**DO PEOPLE SEARCH FOR COVID-19 RELATED WORDS ON GOOGLE BEFORE GETTING VACCINATED?**

**The Relationship Between Google Searches Based on Google Trends Data and Actual Covid-19 Vaccination Amounts in Turkey**

CSSM 502 Final Project

Batuhan Hökelek

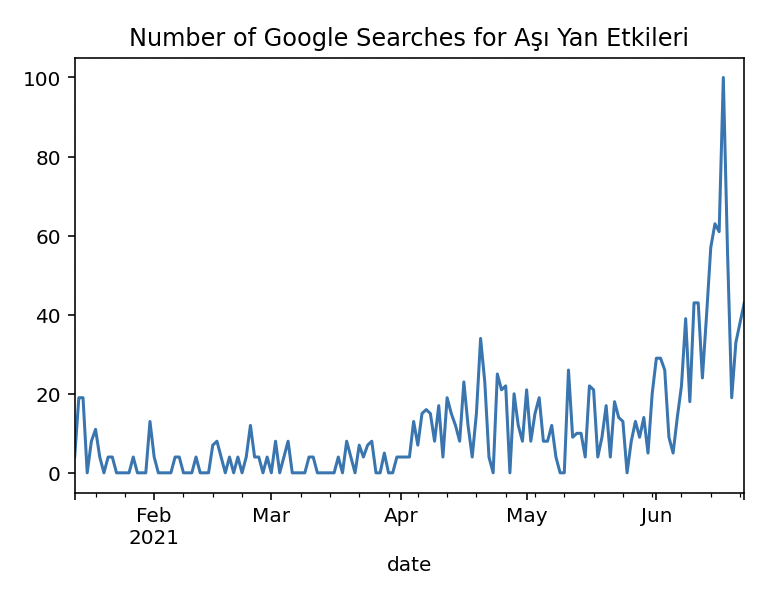
In this study, we examined the relationship between Google search queries related to Covid and the official numbers of vaccination in Turkey. We hypothesize that as people search more covid-19 pandemic-related words on google, they are more likely to get vaccinated since they become more informative about the Covid-19.

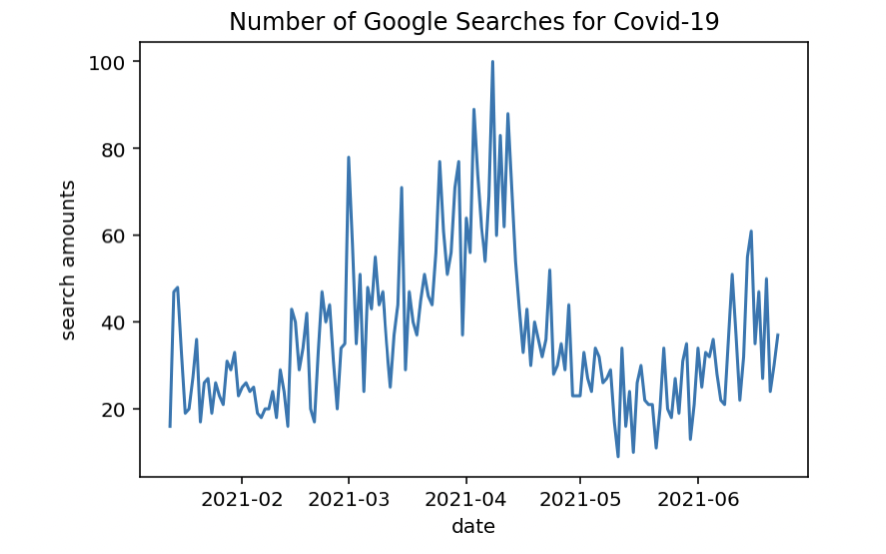
**Data Sets:**

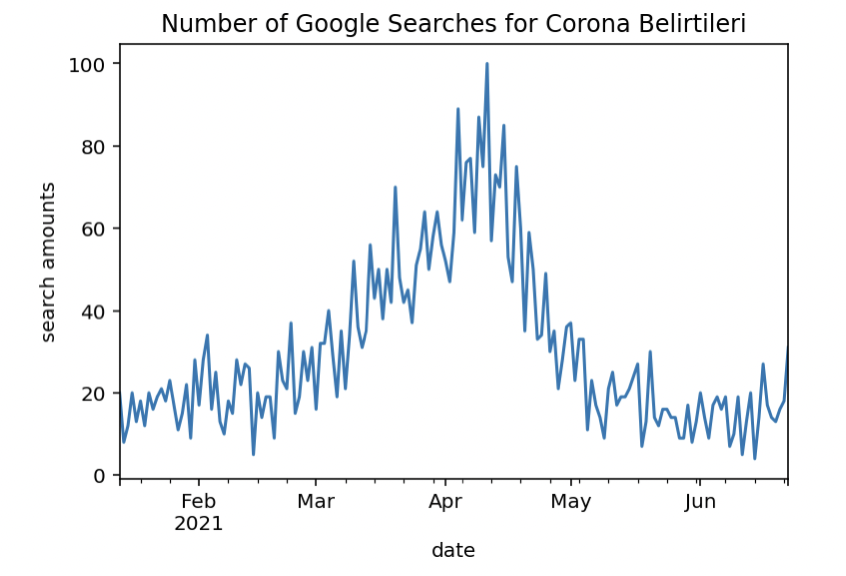
**Google Trends Data Set:** In this study, google trends[[1]](#footnote-2) was used to extract certain queries in google. For the time period, I decided to start with the first day of vaccination in Turkey. We started the time period with the start of vaccination and the 5 months following the first day. So, our time frame is from 13-01-2021 till 22-06-2021. To be able to get the data from google trends, we selected certain keywords by creating three groups which are keywords related to covid, related with vaccines, and related with variances in Turkey. We will add my keyword list in the attachment.

**Actual Vaccination Amount of Turkey Data Set:** For the actual vaccination data in Turkey, which is recently released by the Turkish Health Ministry in TURCOVID19[[2]](#footnote-3). We have gathered the data between 13-01-2021 and 22-06-2021 to create consistency among our two data sets. We cleaned the data and only kept total daily vaccination amounts for our analysis. We have used matplotlib to visualize three keywords selected from our keyword list. We chose “aşı yan etkileri”, “Covid-19”, and “corona belirtileri” since they are the three most searched keywords in google in our time period.

**Visualization of Google Trends Data**

****

****

****

Before starting our analysis, we made visualizations to understand our data better. We have used matplotlib to visualize three keywords selected from our keyword list. The three plots above are the normalized number of queries for three certain keywords. We chose “aşı yan etkileri”, “Covid-19”, and “corona belirtileri” since they are the top three most searched keywords in google in our time period.

**Methodology:**

As our methodology, we have conducted regression analysis. To be able to implement machine learning, we divided our merged data set into two distinct data sets. In the literature, dividing the data as %80 for the train data set and %20 for the test data set is the most recommended strategy in terms of accuracy. Therefore, we divided the data set following this strategy. Train data includes the date between 13-01-2021 till 21-05-2021 which is 129 days in total. Our test data then starts from 22-05-2021 till 22-06-2021.

**Linear Regression**

For the predictive performance of our regression, we have calculated R square measure. R square is found 1.0 for our model. Our Mean Square Error was found 0.0. Our statistical analysis yielded some unexpected outputs. We think that these results are due to the quality of our data set.

**Keyword List:** [“Covid-19", "kovid-19", "corona", "korona", "covid-19 belirtileri", "corona belirtileri", "kovid-19 belirtileri", "covid belirtileri", "korona belirtileri", "corona belirtisi", "korona belirtisi", "corona semptomları", "covid-19 semptomları", "covid-19 virüsü”, "covid virüsü", "corona virüsü", "koronavirüs", “pandemic", "pandemi", "karantina", "quarantine", "covid", "kovid", "Omicron", “delta", "omikron", "omicron belirtileri", "omikron belirtileri", "delta belirtileri", “varyant", "omicron semptomları", "omikron semptomları", "omicron öldürür mü", "omicron virüsü", "delta virüsü", “BioNTech","Turkovac","Sinovac", "Sputnik","biontech aşısı","sinovac aşısı","alman aşısı","çin aşısı”, "türk aşısı","rus aşısı","Covid aşısı","korona aşısı","corona aşısı", "covid 19 vaccine","biontech yan etkileri","alman aşısı yan etkileri", "çin aşısı yan etkileri","sinovac yan etkileri", "kol ağrısı”, "aşı yan etkileri","corona aşısı yan etkileri","korona aşısı yan etkileri”, "covid-19 aşısı yan etkileri", "sinovac aşısı yan etkileri”, "sputnik aşısı yan etkileri", "turkovac aşısı yan etkileri", "turkovac aşısı"]

1. https://trends.google.com/trends/?geo=TR [↑](#footnote-ref-2)
2. https://turcovid19.com/acikveri/ [↑](#footnote-ref-3)