



Big Data & Medicine

Sesiunea de Comunicari Stiintifice

Proiect realizat de:
Stroe Ruxandra Ioana
Tertes Raluca
Tascu Daniel Valentin
Timofte Serban Gabriel

MASURATORI CANTITATIVE IN ERA DIGITALA

Măsurătorile cantitative în era digitală se referă la procesul de colectare, analiză și interpretare a datelor numerice folosind instrumente digitale precum senzori, computere și software specializat.

01 Data Bases / Warehouse

02 Data Analytics

03 Statistics

04 Machine Learning



DATA ENGINEERING



To be completed



MACHINE LEARNING

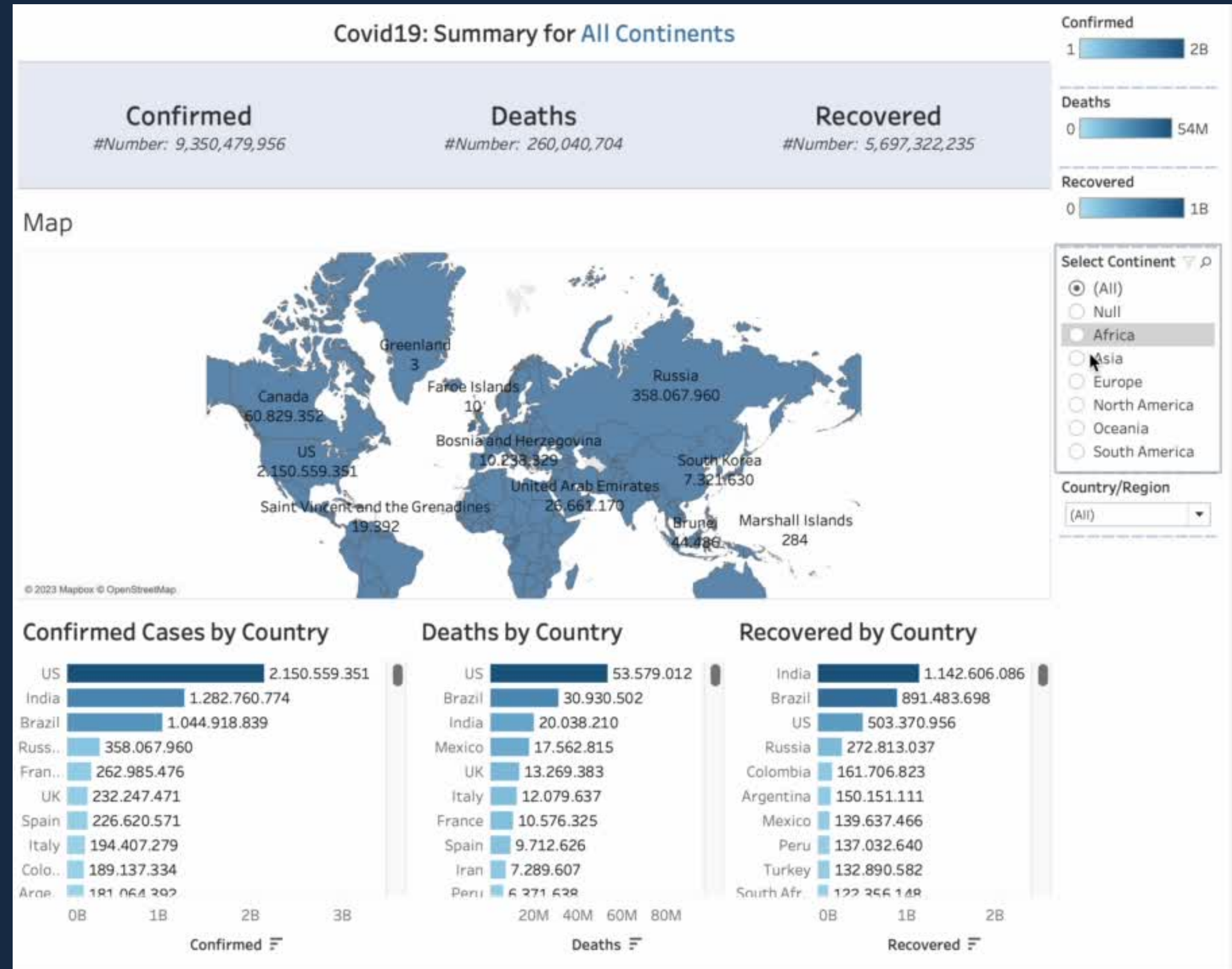


To be completed



Tableau Dashboard

Vizionarea datelor intr-o forma interactiva ajuta specialistii sa analizeze mai usor si mai rapid efectele pandemiei



Teste Statistice



```
# Varsta
# H0: oamenii care au murit sunt mai batrani
dead = subset(data, death_dummy == 1)
alive = subset(data, death_dummy == 0)
mean(dead$age, na.rm = TRUE)
mean(alive$age, na.rm = TRUE)
# Statistic significat?
z.test(alive$age, dead$age, alternative="two.sided", conf.level = 0.99)
# diferenta dintre o persoana care moare si una care traieste este in intervalul -25.52122 si -15.50661
# p-value is 2.2e-16 ~ 0 < 0.05, deci oamenii care au murit de Covid 19 sunt mai batrani
```

```
# Gen
# H0: genul nu are efect
men = subset(data, gender == "male")
women = subset(data, gender == "female")
mean(men$death_dummy, na.rm = TRUE) #8.5%!
mean(women$death_dummy, na.rm = TRUE) #3.7%
# Statistic significat?
z.test(men$death_dummy, women$death_dummy, alternative="two.sided", conf.level = 0.99)
# Barbatii au cu 0,8% pana la 8.8% sanse mai mari de mortalitate in cazul infectarii cu Covid19
# p-value = 0.002 < 0.05, semnificativ statistic
```



```

formatted_gdf = train.groupby(['Date', 'Country_Region'])['ConfirmedCases'].sum()
formatted_gdf = formatted_gdf.reset_index()
formatted_gdf['Date'] = pd.to_datetime(formatted_gdf['Date'])
formatted_gdf['Date'] = formatted_gdf['Date'].dt.strftime('%m/%d/%Y')
formatted_gdf['size'] = formatted_gdf['ConfirmedCases'].pow(0.3)

fig = px.scatter_geo(formatted_gdf, locations="Country_Region", locationmode='country names',
                    color="ConfirmedCases", size='size', hover_name="Country_Region",
                    range_color= [0, 1500],
                    projection="natural earth", animation_frame="Date",
                    title='CORONA: Spread Over Time From Jan 2020 to Apr 2020', color_continuous_scale="portland")
fig.show()

```

Spread Map using Jupyter Notebook



CE AM PUTEA INTELEGE?



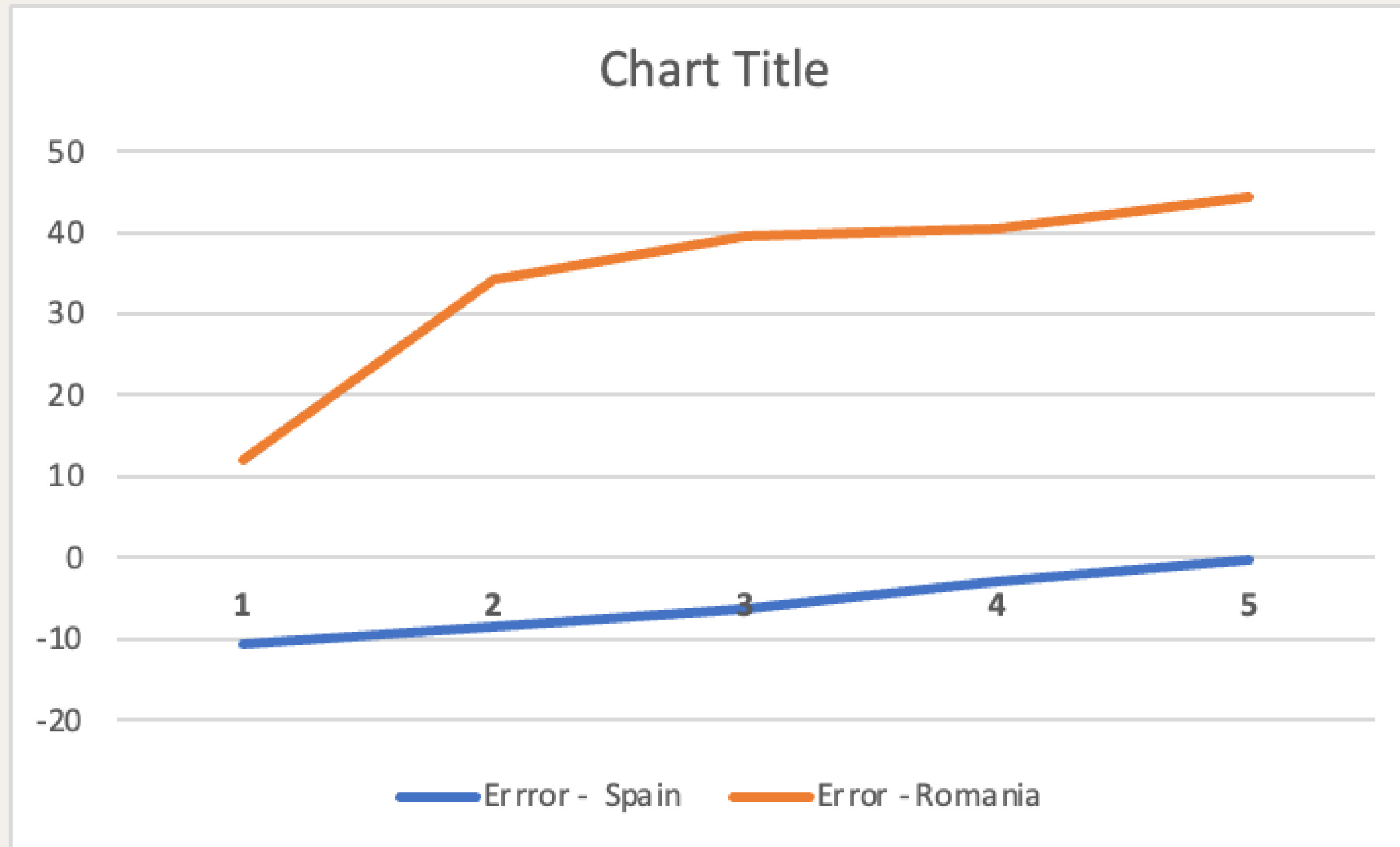
- 01** Viteza de raspandire este ridicata
- 02** Raspandirea este mai accelerata in zonele suprapopulate
- 03** **Virusul s-a raspandit rapid pentru ca noi ne miscam peste tot in lume rapid?**



5 Days Cases Prediction

Spain - Prediction	Spain - Reality	Deta	Error	Ro - Prediction	Ro - Reality	Delta	Error
188699	170537	18162	-10,64988829	2076	2360	-284	12,03389831
189501	174621	14880	-8,52131187	2097	3183	-1086	34,11875589
190303	179143	11160	-6,229660104	2121	3502	-1381	39,43460879
191104	185870	5234	-2,815946629	2148	3613	-1465	40,54802104
191906	191444	462	-0,241323834	2230	4010	-1780	44,38902743

Evolutia Diferentei



Efectul Vaccinului

Vaccin - Confirmari

Raport la 2 luni - May -> Jul 2021

Legatura negativa puternica cu
indicele $r = 61,23\%$

O modificare in randul persoanelor
vaccinate duce la scaderea cu 1,11
persoane confirmate

Vaccin - Decese

Raport la 2 luni - May -> Jul 2021

Legatura negativa ponderata cu
indicele $r = 46,9\%$

O modificare in randul persoanelor
vaccinate duce la scaderea cu 33,46
persoane confirmate

**La 1 persoana in plus vaccinata 33 de persoane nu mor
pentru ca se infecteaza cu o forma mai usoara sau
deloc ?**



Keep In Mind:

**Tehnologia de data analytics poate ajuta
în gestionarea crizei de Covid-19, dar
este important ca raportarea și
colectarea datelor să fie organizată și
transparentă pentru a fi eficientă în
identificarea soluțiilor rapide și eficiente.**