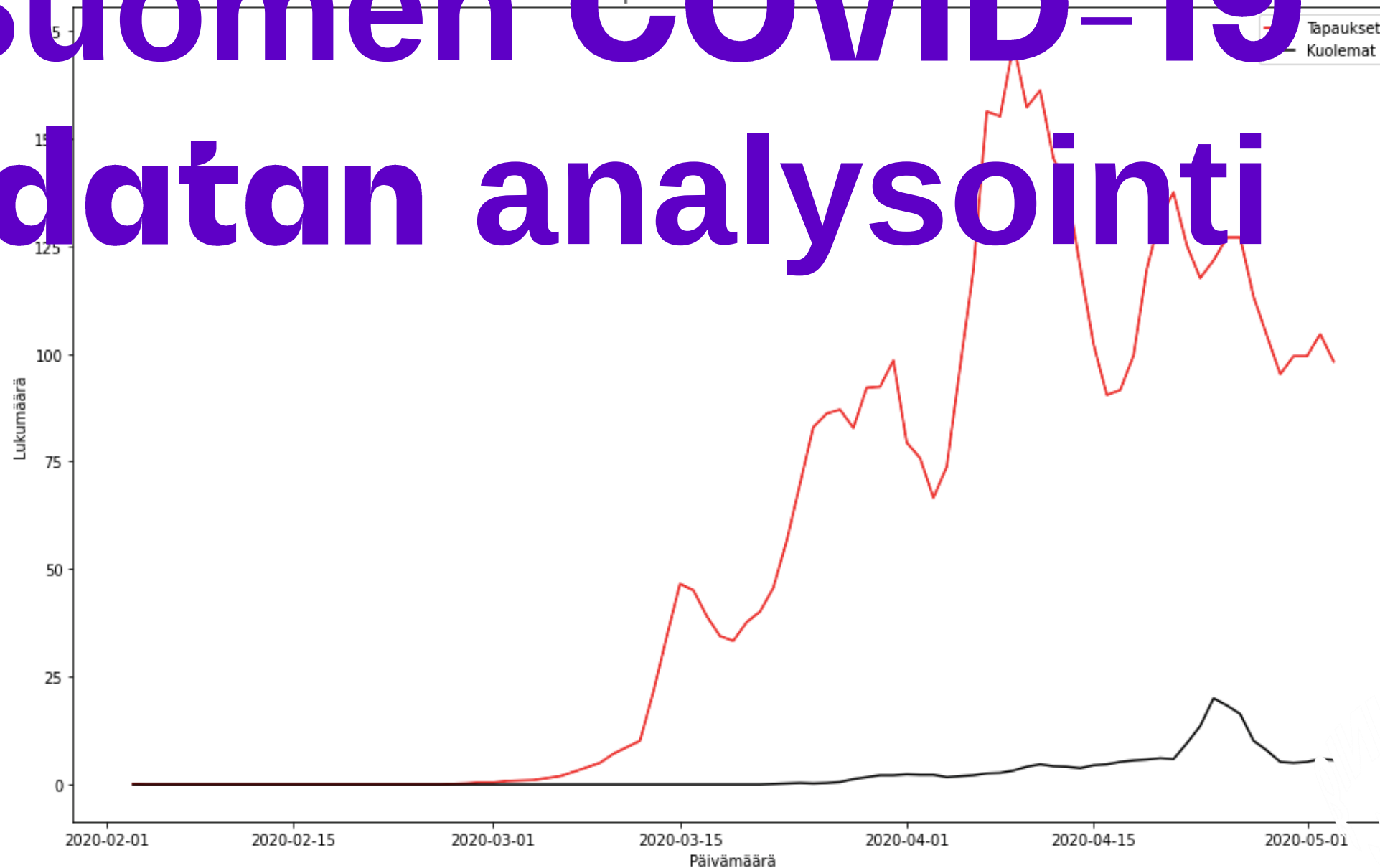


# Suomen COVID-19 -datan analysointi



# Johdanto

**Tässä työssä tutustutaan data-analytiikkaan käyttämällä EU:n julkaisemaa avointa dataa maailman COVID-19 tapauksista.**

**Työ on suunnattu data-analytiikasta, tilastotieteestä ja ohjelmoinnista kiinnostuneille, erityisesti lukiolaisille.**

**Tarvitset tietokoneen, Google-tunnuksen, uteliaan mielen ja 1-2 tuntia keskittynyttä työaikaa.**

**Avainsanat: *Epidemiologia, data-analytiikka, ohjelmointi, Python, Google Colab, terveysteknologia***

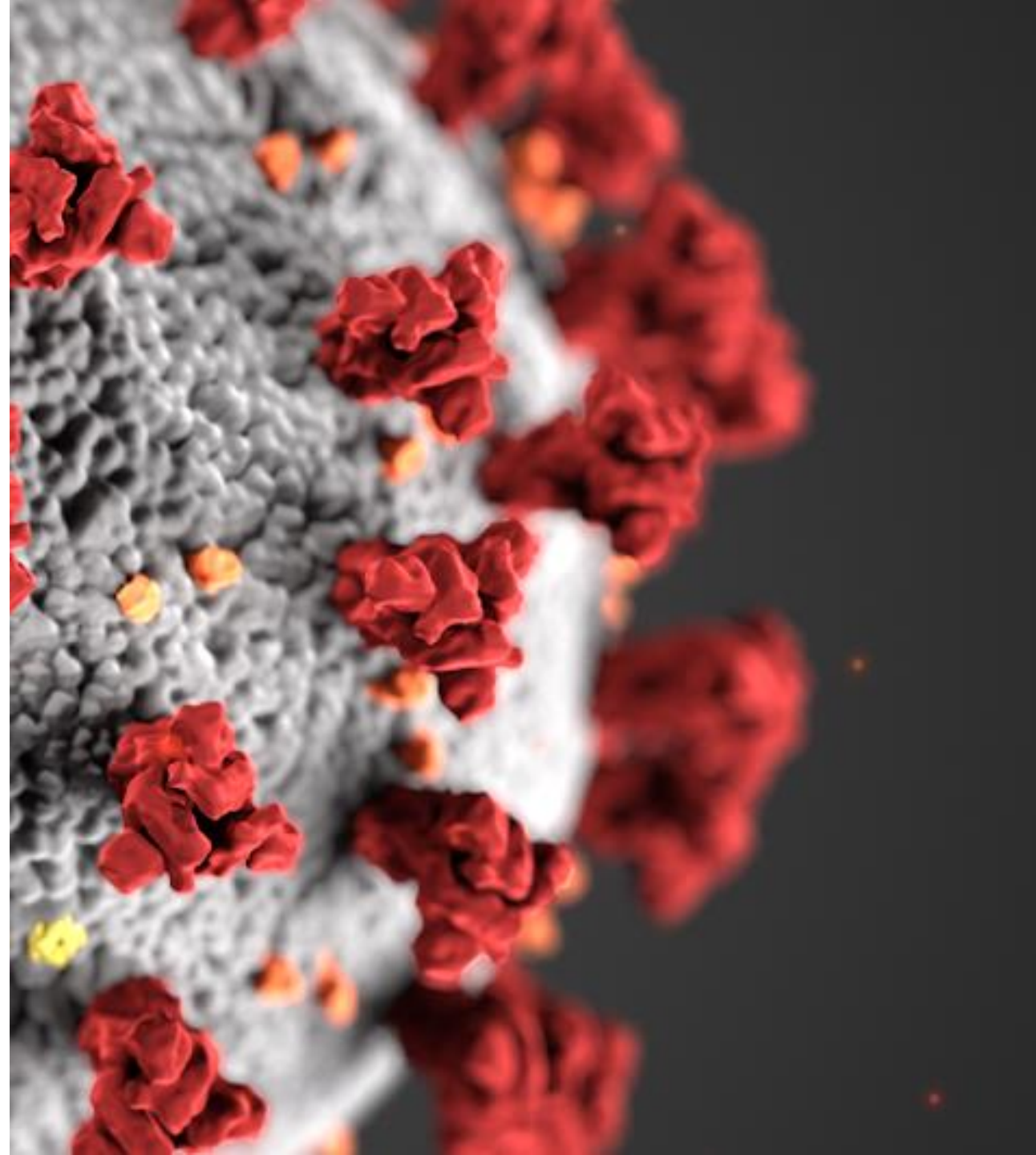
# COVID-19

**COVID-19 on virusperäinen hengitystieinfektio tauti, jonka aiheuttaa koronavirus SARS-CoV-2.**

**Arkikielessä käytetty sana koronavirus viittaa virusheimoon, jonka virukset aiheuttavat ihmisille ja eläimille tauteja.**

**Virus on ensimmäisen kerran todettu Kiinassa Wuhanin maakunnassa marraskuussa 2019 ja maaliskuussa 2020 viruksen levinneisyys todettiin niin suureksi, että WHO julisti taudin pandemiatilaan.**

**Tauti on vaarallinen ja vakava erityisesti vanhoille ihmisille ja pitkäaikaissairauksista kärsiville. Nuoret ja perusterveet voivat sairastaa taudin jopa ilman oireita.**



# Data-analytiikka

**Data-analytiikka voidaan yksinkertaistettuna jakaa kolmeen osaan:**

**Datan eli tiedon valinta, kerääminen ja esikäsittely**

**Valitaan datan lähde, kerätään data ja käsitellään se helposti käytettävään ja analysoitavaan muotoon, esimerkiksi excel-  
taulukon tapainen csv-muoto.**

**Analyysi**

**Lasketaan tilastoarvoja esim. korrelaatio, keskiarvo, keskihajonta**

**Luodaan visualisointeja, esim. taulukkoja ja kuvaajia**

**Luodaan ennusteita**

**Tehdään käytännössä mitä vain, jotta saadaan esitettyä ja  
ymmärrettyä kokonaisuuksia datasta.**

**Johtopäätökset**

**Tehdään ratkaisuja, ehdotuksia ja johtopäätöksiä analyysin  
perusteella. Tässä vaiheessa tarvitaan usein dataan liittyvän  
aihepiirin vahvaa osaamista, jotta pystyy yhdistämään uuden  
tiedon vanhaan.**

# Python- ohjelmointi

**Työssä käytetään data-analytiikassa suosittua Python-ohjelmointikieltä.**

**Vuonna 2019 Python nousi maailman toiseksi käytetyimmäksi ohjelmointikieleksi.**

**Python sopii monenlaiseen ohjelmointiin, mutta kielenä on silti yksinkertainen ja ymmärrettävä.**

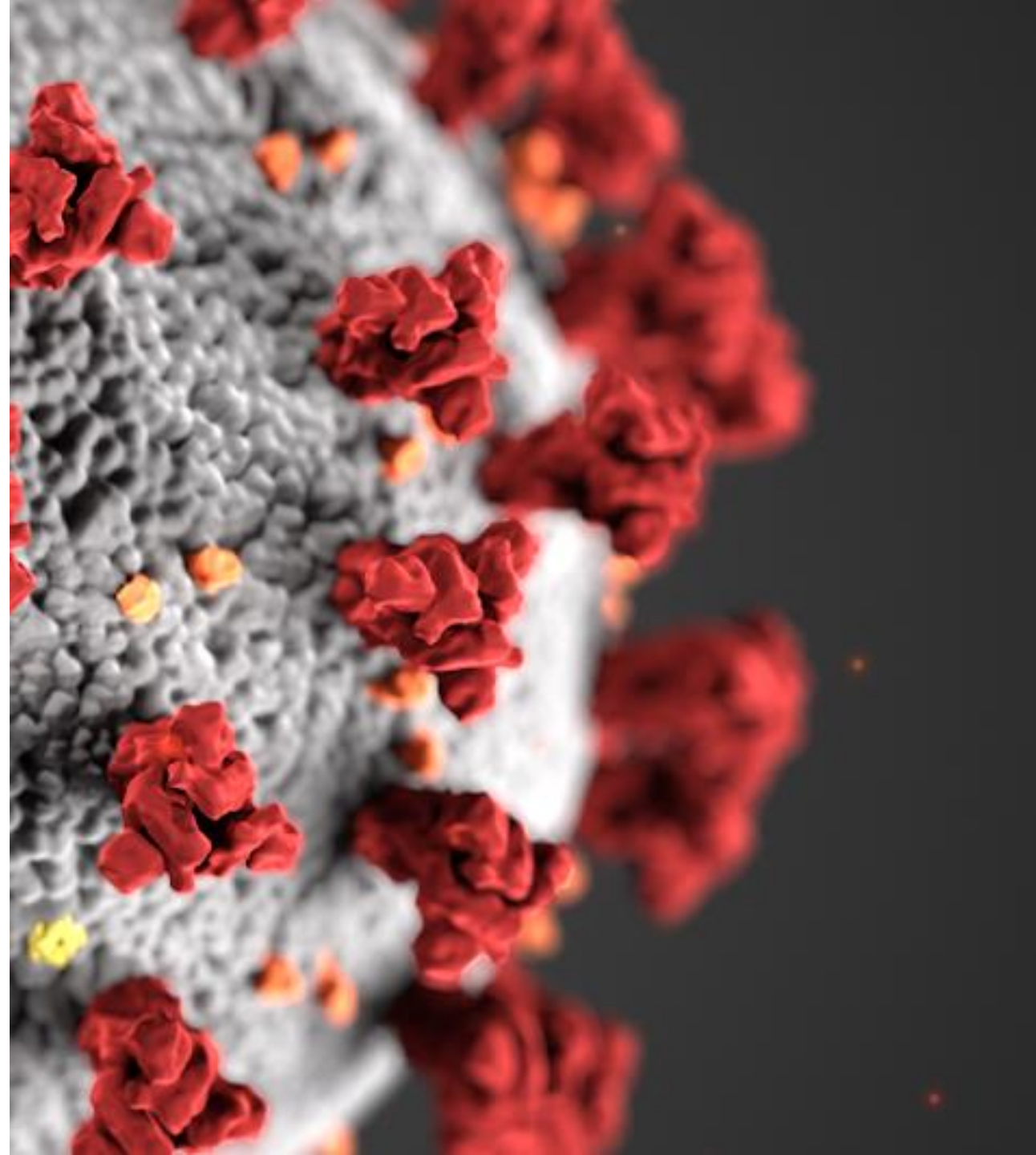
**Näin se sopii hyvin niin aloittavalle ohjelmoijalle kuin alan huippuammattilaiselle.**

# Google Colab

**Ohjelmointiympäristö on ohjelma, jossa koodi kirjoitetaan, koodia muokataan ja usein myös ajetaan eli saadaan toimimaan se, mitä on koodattu.**

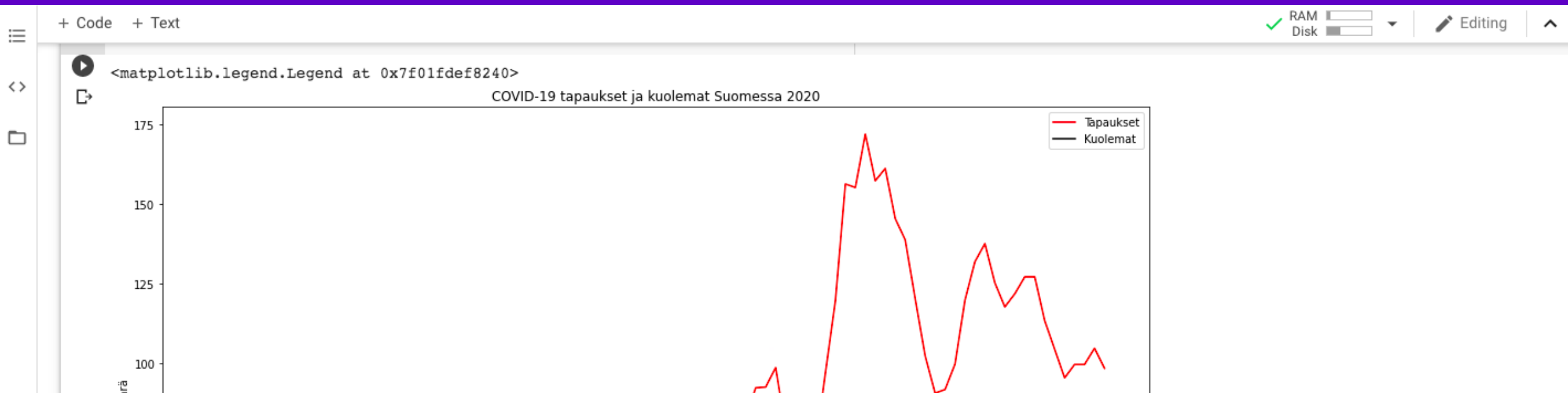
**Python-kieltä voi ohjelmoida lukuisissa eri ohjelmointiympäristöissä. Tässä työssä käytetään helppokäyttöistä Google Colabia.**

**Ympäristö toimii täysin selaimella eli sinun ei tarvitse asentaa tietokoneelle uusia ohjelmia.**



# Työvaiheet

Seuraavaksi pääset tutustumaan itse dataan, laskemaan sille tilastoarvoja ja luomaan kuvaajia epidemian kehityksestä Suomessa. Klikkaa [tästä](#), niin pääset siirtymään tarkempiin työohjeisiin Google Colabissa!



# Tiesitkö että...

**Aalto-yliopistossa on paljon data-analytiikan huippuasiantuntijoita.**

**Aallossa tutkitaan koronaviruksen (COVID-19) leviämistä sekä tilasto-  
että molekyylitasolla.**

**Tässä pari linkkiä tutkimustuloksiin:**  
<https://www.aalto.fi/fi/uutiset/koronan-leviamista-supertietokoneella-mallintaneet-tutkijat-tarkeinta-on-nyt-valttaa>  
<https://www.hs.fi/kotimaa/art-2000006455499.html>

**Kiiitos!**

**#AaltoJunior**

**#kokeilekotona**

# Lähteet

## Kuvat:

Aalto Junior

<https://www.state.gov/wp-content/uploads/2020/03/2871-1024x576.png> (29.4.2020)

## Verkkolähteet:

<https://fi.wikipedia.org/wiki/Data-analyysi> (29.4.2020)

<https://scand.com/company/blog/top-programming-languages-to-use-in-2020/> (29.4.2020)

[https://fi.wikipedia.org/wiki/Python\\_\(ohjelmointikieli\)](https://fi.wikipedia.org/wiki/Python_(ohjelmointikieli)) (29.4.2020)

<https://www.who.int/news-room/q-a-detail/q-a-coronaviruses> (29.4.2020)

<https://data.europa.eu/euodp/fi/data/dataset/covid-19-coronavirus-data/resource/260bbbde-2316-40eb-aec3-7cd7bfc2f590> (29.4.2020)

<https://www.aalto.fi/fi/uutiset/koronan-leviamista-supertietokoneella-mallintaneet-tutkijat-tarkeinta-on-nyt-valttaa> (4.5.2020)

<https://www.hs.fi/kotimaa/art-2000006455499.html> (4.5.2020)