|  |  |
| --- | --- |
| **TECHNICKÁ UNIVERZITA V KOŠICIACH**  **FAKULTA ELEKTROTECHNIKY A INFORMATIKY** | |
| **SPRACOVANIE ÚDAJOV O COVID-19 V KANADE**  **Technológie a systémy spracovania údajov** | |
|  | |
| **2021** | **Adam Burík**  **Dávid Iľaš**  **Daniel Rosina** |

Zdroj údajov

Vhodné údaje na spracovanie ochorenia COVID-19 boli čerpané z: <https://github.com/ccodwg/Covid19Canada?fbclid=IwAR057vz2yp0NY3NaKM3Kqpf9SVQ9nxiviqMae-_ehHaTEAGlYb-kazIFEJQ>. Tieto údaje sú spracovávané dvomi univerzitami v Kanade a to University of Toronto a University of Guelph. Zakladateľmi tohto projektu sú Isha Berry (PhD Candidate, University of Toronto) a Jean-Paul R. Soucy (PhD Candidate, University of Toronto) a celý tím zaoberajúci sa zberom týchto údajov sa skladá z dobrovoľníkov s odbornosťou v oblastiach epidemiológie a verejného zdravia.

Táto skupina dobrovoľníkov zhromažďuje denné prírastky ohľadom COVID-19 prípadoch, úmrtiach, uzdraveniach, testovaní a očkovaní v rámci jednotlivých provincií Kanady. Údaje sa zhromažďujú výlučne z verejne dostupných zdrojov vrátane vládnych správ a spravodajských médií. Aktualizácie sa vykonávajú každú noc medzi 21:30 a 22:00 ET. Zber údajov je väčšinou automatizovaný.

Z githubu boli stiahnuté súbory typu *.csv* z priečinka *timeseries\_prov*, kde sa nachádzajú údaje ohľadom konkrétnych provincií. Pre náš účel boli využité len niektoré dokumenty a to konkrétne:

* **cases\_timeseries\_prov.csv** – obsahuje údaje denných prírastov ochorenia, celkový počet prípadov a to pre konkrétnu provinciu, údaje sú primárne zoradené podľa názvu provincie a sekundárne podľa dátumu, začiatok zaznamenávania je 25.01.2020,
* **mortality\_timeseries\_prov.csv** – obsahuje údaje denných prírastov úmrtí, celkový počet úmrtí a to pre konkrétnu provinciu, údaje sú primárne zoradené podľa názvu provincie a sekundárne podľa dátumu, začiatok zaznamenávania je 08.03.2020,
* **recovered\_timeseries\_prov.csv** – obsahuje údaje denných prírastov uzdravených, celkový počet uzdravených a to pre konkrétnu provinciu, údaje sú primárne zoradené podľa názvu provincie a sekundárne podľa dátumu, začiatok zaznamenávania je 12.02.2020,
* **vaccine\_completion\_timeseries\_prov.csv** – obsahuje údaje denných prírastov zaočkovaných, celkový počet zaočkovaných a to pre konkrétnu provinciu, údaje sú primárne zoradené podľa názvu provincie a sekundárne podľa dátumu, začiatok zaznamenávania je 12.01.2021.

Takisto boli stiahnuté aj súbory typu *.csv* z priečinka *timeseries\_canada*, kde sa nachádzajú údaje ohľadom celej Kanady a to konkrétne:

* **cases\_timeseries\_canada.csv** – obsahuje údaje denných prírastov ochorenia, celkový počet prípadov pre celú krajinu, údaje sú primárne zoradené podľa dátumu, začiatok zaznamenávania je 25.01.2020,
* **mortality\_timeseries\_canada.csv** – obsahuje údaje denných prírastov úmrtí, celkový počet úmrtí pre celú krajinu, údaje sú primárne zoradené podľa dátumu, začiatok zaznamenávania je 08.03.2020,
* **recovered\_timeseries\_canada.csv** – obsahuje údaje denných prírastov uzdravených, celkový počet uzdravených pre celú krajinu, údaje sú primárne zoradené podľa dátumu, začiatok zaznamenávania je 12.02.2020,
* **vaccine\_completion\_timeseries\_canada.csv** – obsahuje údaje denných prírastov zaočkovaných, celkový počet zaočkovaných pre celú krajinu, údaje sú primárne zoradené podľa dátumu, začiatok zaznamenávania je 12.01.2021.

Všetky súbory z údajmi o COVID-19 bolo potrebné čiastočne upraviť a to z dôvodu, že dátum v súboroch *.csv* bol napísaný v nekompatibilnom formáte s MYSQL. Dátum bol zapísaný v tvare DD-MM-YYYY a MYSQL pre dátový typ *date* umožňuje zapísať dátum v tvare YYYY-MM-DD. Zmena prebehla v MS EXCEL, kde bolo upravené rozloženie súboru a zápis dátumu.

Súbory určené pre Kanadu ako celok boli navyše upravené tak, aby všetky údaje boli uložené v jednej tabuľke. Čiže výsledkom spojenia 4 tabuliek (*cases\_timeseries\_canada.csv, mortality\_timeseries\_canada.csv, recovered\_timeseries\_canada.csv, vaccine\_completion\_timeseries\_canada.csv*) je tabuľka, ktorá má tieto stĺpce:

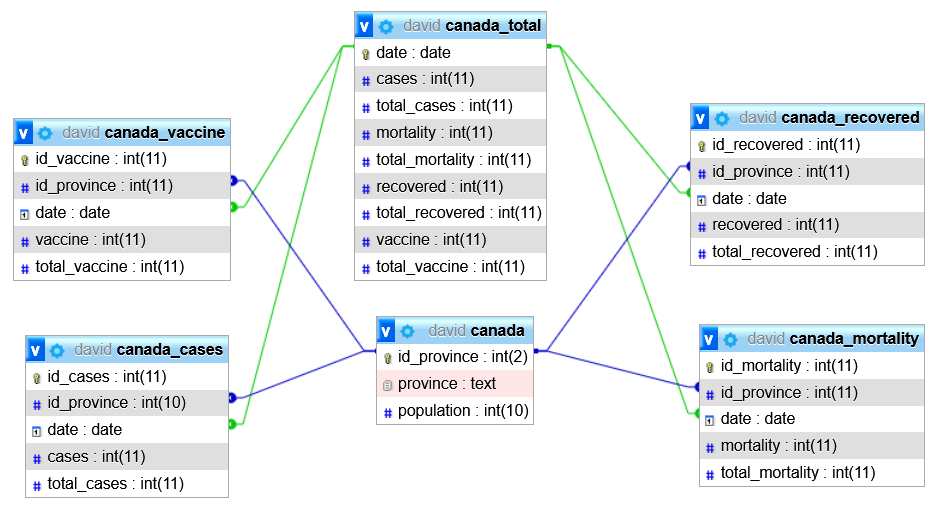
* dátum,
* denný prírastok nakazených,
* celkový počet nakazených,
* denný prírastok úmrtí,
* celkový počet úmrtí,
* denný prírastok uzdravených,
* celkový počet uzdravených,
* denný prírastok zaočkovaných,
* celkový počet zaočkovaných.

Návrh databázy

Na realizáciu databázy bol využitý SQL relačný databázový server MYSQL. Daná databáza slúži na uloženie údajov týkajúcich sa ochorenia COVID-19 v Kanade a v konkrétnych provinciách krajiny. Databáza sa skladá zo 6 tabuliek a to:

* **canada –** slúži na pridelenie ID k danej provincii,
  + **id\_province** – obsahuje označenie provincie, údajový typ *integer*, je primárnym kľúčom,
  + **province** – obsahuje názov provincie, údajový typ *text*,
  + **population** – obsahuje počet ľudí v daných provinciách, údajový typ *integer* ,
* **canada\_total** – údaje sú načítane zo súborov *cases\_timeseries\_canada.csv*, *mortality\_timeseries\_canada.csv*, *recovered\_timeseries\_canada.csv*, *vaccine\_completion\_timeseries\_canada.csv*,
  + **date** – obsahuje konkrétny dátum, údajový typ *date*, je primárnym kľúčom,
  + **cases** – obsahuje denný prírastok nakazených v krajine, údajový typ *integer*,
  + **total\_cases** – obsahuje celkový počet nakazených v krajine, údajový typ *integer*,
  + **mortality –** obsahuje denný prírastok úmrtí v krajine, údajový typ *integer*,
  + **total\_mortality –** obsahuje celkový počet úmrtí v krajine, údajový typ *integer*,
  + **recovered –** obsahuje denný prírastok uzdravených v krajine, údajový typ *integer*,
  + **total\_recovered –** obsahuje celkový počet uzdravených v krajine, údajový typ *integer*,
  + **vaccine –** obsahuje denný prírastok zaočkovaných v krajine, údajový typ *integer*,
  + **total\_vaccine –** obsahuje celkový počet zaočkovaných v krajine, údajový typ *integer*,
* **canada\_recovered** – údaje sú načítané zo súboru *recovered\_timeseries\_prov.csv,*
  + **id\_recovered** – slúži na určenie poradia daných údajov, údajový typ *integer*, automaticky sa inkrementuje, je primárnym kľúčom,
  + **id\_province** – obsahuje označenie provincie, údajový typ *integer*, je cudzím kľúčom pre tabuľku *canada*,
  + **date** – obsahuje konkrétny dátum, údajový typ *date*, je cudzím kľúčom pre tabuľku *canada\_total*,
  + **recovered** – obsahuje denný prírastok uzdravených, údajový typ *integer*,
  + **recovered\_total** – obsahuje počet uzdravených za celú dobu pozorovania, údajový typ *integer*,
* **canada\_cases** – údaje sú načítané zo súboru *cases\_timeseries\_prov.csv,*
  + **id\_cases** – slúži na určenie poradia daných údajov, údajový typ *integer*, automaticky sa inkrementuje, je primárnym kľúčom,
  + **id\_province** – obsahuje označenie provincie, údajový typ *integer*, je cudzím kľúčom pre tabuľku *canada*,
  + **date** – obsahuje konkrétny dátum, údajový typ *date*, je cudzím kľúčom pre tabuľku *canada\_total*,
  + **cases** – obsahuje denný prírastok nakazených, údajový typ *integer*,
  + **total\_cases** – obsahuje počet nakazených za celú dobu pozorovania, údajový typ *integer*,
* **canada\_mortality** – údaje sú načítané zo súboru *mortality\_timeseries\_prov.csv,*
  + **id\_mortality** – slúži na určenie poradia daných údajov, údajový typ *integer*, automaticky sa inkrementuje, je primárnym kľúčom,
  + **id\_province** – obsahuje označenie provincie, údajový typ *integer*, je cudzím kľúčom pre tabuľku *canada*,
  + **date** – obsahuje konkrétny dátum, údajový typ *date*, je cudzím kľúčom pre tabuľku *canada\_total*,
  + **mortality** – obsahuje denný prírastok úmrtí, údajový typ *integer*,
  + **total\_mortality** – obsahuje počet úmrtí za celú dobu pozorovania, údajový typ *integer*,
* **canada\_vaccination** – údaje sú načítané zo súboru *vaccine\_competion\_timeseries\_prov.csv,*
  + **id\_vaccine** – slúži na určenie poradia daných údajov, údajový typ *integer*, automaticky sa inkrementuje, je primárnym kľúčom,
  + **id\_province** – obsahuje označenie provincie, údajový typ *integer*, je cudzím kľúčom pre tabuľku *canada*,
  + **date** – obsahuje konkrétny dátum, údajový typ *date*, je cudzím kľúčom pre tabuľku *canada\_total*,
  + **vaccine** – obsahuje denný prírastok zaočkovaných, údajový typ *integer*,
  + **total\_vaccine** – obsahuje počet zaočkovaných za celú dobu pozorovania, údajový typ *integer*.

V každej tabuľke okrem tabuľky *canada\_total* sú údaje zo všetkých 13 provincií. Tabuľka canada\_total obsahuje údaje všetkých 13 provincií spolu. Na znázornenie databázy bol využitý relačný diagram databázy, ktorý je možné vidieť na Obr. 1.



Obr. 1 Relačný diagram databázy

Vytvorenie databázy a vloženie údajov

Za účelom vytvorenia databázy bola využitá voľne dostupná aplikácia XAMPP. Pomocou tejto aplikácie je možné spustiť MYSQL databázu lokálne na PC a spravovať ju pomocou služby phpmyadmin. Prístup na phpmyadmin je možný pripojením sa na tento link: <http://localhost/phpmyadmin>.

Príkazy na vytvorenie databázy a následne tabuliek boli vkladané v phpmyadmin, keďže je to jednoduché a intuitívne.

**Príkaz na vytvorenie databázy**

CREATE DATABASE CANADA\_COVID19;

**Príkazy na vytvorenie tabuliek**

CREATE TABLE canada (

id\_province int PRIMARY KEY,

province text,

population int );

CREATE TABLE canada\_cases (

id\_cases int AUTO\_INCREMENT PRIMARY KEY,

id\_province int,

date date,

cases int,

total\_cases int );

CREATE TABLE canada\_mortality (

id\_mortality int AUTO\_INCREMENT PRIMARY KEY,

id\_province int,

date date,

mortality int,

total\_mortality int );

CREATE TABLE canada\_vaccine (

id\_vaccine int AUTO\_INCREMENT PRIMARY KEY,

id\_province int,

date date,

vaccine int,

total\_vaccine int );

CREATE TABLE canada\_recovered (

id\_recovered int AUTO\_INCREMENT PRIMARY KEY,

id\_province int,

date date,

recovered int,

total\_recovered int );

CREATE TABLE canada\_total (

date date PRIMARY KEY,

cases int,

total\_cases int,

mortality int,

total\_mortality int,

recovered int,

total\_recovered int,

vaccine int,

total\_vaccine int );

**Pridanie cudzích kľúčov**

ALTER TABLE canada\_recovered ADD FOREIGN KEY (id\_province) REFERENCES canada (id\_province);

ALTER TABLE canada\_mortality ADD FOREIGN KEY (id\_province) REFERENCES canada (id\_province);

ALTER TABLE canada\_cases ADD FOREIGN KEY (id\_province) REFERENCES canada (id\_province);

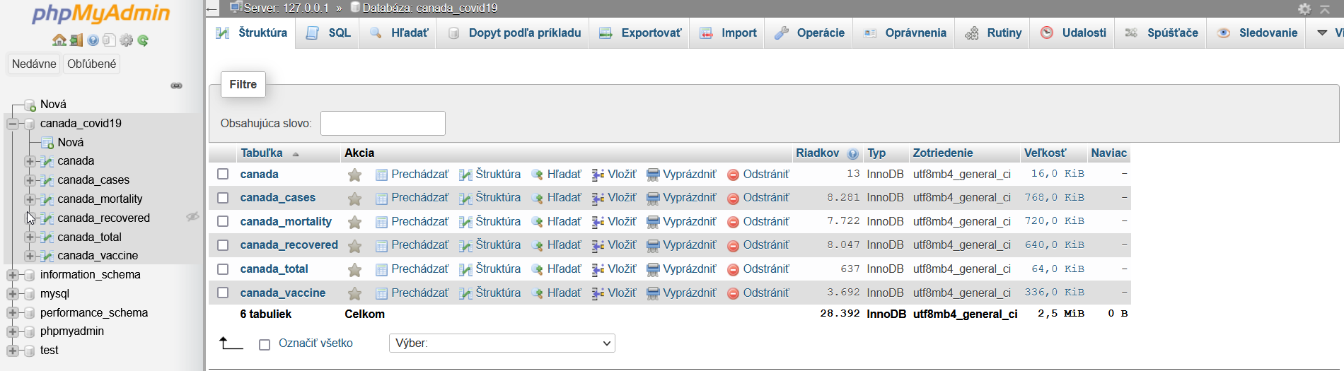
ALTER TABLE canada\_vaccine ADD FOREIGN KEY (id\_province) REFERENCES canada (id\_province);

ALTER TABLE canada\_recovered ADD FOREIGN KEY (date) REFERENCES canada\_total (date);

ALTER TABLE canada\_mortality ADD FOREIGN KEY (date) REFERENCES canada\_total (date);

ALTER TABLE canada\_cases ADD FOREIGN KEY (date) REFERENCES canada\_total (date);

ALTER TABLE canada\_vaccine ADD FOREIGN KEY (date) REFERENCES canada\_total (date);



Obr. 2 Úvodná strana phpmyadmin s tabuľkami

Vloženie údajov bolo realizované funkciou import v phpmyadmin a to sa realizovalo pre každú tabuľku, čiže pre každú tabuľku bol importovaný konkrétny *.csv* súbor.

Spracovanie databáz

Načítanie všetkých databáz bolo realizované v phpmyadmin a to pomocou funkcie import, kde boli nahrané .sql súbory všetkých databáz. Načítanie niektorých databáz bolo problematické alebo nemožné, čiže výber databáz na ďalšie spracovanie bol obmedzený. Po dôkladnej analýze databáz bol výber nasledovný:

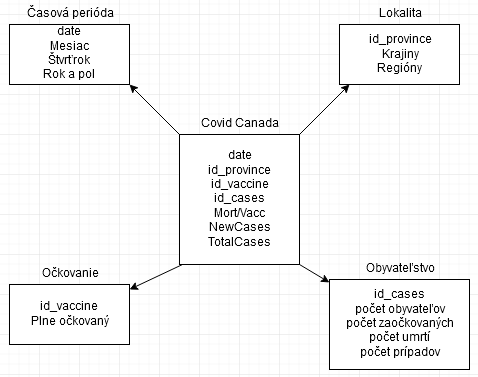
* Slovensko (Fehér, Lechner, Švenk),
* Rakúsko (Kupcová, Maté, Ujházi),
* USA (Bereczová, Hupka, Pribula).

V rámci spracovania týchto databáz budú analyzované nasledujúce tri merateľné fakty:

* Porovnanie úmrtnosti vzhľadom na očkovanie v jednotlivých krajinách v marci a septembri 2021,
* vývoj nových prípadov ochorenia v najľudnatejších regiónoch krajín v zime 2020/2021,
* vývoja celkových prípadov ochorenia v jednotlivých krajinách za rok a pol (od apríla 2020 do septembra 2021).

|  |  |  |  |
| --- | --- | --- | --- |
|  | Information subject: Covid-19 | |  |
|  |  |  |  |
| Časová perióda | Lokalita | Očkovanie | Obyvateľstvo |
| Mesiac | Krajiny | Plne zaočkovaný | počet obyvateľov |
| Štvrťrok | Regióny |  | počet zaočkovaných |
| Rok a Pol |  |  | počet úmrtí |
|  |  |  | počet prípadov |
| Fakty: porovnanie úmrtnosti vzhľadom na očkovanie, vývoj nových prípadov ochorenia, vývoj celkových prípadov ochorenia | | | |
|  |
|  |

Tab. 1 Informačný balíček



Obr. Star schéma

V strednej tabuľke „Covid Canada“ skratky predstavujú:

* Mort/Vacc – 1. merateľný fakt, porovnanie úmrtnosti vzhľadom na očkovanie v jednotlivých krajinách v marci a septembri 2021,
* NewCases – 2. merateľný fakt, vývoj nových prípadov ochorenia v najľudnatejších regiónoch krajín v zime 2020/2021,
* TotalCases – 3. merateľný fakt, vývoja celkových prípadov ochorenia v jednotlivých krajinách za rok a pol (od apríla 2020 do septembra 2021).

1. Porovnanie úmrtnosti v krajine vzhľadom na očkovanie v marci a septembri 2021

Tento merateľný fakt skúma počet úmrtí v jednotlivých krajinách za marec 2021 a september 2021. Úmrtnosť je vyjadrená v prepočte na milión obyvateľov, keďže pri porovnaní úmrtí napríklad medzi USA a Slovenskom, USA má približne 328 miliónov obyvateľov a Slovensko približne 5,5 milióna obyvateľov, je veľmi veľký rozdiel v počtoch úmrtí. Preto boli tieto hodnoty prepočítané na milión obyvateľov, aby boli fakty porovnateľné. Ďalším faktom, ktorý bol meraný, je percentuálna zaočkovanosť obyvateľstva jednotlivých krajín a to pre porovnanie, či zmeny v percentuálnej zaočkovanosti ovplyvnia počty úmrtí.

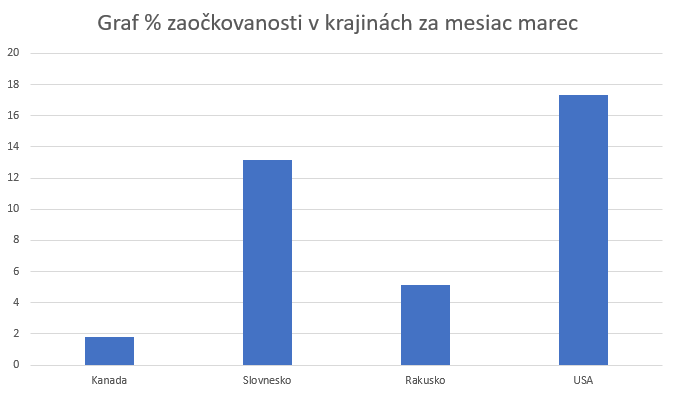
Dopyty pre úmrtnosť a zaočkovanosť pre marec a september:

**Marec:**

SELECT ((SELECT `total\_mortality` FROM `canada`.`canada\_total` WHERE date='2021-03-31') - (SELECT `total\_mortality` FROM `canada`.`canada\_total` WHERE date='2021-03-01'))\*(1000000/(SELECT SUM(`population`) FROM `canada`.`canada`)) AS CanMortality, ((SELECT `total\_vaccine` FROM `canada`.`canada\_total` WHERE date='2021-03-31')/(SELECT SUM(`population`) FROM `canada`.`canada`))\*100 AS CanPomer, ((SELECT `total\_deaths` FROM `slovensko`.`pripady` WHERE date='2021-03-31') - (SELECT `total\_deaths` FROM `slovensko`.`pripady` WHERE date='2021-03-01'))\*(1000000/(SELECT `Population` FROM `slovensko`.`slovensko`)) AS SVKMortality, ((SELECT `people\_vaccinated` FROM `slovensko`.`vakcinacia` WHERE date='2021-03-31')/(SELECT `Population` FROM `slovensko`.`slovensko`))\*100 AS SVKPomer, ((SELECT `cases`.`total\_deaths` FROM `austria`.`fact\_country` INNER JOIN `austria`.`cases` ON `fact\_country`.`cases\_id`=`cases`.`id` WHERE `fact\_country`.`date`='2021-03-31')-(SELECT `cases`.`total\_deaths` FROM `austria`.`fact\_country` INNER JOIN `austria`.`cases` ON `fact\_country`.`cases\_id`=`cases`.`id` WHERE `fact\_country`.`date`='2021-03-01'))\*(1000000/(SELECT `population` FROM `austria`.`country`)) AS AuMortality, ((SELECT `vaccinations`.`people\_fully\_vaccinated` FROM `austria`.`fact\_country` INNER JOIN `austria`.`vaccinations` ON `fact\_country`.`vaccinations\_id`=`vaccinations`.`id` WHERE `fact\_country`.`date`='2021-03-31')/(SELECT `population` FROM `austria`.`country`))\*100 AS AuPomer, ((SELECT `Pocet\_umrti` FROM `usa`.`usa\_covid19\_pripady` WHERE Datum='2021-03-31')-(SELECT `Pocet\_umrti` FROM `usa`.`usa\_covid19\_pripady` WHERE Datum='2021-03-01'))\*(1000000/(SELECT SUM(`Pocet\_Obyvatelov`) FROM `usa`.`usa\_staty`)) AS UsaMortality, ((SELECT SUM(`Pocet\_Plne\_Zaockovanych`) FROM `usa`.`usa\_covid19\_zaockovanost` WHERE `Datum`='2021-03-31' AND `Nazov\_Statu`!='United States')/(SELECT SUM(`Pocet\_Obyvatelov`) FROM `usa`.`usa\_staty`))\*100 AS UsaPomer;

**September:**

SELECT ((SELECT `total\_mortality` FROM `canada`.`canada\_total` WHERE date='2021-09-30') - (SELECT `total\_mortality` FROM `canada`.`canada\_total` WHERE date='2021-09-01'))\*(1000000/(SELECT SUM(`population`) FROM `canada`.`canada`)) AS CanMortality, ((SELECT `total\_vaccine` FROM `canada`.`canada\_total` WHERE date='2021-09-30')/(SELECT SUM(`population`) FROM `canada`.`canada`))\*100 AS CanPomer, ((SELECT `total\_deaths` FROM `slovensko`.`pripady` WHERE date='2021-09-30') - (SELECT `total\_deaths` FROM `slovensko`.`pripady` WHERE date='2021-09-01'))\*(1000000/(SELECT `Population` FROM `slovensko`.`slovensko`)) AS SVKMortality, ((SELECT `people\_vaccinated` FROM `slovensko`.`vakcinacia` WHERE date='2021-09-30')/(SELECT `Population` FROM `slovensko`.`slovensko`))\*100 AS SVKPomer, ((SELECT `cases`.`total\_deaths` FROM `austria`.`fact\_country` INNER JOIN `austria`.`cases` ON `fact\_country`.`cases\_id`=`cases`.`id` WHERE `fact\_country`.`date`='2021-09-30')-(SELECT `cases`.`total\_deaths` FROM `austria`.`fact\_country` INNER JOIN `austria`.`cases` ON `fact\_country`.`cases\_id`=`cases`.`id` WHERE `fact\_country`.`date`='2021-09-01'))\*(1000000/(SELECT `population` FROM `austria`.`country`)) AS AuMortality, ((SELECT `vaccinations`.`people\_fully\_vaccinated` FROM `austria`.`fact\_country` INNER JOIN `austria`.`vaccinations` ON `fact\_country`.`vaccinations\_id`=`vaccinations`.`id` WHERE `fact\_country`.`date`='2021-09-30')/(SELECT `population` FROM `austria`.`country`))\*100 AS AuPomer, ((SELECT `Pocet\_umrti` FROM `usa`.`usa\_covid19\_pripady` WHERE Datum='2021-09-30')-(SELECT `Pocet\_umrti` FROM `usa`.`usa\_covid19\_pripady` WHERE Datum='2021-09-01'))\*(1000000/(SELECT SUM(`Pocet\_Obyvatelov`) FROM `usa`.`usa\_staty`)) AS UsaMortality, ((SELECT SUM(`Pocet\_Plne\_Zaockovanych`) FROM `usa`.`usa\_covid19\_zaockovanost` WHERE `Datum`='2021-09-30' AND `Nazov\_Statu`!='United States')/(SELECT SUM(`Pocet\_Obyvatelov`) FROM `usa`.`usa\_staty`))\*100 AS UsaPomer;

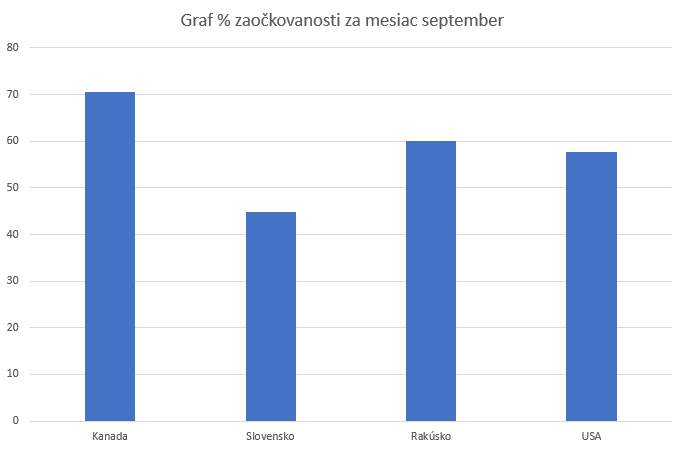


Obr. 4 Graf percentuálnej zaočkovanosti za marec 2021



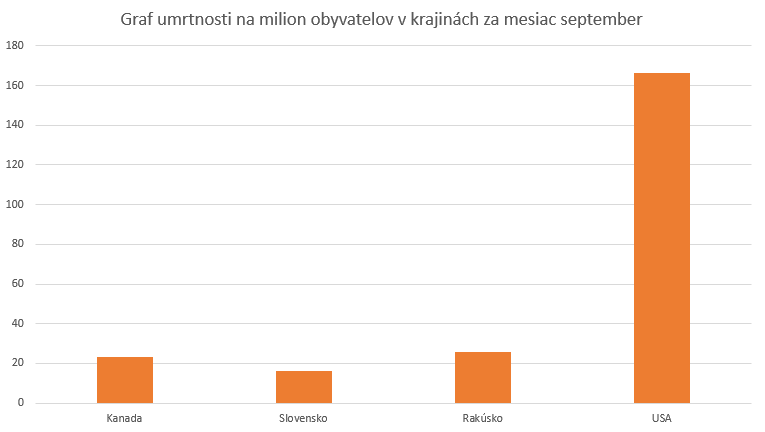
Obr. 5 Graf úmrtnosti za marec 2021

Na Obr. 4 je možné vidieť percentuálnu zaočkovanosť v marci 2021. V tomto období sa očkovanie ešte len rozbiehalo a tomu svedčí aj graf, kde je vidieť, že najvyššia hodnota zaočkovanosti populácie dosiahla len necelých 18% a to pre USA. Na druhej pozícii je Slovensko a to s hodnotou približne 13%, na tretej pozícii je Rakúsko s priližne 5% zaočkovanosti a na poslednom mieste je Kanada s necelými 2%. Z tohto je možné usúdiť, že Slovensko a USA mali lepší štart v očkovaní oproti Rakúsku a Kanade. Pomocou Obr. 5 je možné zistiť určité súvislosti medzi očkovaním a úmrtnosťou. Keďže bolo očkovanie ešte tzv. v plienkach, tak nie je vidieť, aby vyššia percentuálna zaočkovanosť prispela k nižším počtom úmrtí. Napríklad, Slovensko malo ako druhé najvyššiu hodnotu zaočkovanosti národa a to približne 13%, no počet úmrtí na milión obyvateľov je najhoršia spomedzi porovnávaných krajín. Prekvapivo Kanada, ktorej zaočkovanosť bola najnižšia z porovnávania, konkrétne necelé 2%, mala aj najnižší počet úmrtí prepočítaných na milión obyvateľov.



Obr. 6 Graf percentuálnej zaočkovanosti za september 2021

Veľké zmeny nastali o pol roka neskôr. V septembri 2021 sa zaočkovanosť krajín rázne zvýšila, čo je vidieť na Obr. 6. Prekvapením bola Kanada, pre ktorú sa zaočkovanosť oproti marcu zvýšila o necelých 70%, čo je najväčším skokom oproti ostatným krajinám. Naopak, zaočkovanosť na Slovensku a v USA vzrástla len o približne 40%. V Rakúsku sa zaočkovanosť zvýšila o prilbížne 55%.



Obr. 7 Graf úmrtnosti za september 2021

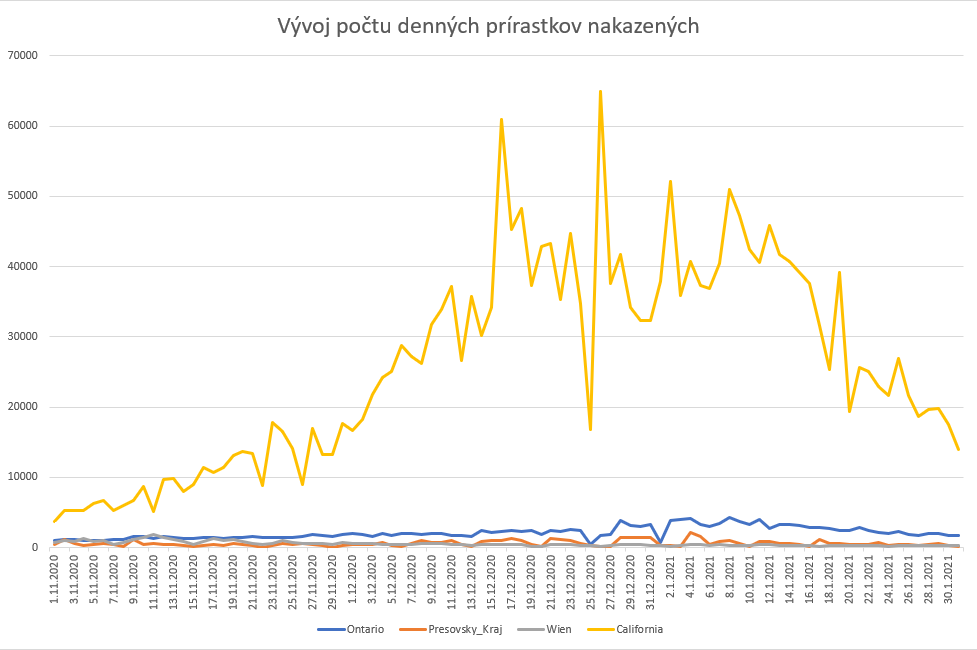
Hlavnou zmenou však bol počet úmrtí v jednotlivých krajinách. Ako prvé, pre Kanadu sa nezmenilo takmer nič. Pri zaočkovanosti necelé 2% a 70% bol počet úmrtí na milión takmer rovnaký a to približne 24. Pozitívnu zmenu zaznamenalo Slovensko a Rakúsko. Na Slovensku sa počet úmrtí na milión rapídne znížil z hodnoty približne 448,5 na 16,3 a v Rakúsku sa počet úmrtí na milión znížil z približne 84,5 na 26. Týmito faktami je možné usúdiť, že očkovanie má pozitívny vplyv na úmrtnosť na Covid-19, keďže po zvýšení percentuálnej zaočkovanosti sa počet úmrtí pomerne dosť znížil. Výnimkou je USA, kde sa úmrtnosť naopak zvýšila zo 112 na 166,5. To znamená, že samotné očkovanie nie je rozhodujúcim faktorom pri úmrtnosti, dôležitými faktormi sú aj napríklad podmienky a opatrenia v krajine.

2. Vývoj nových prípadov ochorenia v najľudnatejšom regióne krajiny v zime 2020/2021 (november, december, január)

Tento merateľný fakt skúma vývoj počtu denných prírastkov nakazenia na Covid-19 počas zimy 2020, konkrétne počas mesiacov november, december a január. Na získanie faktov boli využité počty nových prípadov ochorenia v určitých regiónoch krajín. Vybrané boli najľudnatejšie regióny jednotlivých krajín, konkrétne Ontario z Kanady, Prešovský Kraj zo Slovenska, Wien z Rakúska a California z USA. Keďže nešlo o porovnanie vývoja nových prípadov medzi regiónmi, počet nových prípadov sa neprepočítaval na milión obyvateľov. Zámerom tohto merateľného faktu je nájsť určitú koreláciu medzi počtom nových prípadov a zimným obdobím, konkrétne obrobím Vianočných sviatkov. Keďže sa ľudia počas Vianočných sviatkov navštevujú viac ako bežne, je možné očakávať vyšší počet nových prípadov ochorenia v tomto období.

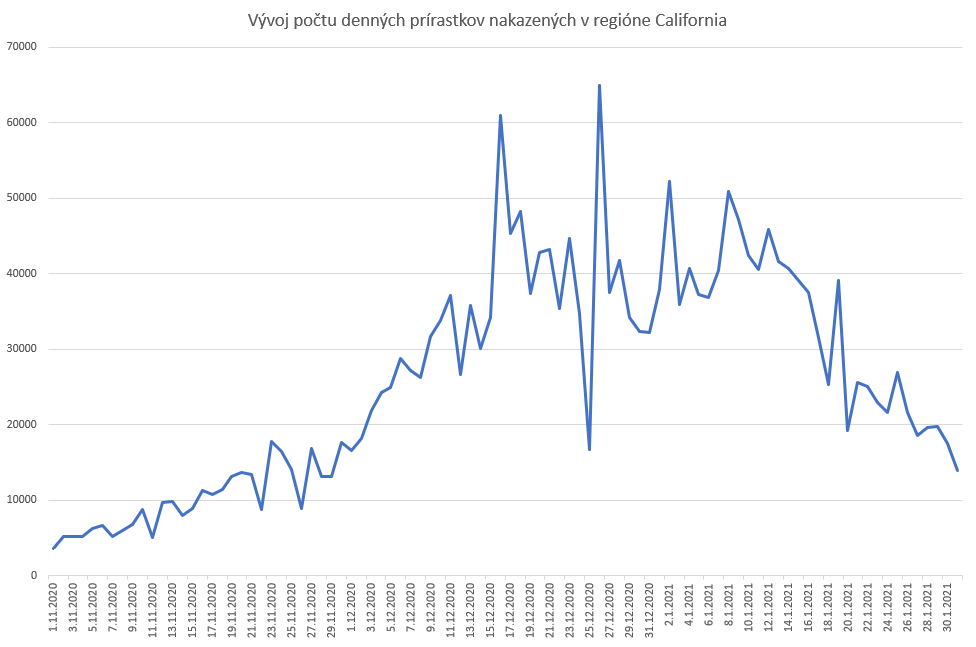
Dopyt pre počet nových prípadov ochorenia počas zimy 2020/2021:

SELECT @dat:=`canada`.`canada\_cases`.`date` AS date, `canada`.`canada\_cases`.`cases` AS Ontario, `slovensko`.`regiony\_data`.`newcases` AS Presovsky\_Kraj, `austria`.`cases`.`new\_cases` AS Wien, ((SELECT `Pocet\_Pripadov` FROM `usa`.`usa\_staty\_pocet\_pripadov` WHERE Datum=@dat AND `Nazov\_statu`='California')-(SELECT `Pocet\_Pripadov` FROM `usa`.`usa\_staty\_pocet\_pripadov` WHERE Datum=date\_add(@dat, interval -1 day) AND `Nazov\_statu`='California')) AS California FROM `canada`.`canada\_cases` INNER JOIN `slovensko`.`regiony\_data` ON `canada`.`canada\_cases`.`date`=`slovensko`.`regiony\_data`.`date` INNER JOIN (`austria`.`fact\_region` INNER JOIN `austria`.`cases` ON `austria`.`fact\_region`.`cases\_id`=`austria`.`cases`.`id` INNER JOIN `austria`.`region` ON `austria`.`fact\_region`.`region\_id`=`austria`.`region`.`id`) ON `canada`.`canada\_cases`.`date`=`austria`.`fact\_region`.`date` INNER JOIN `usa`.`usa\_staty\_pocet\_pripadov` ON `canada`.`canada\_cases`.`date`=`usa`.`usa\_staty\_pocet\_pripadov`.`Datum` WHERE `canada`.`canada\_cases`.`date` BETWEEN '2020-11-01' AND '2021-01-31' AND `canada`.`canada\_cases`.`id\_province`='9' AND `slovensko`.`regiony\_data`.`Region`='Prešovský kraj' AND `austria`.`fact\_region`.`region\_id`='9' AND `usa`.`usa\_staty\_pocet\_pripadov`.`Nazov\_statu`='California';



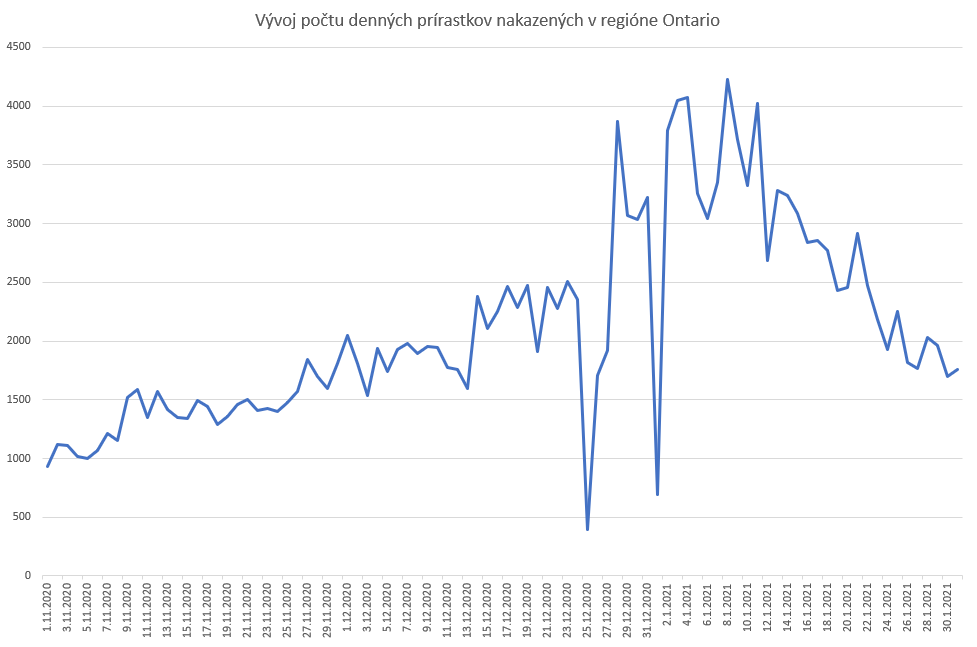
Obr. 8 Graf vývoja nových prípadov ochorenia v regiónoch

Na Obr. 8 je zobrazený čiarový graf pre hodnoty nových prípadov ochorenia všetkých 4 regiónov. Keďže hodnoty neboli prepočítané na milión obyvateľov je veľký skos medzi Californiou a ostatnými regiónmi. Nás však zaujíma vývoj nových prípadov v jednotlivých regiónoch s cieľom objaviť koreláciu medzi počtom denných prírastov ochorenia a Vianočných sviatkov. Už na tomto grafe je vidieť, že v okolí Vianoc sú hodnoty prípadov vyššie.



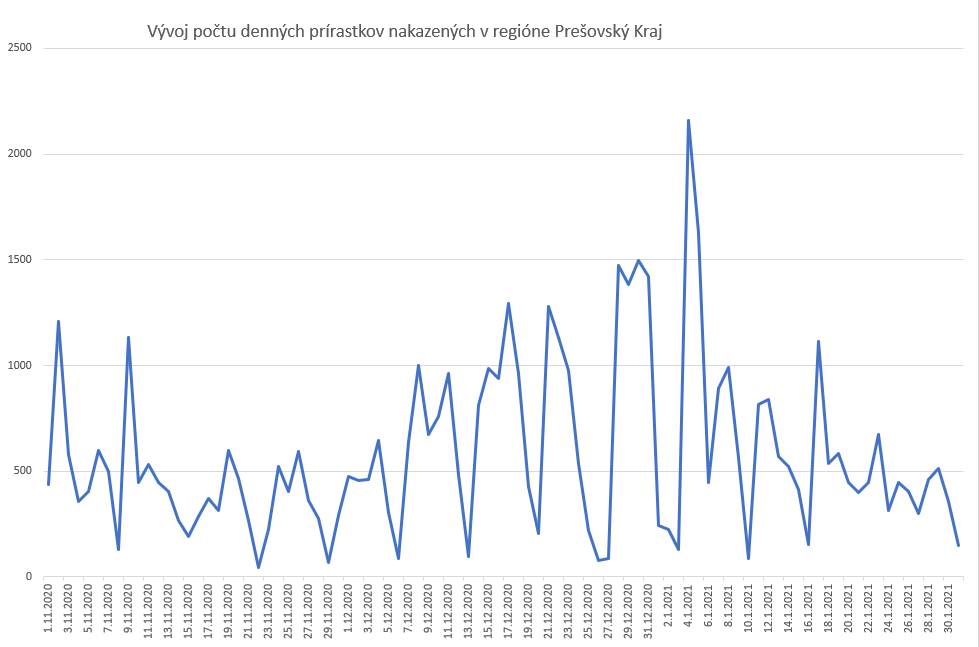
Obr. 9 Graf vývoja nových prípadov ochorenia v regióne California (USA)

V Californii sa graf vyvíja podľa očakávaní, keďže počas decembra boli hodnoty denných prírastkov približne 6-násobné oproti novembru. Počty nových prípadov, podľa grafu, narastali rapídne. Najvyššia hodnota bola zaznamenaná 26.12. a to 64986 nových prípadov. Takisto je vidieť, že po sviatkoch, čiže v januári, hodnoty opäť klesajú.



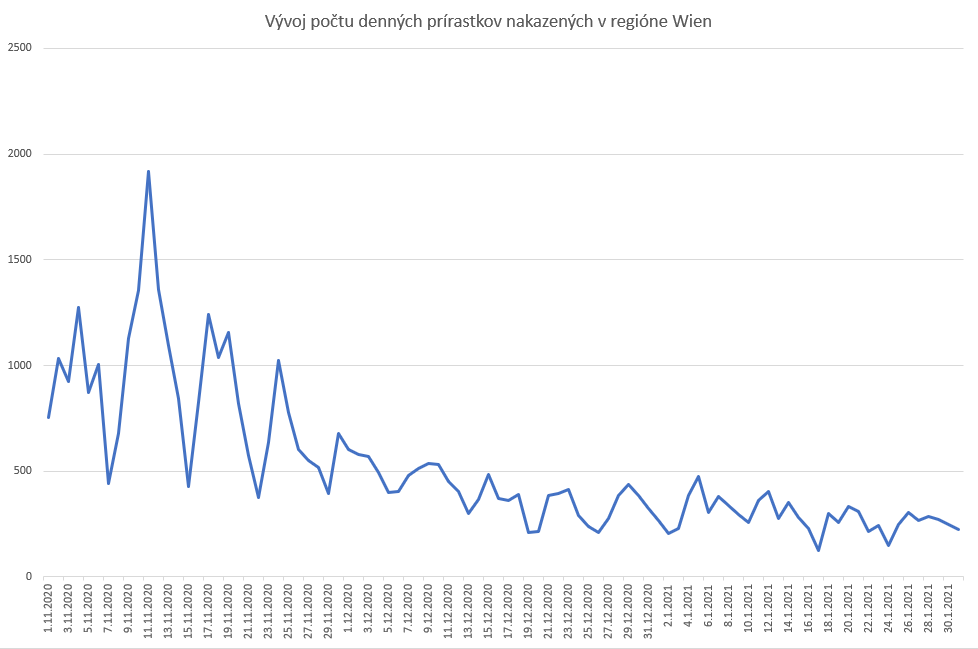
Obr. 10 Graf vývoja nových prípadov ochorenia v regióne Ontario (Kanada)

V Ontariu sa graf vyvíja taktiež podľa očakávaní. Graf však narastá rovnomerne a vrchol dosiahne až 8.1., kde počet nových prípadov dosahuje 4227. Strmý nárast je zaznamenaný už 27.12. Podobne ako v predchádzajúcom grafe sa počet nových prípadov po sviatkoch znižuje.



Obr. 11 Graf vývoja nových prípadov ochorenia v regióne Prešovský Kraj (Slovensko)

V Prešovskom kraji sa graf vyvíja podobne ako predošlé, no počet nových prípadov je na začiatku novembra vyšší a klesá ku koncu mesiaca. Začiatkom decembra však hodnoty začínajú rásť a vrchol dosiahli 4.1. a to konkrétne 2160. Počet nových prípadov ochorenia očakávane klesá po Vianočných sviatkoch.



Obr. 12 Graf vývoja nových prípadov ochorenia v regióne Wien (Rakúsko)

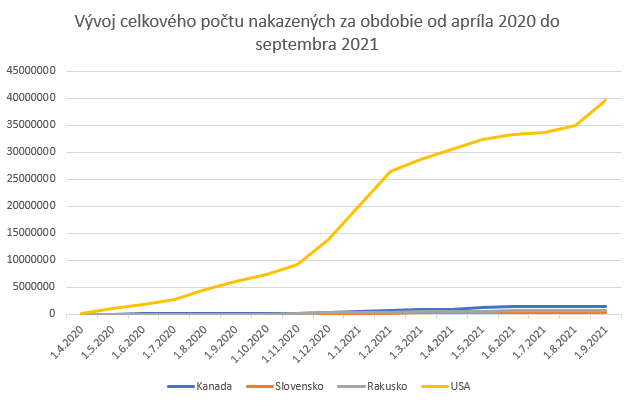
Graf na Obr. 12 nezobrazuje krivku podobnú ostatným grafom. V regióne Wien je vrchol počtu nových prípadov už 11.11., konkrétne 1919 a počas decembra a januára pomaly klesá.

Z týchto grafov je možné usúdiť, že Vianočné sviatky naozaj ovplyvňujú vývoj nových prípadov ochorenia. Ľudia sa viac stretávajú, čo vytvára viac možností nakazenia a šírenia ochorenia Covid-19. Jedinou výnimkou je región Wien v Rakúsku, kde sa vývoj nových prípadov vôbec nepodobal s ostatnými. Regióny Ontario, Prešovský Kraj a California sa v rámci počtu nových prípadov približne zhodovali. V grafoch týchto regiónov sa počty nových prípadov začali zvyšovať už na začiatku decembra, vyvrcholili počas 2 týždňov od Vianoc (24.12.) a následne sa začali znižovať k prijateľnejším hodnotám.

3. Vývoj celkových prípadov ochorenia v krajine za obdobie od apríla 2020 do septembra 2021 (rok a pol)

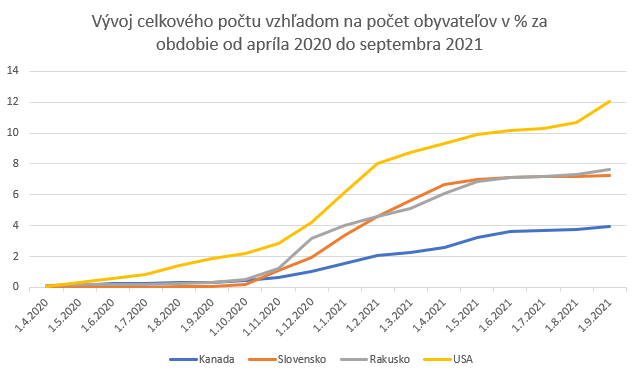
Tento merateľný fakt skúma vývoj celkových prípadov ochorenia v krajinách za rok a pol, od apríla 2020 do septembra 2021. Na získanie informácií boli využité počty celkových prípadov a prepočet celkových prípadov na počet obyvateľov vyjadrených v percentách z dôvodu reálnejšieho porovnania. Počet celkových prípadov ochorenia bol získavaný len z 1. dňa každého mesiaca. Cieľom tohto merania je porovnať vývoj celkových prípadov ochorenia počas dlhej periódy, takmer celého trvania ochorenia.

SELECT `canada`.`canada\_total`.`date` AS date,@canc:=`canada`.`canada\_total`.`total\_cases` AS CanadaCases, @canc/(SELECT SUM(`population`) FROM `canada`.`canada`)\*100 as CanadaPomer, @svkc:=`slovensko`.`pripady`.`total\_cases` AS SvkCases, @svkc/(SELECT `population` FROM `slovensko`.`slovensko`)\*100 AS SvkPomer, @rakc:=`austria`.`cases`.`total\_cases` AS RakuskoCases, @rakc/(SELECT `population` FROM `austria`.`country`)\*100 AS RakuskoPomer, @usac:=`usa`.`usa\_covid19\_pripady`.`Pocet\_Pripadov` AS UsaCases, @usac/(SELECT SUM(`Pocet\_Obyvatelov`) FROM `usa`.`usa\_staty`)\*100 AS UsaPomer FROM `canada`.`canada\_total` INNER JOIN `slovensko`.`pripady` ON `canada`.`canada\_total`.`date`=`slovensko`.`pripady`.`date` INNER JOIN (`austria`.`fact\_country` INNER JOIN `austria`.`cases` ON `fact\_country`.`cases\_id`=`cases`.`id`) ON `canada`.`canada\_total`.`date`=`austria`.`fact\_country`.`date` INNER JOIN `usa`.`usa\_covid19\_pripady` ON `canada`.`canada\_total`.`date`=`usa`.`usa\_covid19\_pripady`.`Datum` WHERE `canada`.`canada\_total`.`date` BETWEEN '2020-04-01' AND '2021-09-31' AND `canada`.`canada\_total`.`date`=DATE\_FORMAT(DATE(canada.canada\_total.date), '%Y-%m-01');



Obr. 13 Vývoj celkového počtu nakazených

Na Obr. 13 je možné vidieť graf zobrazujúci vývoj celkových prípadov nakazených v Kanade, Slovensku, Rakúsku a USA. Z tohto grafu je ťažké porovnať vývoj v jednotlivých krajinách, keďže hodnoty nakazených v USA náramne prevyšujú hodnoty v ostatných krajinách.



Obr. 14 Vývoj celkového počtu nakazených vyjadrený percentuálne

Pre lepšie porovnanie boli všetky hodnoty prepočítane na počet obyvateľov a vyjadrené v percentách, čiže sa jedná o percentuálne premorenie obyvateľstva. Z Obr. 14 je možné vyčítať, že prípady ochorenia sa najhoršie vyvíjali v USA, kde sa prípady za rok a pol zvýšili z 0% na približne 12%. Slovensko a Rakúsko sa vyvíjalo dosť podobne, kde sa prípady za rok a pol vyšplhali na 7-8%. Najlepší vývoj zaznamenala Kanada, ktorá sa za rok a pol dostala len na približne 4%. Z grafu je možné určiť, že najväčší nárast prípadov sa stal počas obdobia od októbra 2020 do mája 2021. Po tomto období, čiže počas letných mesiacov sa počet prípadov ustálil a opätovný nárast začal septembrom 2021. Z týchto informácií je možné usúdiť, že aj ročné obdobia majú vplyv na počet prípadov ochorenia, keďže počas teplých mesiacov sa Covid-19 rozširoval pomalšie a počas chladných mesiac rýchlejšie.