
Beyond the Beer Tent: An Analysis of Oktoberfest's Yearly Visitor, Consumption, and Price Trends

Jana Hoffmann ^{*1} Julia Graf ^{*2} Jessie Midgley ^{*3} Tithi Rakshit ^{*4}

Abstract

Analyzing data spanning 38 years, we explore the evolution of the Oktoberfest and its impact on tourism in Munich and surrounding areas. Our findings highlight fluctuations in visitor numbers, and reveal insights into the festival's resilience in the face of external factors such as weather. Additionally, we investigate the influx of tourists to Upper Bavaria during Oktoberfest and observe deviations in guest arrivals compared to expectations. Furthermore, we analyze trends in beer and chicken consumption and prices, noting a positive correlation between beer price and consumption. We employ regression models to predict beer and chicken prices for 2023, demonstrating varying degrees of accuracy between linear and Gaussian process regression models.

1. Introduction

The Oktoberfest, held in Munich, Germany, is one of the world's most renowned folk festivals, attracting millions annually. In this paper, we analyze nearly four decades of Oktoberfest data, spanning from 1985 to 2023, to uncover insights into its evolution and various influencing factors. Understanding the dynamics of this iconic event is not only of cultural significance but also holds economic and societal implications.

This paper is structured as follows: Section 2 outlines our data sources and analytical methods. Section 3 presents our findings, elucidating key trends and correlations observed within the datasets, and Section 4 discusses the implications of our findings and offers insights into future festivals.

^{*}Equal contribution ¹Matrikelnummer 5760486, jana2.hoffmann@student.uni-tuebingen.de, MSc Bioinformatics ²Matrikelnummer 5656882, jul.graf@student.uni-tuebingen.de, MSc Bioinformatics ³Matrikelnummer 6620875, jessie.midgley@student.uni-tuebingen.de, MSc Bioinformatics ⁴Matrikelnummer 6635972, tithi.rakshit@student.uni-tuebingen.de, MSc Quantitative Data Science.

Project report for the "Data Literacy" course at the University of Tübingen, Winter 2023/24 (Module ML4201). Style template based on the ICML style files 2023. Copyright 2023 by the author(s).

2. Data and Methods

The annual Oktoberfest data from 1985 to 2022 was retrieved from the Open Data Portal of the Statistical Office Munich (opendata.muenchen.de, 2023). This dataset contains the yearly festival duration, visitor numbers, as well as consumption and prices for beer and roast chicken. Additional data for 2023 was retrieved from the State Capital Munich Portal ([Portal München](https://Portal.Muenchen.de), 2023a; [Landeshauptstadt München, Referat für Arbeit und Wirtschaft](https://www.lmu.de)).

We examine the evolution of visitor numbers over the years and use four additional datasets to investigate how these are influenced by precipitation. Weather datasets were retrieved from the Open Data Server of the German Meteorological Service (DWD) and contain data from the weather stations "München-Stadt" and "München (Bavariaring)". Two of the datasets contain the daily precipitation in millimeters, with one spanning from 1954 to 2022, and the other from 2022 to 2024 (DWD, c;d). The remaining two datasets contain the hourly precipitation in millimeters, in ranges from 1997 to 2022, and from 2022 to 2024 (DWD, a;b). We use the Pearson correlation coefficient (PCC) to investigate the correlation between the average number of daily visitors and the average daily precipitation during each Oktoberfest, and also constrain daily precipitation between 10am and 10pm. The PCC ranges from -1 to 1, with a high absolute value indicating a strong correlation between two variables (Nield, 2022).

To investigate the impact of Oktoberfest on visitor numbers in the surrounding districts of Munich, we obtained monthly visitor numbers for February, March, April, August, September, and October from 2006 to 2023 for all districts in Bavaria from the GENESIS-Online database of the Bavarian State Office for Statistics ([Bayerisches Landesamt für Statistik](https://www.baystat.de)). To generate a heat map showing where Oktoberfest has the greatest impact on visitor numbers, we need to calculate how many visitors to each district also attend the festival. One approach is to calculate the average number of visitors for the preceding and following month (August and October, respectively) to estimate the expected number of visitors without the influence of the Oktoberfest. The expected number can then be subtracted from the actual number of visitors to determine the number of visitors to the Oktoberfest. The percentage deviation between the expected

and actual number of visitors to each district accounts for their varying size and popularity.

To validate the method, we applied it to three consecutive months (February, March, and April), assuming similar visitor numbers as no special events occurred during these months. If this assumption holds, our method should accurately estimate the number of visitors in March by using the average number of visitors in February and April. This validation approach is applicable to any three consecutive months without unusual visitor fluctuations, such as during Christmas markets or similar events.

We investigate visitor, price, and consumption trends over the years, and use the PCC to explore the correlation between beer price and per capita beer consumption. To adjust for inflation over time, beer prices are also divided by the Consumer Price Index (CPI) which measures the monthly average price development of goods and services purchased by German households. The CPI data was collected from the Federal Statistical Office of Germany (Destatis) ([Statistisches Bundesamt](#)).

We utilized a Linear regression model to predict the price for the year 2023, and subsequently, compared it with the empirically observed price for the same year. To achieve an optimal fit, we assessed polynomial degrees ranging from one to ten. Additionally, we explored Gaussian process regression using the Radial Basis Function (RBF) and Dot-Product kernels to estimate the price with a probabilistic approach. This same methodology was applied for forecasting the chicken price for the year 2023.

3. Results

3.1 Visitor trend over the years

From the 1985 Oktoberfest until the most recent one in 2023, we observe a considerable fluctuation in yearly visitor numbers (Figure 1). We also observe a dramatic drop in visitor numbers, from 6.9 million visitors in the year 2000 to only 5.5 million in 2001. Focusing solely on 16-day Oktoberfests, the 2001 visitor count stands out as an outlier, falling below the mean by over 2 standard deviations. The Oktoberfest was cancelled in 2020 and 2021 due to the COVID-19 pandemic, and a new record number of 7.2 million visitors was set in 2023.

3.2 Influence of precipitation on visitor numbers

The PCC for the mean daily visitors and the mean daily precipitation each year is -0.1913. The negative correlation indicates that visitor numbers decrease as precipitation increases. Since the absolute value of the PCC is low, the correlation between the variables is weak. Since mean daily precipitation includes hours outside of festival visiting hours, we also investigate whether precipitation at daytime has a stronger correlation with visitor numbers. We there-

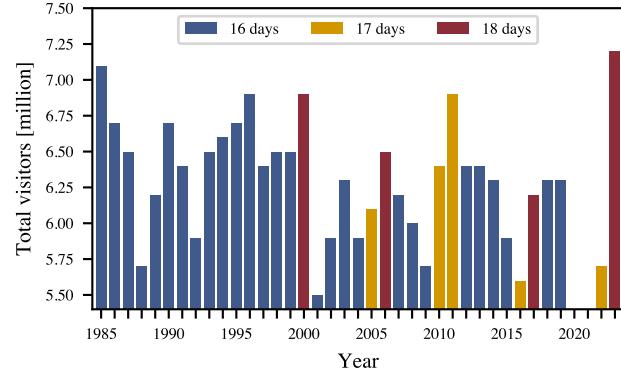


Figure 1. The number of visitors to the Oktoberfest from 1985 to 2023. Bars are coloured depending on the duration of the festival that year.

fore considered precipitation only between 10am and 10pm, and found a PCC of -0.2276, indicating a slightly stronger negative correlation. The persistently weak correlation is reflected in the scattered distribution of the data points in Figure 2.

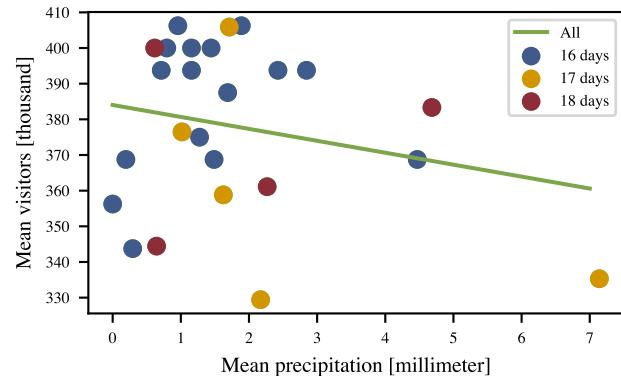


Figure 2. The mean daily visitors as a function of the mean daily precipitation between 10am and 10pm during the Oktoberfest each year. The green line is the linear regression line of all data points. The dots are coloured depending on the duration of the corresponding festival.

3.3 Tourists to Upper Bavaria during Oktoberfest

The number of visitors to Munich in September is boosted by the Oktoberfest (Figure 3). In the nearby districts around Munich (such as Ebersberg), there are smaller but still noticeable deviations between the expected and actual number of guest arrivals. Importantly, the number of guest arrivals consistently surpasses expectations, indicating the impact of Oktoberfest in these areas.

This information can be neatly presented using a heat map to show how the number of visitors in different counties around Munich is affected by the Oktoberfest (Figure 4).

Our validation shows that in almost all cases the March predictions from our method are very close to the actual number of visitors, demonstrating the reliability of our heatmap.

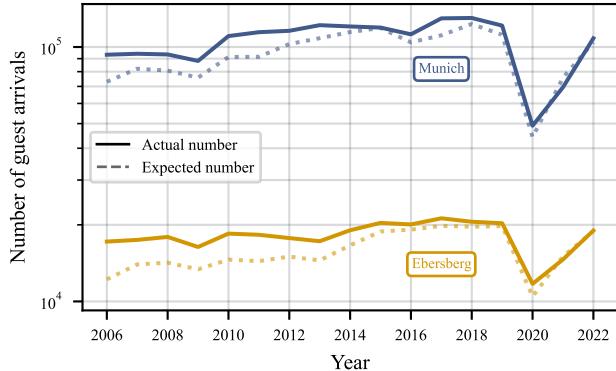


Figure 3. Expected and actual number of guest arrivals in the month of September in the districts of Munich and Ebersberg from 2006 to 2023.

Figure 4. Percentage change between the expected and observed number of tourists in September for the individual counties of Upper Bavaria from 2006 to 2022. The areas in darker purple are those where the number of visitors has been higher than expected. If the GIF is not displayed correctly, please visit [GitHub](#).

3.4 Price and consumption trends

We see an almost linear increase in beer price over the years (Figure 5). The per capita consumption also shows an overall increase, but fluctuates over the years. We observe a low per capita beer consumption in 2023, likely due to the record number of visitors that year. The correlation

plot shows that beer consumption continues to increase, despite rising beer prices. This strong positive relationship between beer price and consumption is confirmed by the PCC of 0.8452. Adjusting beer prices by the CPI (and excluding 2023 outlier) slightly increases the correlation between beer price and consumption from 0.9473 to 0.9548. For roast chicken, a PCC coefficient of -0.8679 shows a strong negative correlation between price and consumption.

3.5 Beer & chicken price prediction

The linear regression model utilized for beer price prediction exhibited impressive performance, forecasting a value of 14.29€ in 2023, which closely mirrors the observed price of 14.33€. The Gaussian Process Regression estimated the price to be 13.13€, accompanied by an uncertainty range of approximately ± 2.52 . Shifting the focus to chicken prices, the linear regression model projected a value of 13.82€ in 2023, deviating from the observed value of 16.49€. On the other hand, the Gaussian Process Regression provided a mean projected price of 13.08€, accompanied by an uncertainty range of approximately ± 1.96 .

4. Discussion & Conclusion

The sudden drop in visitors in 2001 could potentially be explained by the 9/11 terrorist attacks that occurred a few weeks prior and almost resulted in the cancellation of the festival ([Portal München, 2001](#)). Perhaps surprisingly, visitor numbers were still very low during the first festival after the pandemic in 2022. This could be due to the public's remaining fears of contracting the virus in large crowds.

The weak correlation between mean precipitation at daytime and mean daily visitors indicates that precipitation does not have a great influence on festival attendance. This outcome is contrary to our expectations and can be attributed to various factors. One reason is the lack of daily visitor numbers, forcing us to rely on the mean over all days. However, examining the correlation of these means overlooks the daily distributions, potentially leading to a weaker correlation. Nevertheless, this method still demonstrates that the total precipitation has only a minor effect on visitor numbers, with other factors, such as terrorist attacks, likely having a greater impact on festival attendance.

As daily visitor numbers are not provided, we rely on monthly data to analyze Upper Bavaria's visitors. We assume that the majority of Oktoberfest visitors are counted in September's figures, as the festival begins in mid-September and most visitors likely arrive during this month. The purple shading surrounding Munich suggests that the Oktoberfest influences visitor numbers beyond the city. Smaller districts near Munich experience a higher

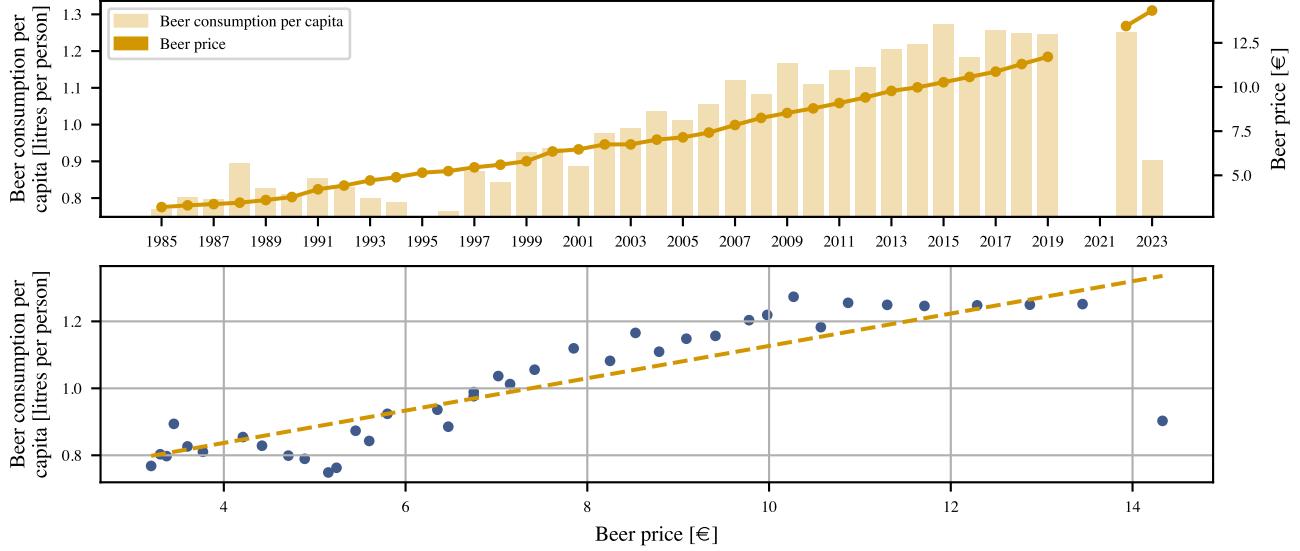


Figure 5. Top: Beer price and beer consumption per capita between 1985 and 2023. *Bottom:* Correlation between beer price and beer consumption per capita.

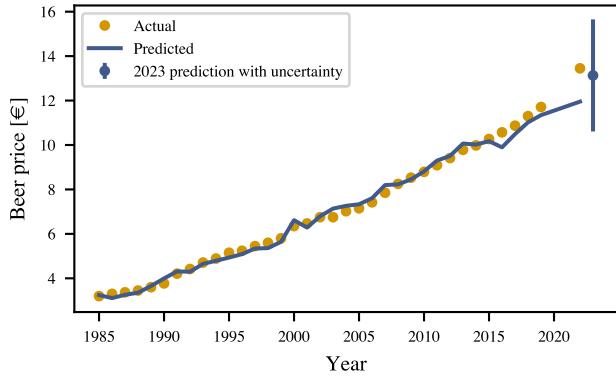


Figure 6. Beer Price Prediction from 1985 till 2023 using Gaussian process regression with an uncertainty level of one standard deviation for 2023.

percentage increase in visitors, indicating the festival's broader impact. Interestingly, one district (Traunstein) consistently experiences fewer tourists than expected. There may be an event in Traunstein in August or October that drives up visitor numbers in that month.

Rising beer prices do not deter beer consumption, as there is a positive correlation between beer price and per capita consumption. The Oktoberfest has grown in popularity since 1985 and now also attracts many international tourists. This could explain why people are willing to pay more for a "Maß" beer, particularly international tourists accustomed to similar prices in their home countries. Adjusting beer prices for inflation using the CPI further confirms the correlation's independence from inflation rates, as it isolates changes in real beer prices from general price increases in the economy.

The 2023 Oktoberfest was the first to introduce free water fountains ([Portal München, 2023b](#)), which could have contributed to the decrease in overall beer consumption we see that year. Interestingly, while beer consumption shows a positive correlation with price, chicken consumption trends in the opposite direction. However, the decrease in chicken consumption might have less to do with the increase in price, but rather with the rise in popularity of other foods. The notable accuracy of the linear regression model, in contrast to the Gaussian process regression in predicting beer prices, may be attributed to the predominantly linear relationship of the beer prices. Consequently, for simplicity and computational efficiency, opting for the linear regression model would be a pragmatic choice for predicting beer prices. On the other hand, both models perform poorly when predicting chicken prices. This could be attributed to the unusual increase in chicken prices around the year 2000, posing challenges for the models in capturing and adapting to such irregularities in the data.

Contribution Statement

Jana Hoffmann analyzed the influence of precipitation on visitor numbers, Julia Graf investigated the influx of tourists to the surrounding area, Jessie Midgley looked at the correlation between consumption and prices, and Tithi Rakshit employed models to predict future prices. All authors contributed to the writing of the report.

Data Availability

The source code for this paper is available on [GitHub](#).

References

- Bayerisches Landesamt für Statistik. Tabelle 45511-002: Tourismus: Gemeinden, Übernachtungen, Ankünfte, Herkunft der Gäste, Monat, Jahr (ab 2006). <https://www.statistikdaten.bayern.de/genesis//online?operation=table&code=45511-002&bypass=true&levelindex=1&levelid=1705932882341#abreadcrumb>. Accessed: 2024-12-6.
- DWD. Aktuelle tägliche Stationsbeobachtungen (Temperatur, Druck, Niederschlag, Sonnenschein-dauer, etc.) für Deutschland, Qualitätskontrolle noch nicht vollständig durchlaufen, Version v23.3. https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/recent/tageswerte_KL_03379_akt.zip, a. Accessed: 2024-01-15.
- DWD. Historische tägliche Stationsbeobachtungen (Temperatur, Druck, Niederschlag, Sonnenschein-dauer, etc.) für Deutschland, Version v23.3. https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/historical/tageswerte_KL_03379_19540601_20221231_hist.zip, b. Accessed: 2024-01-08.
- DWD. Stündliche Stationsmessungen des Niederschlags für Deutschland, Version v23.3. https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/precipitation/recent/stundenwerte_RR_03379_akt.zip, c. Accessed: 2024-01-16.
- DWD. Stündliche Stationsmessungen des Niederschlags für Deutschland, Version v23.3. https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/precipitation/historical/stundenwerte_RR_03379_19970707_20221231_hist.zip, d. Accessed: 2023-12-13.
- Landeshauptstadt München, Referat für Arbeit und Wirtschaft. Oktoberfest 2023. <https://stadt.muenchen.de/dam/jcr:2117b3fe-760c-4a9d-8123-38fe54bf49e4/ge-trraenke-wiesn23.pdf>. Accessed: 2023-11-19.
- Nield, T. *Essential Math for Data Science*. O'Reilly Media, Inc., 2022.
- opendata.muenchen.de. Oktoberfest. <https://opendata.muenchen.de/dataset/8d6c8251-7956-4f92-8c96-f79106aab828>, 2023. Accessed: 2023-11-18.
- Portal München. Weniger Bier wurde letztmals in 1983 getrunken. <https://stadt.muenchen.de/dam/jcr:dd04156b-af36-4f12-bed7-7f172908b1b5/mbo10601.pdf>, 2001. Accessed: 2023-01-11.
- Portal München. Eine Wiesn wie aus dem Bilderbuch. <https://www.muenchen.de/veranstaltungen/oktoberfest/aktuell/wiesn-2023-das-ist-die-abschlussbilanz>, 2023a. Accessed: 2023-12-10.
- Portal München. Kostenloses Trinkwasser auf dem Oktoberfest. <https://www.muenchen.de/veranstaltungen/oktoberfest/aktuell/kostenloses-trinkwasser-auf-der-wiesn#:~:text=Auf%20der%20Wiesn%20gab%20es,2024%20auch%20wieder%20Trinkwasserbrunnen%20geben.>, 2023b. Accessed: 2023-01-11.
- Statistisches Bundesamt. Verbraucherpreisindex: Deutschland, Jahre. <https://www-genesis.destatis.de/datenbank/beta/url/c886863d>. Accessed: 2023-12-15.