Customer Satisfaction of the German Energy Supply Chain

Predicting star ratings and company responses on Trustpilot

Rendering 1

Data Science Graduation Project @DataScientest, Paris
Oct 26, 2023

Matthias Isele - Stefanie Arlt

Outline

Abbreviations

□ i.e. = that is

□ e.g. = for example

□ etc. = and others

□ max. = maximum

□ min. = minimum

□ IQR = inter quartile range

□ ML = machine learning

□ PCA = principal component analysis

□ MAE = Mean Absolute Error

□ MSE = Mean Squared Error

□ RMSE = Root Mean Squared Error

□ R^2 = R-squared

Summary of Rendering 1 and Next Steps

Our goal is to investigate **energy suppliers in Germany on the Trustpilot web site**. The source for our investigation is scraped content from September 2023 in German language.

First, we investigated the **overall ranking** for energy suppliers in Germany based on the Trustpilot scores: We have found information on currently 37 suppliers in Germany about their overall score and number of votes, as well as supported energy supplier categories, e.g. eco power, solar energy etc.

It could be established that neither number of votes nor supported categories have a direct impact on the overall ranking.

Second, we investigated the **customer reviews and company answers** for these suppliers. The content of more than 3000 pages was consolidated in a data set and investigated in terms user activity and mood, correlations between numerical variables and a first analysis of customer comments and supplier answers.

Our findings include that **user activity on Trustpilot** for German energy suppliers skyrocketed in the last four years, from 2,000 to over 25,000 reviews; The years 2018 and prior should be neglected due to comparatively scarce data. 83% of customer experiences are either bad (1 star) or excellent (5 stars). That is, users tend to write reviews only, if their experience is on the extremes. The average customer star rating grew during the pandemic years 2020 – 2022 (2.6 - 4.2 stars), getting a hit due to the energy crisis 2023 (3.6 stars). Over all years, customers tend to be content with German energy suppliers, as the average star rating never dipped below 2.5.

The **number of words of a customer comment** is a good indicator for the star rating, with a Pearson coefficient of 0.5. The more elaborate the comment, the worse the rating. Companies tend to write longer answers if a comment is more elaborate. The year of experience is uncorrelated to star rating or comment or answer length. The response time of a company to a review is uncorrelated to any of the variables. In particular, companies answer good and bad reviews in the same time frame and an answer does not impact the given star rating.

Grouped by star ratings, the number of words of customer comments is lognormal distributed. The distributions increase in variance and mean for lower star ratings. Overall star ratings, there is a significant overlap. The number of words of company answers are pointwise discontinuous, i.e. there are company specific standard answers for certain star ratings; A segmentation in terms of companies is necessary.

Analyzing the comments, we have **identified for further investigation**:

- Key words indicating customer satisfaction and company responses.
- Customer satisfaction areas except pricing, e.g., billing, complaint management.

For the next step in Machine Learning, we see the following options:

- Rating prediction based on comment and answer word counts, also segmented by companies.
- Predicting the length of company answers based on star rating and comment word count.

Understanding and manipulation of data

Dataset

Target of this analysis are ratings, customer votes and supplier feedback of energy suppliers in Germany on the Trustpilot web site[1]. The source for our investigation is scraped content from September 2023 in German, as the language settings on the portal define the respective market. We tried out English language as well but then we only got UK / American energy suppliers, or comments in English language only from English speaking expats currently living in Germany, which is only a small percentage of the local customer base.

We scraped content on two levels: File 1 contains general information of energy suppliers in Germany (supplier level), while File 2 collects for each energy supplier customer reviews and supplier answers (customer level). The total data may be obtained by joining File 1 and File 2 on the supplier column.

File 1 - Energy suppliers in Germany

Web scraping

At the time of investigation, 37 distributors in Germany were listed on two pages on the Trustpilot web site for the category of 'power supply company' with 20 and 17 entries per page respectively.[2]

The standard sorting of the suppliers is "according to relevance". This filter showed all energy suppliers sorted by highest score and numbers of votes. In addition, the companies had to fulfill the following conditions, to ensure that the companies with the highest votes receive up-to-date customer feedback:

□ The supplier needs to have received at least 25 evaluations in the last 12 months.

□ The supplier must have the status "asking for evaluation".

We extracted the name, score, number of votes and business location of each energy distributor by using Beautiful Soup and URL Lib in addition to the classical pandas and NumPy packages.

On the page, the energy suppliers were all represented in a similar way, so we assumed logically that the structure of the respective web code was the same. The page was not set up as a table, but as a sequence of div containers, to whom specific key words in style or span tags helped identify the important information. Therefore, it was possible to iterate on the “ener_tp” variable of the first page, containing all the pertinent information on the energy suppliers.

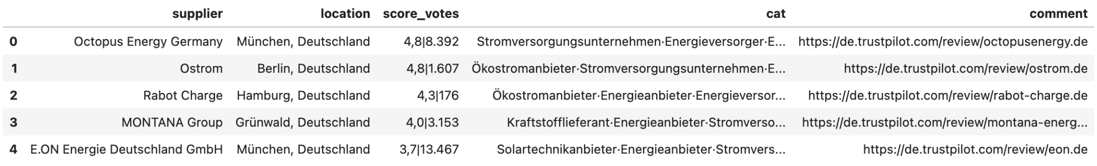
The information was then stored in lists, which could be transformed in a data frame with the zip function.

These steps could be repeated for the second page as well, stored into the “ener2_tp” variable.

Finally, both page data frames were concatenated to create one data frame, as shown in the table below.

[1] See the following link: https://de.trustpilot.com/categories/electric_utility_company?

[2] See screenshots of the Trustpilot website in the chapter “Additional” in the appendix.



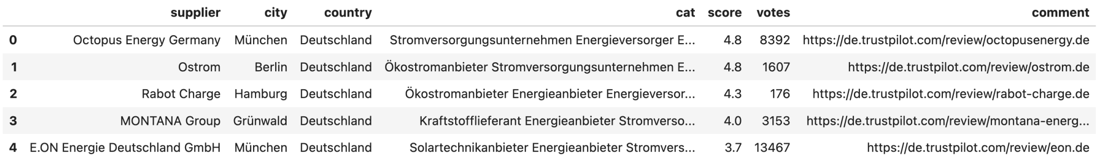
	supplier	location	score_votes	cat	comment
0	Octopus Energy Germany	München, Deutschland	4,8|8.392	Stromversorgungsunternehmen·Energieversorger·E...	https://de.trustpilot.com/review/octopusenergy.de
1	Ostrom	Berlin, Deutschland	4,8|1.607	Ökostromanbieter·Stromversorgungsunternehmen·E...	https://de.trustpilot.com/review/ostrom.de
2	Rabot Charge	Hamburg, Deutschland	4,3|176	Ökostromanbieter·Energieanbieter·Energieversor...	https://de.trustpilot.com/review/rabot-charge.de
3	MONTANA Group	Grünwald, Deutschland	4,0|3.153	Kraftstofflieferant·Energieanbieter·Stromverso...	https://de.trustpilot.com/review/montana-energ...
4	E.ON Energie Deutschland GmbH	München, Deutschland	3,7|13.467	Solartechnikanbieter·Energieanbieter·Stromvers...	https://de.trustpilot.com/review/eon.de

Figure 1: Newly scraped data set "Energy suppliers in Germany"

To make further processing easier, some basic cleaning and simple column transformations have been applied:

- □ Separation of scores and votes in 2 separate columns
- □ Change of data types to float for ‘score’ column
- □ Overall change of decimal separator to English notation and removal of German thousand separator
- □ Change of data type to integer for ‘votes column.
- □ Split of location information into two columns ‘city’ and ‘country’.

The final scraping result was stored with a new order of columns as a csv-file to be used for further investigation, as follows:



	supplier	city	country	cat	score	votes	comment
0	Octopus Energy Germany	München	Deutschland	Stromversorgungsunternehmen Energieversorger E...	4.8	8392	https://de.trustpilot.com/review/octopusenergy.de
1	Ostrom	Berlin	Deutschland	Ökostromanbieter Stromversorgungsunternehmen E...	4.8	1607	https://de.trustpilot.com/review/ostrom.de
2	Rabot Charge	Hamburg	Deutschland	Ökostromanbieter Energieanbieter Energieversor...	4.3	176	https://de.trustpilot.com/review/rabot-charge.de
3	MONTANA Group	Grünwald	Deutschland	Kraftstofflieferant Energieanbieter Stromverso...	4.0	3153	https://de.trustpilot.com/review/montana-energ...
4	E.ON Energie Deutschland GmbH	München	Deutschland	Solartechnikanbieter Energieanbieter Stromvers...	3.7	13467	https://de.trustpilot.com/review/eon.de

Figure 2: Final scraped data set “Energy suppliers in Germany”

Target and explanatory variables

Although data types were looking okay, it was clear that some values were missing or needed to be checked for clarification.

- □ If there were 0 votes, the data row was to be deleted after relevance clarification.
- □ For missing city information valid company information was researched and replaced accordingly.

After a reset of index and uneventful check for special chars, the data was deemed clean.

Features and Limitations

The target of this exploration is the rating and comments power suppliers on the German market where some companies offer energy with different company names, all synonyms in German language.

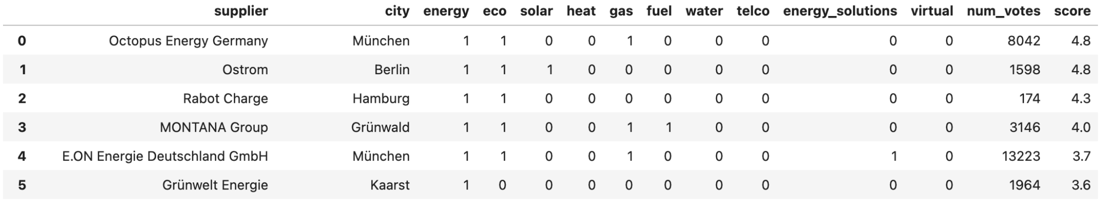
The Trustpilot search was for “energy supplier”, i.e. “Stromversorgungsunternehmen” in German, but with the result it became clear that some companies also render additional services and are listed in several categories.

This information was consolidated in the ‘cat’ column: To access its content, all information was exported to text for a comprehensive list of unique categories, which were then added as separate columns into the data frame according to the following strategy:

Finding	Action	Key words in category and translations
Synonyms for energy suppliers	To be consolidate into one common category	"Energieversorger", "Energieanbieter", "Energieversorgungsunternehmen", "Stromversorgungsunternehmen", "Stadtwerke“
Specialized power labels	Filter in separate columns	□ 'eco energy' for 'Ökostromanbieter' □ 'solar energy' for 'Solarenergieunternehmen' □ 'heat flow' for 'Wärmeenergie-Unternehmen'
Diversification labels	Filter in separate columns	□ 'gas supplier' for 'Gasversorgungsunternehmen' □ 'fuel supplier' for 'Mineralölunternehmen', 'Kraftstofflieferant', □ 'water supplier' for 'Wasserversorgungsunternehmen' □ 'telecommunications provider' for 'Telekommunikationsanbieter', 'Internetanbieter', 'Telefon- und Internetdienst' □ 'energy solutions' for 'Energieanlagen und -lösungen', 'Solartechnikanbieter', 'Heizungsanlagenanbieter', 'Anbieter von Elektronikbauteilen', 'Technischer Kundendienst', 'Elektronikunternehmen' □ virtual' for 'Reiseanbieter', 'Online-Marktplatz'

Table 1: Categories for power labels and diversification

The data set was now ready for exploration and visualization:



	supplier	city	energy	eco	solar	heat	gas	fuel	water	telco	energy_solutions	virtual	num_votes	score
0	Octopus Energy Germany	München	1	1	0	0	1	0	0	0	0	0	8042	4.8
1	Ostrom	Berlin	1	1	1	0	0	0	0	0	0	0	1598	4.8
2	Rabot Charge	Hamburg	1	1	0	0	0	0	0	0	0	0	174	4.3
3	MONTANA Group	Grünwald	1	1	0	0	1	1	0	0	0	0	3146	4.0
4	E.ON Energie Deutschland GmbH	München	1	1	0	0	1	0	0	0	1	0	13223	3.7
5	Grünwelt Energie	Kaarst	1	0	0	0	0	0	0	0	0	0	1964	3.6

Figure 3: Data Set “Energy suppliers in Germany” after Clean-up

File 2 – Customer reviews and supplier answers

Web scraping

The reader may have a look at Figure 20 and Figure 21 in the appendix to see a typical customer feedback and an optional supplier response on Trustpilot. The number of reviews for each energy supplier varies widely, from less than 10 to over 10.000. For each review, there is either none or one supplier response. For a fixed energy supplier, the reviews are distributed over several pages such that each page contains exactly 20 reviews except the last page, where review numbers from 1 to 20 are possible. In the scraping phase we tried to get as much data as possible from customer reviews and supplier answers, i.e. we wanted to gather for each German energy supplier all reviews and answers there were (september 2023). This could be accomplished using Beautiful Soup and the following methodology.

From the supplier column in File 1 we generated for each supplier their respective start page URL in Trustpilot. The list of start page URL’s is used to iterate over in a for-loop. On the respective start pages one can read out the total number of review pages (per supplier). A second sub-for-loop now iterates over all pages and reads out all review information and answers and stores them in respective lists. These lists are concatenated to create the data set File 2. In practice, Trustpilot noted that we are scrapers and denied access to the page each time after scraping around 300 pages. We had to take measures to disguise ourselves as casual users, including random time delays during iterations and rotating User-Agents. This improved the situation to scrape up to 700 pages without getting blocked. Still, this is not enough for big suppliers like Octopus Energy with over 1000 pages, so we had to generalize the algorithm to allow for selection of iteration intervals manually. The obtained sections could be concatenated to File 2 containing all possible customer reviews and supplier answers of German energy suppliers there were September 2023. The anti-blocking measures could have been improved by including e.g. rotating Proxies, but the pragmatic manual section-wise approach got us to the goal faster.

Target and explanatory variables

The following features were directly scraped from customer reviews and supplier answers. See Figure 20 and Figure 21 in the appendix for screenshots of typical reviews and answers on Trustpilot.

Description of variable	Label	Appearance	Dtype
Customer nickname	“Nickname”	*optional*	*object*
Location of customer	”Location”	*mandatory*	*object*
Star rating of customer	“Stars”	*mandatory*	*Int64*
Headline of post	“Headline”	*mandatory*	*object*
Date of post	“DoP”	*mandatory*	*datetime64[ns, UTC]*
Date of experience	“DoE”	*mandatory*	*datetime64[n]*
Comment of customer	“Comment”	*optional*	*object*
Answer of energy supplier	“Answer”	*optional*	*object*
Date of answer	“DoA”	*if, and only if answer exists*	*datetime64[ns, UTC]*
Company name, i.e. energy supplier)	“Company”	*mandatory*	*object*

Table 2: Feature description "Customer feedback and supplier answers"

After scraping we computed and included the following variables.

Description(s) of variable(s)	Label(s)	Dtype
Splits of DoP, DoE, DoA into day, month, year	“DoP.day”, “DoP.month”, …	*int64*
Is there a comment? (value: 1 or 0)	“Comment_TF”	*int64*
Is there an answer? (value: 1 or 0)	“Answer_TF”	*int64*
Number of words of headline	“Words_Headline”	*int64*
Number of words of comment	“Words_Comment”	*int64*
Number of words of answer	“Words_Answer”	*int64*
Response time of energy supplier to review in days (DoA-DoP)	“Response_time”	*float64*

Table 3: Newly engineered features for data set 2

The two target variables are the star ratings ‘Stars’ and the supplier answers ‘Answer’, which will be predicted by comments and related variables. In a first approach, instead of predicting answers, we will predict answer length “Words_Answer”.

Features and Limitations

Some variables of the data set may be negligible, e.g. location, which carries the value “DE” in 98% of cases or user nicknames, which are hard to classify, e.g. for gender. Other explanatory variables, e.g. key words of comments to predict star ratings, are still left to be derived.

Visualizations and data exploration

File 1 - Energy suppliers in Germany

Relations between explanatory variables and target

To investigate the relationship between the variables of the data set, both Spearman and Pearson tests have been conducted and visualized by a heat map: Red colors indicate a high positive correlation between the variables, whereas blue colors show a poor and even negative correlation, see Figure 4.

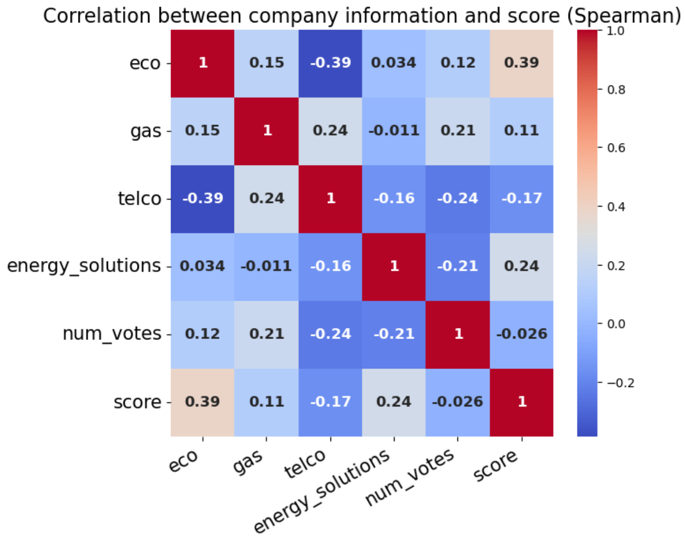


Figure 4: Heat map (Spearman r) for data set 1

It has been observed that eco-friendly energy seems positively correlated to the rating score, also solar energy, and the offering of energy solutions. Heat flow shows highest negative correlation to the score. Diversification in gas and fuel supply is only mildly represented as positive. Number of Votes and telco are positively correlated as well.

So, we can postulate that that the supported categories have no direct impact on the overall ranking.

Checking with a Spearman-r test, it could also be concluded by a p-value of 0.887 that the number of votes and the score are not correlated as well.

Indeed, if we look at the distribution of scores and the number of votes vs. the rating score, we can see that for every rating there are suppliers with low and high number of votes. Even if there are some extreme values for the middle scores, most supplier ratings are based on less than 500 votes (Figure 5).

Although we see limited data points, we can observe a tendency of higher number of votes for low ranking and higher ranking. An explanation might be that customers tend to express more feedback when they are exceptionally happy or exceptionally unhappy.

To consider is also the history as some suppliers, like e.g. Vattenfall, have been in business very long and their average score is changing over time.

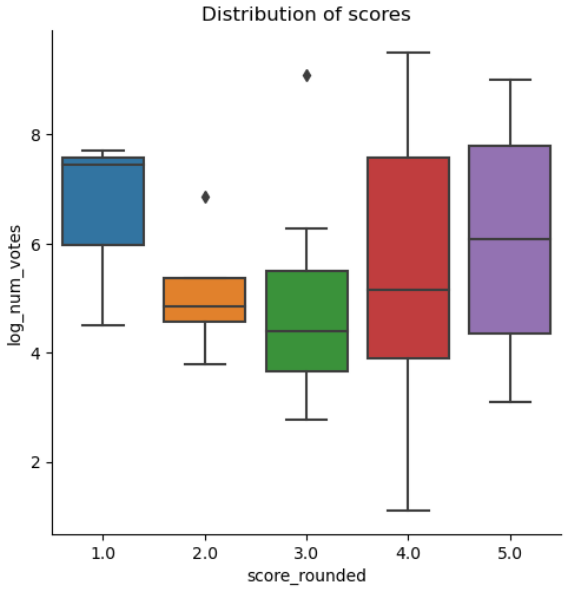


Figure 5: Distribution of number of votes per score for German Energy suppliers on TrustPilot

Overall ratings presented as the following energy provider as the top and bottom five in the Trustpilot ranking:

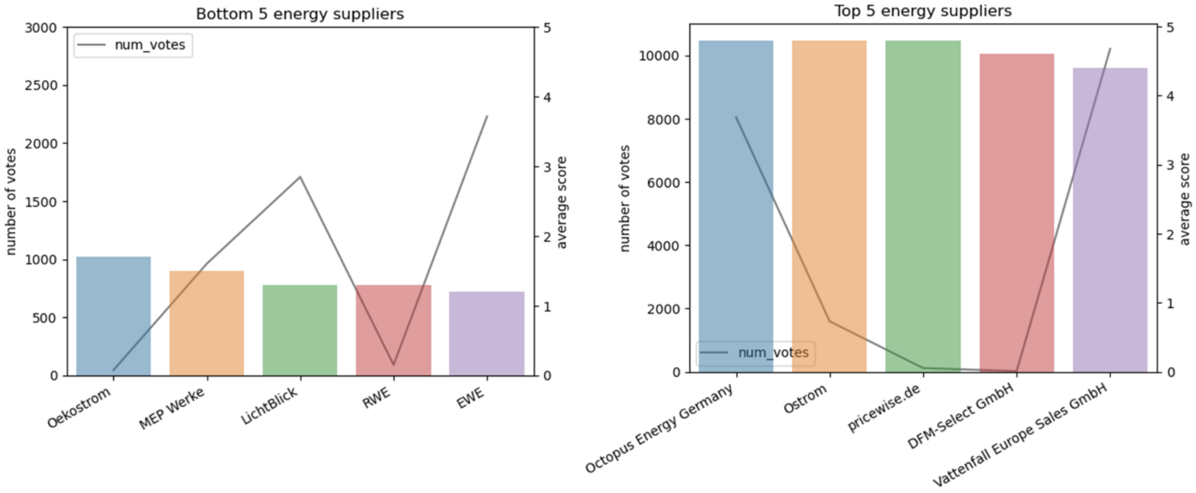


Figure 6: Bottom 5 and Top 5 energy suppliers in Germany

As we can see, the number of votes is spread all over the spectrum: Small suppliers with only a few comments, like Ostrom or pricewise, are neck-and-neck as top performers with big international corporations like Octopus Energy or Vattenfall with many votes. While new eco-oriented suppliers like LichtBlick have nearly as many votes as traditional suppliers such as EWE with the same low average score, we can also see the opposite: new supplier Oekostrom and traditional RWE with only a few votes and equally low ranking.

Relationships between variables

Having a look at the diversified offering mentioned before[3] we can see the following situation:

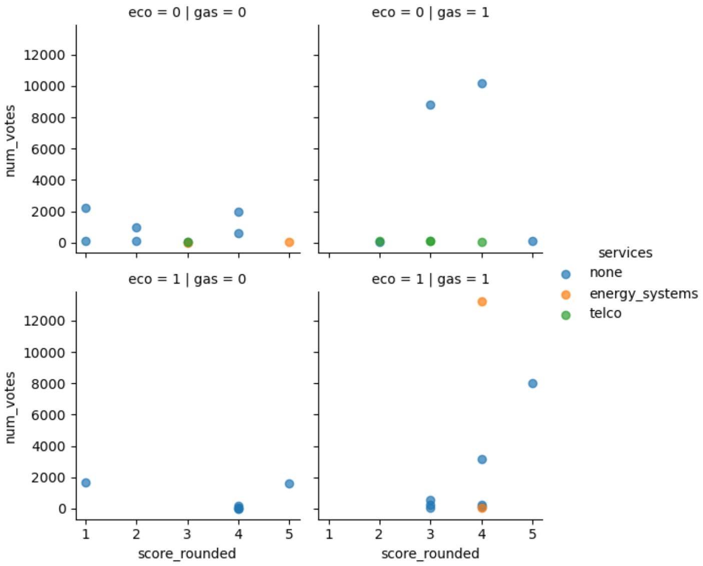


Figure 7: Impact of diversification on energy suppliers in Germany

The offering of eco power is no guarantee for many (positive) votes or a high rating. However, suppliers who are also delivering gas, tend to have more votes but are distribute over scores 2 to 5.

[3] See classification in chapter “Features and Limitations” for file 1.

In the overall ranking, telecommunication has no impact, but suppliers also offering energy systems like photovoltaic technologies tend to have higher ranking.

Could this be an indication, that functioning business processes and customer orientation are the most important factors for a good score? We should analyze the comments for feedback.

File 2 – Customer reviews and supplier answers

Overview of user activity and mood

The number of comments on German energy suppliers in Trustpilot saw a strong increase in the recent years, see Figure 8. The threshold of more than 1500 comments per year was first reached 2019, reaching 25,000 comments by 2023. The website saw a strong boost in the pandemic years 2020-2022, as more people were forced to go digital. One can expect that the war in Ukraine increased the number of comments in 2023 additionally, as a lot of people were affected by the resulting energy crisis. The true number of comments at the end of 2023 is expected to be higher, as the figure on the right is to date September 2023. The users are mostly from Germany (98.1%), followed by Austria (0.3%), Netherlands (0.2%), US (0.2%) and Spain (0.1%). In total, users state to be from 78 different countries. One can expect that a portion of country labels stem from accidental miss clicks or are wrong by choice to protect customer privacy.

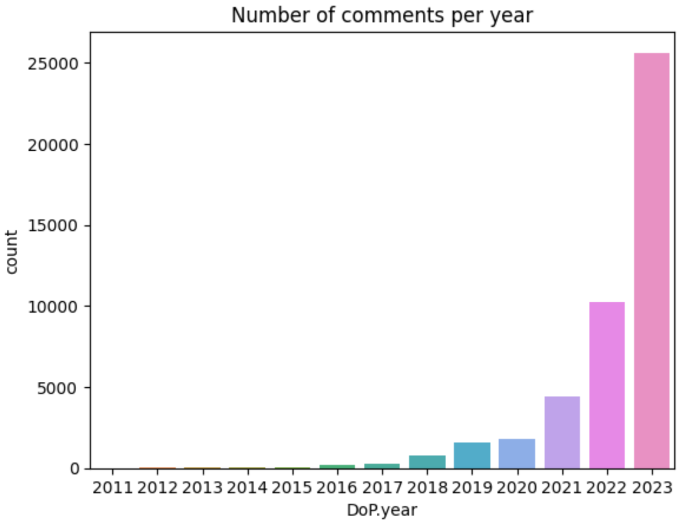


Figure 8: Number of comments on German energy suppliers to date September 2023

Figure 9 shows counts of the five possible star ratings on German energy suppliers. Customers tend to review only if their experience is on the extremes, i.e. bad (1 star) or great or excellent (4 or 5 stars).[4] In particular, 69% of reviews are great (11%) or excellent (58%), 25% of reviews are bad, and 6% of reviews are poor or average.

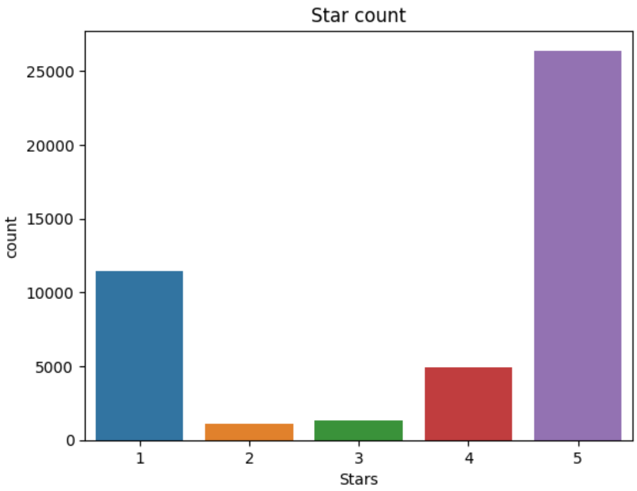


Figure 9: Count of star ratings

[4] For an explanation oft he star rating see: https://support.trustpilot.com/hc/en-us/articles/201748946-TrustScore-and-star-rating-explained#:~:text=A%20TrustScore%20is%20the%20overall,how%20they're%20calculated%20here.

The average number of stars for German energy suppliers varies drastically over the years from 2011 to 2023, as shown ine Figure 10. Starting from the global maximum of 4.8 stars in 2013 we see a steep downward trend to the global minimum of 2.6 stars in 2018. From here a steady upward trend to 2 stars in 2022, followed by a dip to 3.7 stars onto 2023. The dip from 2022 to 2023 is in line with the energy crisis. The pandemic years 2020-2022 seem to not bother customers negatively. In contrary, a lot of customers seem to have found time to write positive reviews. The liberalization of the German energy market 2019[5] led to a reduction of monopoles which is in favor of customers. A comparison with the number of comments per year, i.e. Figure 8, shows that the average number of stars before 2019 is rather insignificant, due to low number of reviews. Special effects should be considered to explain the trend from 2013 to 2018. When the website was new and unknown, workers of registered companies could dominate the reviews leading to high ratings. As more external customers entered the platform, more realistic reviews were established. Another effect is high volatility, i.e. early companies could potentially easily dominate the data, as so few data was amenable.

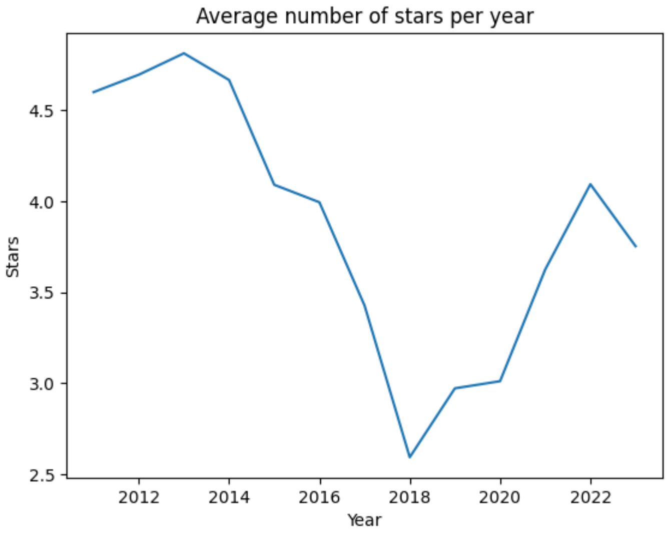


Figure 10: Average star rating per year

Relationship between variables

Pearson correlations of selected numerical variables are visualized in Figure 11. The year of experience is almost completely uncorrelated to any of the variables (|r_y| < 0.11). The number of words of a comment and the number of words of the answer are good indicators for the number of stars (r_cs=-0.49, r_as=-0.55). Lesser words mean more stars. The number of words of the headline is slightly correlated to the number of stars (r_hs=-0.17). The number of words of comments and of answers is correlated by r_ca=0.31. If customers write more text then companies tend to answer more elaborately. The response time of the company is almost completely uncorrelated to any of the variables

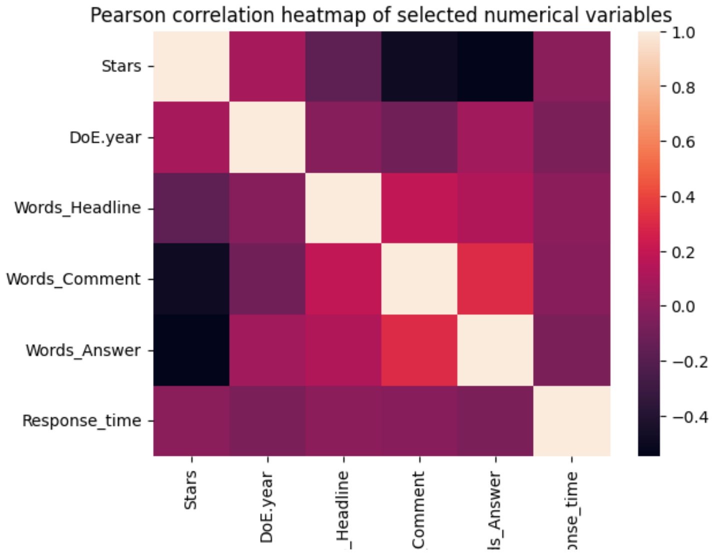


Figure 11: Correlation heatmap

[5] See: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://static.agora-energiewende.de/fileadmin/Projekte/2019/Liberalisation_Power_Market/Liberalisation_Electricity_Markets_Germany_V1-0.pdf

(|r_rt| < 0.06). This means that a company will answer good and bad reviews in around the same time frame. The year of experience is almost completely uncorrelated to any of the variables (|r_y| < 0.11).

Analyzing comments and answers

In average, the number of words of a comment is a very strong indicator for the number of stars (r_cs=-0.98) as well as the number of words of an answer (r_as=-0.90), see Figure 15. However, the distributions of comment-word count per star rating overlap. The variance of word count per comment is higher for lower ratings, the most extreme outliers being for a star rating of 1. The median and quantiles tend to shrink for higher star ratings (Figure 14). The Kernel Density Estimations (KDE's) have higher variance and skew to the right for lower star ratings (Figure 12). This is a strong hint that word counts of comments are lognormal distributed. Indeed, after taking the natural logarithm, centering and normalizing, the data follows a standard normal distribution, as seen in Figure 13. Regarding word counts of answers the situation is not that straightforward. Two peaks for each rating in Figure 16 indicate that companies tend to have standard answers for each rating and that the statistic is dominated by the two companies with the most comments. Indeed, separate answer counts for the

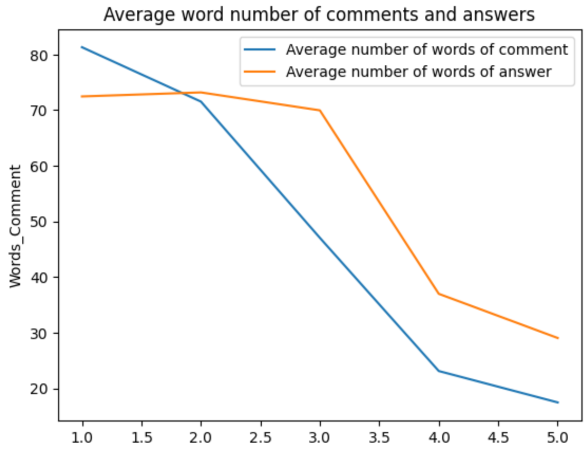


Figure 15: Average word count of comments and answers

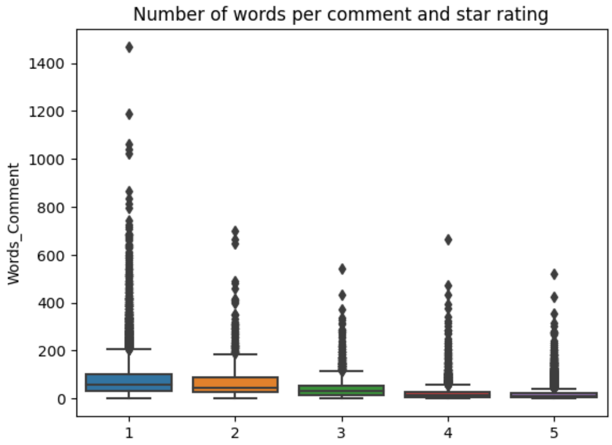


Figure 14: Boxplots: Word counts of comments per star rating

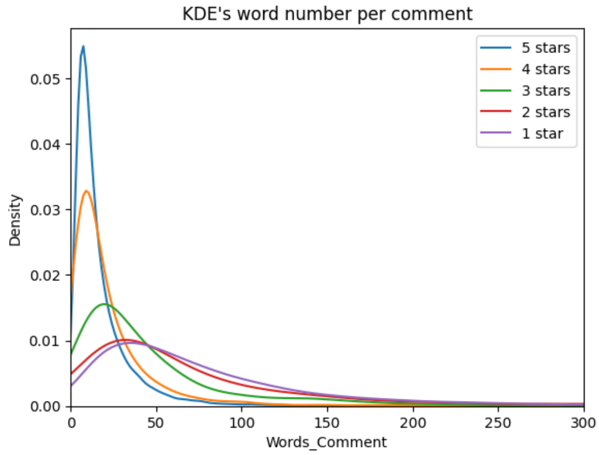


Figure 12: KDE: Word count of comments per star rating

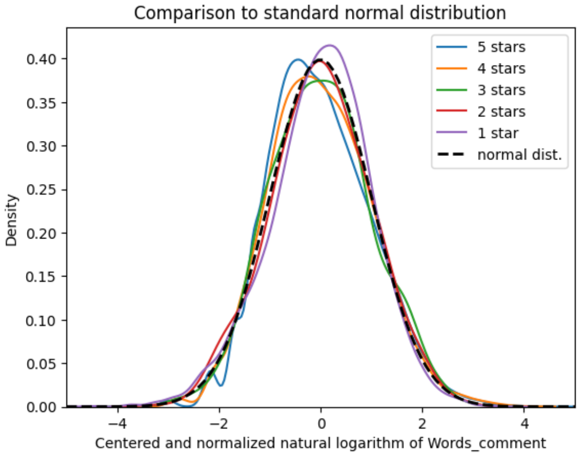


Figure 13: Word count of comments is lognormal distributed

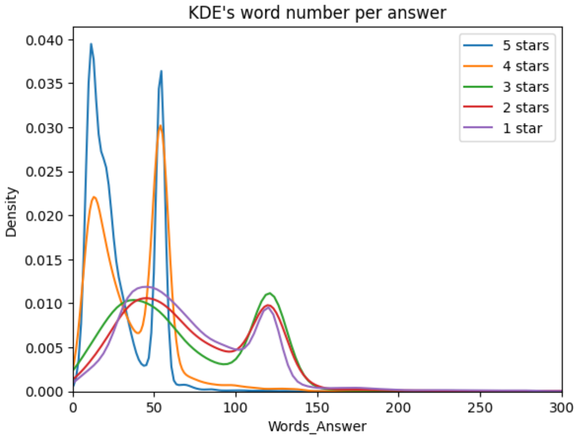


Figure 16: Word count of company answers per star rating

two biggest players “E.ON Energy” and “Octopus Energy”, Figure 18 and Figure 17, support the claim. A second peak in the 5-star rating of “Octopus Energy” may be explained by a change in the policy of standard answer length.

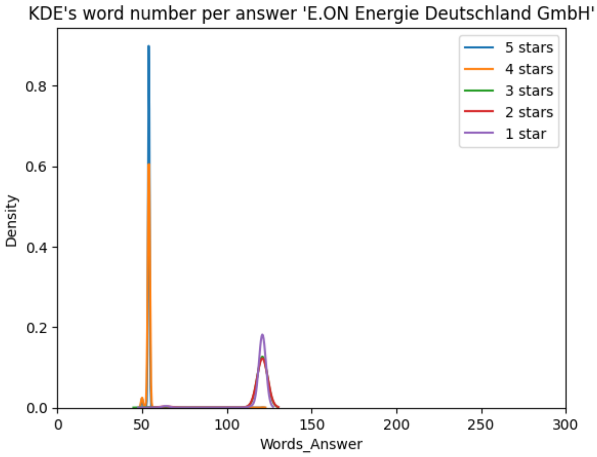


Figure 18: Word count of E.ON Energy's answers

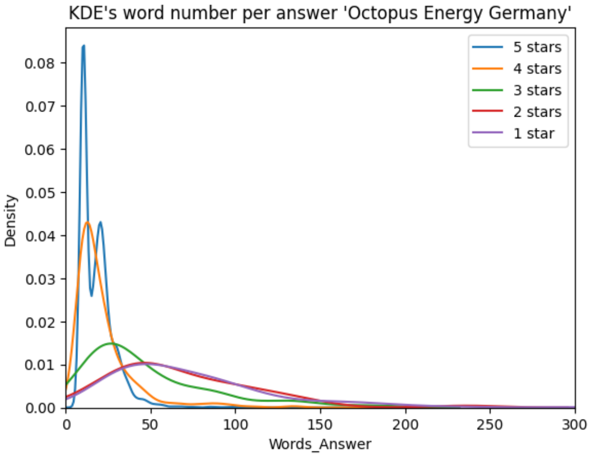


Figure 17: Word count of Octopus Energy's answers

Outlook: Data Preparation for Machine Learning

For a first step in Machine Learning, we see the following options:

- Prediction of star ratings based on comment and answer word counts, also segmented by companies.
- Predicting the length of company answers based on star rating and comment word count.

Sufficient data is provided by File 2 “customer reviews and supplier answers”, as we want to work on the customer comment level. The analysis shows that it is useful to apply the logarithm on word count variables to obtain normal distributions. The standard scaler is to be applied. Outliers in word count are significant indicators for star rating, so they should not be deleted from the data set. Not every customer review possesses a company answer. Further analysis has to show, whether we replace missing values by standard answers.

However, just restricting the data set to rows which contain answers will leave 28,533 rows left from 45,135 which is expected to be sufficient. A min-max-scaling is to be applied on the star rating column. It has to be investigated whether it makes sense to transform the star rating problem to a binary problem, as 83% of ratings are either 1 star or 5 stars.

In further steps we will refine the prediction of star ratings by identifying key words of comments and headlines. Company answers may be separated into standard answers and personalized answers to refine their predictions as well.

Appendix

Bibliography

This is an annotated list of literature and links supporting the research.

Understanding the dataset: General information energy industry

- □ Stromversorgung und Netzbetrieb in Deutschland: https://www.verbraucherzentrale.de/wissen/energie/preise-tarife-anbieterwechsel/wer-macht-was-stromanbieter-netzbetreiber-messstellenbetreiber-38444#:~:text=Das%20Wichtigste%20in%20K%C3%BCrze%3A,k%C3%B6nnen%20Sie%20diesen%20also%20nicht.
- □ Regularien Energiewirtschaft: https://www.ewerk.com/digitalhappen/energie/regulierung-gesetze-energiebranche
- □ EU Binnenmarkt Energiewirtschaft: https://www.europarl.europa.eu/factsheets/de/sheet/45/energiebinnenmarkt

Webscraping Target Site

- □ Trustpilot Website Search categories: https://support.trustpilot.com/hc/de/articles/360022026634
- □ Trustpilot transparency report: https://assets.ctfassets.net/b7g9mrbfayuu/tHyJSsKiNJxZvAuGPr6hz/5c6a42f3719debd02ca25989a7225222/Trustpilot_Transparency_Reporting__Active_Recipients__under_the_Digital_Services_Act_-__21_March_2023.pdf
- □ Trustpilot energy suppliers: https://de.trustpilot.com/categories/electric_utility_company
- □ Official overview energy market actors (Bundesnetzagentur): https://www.marktstammdatenregister.de/MaStR/Akteur/Marktakteur/IndexOeffentlich

Visualization

- □ GeoPandas User Guide: https://geopandas.org/en/stable/docs/user_guide.html
- □ Hex code colors and palettes: https://www.color-hex.com/
- □ Seaborn color palettes: https://seaborn.pydata.org/tutorial/color_palettes.html
- □ Controlling figure aesthetics: https://seaborn.pydata.org/tutorial/aesthetics.html
- □ Pyplot Tutorial in Matplotlib: https://matplotlib.org/stable/tutorials/introductory/pyplot.html
- □ Matplotlib Examples: https://matplotlib.org/stable/gallery/index.html#subplots-axes-and-figures

Data Cleaning

- □ Datacamp.com: Data Cleaning Tutorial, Feb 2022: https://www.datacamp.com/tutorial/tutorial-data-cleaning-tutorial
- □ Ngo, Huong: How to Clean Your Data in Python, Jul 30, 2022: https://towardsdatascience.com/how-to-clean-your-data-in-python-8f178638b98d

Outliers

- Grace-Martin, Karen: Outliers: to drop or not to drop, 2018: https://www.theanalysisfactor.com/outliers-to-drop-or-not-to-drop/
- Sharma, Natasha: Ways to Detect and Remove the Outliers, May 22, 2022: https://towardsdatascience.com/ways-to-detect-and-remove-the-outliers-404d16608dba
- Dey, Akash: How to handle outliers, Feb 2022: https://www.kaggle.com/code/aimack/how-to-handle-outliers/notebook
- IQR score: https://en.wikipedia.org/wiki/Interquartile_range
- Z-Score: https://en.wikipedia.org/wiki/Standard_score

Missing Values:

- Narang, Mohita: Handling Missing Values. Beginners Tutorial, Apr 4, 2022: https://www.naukri.com/learning/articles/handling-missing-values-beginners-tutorial/
- Kumar, Satyam: 7 Ways to Handle Missing Values in Machine Learning, Jul 24,2020: https://towardsdatascience.com/7-ways-to-handle-missing-values-in-machine-learning-1a6326adf79e

Statistics

- Bevans, Rebecca: Choosing the Right Statistical Test. Types & Examples, Jan 28, 2020: https://www.scribbr.com/statistics/statistical-tests/

List of figures

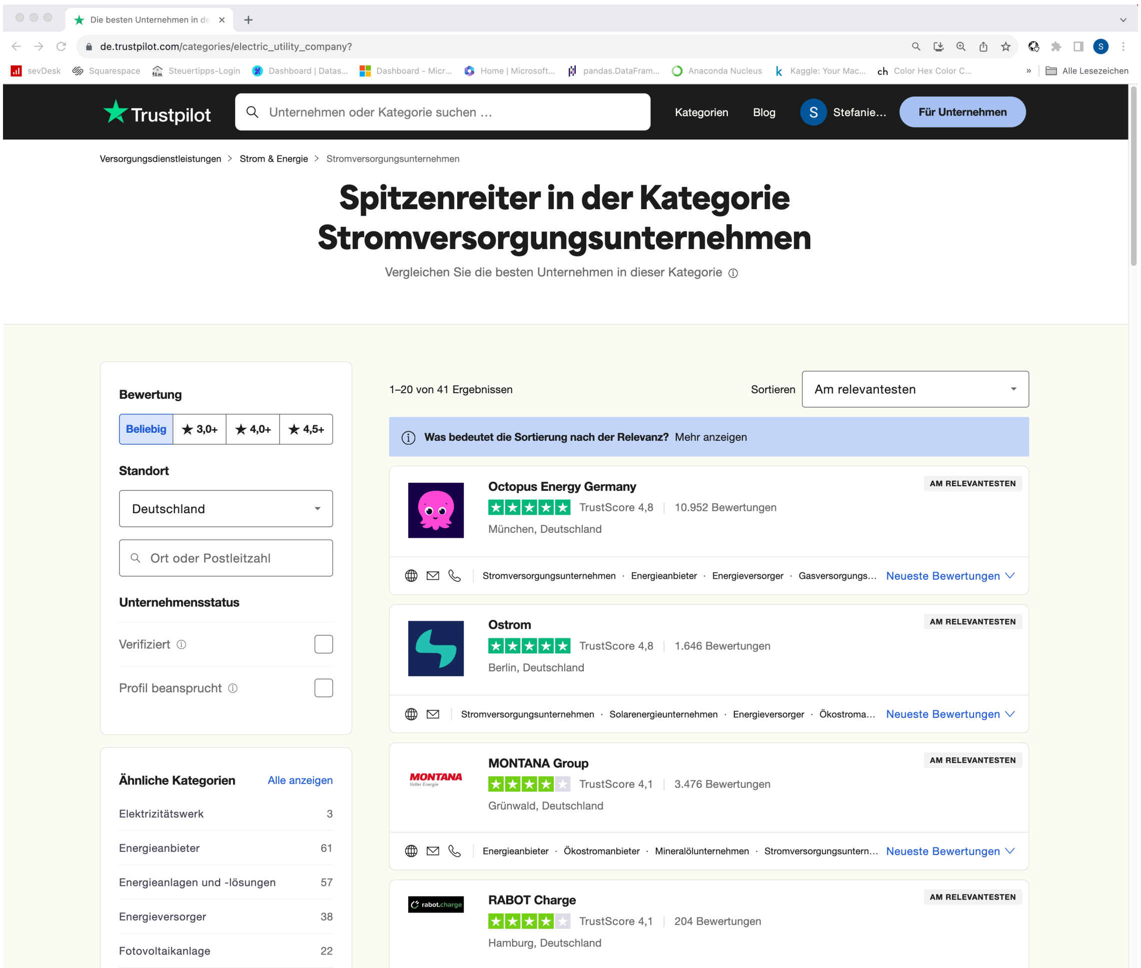
List of tables

Data Files and Contributions

All data files are available for download on GitHub.

Additional Figures

TrustPilot Website



Versorgungsdienstleistungen > Strom & Energie > Stromversorgungsunternehmen

Spitzenreiter in der Kategorie Stromversorgungsunternehmen

Vergleichen Sie die besten Unternehmen in dieser Kategorie ⓘ

Bewertung

Beliebig | ★ 3,0+ | ★ 4,0+ | ★ 4,5+

Standort

Deutschland

Ort oder Postleitzahl

Unternehmensstatus

Verifiziert ⓘ □

Profil beansprucht ⓘ □

Ähnliche Kategorien Alle anzeigen

Elektrizitätswerk	3
Energieanbieter	61
Energieanlagen und -lösungen	57
Energieversorger	38
Fotovoltaikanlage	22

1–20 von 41 Ergebnissen

Sortieren: Am relevantesten

ⓘ **Was bedeutet die Sortierung nach der Relevanz?** Mehr anzeigen

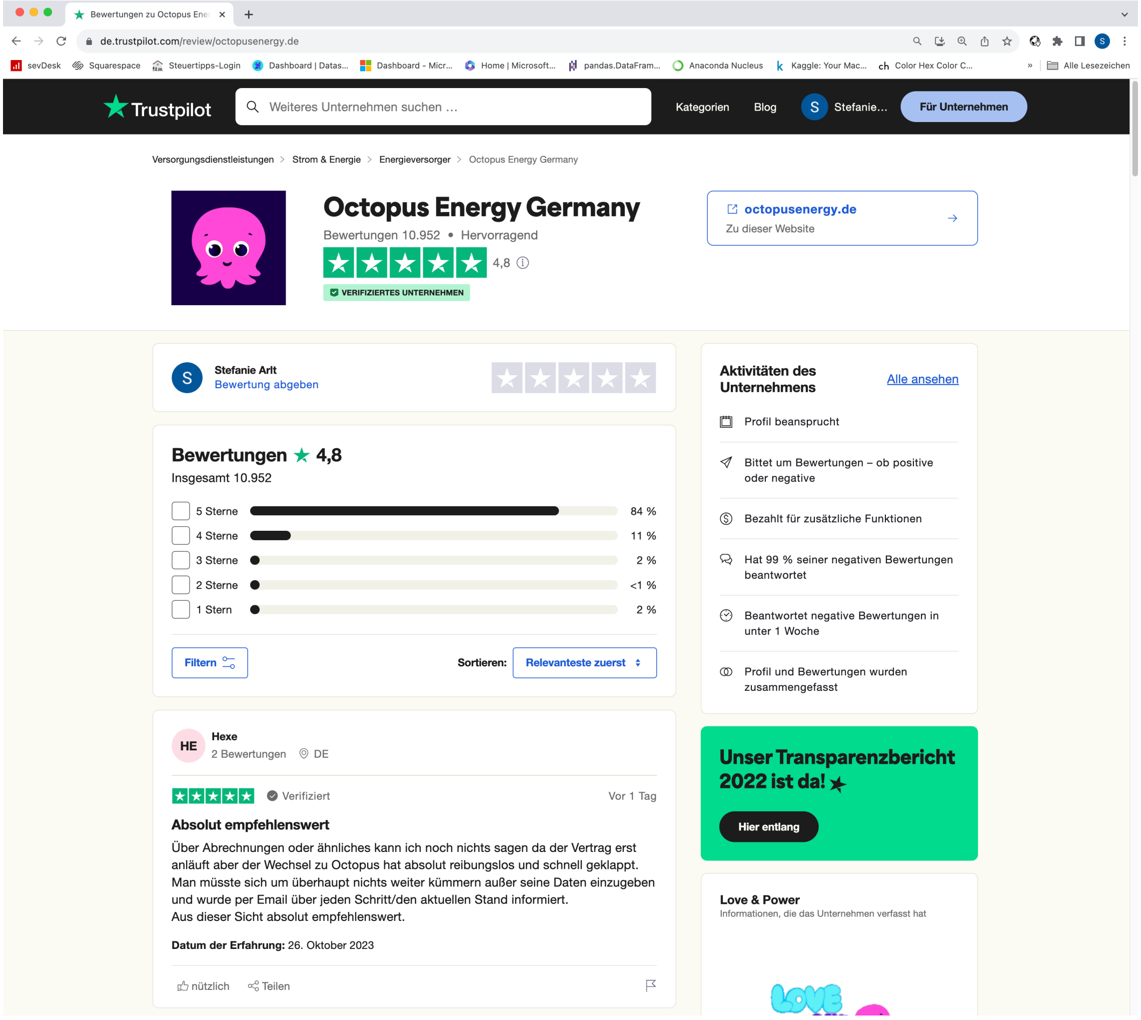
Octopus Energy Germany — AM RELEVANTESTEN
TrustScore 4,8 | 10.952 Bewertungen
München, Deutschland
Stromversorgungsunternehmen · Energieanbieter · Energieversorger · Gasversorgungs... Neueste Bewertungen

Ostrom — AM RELEVANTESTEN
TrustScore 4,8 | 1.646 Bewertungen
Berlin, Deutschland
Stromversorgungsunternehmen · Solarenergieunternehmen · Energieversorger · Ökostroma... Neueste Bewertungen

MONTANA Group — AM RELEVANTESTEN
TrustScore 4,1 | 3.476 Bewertungen
Grünwald, Deutschland
Energieanbieter · Ökostromanbieter · Mineralölunternehmen · Stromversorgungsuntern... Neueste Bewertungen

RABOT Charge — AM RELEVANTESTEN
TrustScore 4,1 | 204 Bewertungen
Hamburg, Deutschland

Figure 19: Energy suppliers on TrustPilot website



Versorgungsdienstleistungen > Strom & Energie > Energieversorger > Octopus Energy Germany

Octopus Energy Germany

Bewertungen 10.952 • Hervorragend

★★★★★ 4,8

VERIFIZIERTES UNTERNEHMEN

octopusenergy.de
Zu dieser Website

Stefanie Arlt
Bewertung abgeben

Bewertungen ★ 4,8

Insgesamt 10.952

□ 5 Sterne	84 %
□ 4 Sterne	11 %
□ 3 Sterne	2 %
□ 2 Sterne	<1 %
□ 1 Stern	2 %

Filtern

Sortieren: Relevanteste zuerst

HE Hexe
2 Bewertungen DE

★★★★★ Verifiziert Vor 1 Tag

Absolut empfehlenswert

Über Abrechnungen oder ähnliches kann ich noch nichts sagen da der Vertrag erst anläuft aber der Wechsel zu Octopus hat absolut reibungslos und schnell geklappt. Man müsste sich um überhaupt nichts weiter kümmern außer seine Daten einzugeben und wurde per Email über jeden Schritt/den aktuellen Stand informiert.
Aus dieser Sicht absolut empfehlenswert.

Datum der Erfahrung: 26. Oktober 2023

nützlich Teilen

Aktivitäten des Unternehmens Alle ansehen

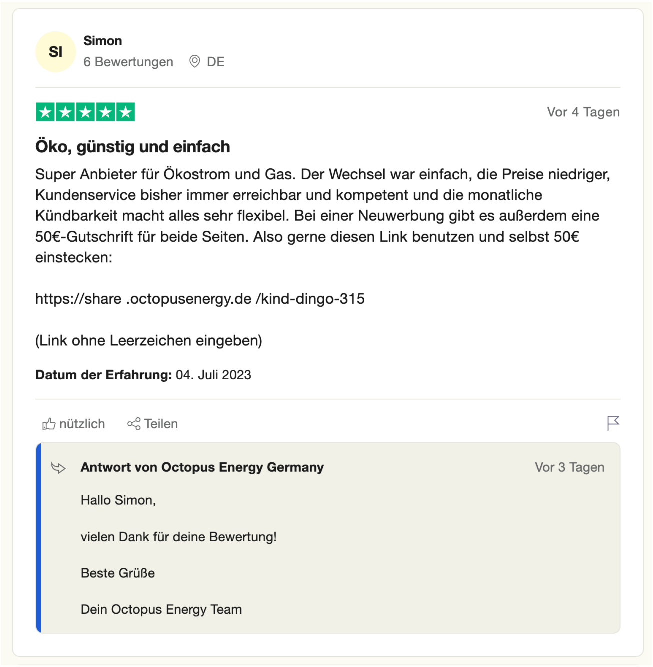
- Profil beansprucht
- Bittet um Bewertungen – ob positive oder negative
- Bezahlt für zusätzliche Funktionen
- Hat 99 % seiner negativen Bewertungen beantwortet
- Beantwortet negative Bewertungen in unter 1 Woche
- Profil und Bewertungen wurden zusammengefasst

Unser Transparenzbericht 2022 ist da!

Hier entlang

Love & Power
Informationen, die das Unternehmen verfasst hat

Figure 20: TrustPilot rating site "Octopus Energy"



SI Simon
6 Bewertungen DE

★★★★★ Vor 4 Tagen

Öko, günstig und einfach

Super Anbieter für Ökostrom und Gas. Der Wechsel war einfach, die Preise niedriger, Kundenservice bisher immer erreichbar und kompetent und die monatliche Kündbarkeit macht alles sehr flexibel. Bei einer Neuwerbung gibt es außerdem eine 50€-Gutschrift für beide Seiten. Also gerne diesen Link benutzen und selbst 50€ einstecken:

https://share .octopusenergy.de /kind-dingo-315

(Link ohne Leerzeichen eingeben)

Datum der Erfahrung: 04. Juli 2023

nützlich Teilen

Antwort von Octopus Energy Germany Vor 3 Tagen

Hallo Simon,

vielen Dank für deine Bewertung!

Beste Grüße

Dein Octopus Energy Team

Figure 21: Customer vote with supplier feedback on TrustPilot