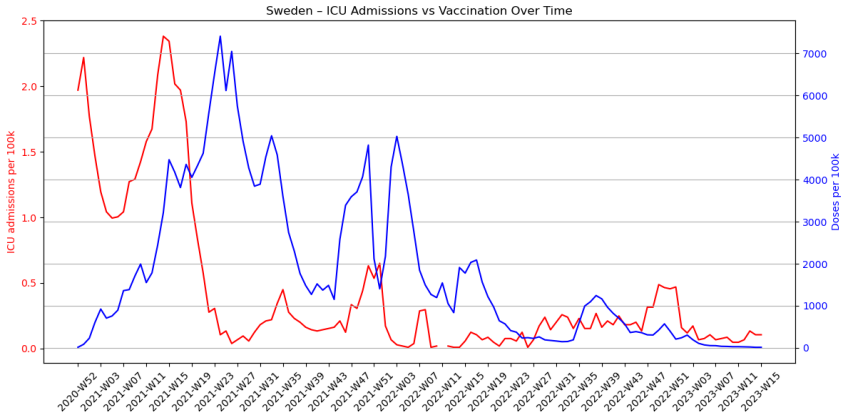
Assignment 2

Rasmus Eliasson & Oskar Flodin



rael23@student.bth.se
osfl22@student.bth.se

31 maj 2025

Blekinge Institute of Technology

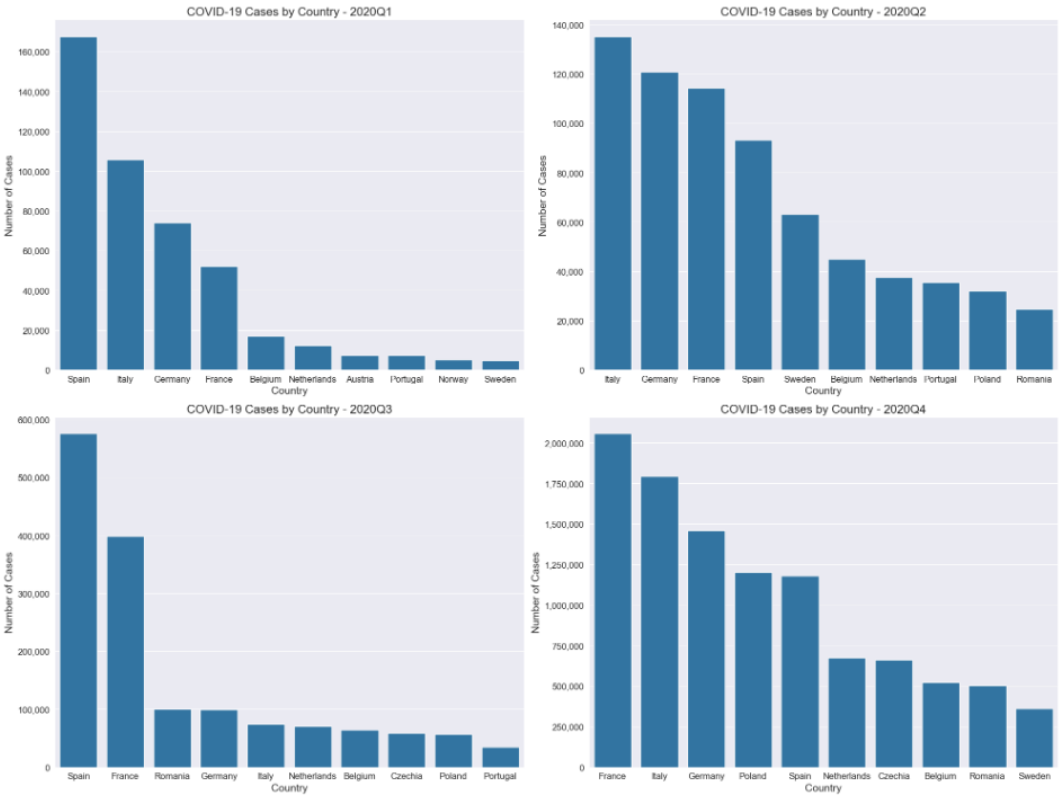
1 Introduction

This report presents an explanatory data analysis of COVID-19 in EU/EES countries, focusing on the spread of cases, deaths, vaccination coverage, and hospitalization trends. The analysis is based on datasets provided by the European Centre for Disease Prevention and Control (ECDC) and aims to identify patterns and insights related to vaccine uptake, age-based coverage, and healthcare system impact. The findings are supported by statistical methods and visualizations.

While not all visualizations generated during the exploratory analysis are included as figures in this report, the conclusions and insights presented are drawn from comprehensive examination of these plots within the accompanying Jupyter notebook.

2 Mandatory questions

2.1 Question 1



Figur 1. Amount of cases ranked in descending order by country in each year and quatre

Between 2020 and 2022, the countries that most frequently reported the highest number of COVID-19 cases per quarter were France, Germany, Italy, and Spain. These countries appeared consistently in the top 5 rankings across nearly every

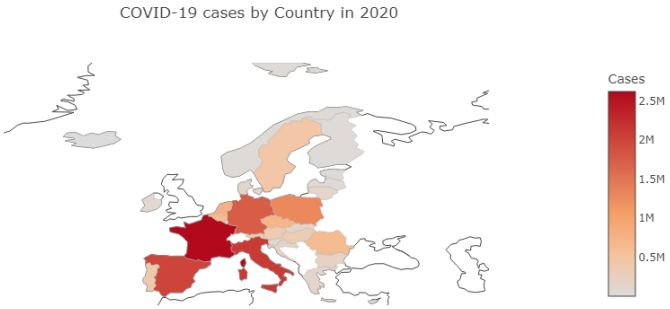
quarter. Smaller countries such as Belgium, the Netherlands, and Portugal appeared more sporadically and typically in the lower half of the top 10.

The high case counts in absolute terms can be partially explained by population size. However, in some quarters (especially Q1 of 2022) even smaller countries like Austria, Portugal, and Denmark reported several million cases, which is disproportionately high given their populations. This likely correlates with the rapid spread of the Omicron variant, which became dominant in early 2022.

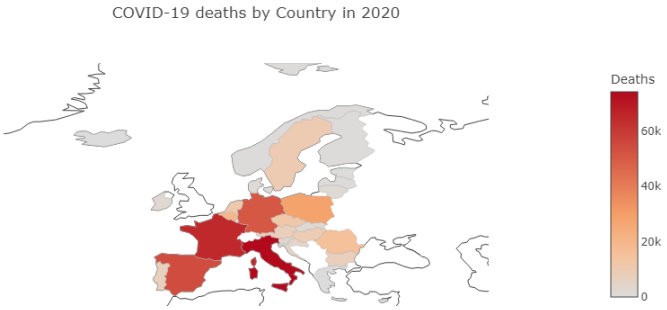
A geographical pattern is also noticeable. Most of the countries consistently in the top 10 are located in Western and Southern Europe, and many share borders (e.g., France–Germany, Germany–Austria). This regional clustering suggests that geographic proximity and mobility between these nations may have contributed to case transmission.

In summary, both population size and geographical location appear to be important factors in understanding the quarterly distribution of COVID-19 cases across Europe during this period.

2.2 Question 2



Figur 2. COVID-19 cases of each country in EU/EES (2020).



Figur 3. COVID-19 deaths of each country in EU/EES (2020).

The maps showing total COVID-19 cases and deaths for each year from 2020 to 2022 reveal several clear patterns:

- Central and Southern European countries such as France, Germany, Italy, and Spain consistently show higher numbers of both cases and deaths.
- Countries that are geographically more isolated, such as the Nordic countries and Ireland, tend to have lower case numbers across all three years. This might be due to factors like lower population density, stricter border control, or earlier lockdown measures.
- There is a strong correlation between the number of cases and deaths. Countries with the highest number of cases also tend to have the highest death toll. For example, in 2020 and 2021, France and Italy show both high case and death counts, which aligns with their overwhelmed healthcare systems during peaks of the pandemic.

2.3 Question 3

In Europe, the three most commonly administered COVID-19 vaccines were COM (Pfizer-BioNTech), MOD (Moderna), and AZ (AstraZeneca), in that order. Figure 12 presents a heatmap showing the distribution of these vaccines across EU/EEA countries as a percentage of total doses administered.

The data reveals that Pfizer (COM) dominated vaccine distribution in most countries, with between 59.6% and 85.7% of doses coming from this brand. An exception was Liechtenstein, where only 33.8% of doses were Pfizer, and Moderna (MOD) was the most used vaccine. The second most used vaccine varied more widely across countries, although Moderna generally had greater uptake than AstraZeneca.

It is important to note a limitation in the dataset: for Germany and France, a large proportion of the vaccine data is categorized under "unknownbrands. Since these two countries represent a significant share of the EU population, this uncertainty limits the precision of cross-country comparisons.

2.4 Question 4

The three heatmaps figure figure 13 shows doses for each vaccine in top 3 per 100,000 inhabitants by target group and country, where the target groups are children(below 18), adult(18-59) and senior(above 60). The colour scale is logarithmic to highlight both small and large differences.

- Pfizer is in general for all different target groups but is heavily dominated for children and the vaccine most children used.
- Moderna has a similar pattern as COM in terms of all three target groups got the vaccine and the difference is the amount of doses.
- AstraZeneca had a patter on targeting the groups senior and adults.There where only two countries where a bigger amount of children vaccinated with AstraZeneca.

2.5 Question 5

In identifying vaccine sceptical countries, two criteria were examined, due to lacking data:

- Late start of vaccination campaigns
- Low total vaccine doses administered per capita

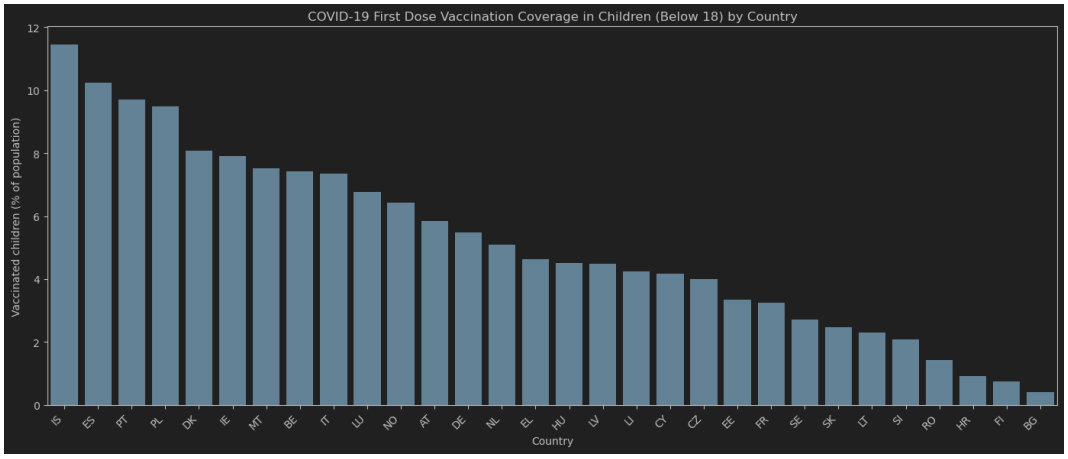
The countries that started vaccinations latest were Bulgaria, Romania, and Croatia. The lowest vaccination rates per capita were seen in Bulgaria, Romania, and Liechtenstein.

Statistical tests showed that in vaccine sceptical countries, higher vaccination rates were generally linked to lower hospitalization levels. However, Ireland despite high vaccination rates showed a different pattern. They showed a positive correlation between hospitalization and vaccination. Meaning they had increase hospitalization together with increase vaccination. And for ICU admissions, their correlation was negative but weaker. This suggests that variations in healthcare systems or the emergence of new COVID-19 mutations may affect how vaccination influences hospitalization outcomes across countries.

In conclusion, the data indicates that vaccination helped overall to reduce hospitalizations, especially in countries with initially low vaccine uptake and potentially weaker healthcare systems.

2.6 Question 6

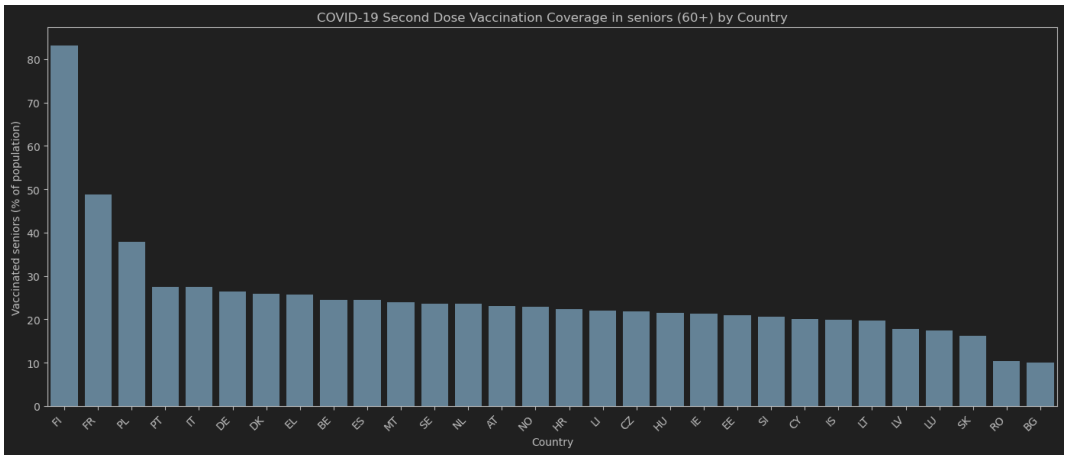
Based on the bar plot in figure 4, Iceland, Spain, and Portugal have the highest proportion of children under 18 vaccinated with a first COVID-19 dose, representing up to 11.5% of their total populations. In contrast, Bulgaria, Finland, and Hungary show the lowest shares, with Bulgaria at just 0.5%.



Figur 4. Amount of first doses for children below 18 divided by population of the country in the EU/EES.

2.7 Question 7

The bar plot in figure 5 shows anomalies in Finland (83%), France (49%), and Poland (38%). These values are likely inaccurate, as they imply over 100% of the senior population received a second dose. This may result from double-counting or differences in national vs regional reporting. Excluding these, the top countries for vaccinating seniors are Portugal (28%), Italy (27%), and Denmark (26%). A pattern emerges where Northern and Western European countries show higher coverage, while Eastern countries like Romania and Bulgaria have significantly lower rates, below 15%.

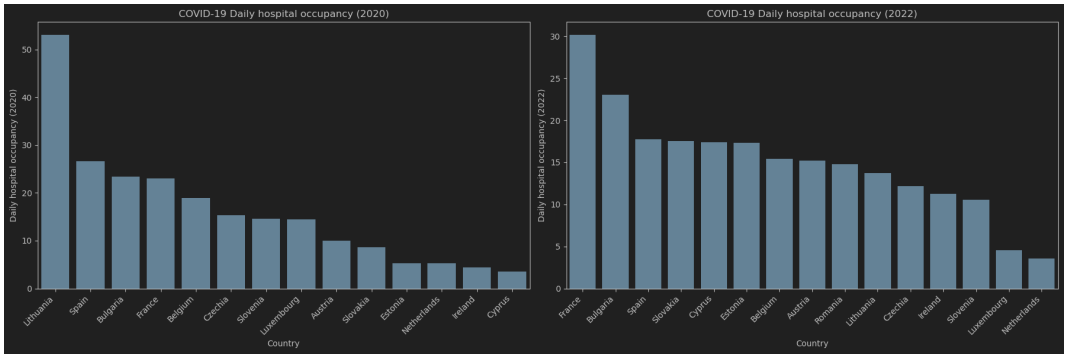


Figur 5. Amount of second doses for seniors 60 and above divided by population of the country in the EU/EES.

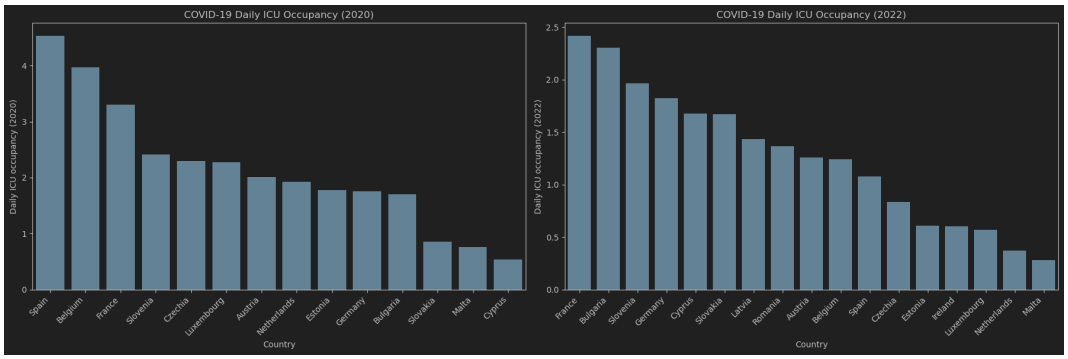
2.8 Question 8

In 2020, Spain, Belgium, and France had the highest mean ICU occupancy, while Lithuania, Spain, and Bulgaria led in the mean hospital occupancy. In 2022, France, Bulgaria, and Slovenia topped ICU, and France, Bulgaria, and Spain had the highest hospital loads.

France's ICU occupancy dropped by 27%, while hospital occupancy rose by 30%, reaching the highest level in 2022. The largest ICU drops were in Spain (-76%) and Belgium (-70%), while Cyprus (+360%) and Slovakia (+105%) had the biggest hospital increases.



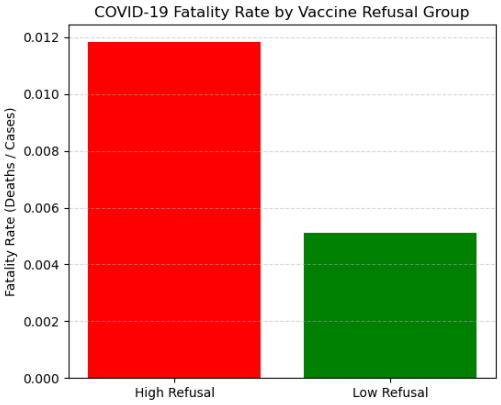
Figur 6.



Figur 7.

3 Further questions

3.1 How did the COVID-19 fatality rate differ between countries with high vs. low vaccine refusal rates?



Figur 8. COVID-19 fatality rates by vaccine refusal group

Countries were categorized into *high* and *low* vaccine refusal groups based on the median number of vaccine doses administered per capita. The average COVID-19 fatality rate (deaths per reported cases) in the high refusal group was **1.17%**, more than twice the **0.52%** observed in the low refusal group.

A Mann–Whitney U test confirmed this difference to be statistically significant ($U = 166.000$, $p = 0.0087$).

Interpretation

Higher vaccine refusal rates are associated with significantly higher fatality outcomes, suggesting that vaccine uptake likely played a key role in reducing COVID-19 mortality. However, vaccine coverage is also intertwined with broader structural factors such as:

- Stronger healthcare infrastructure
- Faster and more organized pandemic response
- Socioeconomic stability and public trust in institutions

These factors may confound a direct causal interpretation, but they underscore the complexity of public health outcomes.

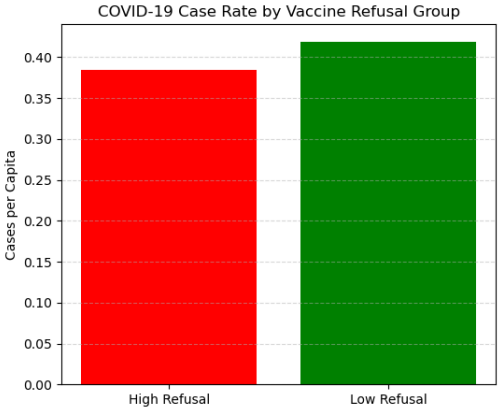
Additional Factors to Consider

- Timing and prevalence of viral variants
- Differences in reporting accuracy and testing availability
- Public health compliance and societal behaviour
- Trust in government and vaccine communication efforts

Conclusion

While increased vaccination rates appear to be strongly linked to lower fatality rates, this relationship is part of a wider network of public health influences. Crucially, the populations most vulnerable to severe outcomes were often those facing the greatest barriers to vaccine access or the highest levels of scepticism. Addressing both logistical and perceptual challenges remains essential for future preparedness.

3.2 How effective was vaccination in mitigating the spread of COVID-19 and reducing ICU utilization?



Countries were categorized into *high* and *low* vaccine refusal groups based on the median number of vaccine doses administered per capita. The average COVID-19 case rate in the high refusal group was **38.42%** of the population, compared to **41.91%** in the low refusal group.

A Mann–Whitney U test was used to assess whether the difference in case rates between groups was statistically significant. The result was a failure to reject the null hypothesis ($U = 88.000$, $p = 0.4911$), indicating no statistically significant difference in case rates between the groups.

Figur 9. COVID-19 case rates by vaccine refusal group

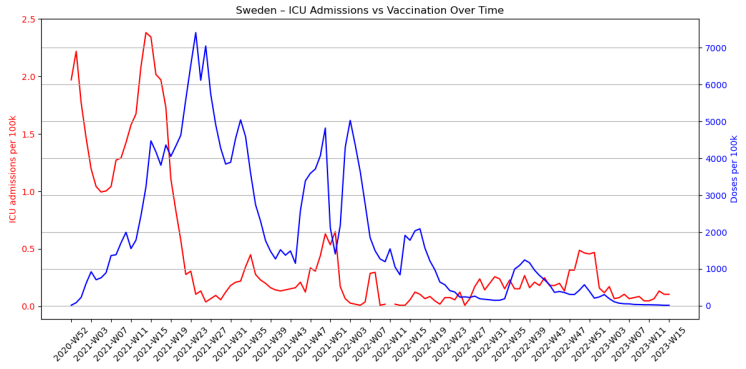
Interpretation: Impact on COVID-19 Spread

Based on these static group comparisons, we cannot confidently associate high vaccine refusal with increased COVID-19 spread. This suggests that vaccination had limited or no detectable effect on transmission rates in this dataset.

Visual Trends: ICU Admissions and Vaccine Coverage

Since vaccination did not show a clear effect on case incidence, we investigated whether vaccines reduced severe outcomes. This was done by normalizing ICU admissions for countries by rates per 100 000 population.

As can be seen in the plots, countries such as France and Sweden revealed a consistent trend: as vaccine coverage increased, ICU admissions declined. For example, Sweden experienced a marked drop in ICU admissions after mid 2021. This suggests vaccines had the expected effect. That they were effective in reducing the severity of infections.



Figur 10.

Conclusion

From this analysis vaccination levels did not appear to prevent the spread of COVID-19. However, temporal data strongly suggest that vaccines effectively reduced ICU utilization. This highlights the positive impact and necessity of vaccination.

3.3 Did some countries experience a continuous increase in daily COVID-19 deaths, despite increased vaccination coverage?

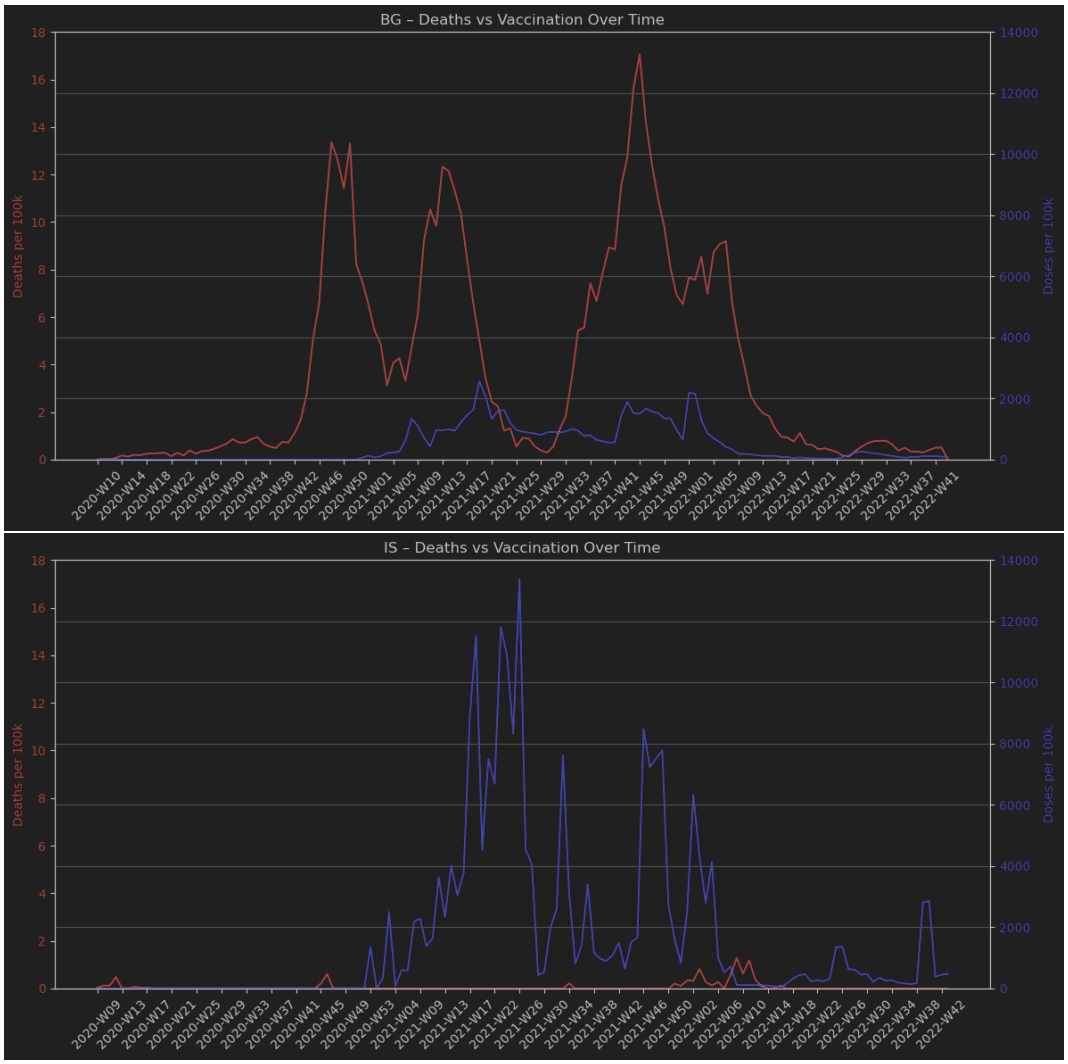
This question was chosen because it explores whether increased vaccination alone was sufficient to reduce COVID-19 death rates, a topic of high public health relevance. It helps explain why some countries continued to experience high mortality despite vaccination and highlights the importance of broader healthcare capacity and vaccine uptake. To address this, both visual time series analysis figure 11 and a statistical correlation test were used.

Visual Interpretation

Time series plots were created for selected countries, showing deaths per 100.000 (red) and vaccine doses per 100.000 (blue) over time. This visualization technique clearly highlights the temporal relationship between vaccination coverage and mortality. Countries such as Bulgaria, Hungary, and Slovakia show persistently high or repeated peaks in deaths, despite increases in vaccination coverage. In contrast, countries like Iceland, Norway, and the Netherlands demonstrate a decline in death rates shortly after the start of mass vaccination campaigns.

Statistical Correlation

To validate these visual patterns, a Spearman correlation test was applied per country. This test was chosen because the data was not normally distributed and included extreme values, making non-parametric methods more appropriate for detecting monotonic trends.
Statistical results confirmed this variation:



Figur 11. Graphs showing Deaths and vaccination doses per 100.000 for Bulgaria and Iceland

- Bulgaria rejected the null hypothesis and showed a positive correlation between vaccination and deaths, likely due to delayed rollout and rising deaths before vaccine impact.
- Iceland did not reject the null hypothesis, indicating no significant relationship between vaccine coverage and death rate over time.

Conclusion

Visual patterns indicate that vaccination tended to reduce deaths, but this effect was not consistent across countries. In nations with low vaccine uptake or weaker healthcare systems, such as Hungary and Bulgaria deaths remained high despite vaccine rollout. This suggests that while vaccines were critical, they were not sufficient alone. Healthcare system strength, timing of rollout, and public trust were also key to reducing COVID-19 mortality

4 Discussion and Conclusion

This report presented an explanatory data analysis of the COVID-19 pandemic across EU/EEA countries using multiple datasets from ECDC. Our analysis addressed key aspects such as infection patterns, vaccination rates, age group coverage, and healthcare burden. We observed clear geographical and temporal patterns in case counts and fatality rates, with Western and Southern European countries experiencing the highest burden, especially during the early phases of the pandemic.

Vaccination clearly correlated with reduced ICU admission and lower fatality rates, particularly in countries with early and extensive vaccine rollout. However, the data also revealed exceptions. Some countries showed inconsistent or even paradoxical trends, such as increased hospitalizations despite high vaccine coverage. These inconsistencies may be attributed to a range of factors, including:

- Rapid vaccine rollout that may not allow time for full immunity to develop before exposure.
- Emergence of new viral variants with higher transmissibility or partial immune escape.
- Variations in the overall quality of healthcare systems.
- Inconsistencies or delays in data reporting and classification.

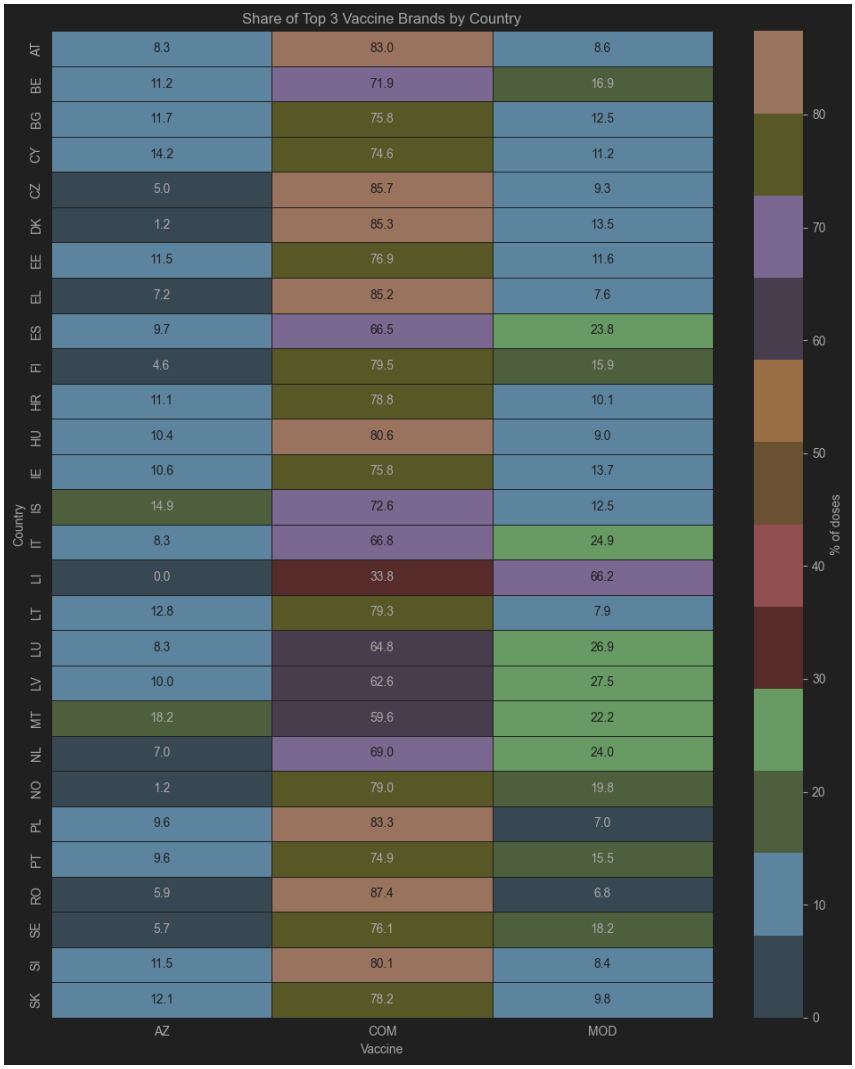
A combination of these factors may have contributed to apparent increases in ICU occupancy that do not accurately reflect vaccine efficacy, leading to potentially misleading or counterintuitive results.

4.1 Limitations and Future Work

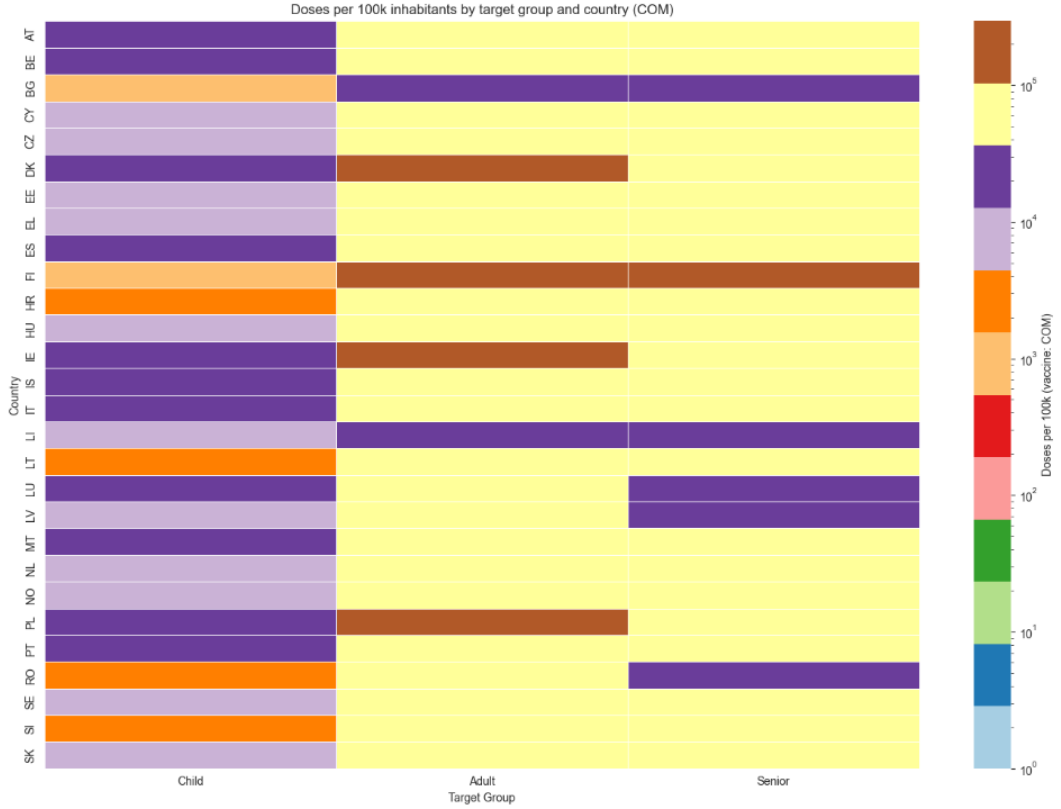
Some data limitations, such as unknown vaccine brands in large countries (e.g., France, Germany), and population inconsistencies in age based vaccine metrics, affected our ability to generalize certain trends. Future analyses could be strengthened by incorporating additional factors such as healthcare infrastructure, public trust indicators, and cultural habits. For instance, in Nordic countries like Sweden, lower population density and greater interpersonal distancing may influence transmission patterns. Including such factors would provide deeper, more nuanced insights.

In conclusion, while the data strongly supports the effectiveness of vaccination in reducing severe outcomes, a multifactorial view is essential to fully understand the variation in pandemic responses across Europe.

Appendix A – Full-size Figures



Figur 12. Amount of cases ranked in descending order by country in each year and quatre



Figur 13. Amount of Pfizer doses per 100,00 by each target group, Child(below 18), adult(18-59), senior(above 60).