

Final Project

Abstract

Introduction

Life in the city and countryside of the 19th century has, in modern times, often been described in both positive and negative terms. However, historians seldom focus on how contemporary people themselves saw it.

Here, contemporary literature and news articles can be of great help, and they have been for a number of historians. These historians, however, have mostly been limited to their own understanding, and analysis of the literature they selected. Very few of them have focused on Denmark, and almost none of them exclusively. This is highly relevant because different countries entered the Industrial Revolution at different times.

This paper will therefore try to answer how good or bad the contemporary authors portrayed the countryside and the city in the 19th-century, and if this changed throughout the 19th-hundreds. To answer this we will use digital methods to get a more neutral and transparent analysis of Danish contemporary literature. Our inspiration for this question comes from the book *Nineteenth-Century Europe – A Cultural History* by Hannu Salmi. In this book, he describes how people inside the city began to idolize the countryside¹, and how people with the capital to do so often took vacations there. After reading the book, we became very interested in whether that was actually true or just an interpretation by Hannu Salmi—and whether it was different in Denmark, where industrialization happened a bit later and somewhat differently, for example with more focus on industrialization in the agricultural sector².

Baggrund

The 19th century was defined by urbanization like never before. Cities were not new, but industrialization created an almost unseen demand for people in them, while at the same time reducing demand in the countryside. In Denmark, 20% of the population lived in cities in 1814, a number that had risen to 49% by 1911.

The Industrial Revolution and the rise of capitalism led to a growth in tertiary professions, which in turn led to the production of new products and services. Cities soon became home to a new middle class who sought to define themselves through these new products, creating a new culture and lifestyle that differed significantly from that of the countryside³. These new lifestyles and cultures are often discussed from a modern viewpoint. However, this often overlooks how people back then actually experienced these changes. Did people see the city as a shining place of invention, or a noisy place of overwhelming speed and impressions? And did these inventions change how people viewed the countryside in comparison? As Hannu Salmi states in the book *Nineteenth-Century Europe – A Cultural*

¹ Hannu Salmi, *Nineteenth-Century Europe* (Polity Press, 2008), 79.

² Claus Møller Jørgensen, "Fra enevældig helstat til nationalstat, 1814-1914", i *En Danmarks Historie fra vikingetid til nutid*, af Thorsten Borring Olesen og Paulsen Bjørn, 2. udg. (Aarhus universitetsforlag, u.å.), 276–79.

³ Jørgensen, "Fra enevældig helstat til nationalstat, 1814-1914", 276 and 288–89.

History: - A cultural History “In historical literature and imagination, the nineteenth century is indeed portrayed more often as black and white or grey”⁴, which of course with almost certainty can't be the same way people back then saw it. Understanding how people in general felt in the 19th-century can however be a difficult task, luckily these thoughts and emotions are often picked up by contemporary authors, both in news articles and especially in literature. From a purely historical standpoint, this is interesting because it allows us to understand what emotions people felt in a century of such innovation. These results should also enable us to understand how people back then reacted to a whole new world of invention—and thereby enable us to see if there are any modern comparisons. It can, however, be questioned to what degree fictional literature can reflect contemporary culture. It is also important to consider which literary genres are used in digital analysis. The 19th century, after all, contained a number of literary periods such as *Romantikken*, *Romantismen*, and *Det moderne gennembrud*, which each had different focuses and goals—and may have ignored certain parts of contemporary culture⁵.

We will, however, try our best to solve this issue by using a fair mix of these different genres, the selection of which will be explained later in the paper.

Related work

The work of determining culture with the help of the help of literature, is not a new or unique task, our study is in this regard inspired by Hannu Salmi who, in his work *Nineteenth-Century Europe – A Cultural History*, shows that literary analysis can lead to a new understanding of culture history, his work is however, made only using close reading. Sentimental analysis as a way to understand culture is a discipline that has moved a lot throughout recent years. Thomas Schmidt and Manuel Burghardts paper “An evaluation of Lexicon-based Sentiment Analysis Techniques for the plays of Gotthold Ephraim Lessing” is a shining example of this. The focus of the paper is both on the limitations and abilities of computer based sentiment analysis, through German plays in the 1600-1800 hundreds. Which has been especially useful for us to learn about how it was smart to use the sentimental analysis on our text⁶. “Sentiment Classification of historical Danish and Norwegian Literary Text” from Kristine Nielsen Degn et al. where they use sentimental analysis to see the different expression from female and male authors in danish and norwegian culture, is a clear example that shows how result full computer based sentiment analysis can be in determining culture. their paper just like Thomas Schmidt and Manuel Burghardts, also shows the limits and workarounds that are required to make the sentimental analysis work. The journal “Samlaren” from Sigrd Schottenius Cullhed and Jon Viklund, has a focus on the different options when it comes to sentimental analysis, and what weakness the different approaches have.

The limits of lexicon based sentimental analysis

One of the clear limits of sentimental analysis, especially when they are lexicon based as ours is, is the difficulty to understand more complex emotions such as irony, sadness, anger and so on. like Thomas Schmidt and Manuel Burghardt says in “An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing” “common

⁴ Salmi, *Nineteenth-Century Europe*, 1.

⁵ Ukendt, “Litteraturhistoriske oversigter”, Dansksiderne.dk, u.å., <https://dansksiderne.dk/index.php?id=3440>.

⁶ Thomas Schmidt og Manuel Burghardt, “An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing”, Sentimental Analysis, u.å.

problems of lexicon-based methods like the handling of irony and negations are certainly additional reasons for the mediocre accuracies. Based on our results, we consider the usage of general purpose lexicons alone as not sufficient to achieve acceptable accuracy scores”⁷ Which is also why our focus only lies in if the countryside is portrayed good or bad, and aren't sorting things into more complex emotions. We however are still vulnerable to irony, and figurative language, especially because we are using text as data⁸, which means our analysis isn't able to see the difference between how words are in relation to each other. Sentinel analysis, if based on a lexicon, is also limited to the words that lexicon holds. Which can be especially limiting if you are, use data which is from a different time period than the lexicon is decided for. you can of course teach an AI to understand the language and thereby make it more robust when it comes to new words, this is however resource intensive, and is thereby unavailable for us in this paper.

expectation/Hypothesis

As a result of our sentimental analysis we expect to see a clear tendency in contemporary literature to idealize the countryside and generally portray it as better than the city—especially in genres like *Romantikken*. We also strongly suspect that this tendency toward a more idealized view of the countryside increased later in the 19th century

Through sentiment analysis of danish 19th-century fictional literature we expect to find a more positive description of rural life compared to city life. The increased urbanisation of society was criticized by influential authors at the time, as noted by Hannu Salmi⁹

Methods

Research Design

This paper aims to achieve a nomothetic and quantitative result. To do this, we will use n words around our predefined urban and rural keywords, respectively. This allows us to use the Danish sentiment analysis tool Sentida to transform these qualitative values into either negative or positive numbers, which we can then use to describe the general sentiment of city and rural life in the literature.

We have chosen to test our hypothesis on Danish literature from the 19th century, spanning different literary periods. We hope literature can provide insight into how people in the 19th century viewed and perhaps even idealized the countryside or the city. We will compare and contrast works spanning different authors, genres and themes to get a broad view. In this way, we can attempt to see patterns and potentially groupings of different views based on these factors.

N-words from keyword and Sentida

⁷ Schmidt og Burgardt, “An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing”.

⁸ Nina Tahmasebi og Simon Hengchen, “Samlaren” 140 (2019): 200–201.

⁹ Hannu Salmi, “19th Century

Our project is using N-words from keywords to find a negative or positive sentimental value. To get our key-word we have asked chatGPT to make us a list of words associated with the city, and the countryside. This list will be stated in the paper. To then determine the negative or positive value of the words around our keyword, we are using the danish lexicon based sentiment analysis software called sentida. We have not translated Sentido into the 19-hundred danish language, because we lack the resources to do so, and we are therefore only using the standard version of the program, there may therefore be certain words that the program is unable to understand.

Software Framework and Material

Hardware: Lenovo Yoga laptop

Processor AMD Ryzen 7 7735HS with Radeon Graphics 3.20 GHz

Installed RAM 16,0 GB (13,7 GB usable)

System type 64-bit operating system, x64-based processor

Software: Project coded in R version 4,4,2 on Windows 11. We used the tidyverse (Hadley et al 2019), tidytext, tokenizers , purrr, here, pdfutils, stringr, tidyr, sentida packages.

Data Acquisition and Processing:

We have used a combination of books and poems from 19th-century Denmark as sources for our analysis, which can be found under "Analyzed Sources." Our goal was to get the most representative view of literature from the 19th century. To do this, we focused mainly on different literary periods in the century. We selected the literary periods Romantikken, Romantismen, and Det moderne gennembrud, which were the main dominant periods in Denmark at the time. To ensure equal representation, we tried to use an equal number of texts from each period. We mainly used short stories as our sources because we believe they have a greater chance of casually expressing an opinion on either the city or the countryside. Romantikken is, however, a special case because this period is especially defined by poems. To compensate for this, we included more texts from this period due to their shorter length. The decision on how to select the most representative literature was not easy. Other options—such as focusing on when in the 19th century the literature was written—were considered. However, we chose to focus on literary periods instead, as we believe literature changes more significantly based on literary movements than on the exact year of publication.

Additionally, the literary periods are already spread throughout the 19th century, which we hope will still make our results representative for the century as a whole.

. μ

Analyzed sources

We have used the following books/poems in our analysis:

- 1) *Kun en spillemand* – <https://tekster.kb.dk/text/adl-texts-andersen05val-root>
- 2) *Et dukkehjem* – https://www.ibsen.uio.no/DRVIT_Du%7CDuht.pdf

- 3) *Lykke-Per* – <https://www.ronlev.dk/danske-klassikere/lykke-per-af-henrik-pontoppidan.html>
- 4) A collection of Danish romantic poems and works – <https://dansksiderne.dk/index.php?id=3441#c11398>
- 5) *Begrebet Angst* – <https://www.kb.dk/e-mat/dod/111408014975-bw.pdf>
fra hytterne - <https://www.kb.dk/e-mat/dod/130020755076.pdf>
Det forjættede land - https://books.google.com/books/download/Det_forj%C3%A6ttede_land.pdf?id=SRwwAAAAMAAJ&output=pdf