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
Berra Karayel
0054477
David Carlson

In this study, we aim to understand the relationship between google search queries related with Covid-19 pandemic and official statistics of vaccination amounts in Turkey. Our hypothesis is 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] used to extract certain queries in google. For the time period, we decided to start with the first day of vaccination in Turkey. We started time period with the start of vaccination and the 5 months following the first day. So, our time period is from 13-01-2021 till 22-06-2021. To be able to get the data from google trends, we decided certain keywords by creating three groups which are keywords related with covid, related with vaccines, and related with variances in Turkey. We will add my keyword list in the attachment.

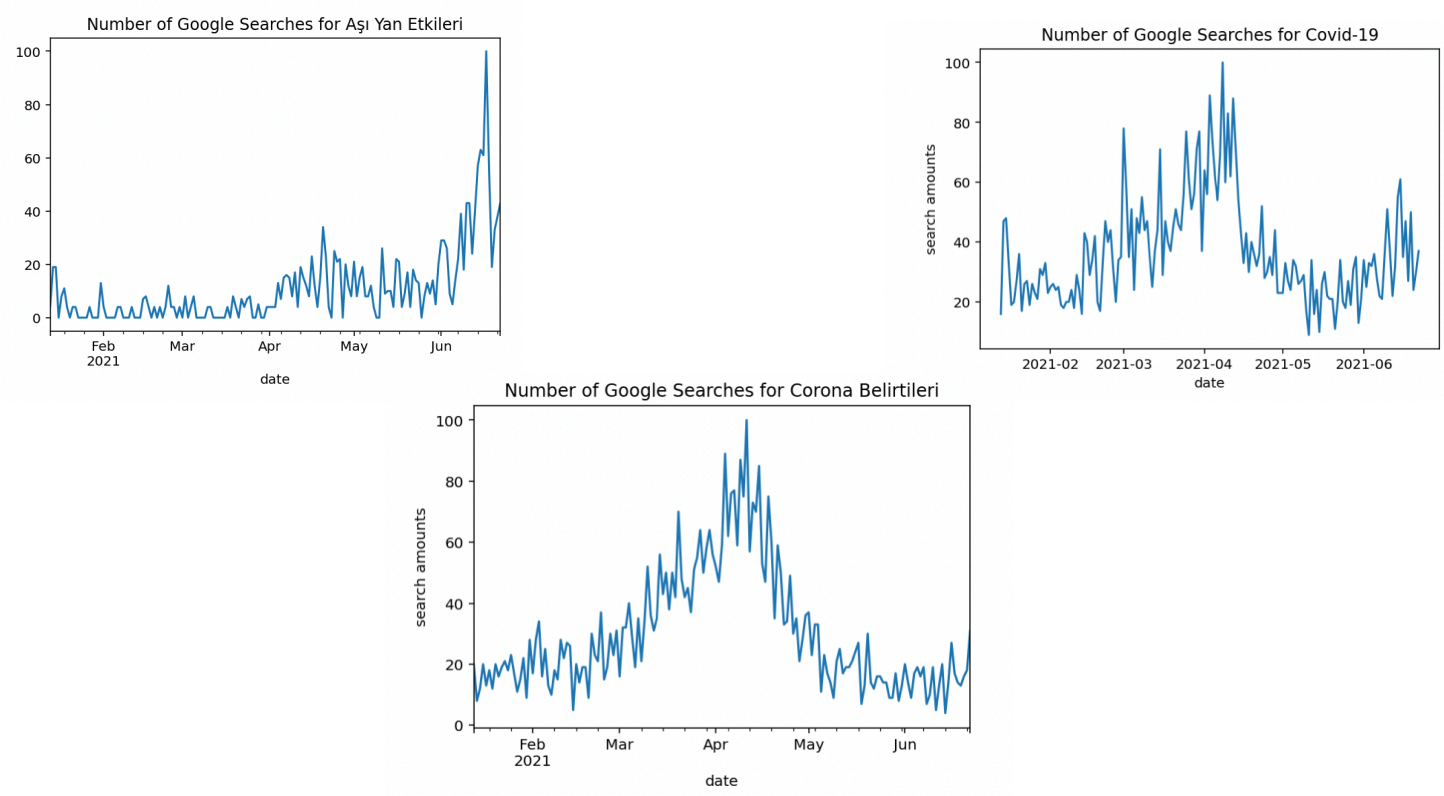
Actual Vaccination Amount of Turkey Data Set: For the actual vaccination data in Turkey, which is recently released by Turkish Health Ministry in TURCOVID19[2]. We have gathered the data

[1] https://trends.google.com/trends/?geo=TR

[2] https://turcovid19.com/acikveri/

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.

Visualization of the Data Collected from Google Trends



Before starting our analysis, we wanted to understand our data better and made visualizations. We have used matplotlib to visualize three keywords selected from our keyword list. These three plots 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 on google in our time period.

Methodology

As our methodology, we have conducted regression analysis. To be able to implement machine learning and measure the performance and accuracy of our model, we divided our merged

data set into two distinct data sets. In the literature, dividing the data as 80% and 20% is the most accurate one, so we decided to follow those ratio. Therefore, we have created our training and test data sets. 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 our analysis, we used linear regression analysis. For the predictive performance of our regression, we have calculated R2 measure. R2 is 1.0 for our model. Our Mean Square Error yielded as 0.0. We have reached some unexpected outputs as a result of our statistical analysis due to the quality of our Google Trends data. We think that these results are due to the unsuitability of our data sets.

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"]