Simen Logstein

[Company name]  [Company address]

Matte rapport

Effekten av vaksinen i Italia

Table of Contents

[Innlending 1](#_Toc104897946)

[Problemstilling 2](#_Toc104897947)

[Koder/algoritmer 2](#_Toc104897948)

[Biblioteker 2](#_Toc104897949)

[Plotte Grafer 2](#_Toc104897950)

[Lage derivert graf 2](#_Toc104897951)

[Regresjon i python 4](#_Toc104897952)

[Finne Start-Verdier 4](#_Toc104897953)

[Curvefit-funksjonen 5](#_Toc104897954)

[Resultat 5](#_Toc104897955)

[Grafene ut ifra datasettet 5](#_Toc104897956)

[Deriverte 7](#_Toc104897957)

[Regresjoner 8](#_Toc104897958)

[Døde som funksjon av Dager 8](#_Toc104897959)

[Friske som funksjon av dager 10](#_Toc104897960)

[Regresjons modell av Friske som funksjon av Døde 11](#_Toc104897961)

[Diskusjon 12](#_Toc104897962)

[Effekten av vaksinen 12](#_Toc104897963)

[Modeller av pandemien 13](#_Toc104897964)

[Feilkilder 14](#_Toc104897965)

[Konklusjon 14](#_Toc104897966)

[Kilder 14](#_Toc104897967)

# Innlending

Covid har forårsaket mange dødsfall i mange land. Noen land håndterte denne globale pandemien bra, mens andre gjorde ikke det. Italia, et land med masse gamle mennesker over 65 år ble sterkt rammet av Covid-19. Hvis man ser på død per kapital, så kommer Italia på 4 plass, selv om de ligger på 37 plass når det gjelder antall smittede.  
(weforum.org)

I 2020 11 måneder etter Covid ble oppdaget kom Covid-19 vaksinen og ble kvalifisert for å kunne bli tatt i bruk på folk. Dette var positivt ettersom folk kunne endelige gå ut i dagliglivet akkurat som før uten å bekymre seg for å bli syk og død. Studier viser at vaksinen har reddet 500 tusen folk på bare under et år. (ecdc.europa.eu).  
Dette temaet ble valgt fordi det virker interessant å se på statestikken på hvordan koronasituasjonen utvikler seg i land og hvordan helsevesenet gjør simulasjoner og antar hvordan situasjonen utvikler seg.

# Problemstilling

Denne rapporten har som formål å finne ut av 2 ting, nemlig å finne ut hvor stor effekt vaksinen hadde på dødsfallene og de som ble friske i Italia. Den skal også finne ut hvor gode modeller det går an å lage og hvor nøyaktige de er.

# Koder/algoritmer

## Biblioteker

For å gjøre undersøke datasettet trengs det mange forskjellige koder, algoritmer og funksjoner som må brukes for å manipulere dataen og framstille data’en på en god og enkel måte. Noen av funksjonene som blir brukt finnes ikke i normal python og må derfor bli importert i starten av koden. I de følgende kodene, så er de følgene bibliotekene lastet ned.

* Pylab
* Pandas
* Numpy
* Matplotlib.pyplot
* Scipy.optimize

## Plotte Grafer

For å finne ut av hvor stor effekt vaksinen hadde på Italia, så må først dataen til de friske bli tegnet på en graf. For å tegne den brukes kommandoen «Plot(x, y)». Dette vil da lage en graf med de gitte x-verdiene og y-verdiene.

For å sette X-verdiene og Y-verdiene til data’ene i excel dokumentet, så brukes pandas.

Text

Description automatically generated with low confidence  
Bildetekst: Program som tegner dataen i excel dokumentet

## Lage derivert graf

For å finne stigningstallet, så må man ta og programmet, må si at dette skal gjøres for hver eneste X-verdi og Y-verdi.  
Dette gjøres ved å lage 2 lister, en for X-verdier og en for Y-verdier. For å sette inn hver eneste kan vi bruke en «for i in range» funksjon.

Er sånn man kan tenke seg det.

For å nå gjøre kan vi lage en til «for i in range» løkke som kan se sånn ut.

Dette ville da være tankegangen bak hvordan man lager en graf av den deriverte.  
Hvis man da tar dette i praksis, så får man et sånt resultat.

Text

Description automatically generated

Kan lege til Mean() for å få en gjennomsnitt for en mengde x og y verdier.

## Regresjon i python

Regresjon er vanskelig å kode, derfor brukes scipy.optimize for å regne ut regresjonsmodellen.  
Eneste som trengs er å definere typen funksjon og startverdiene. For eksempel:

Denne måten for å lage en regresjonsmodell til dataen kan brukes for å lage logaritmiske modeller og eksponentiell modeller.

Under, så står det både litt om hvordan curvefit fungerer og hvordan man kan få startverdiene.

### Finne Start-Verdier

De er forskjellige metoder og måter å finne start verdiene i en funksjon. Som i en logistisk funksjon som har man variablene a, b og c i formen

Det at C er øverst viser at C-verdien er en horisontal asymptote og det betyr at C kan da være høyeste verdi i grafen. En «bubble sort» kan finne ut hva den største verdien i listen din er.

En startverdi for A formel kan bli funnet ved å ta f(0) og enten finne verdien ved algebra eller Cas.

Startverdi for B kan også bli uttrykt som en formel, men kan ikke ta f(0). Istedet tas F(medianen).

Startverdiene i dette forsøket er derimot ubetydelige siden de er der egentlig bare for å hjelpe programmet litt hvis det skal trengs, men siden datasettet har såpass mye informasjon og data, så kan alle startverdiene settes som standarden, 1.

### Curvefit-funksjonen

«Curvefit» funksjonen fra Scipy er veldig bra for å regne ut en passende modell og gjør dette ved å gjøre «minste kvadraters metode». Det er å ta kvadratet til differansen mellom hver observasjon og gjennomsnittsverdien av alle tallene du har (*arimetiske middeltall*) og summere alle sammen. Det som blir regnet ut er verdien til parameterne som gjør at kvadratsummen er så liten som mulig.

Hvis den skulle lage en graf av antall friske som funksjon av dager, så ville Y være friske og x være dager og den ville prøvd å sette x inn i ulike situasjoner/formler for å gjøre Q, så liten som mulig.

# Resultat

## Grafene ut ifra datasettet

Chart, line chart

Description automatically generated

Figur: 1

Chart, line chart

Description automatically generated

Grafen her virker ulogisk siden alle de friske plutselig blir null. Derfor er en bedre modell en som bare ser på dataen før den plutselig blir null. Dette gjør vi for alle de andre modellene som innebærer friske data’en.

Chart, line chart

Description automatically generated

Figur: 2

Chart, line chart

Description automatically generated

Figur: 3

## Deriverte

Chart, line chart

Description automatically generated

Figur: 4

Chart, line chart

Description automatically generated

Figur: 5

Chart, histogram

Description automatically generated

Figur: 6

## Regresjoner

I utskriftene av funksjonene, så er den tykke blå funksjonen en tegning av dataen i datasettet, mens den rød er en regresjonsmodell Python har foreslått.

### Døde som funksjon av Dager

Chart, line chart

Description automatically generated

Figur: 7

Graphical user interface, text, application

Description automatically generated with medium confidence  
Ut ifra regningen til programmet, så skal A = 19.9, B = 0.01, C = 153 251

Det vil si at funksjonen Python foreslår er

Chart, line chart

Description automatically generated

Figur: 8

Text

Description automatically generated

Python regner ut at A = 33122.5 og B = 1.002. Da blir funksjonen

### Friske som funksjon av dager

Chart, line chart

Description automatically generated

Figur: 9

Text

Description automatically generated with medium confidence

I følge python programmet er funksjonen når den følger en logaritmisk vekst.

Chart, line chart

Description automatically generated

Figur: 10

Text

Description automatically generated

Når python følger en eksponentiell vekst så blir funksjonen

### Regresjons modell av Friske som funksjon av Døde

Chart, line chart

Description automatically generated

Figur: 11

Graphical user interface, text

Description automatically generated

Python anbefaler funksjonen

Chart, line chart

Description automatically generated

Figur: 12

Text

Description automatically generated

Python anbefaler den eksponentielle funksjonen

# Diskusjon

## Effekten av vaksinen

Ut ifra resultatet av plottingen, så er det en rekke informasjon man kan finne ut. Vaksinen ble klargjort for offentligheten etter 11 måneder fra da viruset ble først oppdaget i Kina. Fra grafen over antall friske som funksjon av dager, så er det en drastisk økning på de som blir friske av koronaviruset etter bare rundt 300 dager. Det kan først se ut som at veksten etter 300 dager ikke er et resultat av vaksinen, men siden målingene ble startet i slutten av januar, så matcher datoene fortsatt opp.

(github.CSSEGISandData)

Når det gjelder selve effekten, så kan man se at antall friske i figur 2 plutselig stiger etter 300 dager noe som indikerer at vaksinen har stor effekt, men hvis man går over til figur 1, så vil de døde også øke etter dag 300. Vaksiner er laget for å at kroppen enklere kan bekjempe sykdommen, men grafen reflekterer ikke dette.

Dette kan være fordi vaksinen ikke ennå har blitt gitt ut til alle, men folk er nå mer sosialt aktiv fordi de har snart tilgang til vaksinen. Det betyr at økningen i friske og døde skyldes at folk tar opp sosiale aktiviteter igjen i større grad og er ikke et resultat av vaksinen sin bruk.

En annen mulig grunn til økning i friske og døde i figur 1 og 2 kan være et resultat av mutasjoner av covid. Alpha-mutasjonen kom ut i september som kan også ha en rolle i antall smittede og døde, siden det ble rapportert at den var mer smittbar. Muligens skapte alpha-mutasjonen en smittebølge som kom da vaksinen ble gitt ut til offentligheten og er en mulig faktor for hvorfor døde og friske stiger.

En annen forklaring til hvorfor smitten økte er at Italia åpnet skolene sine i september 2020. Dette kan ha vært en stor grunn til at smitten økte, men burde ikke ha så stor effekt på mengden som dør siden, de som blir smittet ikke er innenfor risikogruppen.

En generell bedre måte å se vaksinen sin effekt er gjennom figur 3.  
Grafen på figur 3, så er det 37500 døde etter 300 dager. Det betyr at etter 37500 døde, så blir vaksinen laget og hvis vaksinen har en effekt, så vil den bli sett etter 37500.

Ut ifra figur 3, så ser det ut som det begynner å stige mer og mer etter x-verdien 37500 noe som betyr at det er flere og flere som blir friske etter at vaksinen blir laget.  
Dette ser ut til å bli reflektert i funksjonen sin deriverte. Stigningstallet starter veldig lavt, men går sakte opp etter x-verdien 37500 noe som reflekterer at flere blir vaksinert og blir friske etter korona. Fra figuren, så har ikke vaksinen en stor effekt med en gang, men dette kan være fordi alle vaksinene på starten ble kjøpt opp av de rike landene og ikke lot noen være igjen til de andre(statnews.com).

## Modeller av pandemien

Modeller av de forskjellige funksjonene var overaskende nøyaktige. Ikke alle passet like bra som andre av forskjellige grunner.

De fleste modellene som fulgte en logaritmisk vekst hadde en nøyaktig modell for datasettet og klarte å representere det faktiske datasettet med stor nøyaktighet i forhold til de eksponentiell funksjons modellene som også ble laget.

En mulig grunn til dette er at logaritmisk vekst er som oftest mer passende som en modell er at den har en mye treigere vekst enn den eksponentielle. Den eksponentielle har som natur å vokse raskt, men det er ikke såpass stor vekst i de fleste modellene. Dette forårsaker at Python prøver å lage mindre vekst, men må kompensere for dette med en større start-verdi.

En annen mulig forklaring er at den eksponentielle funksjonen ikke har en horisontal asymptote etter at x er større enn 0. Det betyr at Python lager en funksjon som kan ligge så nærme den ordentlige grafen, så lenge som mulig uansett hvor lenge datasettet fortsetter. Det er derfor den vil ha en treig stigning i tilfelle grafen til datasettet skal fortsette lengere enn det den gjør. Dette skjer ikke med logaritmisk vekst.

Logaritmisk vekst har en horisontal asymptote som stopper den å vokse mer. Denne funker bra for modellene, fordi den vet allerede at asymptoten er høyeste verdi i datasettet. Dette gjør at den kan lage en vekst som passer bra til å møte denne asymptoten på mengden x-verdier som er spesifisert.

## Feilkilder

En mulig feilkilde til funksjonene er at dataene ikke er helt nøyaktige. Funksjonene er blitt laget å lese av hver eneste x og y verdi, men hvis det plutselig blir stort avvik i dataen, kan dette forårsake feil i modeller og funksjonen. I et forsøk for å fikse dette, så tar rapporten gjennomsnitt av 30 verdier og plotter dem, men det kan fortsatt hende at det er noen avvik i datasettet som da skaper unøyaktigheter.

# Konklusjon

Ut ifra det som har blitt diskutert, så kan man konkludere at vaksinen har reddet liv, siden stigningstallet for folk som overleverer blir større og større jo lengere vaksinen har vært offentlig tilgjengelig. Stigningen er ikke det største, men det kan være fordi Italia ikke kjøpte opp mange vaksiner.

Vaksinen kan også ha hatt en negativ effekt at folk er mer avslappet når det gjelder korona regler fordi den har blitt laget og derfor stiger smitten, men er vanskelig å avgjøre om dette er grunnen til stigningen.

Når det gjelder modeller, så har det blitt funnet ut at logaritmisk vekst normalt passer bedre enn en eksponentiell vekst for å simulere veksten til et datasett.

# Kilder

<https://www.medicalnewstoday.com/articles/how-did-we-develop-a-covid-19-vaccine-so-quickly#MNT-takeaways>

<https://www.sciencenews.org/article/covid-coronavirus-vaccine-development-speed>

Hvor lang tid det tok for korona vaksinen ble laget.

<https://www.weforum.org/agenda/2020/12/italy-death-toll-pandemic-covid-coronavirus-health-population-europe/>

Statestikk om italia dødsfall

<https://www.news-medical.net/health/History-of-COVID-19.aspx>

timeline om korona

<https://www.brookings.edu/blog/order-from-chaos/2020/06/16/reopening-the-world-italys-reopening/>

<https://www.pnas.org/doi/10.1073/pnas.2019617118>

italia som åpner opp

<https://www.statnews.com/2021/12/13/we-have-enough-covid-vaccines-for-most-of-world-but-rich-countries-stockpiling-more-than-they-need/>

rike land som kjøper opp alle vaksinene