Dobrý den,

děkuji za projevený zájem a na základě Vámi zadaných kritérií jsem sestavila tabulku s daty, které Vám mohou sloužit jako podklad pro Vaši analýzu. Pro lepší pochopení tabulky přikládám bližší vysvětlení jednotlivých proměnných.

K dispozici máte tedy jednotlivé země (*country*), datum (*date*), přírůstek nemocných (*confirmed*) a:

1. Časové proměnné

* Víkend/pracovní den - „*week\_vs\_weekend*“ zde si lehce vyfiltrujete požadovanou proměnnou – week (dny v týdnu), weekend (SO, NE).
* Roční období daného dne - „*seasons*“ podobně i zde si můžete navolit jednotlivá roční období **0 = jaro** (21.3.-20.6.), **1 = léto** (21.6.-22.9.), **2 = podzim** (23.9.-20.12) a **3 = zima** (21.12. – 20.3).

1. Proměnné specifické pro daný stát

V rámci této skupiny máte k dispozici sloupce:

* Hustota zalidnění – „*population\_density*“
* Medián věku obyvatel ve věku 2018 – „*median\_age\_2018*“
* GINI koeficient – „*gin*i“
* dětská úmrtnost – „*mortality\_under5*“
* HDP na obyvatele – „*GDP\_per\_capita*“

V rámci těchto ukazatelů bych vás chtěla upozornit, že jsou vztažena k roku 2020. Bohužel se mi nepodařilo získat novější údaje. Na druhou stranu po analýze historických dat se žádný z těchto ukazatelů nemění skokově a pro Vaši analýzu vývoje covidu nejsou tyto nepatrné diference relevantní. Pokud tedy budete srovnávat vývoj nemoci v rámci různých zemí lze tyto proměnné využít. Pro srovnání vývoje přírůstků v rámci jednoho státu v čase je ale použít nelze.

* Rozdíl mezi očekávanou dobou dožití v roce 1965 a 2015 – „*life\_expect\_diff*“
* Podíly jednotlivých náboženství – „*MAX\_religion\_in\_perc*“ – číslo **je již v procentech**. Data pro výpočet této proměnné jsou vzhledem k nedostatku a nesrovnalosti dat vztažena k roku 2010. Nemyslím si, že by to výrazným způsobem zkreslilo Vaši analýzu. Opět na základě analýzy historických dat nedochází ke změně tohoto koeficientu skokově, vývoj je lineární a pro účely Vaší práce jsou tyto hodnoty dostačující.

1. Počasí

* Průměrná denní teplota – „*AVG\_day\_temp*“ – den je počítán od 6:00 – 18:00
* Počet hodin v daném dni, kdy byly srážky nenulové – „*rain\_hours*“
* Maximální síla větru v nárazech během dne – „*gust\_MAX*“

V rámci těchto proměnných jsou k dispozici pouze hodnoty za evropské státy a Rusko a posledním datumem je 2021-04-30. Tudíž pro srovnávání i jiných států a pozdějšího data nelze ve Vaší analýze tyto proměnné použít.