datasets, code, and plots, are available on the project repository at: `https://tubcloud.tu-berlin.de/s/sApsFLrP24d8Ax2`