

Univerzita Pardubice
Fakulta elektrotechniky a informatiky

Zpracování dat pro předmět NMAST

Bc. Lukáš Milar, Bc. Tomáš Prudký

Semestrální práce

2021

OBSAH

Seznam obrázků	4
Seznam tabulek	6
Úvod	7
1 Popis dat	8
2 Popisná statistika	10
3 Základní grafy	12
3.1 Histogram	12
3.2 Bodový graf	16
3.3 Boxplot	21
3.4 3D graf	23
3.5 Hexbin	25
3.6 Chernoff faces	26
3.7 QQPlot	29
4 Testování statistických hypotéz	31
4.1 Jednovýběrový Studentův test vůči střední hodnotě	31
4.2 Dvouvýběrový Studentův test	35
4.3 Wilcoxon test	39
4.4 Fisherův test	40
4.5 Shapiro Wilk test	40
5 ANOVA	42
6 Variance	46
7 Korelace	47
7.1 Korelační matice	47

8 Kovariance	52
8.1 Kovarianční matice	52
9 Testování v kontingenčních tabulkách	58
9.1 Pearsonův Chí-kvadrát test	58
10 Regrese	59
10.1 Lineární regrese	59
10.2 Kvadratická regrese	60
Závěr	61
Použitá literatura	62
Seznam příloh	63
Příloha A	64

SEZNAM OBRÁZKŮ

1	Klouzavý průměr nových případů v ČR od 7. 3. 2020	12
2	Nové případy na milion v ČR od 7. 3. 2020	13
3	Klouzavý průměr nových případů na milion v ČR od 7. 3. 2020	13
4	Hospitalizovaní pacienti v ČR od 7. 3. 2020	14
5	Nově testování v ČR od 7. 3. 2020	14
6	Nové případy v ČR od 7. 3. 2020	15
7	Klouzavý průměr nových případů na milion pro Česko a Rakousko od 7. 3. 2020	15
8	Bodový graf zlogaritmovaných nových případů	16
9	Bodový graf nových testů	16
10	Bodový graf reprodukčního čísla	17
11	Bodový graf pacientů na icu	17
12	Bodový graf hospitalizovaných pacientů	18
13	Bodový graf týdenních přírůstků na icu	18
14	Bodový graf týdenních hospitalizací	19
15	Bodový graf pozitivitu testů	19
16	Bodový graf nových očkování	20
17	Bodový graf smrtnosti	20
18	Boxplot graf pro nové případy na milion	21
19	Boxplot graf pro reprodukční číslo	21
20	Boxplot graf pro zlogaritmované nové smrti	22
21	3D graf počtu případů a počtu testů	23
22	3D graf zlogaritmovaných počtu případů a počtu testů	23
23	3D graf počtu případů a počtu nových očkování	24
24	3D graf počtu nových případů	24
25	3D graf reprodukčního čísla	25
26	Hexbin graf nových zlog. nových případů a nových úmrtí	25
27	Legenda Chernoff faces grafu tabulky popisné statistiky	27
28	Chernoff faces graf tabulky popisné statistiky	28
29	QQPlot graf nových případů a nových úmrtí	29
30	QQPlot graf nových testů a nových případů	30

31	Anova graf nových testů, případů a úmrtí	42
32	Anova graf nových testů, případů a úmrtí	43
33	Anova graf nových testů, případů a úmrtí	43
34	Anova graf nových testů, případů a úmrtí	44
35	Anova graf nových testů, případů a úmrtí	44
36	Anova graf nových testů, případů a úmrtí	45
37	Heatmap graf korelační matice	51
38	Heatmap graf kovarianční matice	56
39	Graf kovarianční matice	56
40	GGQQPlot graf korelační matice	57
41	Graf lineární regrese	59
42	Graf kvadratické regrese	60

SEZNAM TABULEK

1	Části popisné statistiky aplikované na data nových případů a jejich 7denního klouzavého průměru v ČR od 7. 3. 2020	10
2	Části popisné statistiky aplikované na data nových případů na milion a jejich 7denního klouzavého průměru v ČR od 7. 3. 2020	10
3	Části popisné statistiky aplikované na data nových hospitalizací a nových hospitalizací na milion v ČR od 7. 3. 2020	11
4	Hodnoty dat znázorněných pomocí Chernoffových obličejů	26
5	Kontingenční tabulka nových případů na milion	58

ÚVOD

Tato semestrální práce se zabývá analýzou vývoje epidemie nemoci Covid-19 v ČR. Za tímto účelem jsou srovnány přírůstky nových případů s našimi sousedy, efektivita testů při odhalování nových případů, úmrtnost nakažených, střední hodnota hospitalizovaných a vývoj střední hodnoty nových případů v čase. Dále je provedeno srovnání středních hodnot nových případů na milion s našimi sousedy a zkoumáno jaké rozdělení pravděpodobnosti data následují. Nakonec je pomocí regrese analyzováno podle jaké funkce se řídí přírůstek nových pacientů na ICU v závislosti na nových hospitalizacích. Použitá data čerpají ze zdroje [1].

1 POPIS DAT

Data použitá v této práci se zabývají veličinami ohledně nemoci Covid-19 a pochází od společnosti Our World in Data. Tato data jsou denně aktualizována a obsahují například informace o očkování, testech, hospitalizacích, nových případech, nových úmrtích či reprodukčním čísle. Veškeré hodnoty jsou pozorovány napříč mnoha státy. Pro bližší popis těchto dat vizte zdroj [1].

Covid-19 (též COVID-19; z anglického spojení coronavirus disease 2019, což česky znamená koronavirové onemocnění 2019; výslovnost: [kovid devatenáct]; podle ICD-11 označené XN109) je vysoce infekční onemocnění, které je způsobeno koronavirem SARS-CoV-2. První případ byl identifikován v čínském Wu-chanu v prosinci 2019. Od té doby se virus rozšířil po celém světě, což způsobilo přetrvávající pandemii.

Příznaky nemoci covid-19 jsou různé, od bezpříznakového stavu až po závažné onemocnění, ale často zahrnují horečku, kašel, únavu, dýchací potíže a ztrátu čichu a chuti. Příznaky začínají jeden až čtrnáct dní po vystavení viru. U přibližně jednoho z pěti infikovaných jedinců se neobjeví žádné příznaky. Zatímco většina lidí má mírné příznaky, u některých lidí se vyvine syndrom akutní dechové tísně. Tento syndrom může být přivoděn cytokinovými bouřemi, víceorgánovým selháním, septickým šokem a krevními sraženinami. Bylo pozorováno dlouhodobější poškození orgánů (zejména plic a srdce). Existuje obava z významného počtu pacientů, kteří se zotavili z akutní fáze onemocnění, ale nadále pocítují řadu následků – známých jako dlouhodobý covid-19 – i několik měsíců poté. Mezi tyto účinky patří silná únava, ztráta paměti a další kognitivní problémy, slabá horečka, svalová slabost a dušnost.

Virus, který způsobuje covid-19, se šíří hlavně vzdušným přenosem, když je infikovaná osoba v blízkém kontaktu s jinou osobou. Malé kapičky a aerosoly obsahující virus se mohou šířit z nosu a úst infikované osoby při dýchání, kašlání, kýchání, zpěvu nebo mluvení. Ostatní lidé se mohou nakazit, pokud se virus dostane do jejich úst, nosu nebo očí. Virus se může šířit také kontaminovaným povrchem, i když to není považováno za hlavní cestu přenosu. Přesná cesta přenosu je zřídka přesvědčivě prokázána, ale k infekci dochází hlavně tehdy, když jsou lidé dostatečně blízko sebe. Virus se může šířit až dva dny předtím, než infikované osoby projeví příznaky, a od jedinců, kteří nikdy nepocítují příznaky. Lidé zůstávají infekční po dobu až deseti dnů při středně závažných případech a dva týdny ve

vážných případech. Virus se šíří snadněji ve vnitřních prostorech a v davu. Pro diagnózu onemocnění byly vyvinuty různé testovací metody. Standardní diagnostickou metodou je reverzní transkripční polymerázová řetězová reakce v reálném čase (PCR test) výtěrem z nosohltanu.

Preventivní opatření zahrnují fyzický či společenský odstup, umístění ohrožených osob do karantény, větrání vnitřních prostor, zakrývání úst a nosu při kašli a kýchání, mytí rukou a udržování neumytých rukou pryč od obličeje. Aby se minimalizovalo riziko přenosu, bylo na veřejnosti doporučeno použití roušek, obličejových masek nebo jiného zakrytí dýchacích cest. Bylo vyvinuto několik vakcín proti covidu-19, načež většina států světa zahájila očkovací kampaně a samotné očkování, jehož rozsah je ovšem závislý na přístupnosti dostatečného množství vakcín.

Ačkoli probíhají práce na vývoji léků, které zpomalují a zastavují virus, primární léčba je v současnosti symptomatická. Zahrnuje léčbu příznaků, podpůrnou péči, izolaci a některá experimentální opatření.

Vzhledem k velkému množství dat jsou v této práci použity zpravidla údaje pro Českou republiku, ze kterých je dále využit užší výčet dostupných veličin.

2 POPISNÁ STATISTIKA

V tabulkách níže jsou zobrazeny hodnoty popisné statistiky pro veličiny nových případů, 7denního klouzavého průměru nových případů, nových případů na milion, 7denního klouzavého průměru nových případů na milion, hospitalizovaných pacientů a hospitalizovaných pacientů na milion v České republice. Hodnoty 7denního klouzavého průměru lépe zachycují tyto veličiny v rámci dlouhodobých trendů, jelikož je eliminováno zkreslení v podobě menšího počtu uskutečněných testů například během víkendů.

	Nové případy	7denní klouzavý průměr nových případů
průměr	2940.58	2936.89
modus	75.00	57.57
medián	416.00	422.29
max	17773.00	12954.86
min	-2214.00	2.71
šikmost	1.55	1.16
špičatost	1.38	-0.07
odchylka	4277.10	3876.20
variance	18293577.15	15024928.55

Tabulka 1: Části popisné statistiky aplikované na data nových případů a jejich 7denního klouzavého průměru v ČR od 7. 3. 2020

	Nové případy na milión	7denní klouzavý průměr nových případů na milión
průměr	274.19	273.85
modus	6.99	5.37
medián	38.79	39.38
max	1657.22	1207.96
min	-206.44	0.25
šikmost	1.55	1.16
špičatost	1.38	-0.07
odchylka	398.81	361.43
variance	159052.40	130633.34

Tabulka 2: Části popisné statistiky aplikované na data nových případů na milión a jejich 7denního klouzavého průměru v ČR od 7. 3. 2020

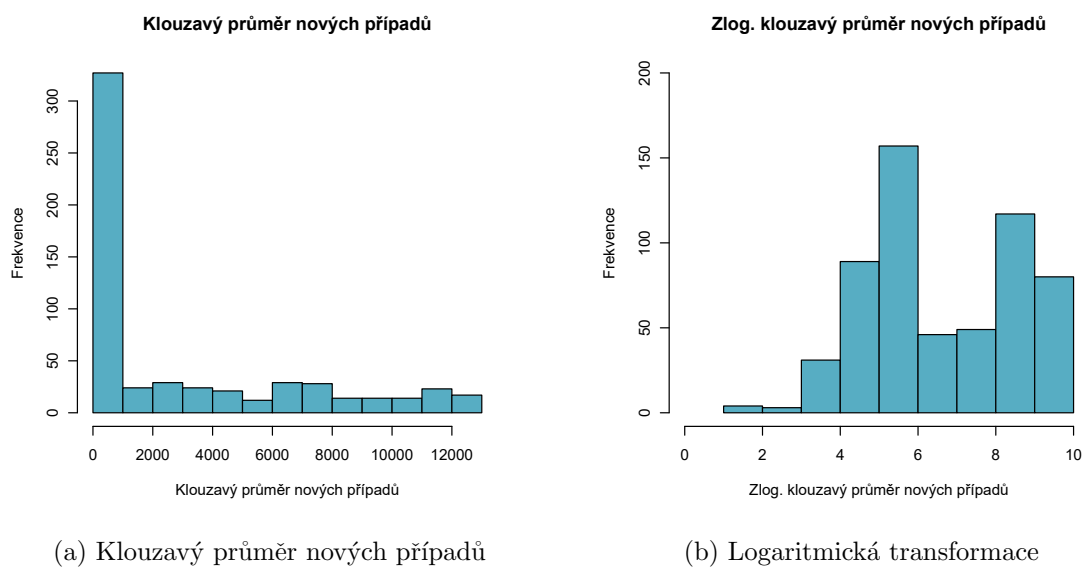
	Hospitalizovaní pacienti	Hospitalizovaní pacienti na milión
průměr	2370.49	221.03
modus	69.00	6.43
medián	339.00	31.61
max	9509.00	886.66
min	0.00	0.00
šikmost	0.86	0.86
špičatost	-0.86	-0.86
odchylka	2973.01	277.22
variance	8838793.84	76848.36

Tabulka 3: Části popisné statistiky aplikované na data nových hospitalizací a nových hospitalizací na milión v ČR od 7. 3. 2020

3 ZÁKLADNÍ GRAFY

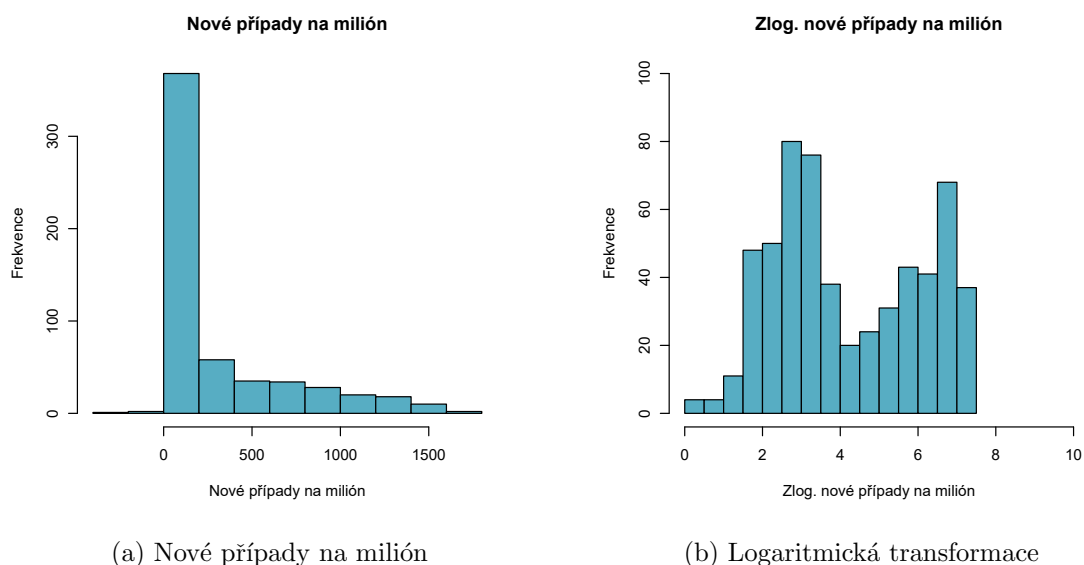
3.1 Histogram

Následující histogram zobrazuje četnost hodnot klouzavého průměru nových případů v ČR od 7. 3. 2020. Vzhledem k očividnému zešikmení dat vlevo byla pro lepší přehlednost data zlogaritmována.



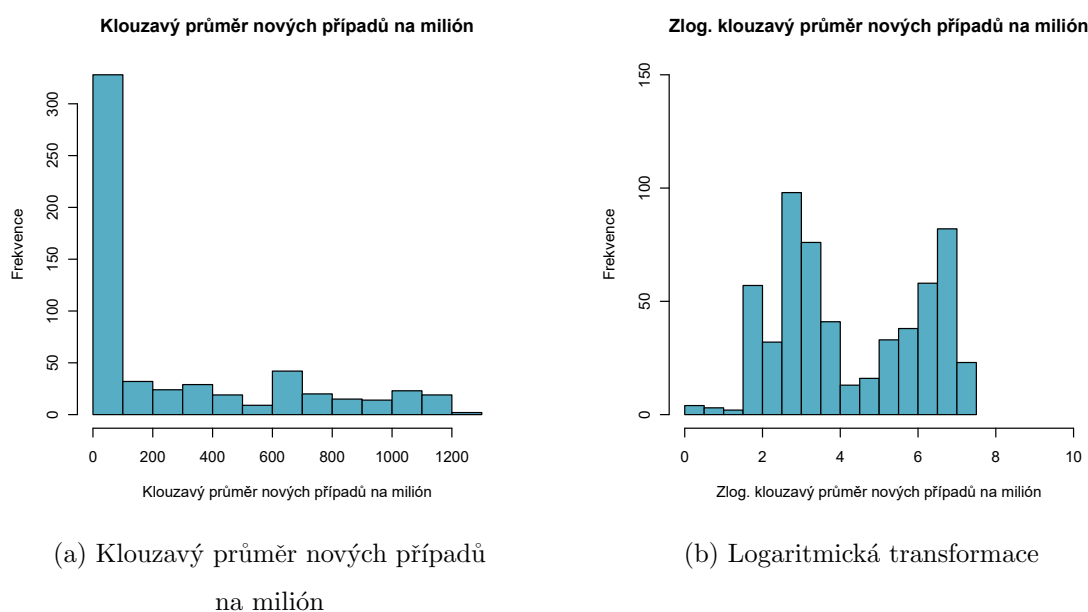
Obrázek 1: Klouzavý průměr nových případů v ČR od 7. 3. 2020

Následující histogram zobrazuje četnost hodnot nových případů na milión obyvatel v ČR od 7. 3. 2020. Vzhledem k očividnému zešikmení dat vlevo byla pro lepší přehlednost data opět zlogaritmována.



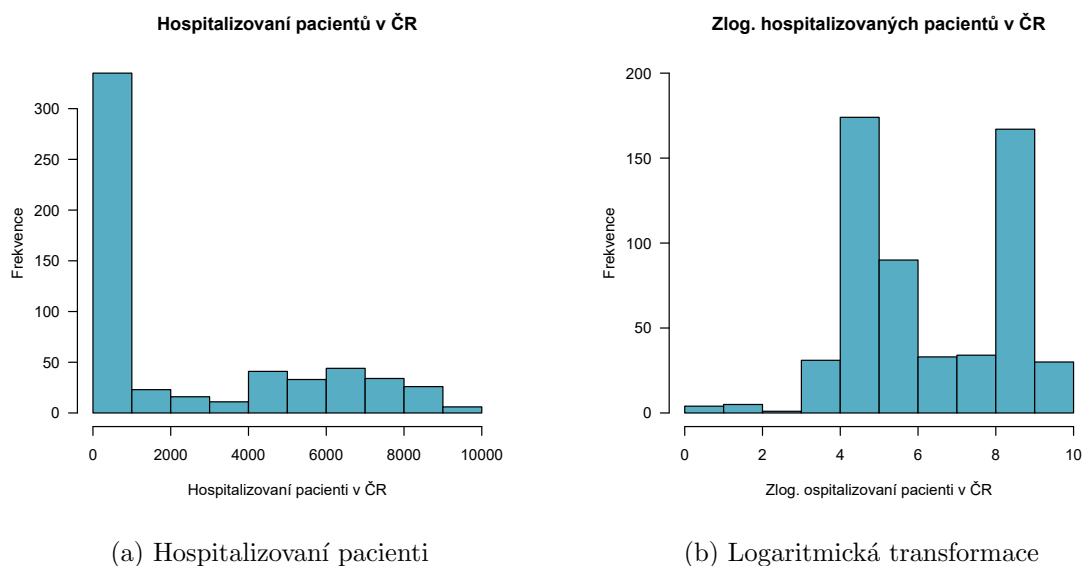
Obrázek 2: Nové případy na milión v ČR od 7. 3. 2020

Následující histogram zobrazuje četnost hodnot 7denního klouzavého průměru nových případů na milión obyvatel v ČR od 7. 3. 2020. Vzhledem k očividnému zešikmení dat vlevo byla pro lepší přehlednost data opět zlogaritmována.



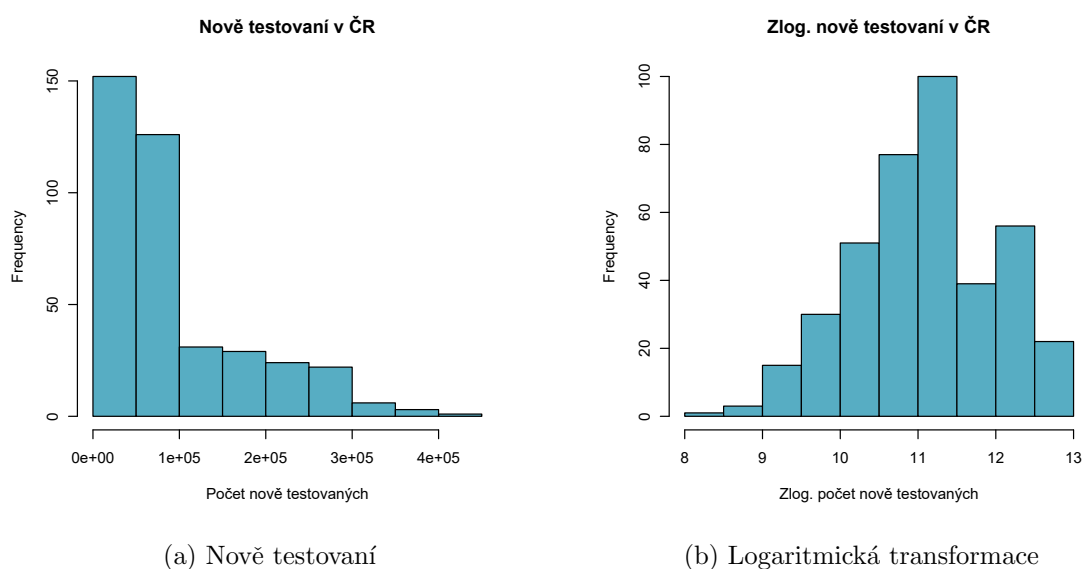
Obrázek 3: Klouzavý průměr nových případů na milión v ČR od 7. 3. 2020

Následující histogram zobrazuje četnost hodnot hospitalizovaných pacientů v ČR od 7. 3. 2020. Vzhledem k očividnému zešikmení dat vlevo byla pro lepší přehlednost data opět zlogaritmována.



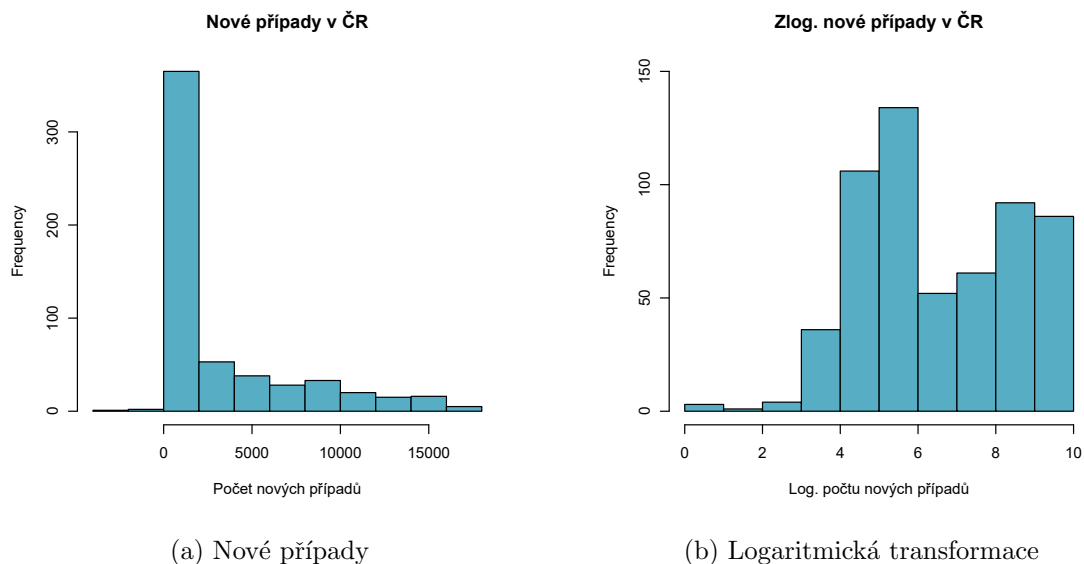
Obrázek 4: Hospitalizovaní pacienti v ČR od 7. 3. 2020

Následující histogram zobrazuje četnost hodnot nově testovaných v ČR od 7. 3. 2020. Vzhledem k očividnému zešikmení dat vlevo byla pro lepší přehlednost data opět zlogaritmována.



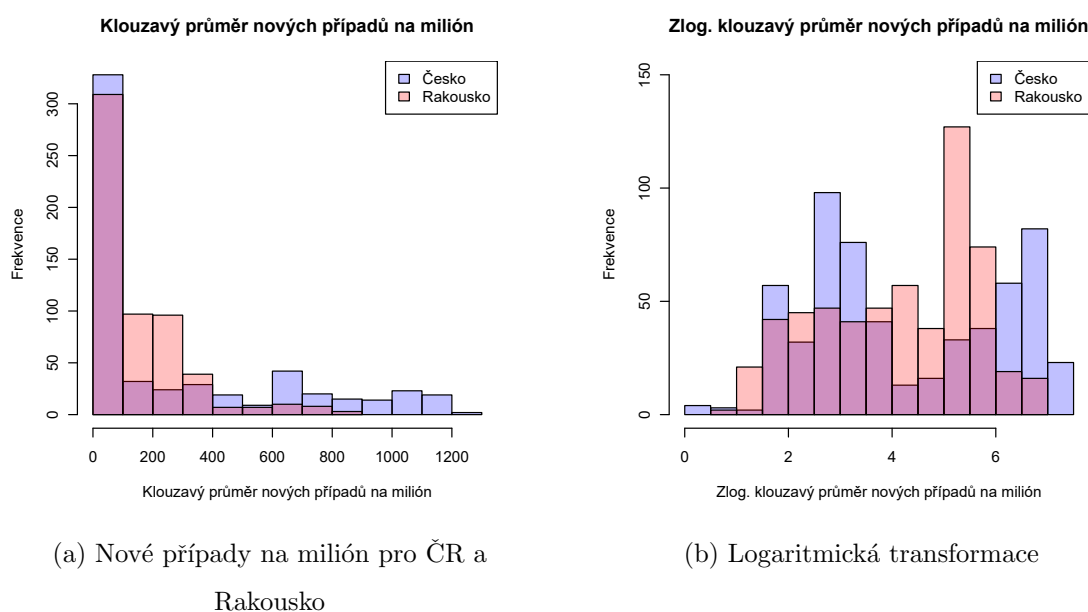
Obrázek 5: Nově testování v ČR od 7. 3. 2020

Následující histogram zobrazuje četnost hodnot nových případů v ČR od 7. 3. 2020. Vzhledem k očividnému zešikmení dat vlevo byla pro lepší přehlednost data opět zlogaritmována.



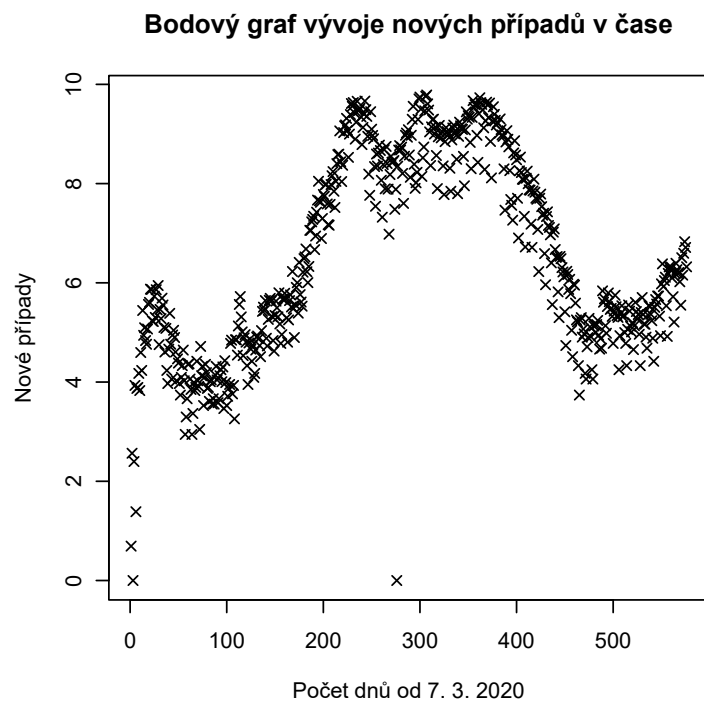
Obrázek 6: Nové případy v ČR od 7. 3. 2020

Následující histogram zobrazuje srovnání četnosti hodnot klouzavého průměru nových případů v ČR a Rakousku od 7. 3. 2020. Vzhledem k očividnému zešikmení dat vlevo byla pro lepší přehlednost data opět zlogaritmována.

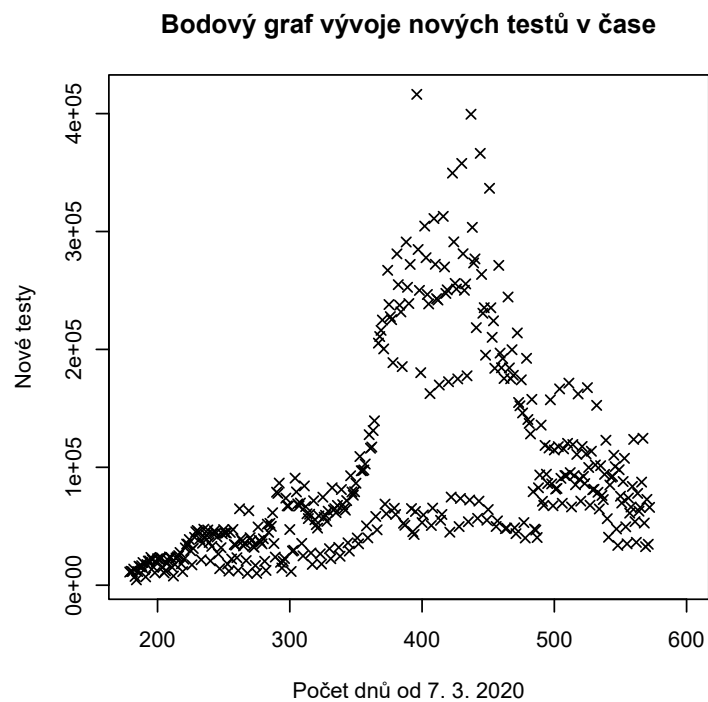


Obrázek 7: Klouzavý průměr nových případů na milión pro Česko a Rakousko od 7. 3. 2020

3.2 Bodový graf

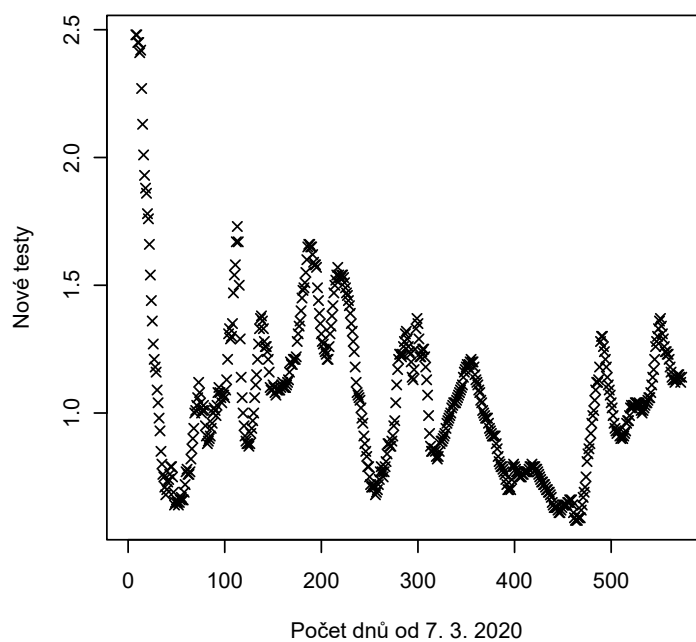


Obrázek 8: Bodový graf zlogaritmovaných nových případů



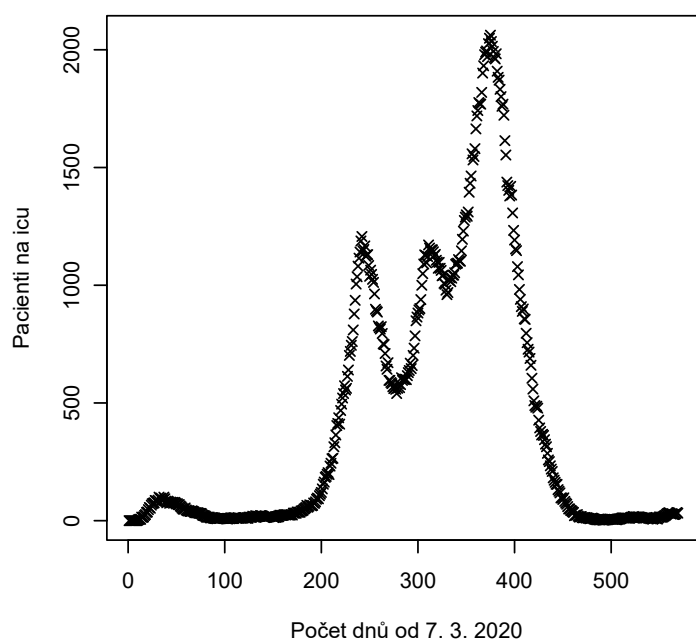
Obrázek 9: Bodový graf nových testů

Bodový graf vývoje reprodukčního čísla v čase



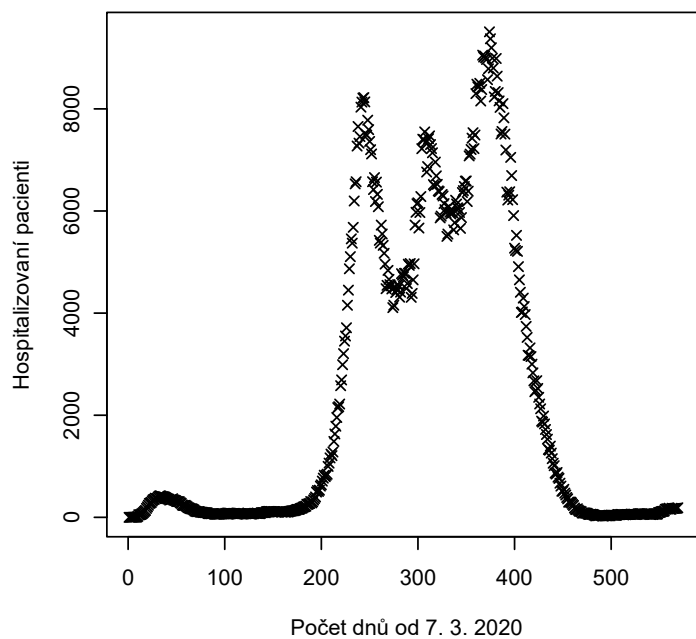
Obrázek 10: Bodový graf reprodukčního čísla

Bodový graf vývoje pacientů na icu v čase



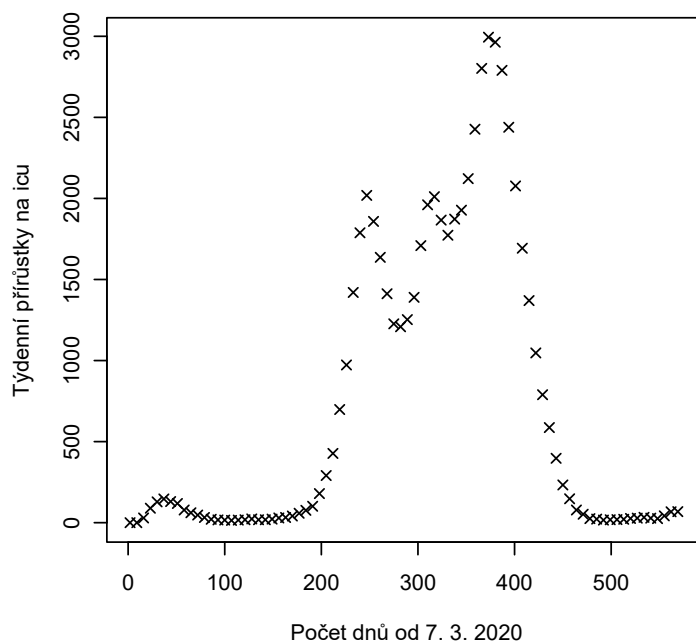
Obrázek 11: Bodový graf pacientů na icu

Bodový graf vývoje hospitalizovaných pacientů v čase

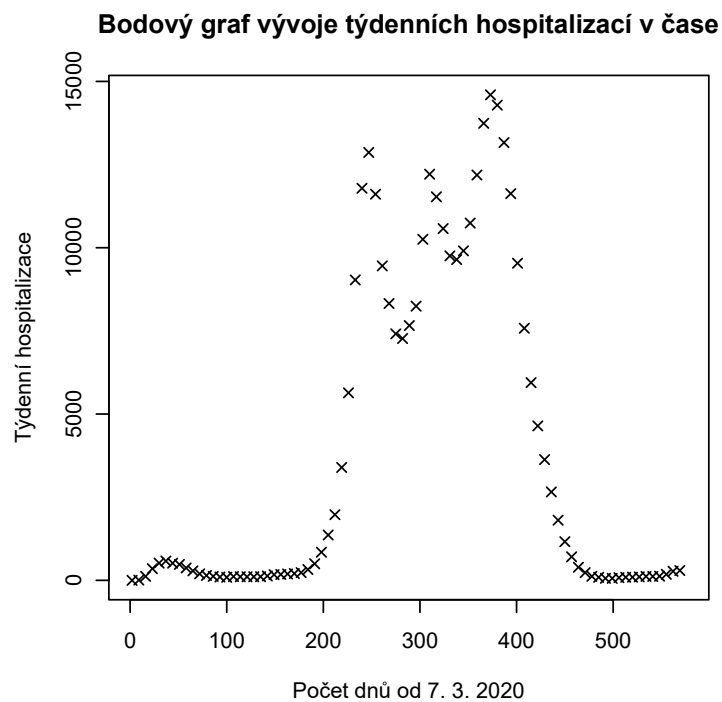


Obrázek 12: Bodový graf hospitalizovaných pacientů

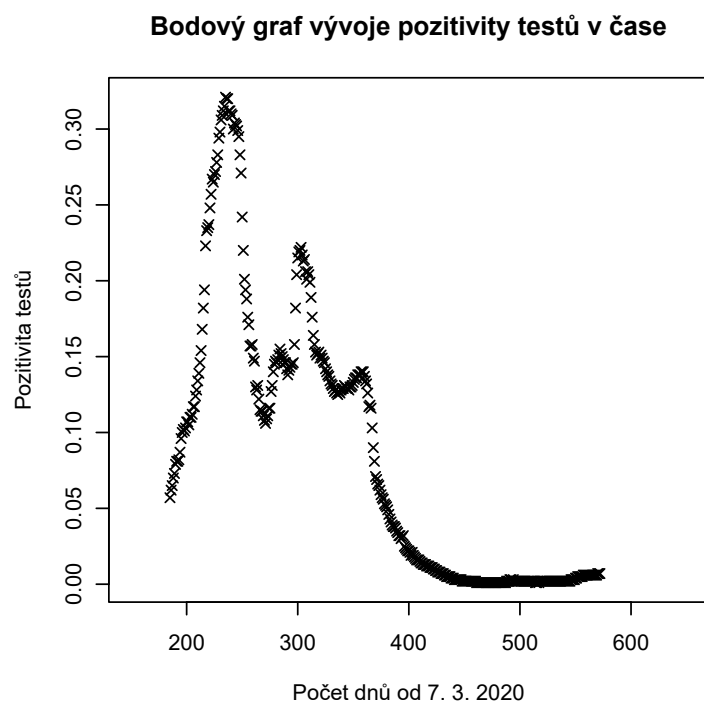
Bodový graf vývoje týdenních přírůstků na icu v čase



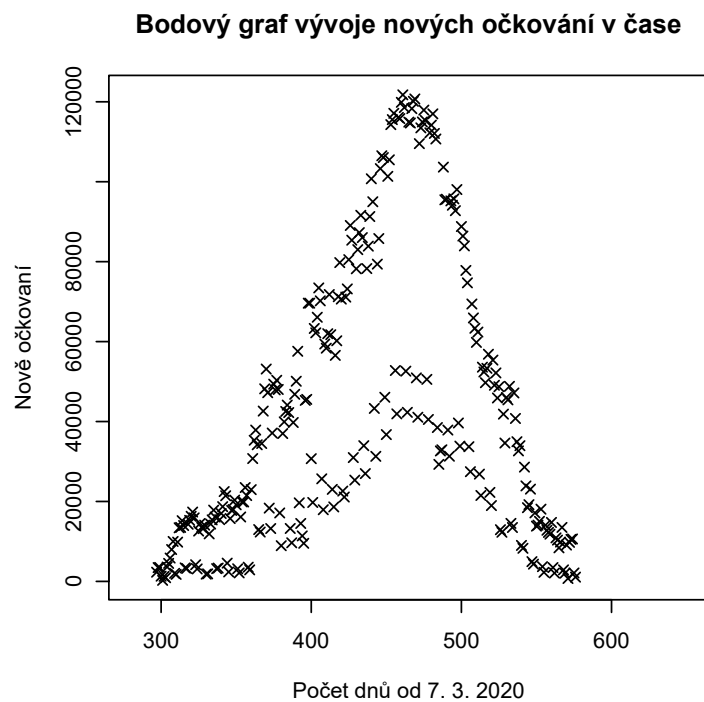
Obrázek 13: Bodový graf týdenních přírůstků na icu



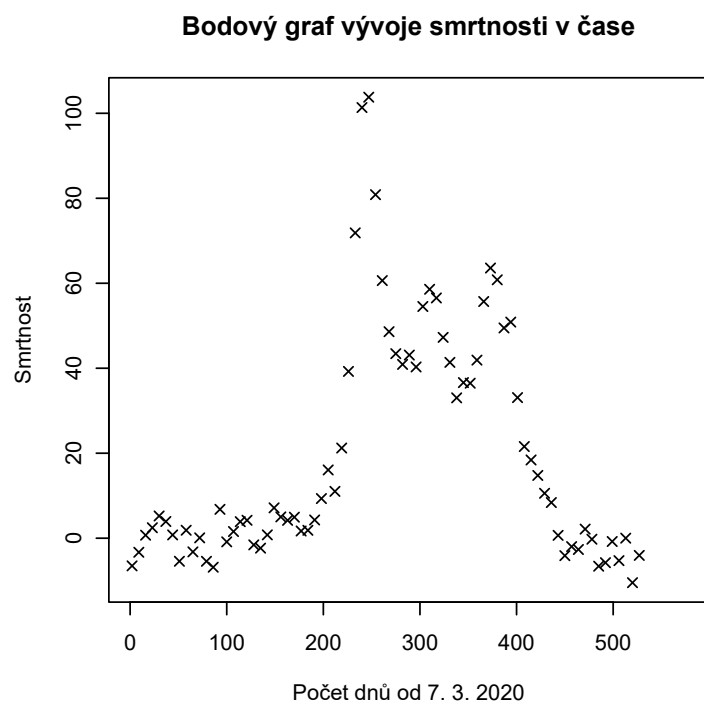
Obrázek 14: Bodový graf týdenních hospitalizací



Obrázek 15: Bodový graf positivity testů

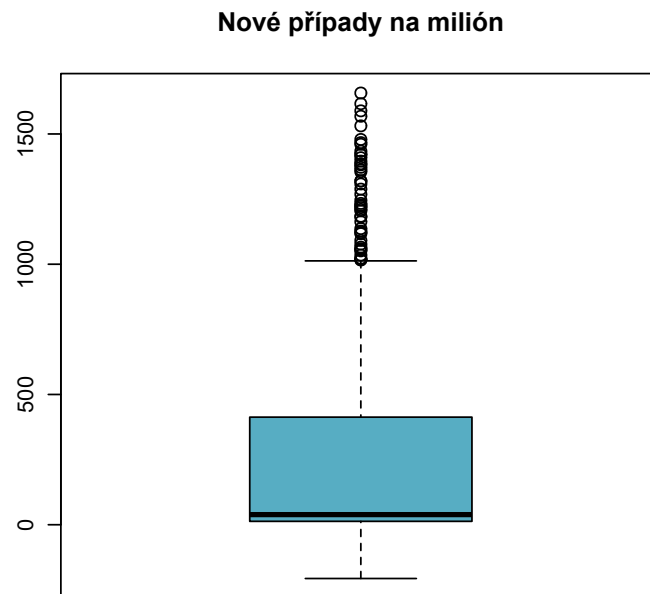


Obrázek 16: Bodový graf nových očkování

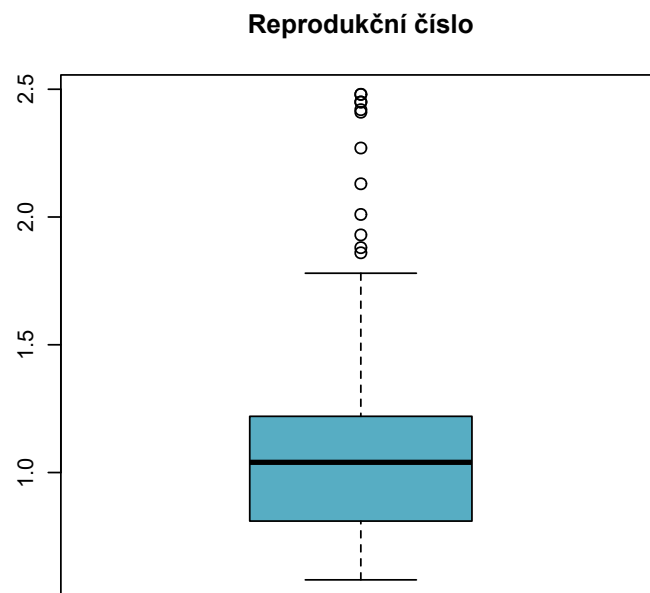


Obrázek 17: Bodový graf smrtnosti

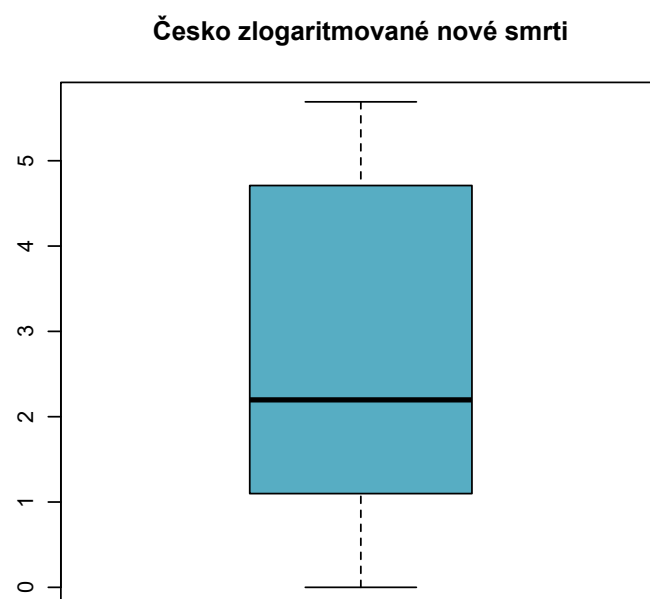
3.3 Boxplot



Obrázek 18: Boxplot graf pro nové případy na milión

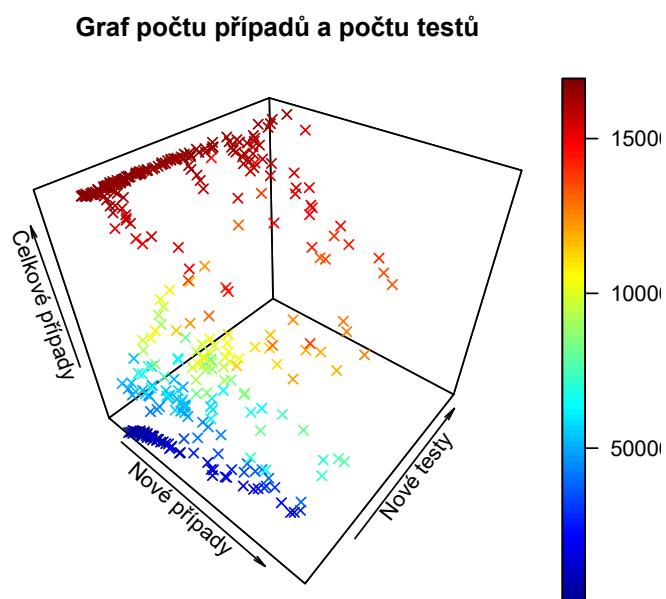


Obrázek 19: Boxplot graf pro reprodukční číslo



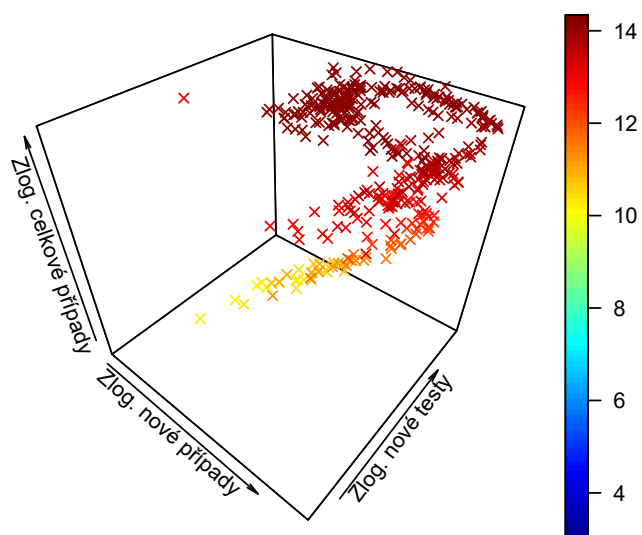
Obrázek 20: Boxplot graf pro zlogaritmované nové smrti

3.4 3D graf



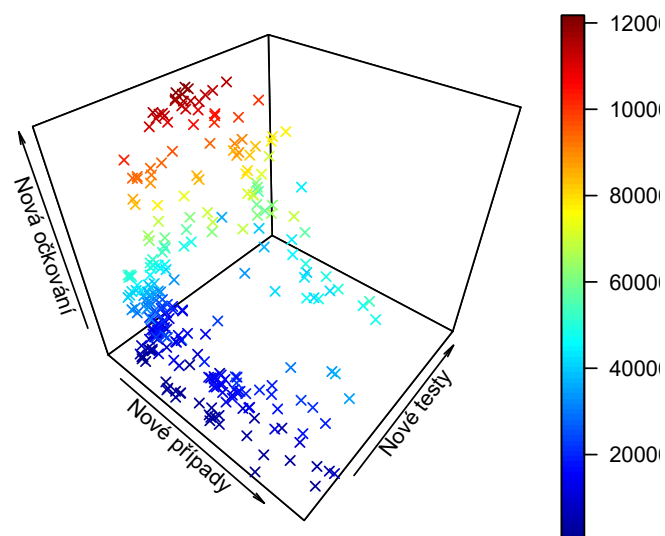
Obrázek 21: 3D graf počtu případů a počtu testů

Graf zlogaritmovaného počtu případů a počtu testů



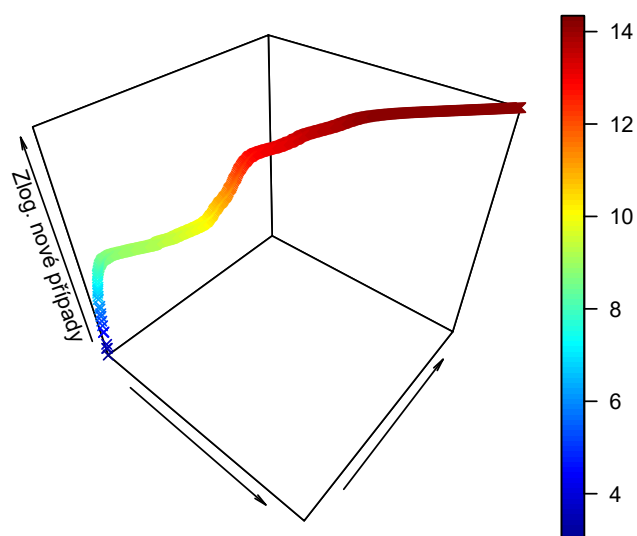
Obrázek 22: 3D graf zlogaritmovaných počtu případů a počtu testů

Graf počtu případů a počtu očkování



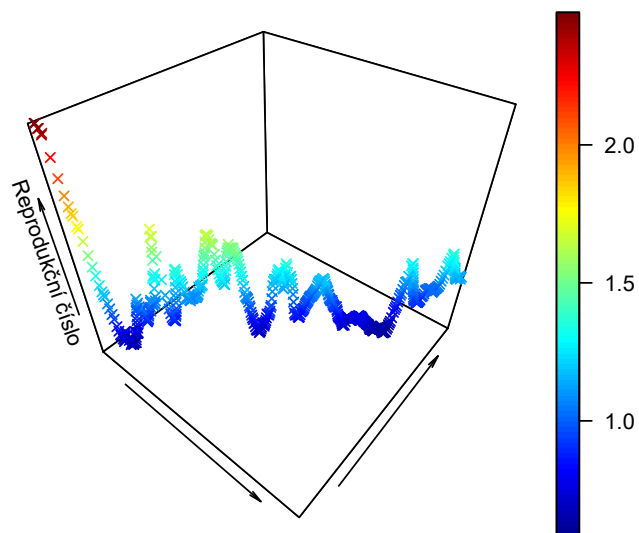
Obrázek 23: 3D graf počtu případů a počtu nových očkování

Graf počtu nových případů



Obrázek 24: 3D graf počtu nových případů

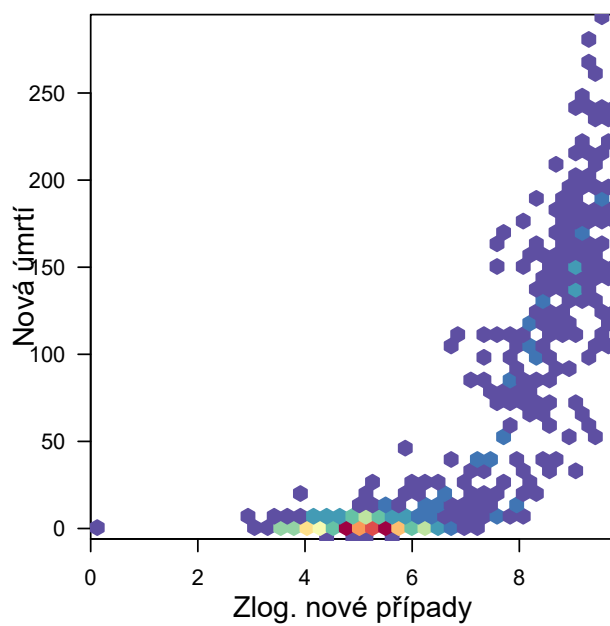
Graf vývoje reprodukčního čísla



Obrázek 25: 3D graf reprodukčního čísla

3.5 Hexbin

Graf zlog. nových případů a nových úmrtí



Obrázek 26: Hexbin graf nových zlog. nových případů a nových úmrtí

3.6 Chernoff faces

Pomocí chernoff faces jsou vyobrazeny obličeje na základě hodnot pro Česko. Hodnoty, které jsou použity jsou *new_cases*, *new_deaths*, *new_tests*, *new_vaccinations*, *icu_patients*. Každý obličej reprezentuje jednu funkci, která je aplikována na zvolené hodnoty. Z grafu je zřejmé zešíkmení dat vzhledem k velkým rozdílům mezi hodnotami průměru a mediánu zobrazovaných veličin.

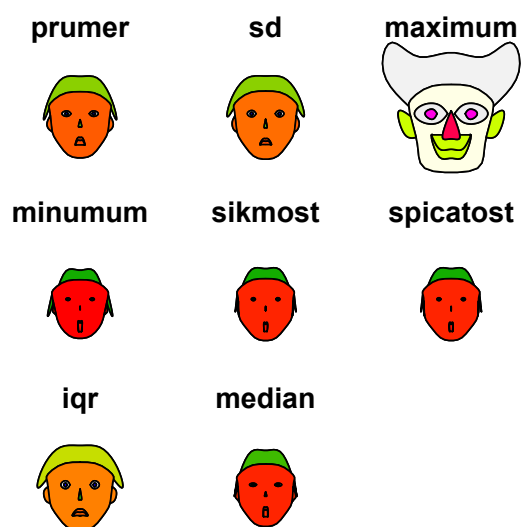
	new_cases	new_deaths	new_tests	new_vaccinations	icu_patients
prumer	2940.58	54.33	94464.09	42220.25	424.51
sd	4277.10	72.32	83099.10	35340.13	558.31
maximum	17773.00	295.00	416333.00	121742.00	2062.00
minumum	-2214.00	-6.00	4537.00	266.00	0.00
sikmost	1.55	1.11	1.39	0.79	1.23
spicatost	1.38	-0.06	1.22	-0.57	0.43
iqr	4275.75	107.00	82303.50	49771.00	799.00
median	416.00	7.00	65211.00	33315.00	74.00

Tabulka 4: Hodnoty dat znázorněných pomocí Chernoffových obličejů

effect of variables:

modified item	Var
"height of face	" "new_cases"
"width of face	" "new_deaths"
"structure of face"	"new_tests"
"height of mouth	" "new_vaccinations"
"width of mouth	" "icu_patients"
"smiling	" "new_cases"
"height of eyes	" "new_deaths"
"width of eyes	" "new_tests"
"height of hair	" "new_vaccinations"
"width of hair	" "icu_patients"
"style of hair	" "new_cases"
"height of nose	" "new_deaths"
"width of nose	" "new_tests"
"width of ear	" "new_vaccinations"
"height of ear	" "icu_patients"

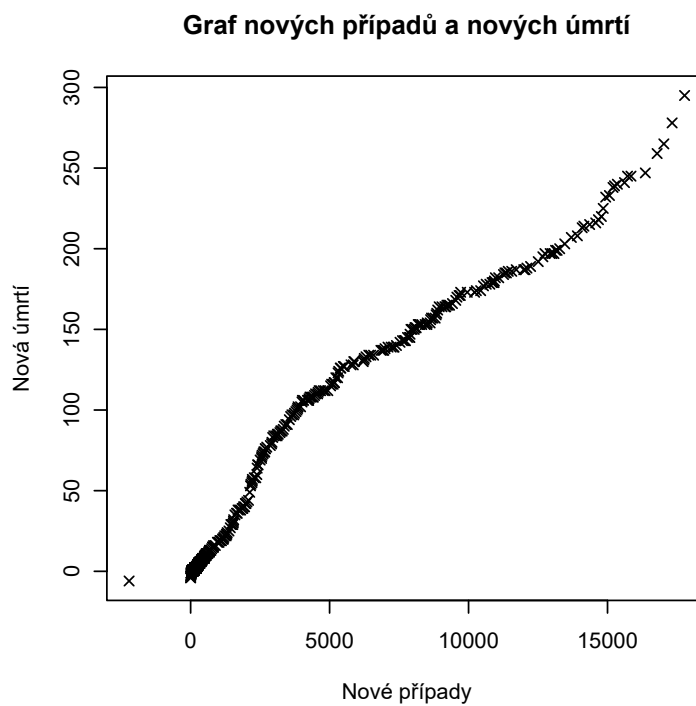
Obrázek 27: Legenda Chernoff faces grafu tabulky popisné statistiky



Obrázek 28: Chernoff faces graf tabulky popisné statistiky

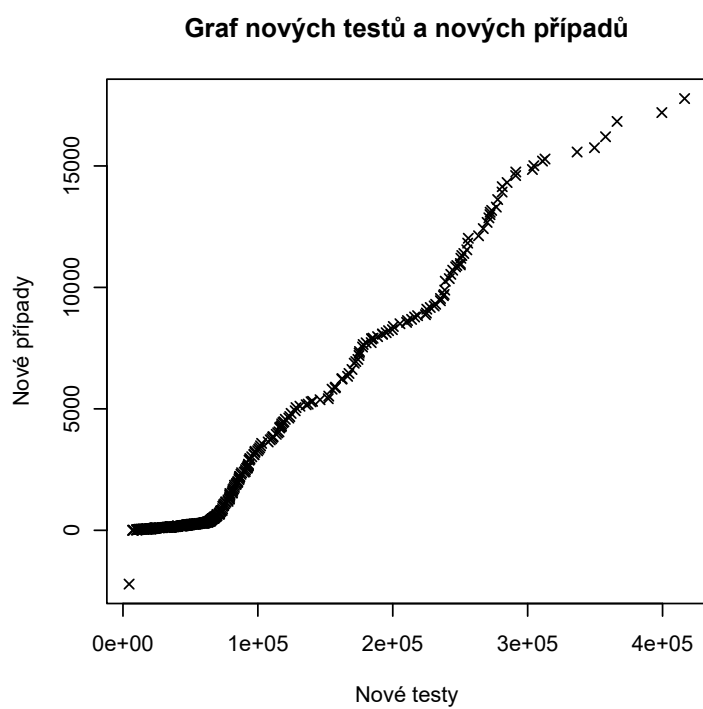
3.7 QQPlot

Na základě následujícího QQPlot grafu je možné dojít k závěru, že počty nových úmrtí a počty nových případů v ČR se řídí dle podobného rozdělení pravděpodobnosti.



Obrázek 29: QQPlot graf nových případů a nových úmrtí

Na základě následujícího QQPlot grafu je možné dojít k závěru, že počty nových případů a počty nových testů v ČR se řídí dle podobného rozdělení pravděpodobnosti.



Obrázek 30: QQPlot graf nových testů a nových případů

4 TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

4.1 Jednovýběrový Studentův test vůči střední hodnotě

Následující test testuje zda se střední hodnota nových případů v ČR rovná hodnotě 3300 s hladinou významnosti $\alpha = 0.05$. Testová statistika nabývá hodnoty -2.0168 při 575 stupních volnosti. Vzhledem k tomu, že hodnota p-value je nižší než hladina významnosti, tuto hypotézu zamítáme ve prospěch hypotézy alternativní, tudíž že se střední hodnota nových případů v ČR nerovná hodnotě 3300.

One Sample t-test

```
data:  new_cases_czechia
t = -2.0168, df = 575, p-value = 0.04418
alternative hypothesis: true mean is not equal to 3300
95 percent confidence interval:
 2590.550 3290.603
sample estimates:
mean of x
 2940.576
```

Následující test testuje zda se střední hodnota 7denního klouzavého průměru nových případů v ČR rovná hodnotě 3300 s hladinou významnosti $\alpha = 0.05$. Testová statistika nabývá hodnoty -2.2482 při 575 stupních volnosti. Vzhledem k tomu, že hodnota p-value je nižší než hladina významnosti, tuto hypotézu zamítáme ve prospěch hypotézy alternativní, tudíž že se střední hodnota nových případů v ČR nerovná hodnotě 3300.

One Sample t-test

```
data: new_cases_smoothed_czechia
t = -2.2482, df = 575, p-value = 0.02494
alternative hypothesis: true mean is not equal to 3300
95 percent confidence interval:
 2619.672 3254.109
sample estimates:
mean of x
 2936.89
```

Následující test testuje zda se střední hodnota nových případů na milión v ČR rovná hodnotě 300 s hladinou významnosti $\alpha = 0.05$. Testová statistika nabývá hodnoty -1.5531 při 575 stupních volnosti. Vzhledem k tomu, že hodnota p-value je vyšší než hladina významnosti, tuto hypotézu nemůžeme zamítnout ve prospěch hypotézy alternativní.

One Sample t-test

```
data: new_cases_per_million_czechia
t = -1.5531, df = 575, p-value = 0.1209
alternative hypothesis: true mean is not equal to 300
95 percent confidence interval:
 241.5531 306.8289
sample estimates:
mean of x
 274.191
```


Následující test testuje zda se střední hodnota 7denního klouzavého průměru nových případů na milión v ČR rovná hodnotě 300 s hladinou významnosti $\alpha = 0.05$. Testová statistika nabývá hodnoty -1.7366 při 575 stupních volnosti. Vzhledem k tomu, že hodnota p-value je vyšší než hladina významnosti, tuto hypotézu nemůžeme zamítnout ve prospěch hypotézy alternativní.

One Sample t-test

```
data: new_cases_smoothed_per_million_czechia
t = -1.7366, df = 575, p-value = 0.08299
alternative hypothesis: true mean is not equal to 300
95 percent confidence interval:
 244.2687 303.4260
sample estimates:
mean of x
 273.8474
```

Následující test testuje zda se střední hodnota nových hospitalizovaných pacientů v ČR rovná hodnotě 2000 s hladinou významnosti $\alpha = 0.05$. Testová statistika nabývá hodnoty 2.9726 při 568 stupních volnosti. Vzhledem k tomu, že hodnota p-value je nižší než hladina významnosti, tuto hypotézu zamítáme ve prospěch hypotézy alternativní, tudíž že se střední hodnota nových hospitalizovaných pacientů v ČR nerovná hodnotě 2000.

One Sample t-test

```
data: hosp_patients_czechia
t = 2.9726, df = 568, p-value = 0.003078
alternative hypothesis: true mean is not equal to 2000
95 percent confidence interval:
 2125.690 2615.294
sample estimates:
mean of x
 2370.492
```

Následující test testuje zda se střední hodnota nových hospitalizovaných pacientů na milión v ČR rovná hodnotě 200 s hladinou významnosti $\alpha = 0.05$. Testová statistika nabývá hodnoty 1.8099 při 568 stupních volnosti. Vzhledem k tomu, že hodnota p-value je vyšší než hladina významnosti, tuto hypotézu nemůžeme zamítnout ve prospěch hypotézy alternativní.

One Sample t-test

```
data: hosp_patients_per_million_czechia
t = 1.8099, df = 568, p-value = 0.07083
alternative hypothesis: true mean is not equal to 200
95 percent confidence interval:
 198.2078 243.8604
sample estimates:
mean of x
 221.0341
```

4.2 Dvouvýběrový Studentův test

Následující dvouvýběrový t-test testuje hypotézu, že střední hodnota nových případů v první části dat z ČR je rovna střední hodnotě v druhé části. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů v první části dat z ČR se nerovná střední hodnotě z druhé části.

Welch Two Sample t-test

```
data:  new_cases_czechia_p1 and new_cases_czechia_p2
t = -4.518, df = 537.03, p-value = 7.683e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  -2272.4359  -895.1752
sample estimates:
mean of x mean of y
 2148.674  3732.479
```

Následující dvouvýběrový t-test testuje hypotézu, že střední hodnota nových případů v ČR je rovna střední hodnotě nových případů v Německu. Testová statistika nabývá hodnoty -11,133 při 843,1 stupních volnosti. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů v Německu se nerovná střední hodnotě nových případů v ČR.

Welch Two Sample t-test

```
data: new_cases_czechia and new_cases_germany
t = -11.133, df = 843.1, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  -5240.394 -3669.509
sample estimates:
mean of x mean of y
  2940.576  7395.528
```

Následující dvouvýběrový t-test testuje hypotézu, že střední hodnota nových případů na milión v ČR je rovna střední hodnotě nových případů na milión na Slovensku. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů na milión na Slovensku se nerovná střední hodnotě nových případů na milión v ČR.

Welch Two Sample t-test

```
data: new_cases_per_million_czechia and new_cases_per_million_slovakia
t = 7.7283, df = 817.84, p-value = 3.194e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  105.8863 177.9857
sample estimates:
mean of x mean of y
  274.191  132.255
```

Následující dvouvýběrový t-test testuje hypotézu, že střední hodnota nových případů na milión v ČR je rovna střední hodnotě nových případů na milión v Německu. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů na milión v Německu se nerovná střední hodnotě nových případů na milión v ČR.

Two Sample t-test

```
data:  new_cases_per_million_czechia and new_cases_per_million_germany
t = 10.844, df = 1150, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 152.3817 219.7075
sample estimates:
mean of x mean of y
274.19103  88.14644
```

Následující dvouvýběrový t-test testuje hypotézu, že střední hodnota nových případů na milión v ČR je rovna střední hodnotě nových případů na milión v Polsku. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů na milión v Polsku se nerovná střední hodnotě nových případů na milión v ČR.

Two Sample t-test

```
data: new_cases_per_million_czechia and new_cases_per_million_poland
t = 7.5404, df = 1150, p-value = 9.477e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 103.9326 177.0431
sample estimates:
mean of x mean of y
 274.1910 133.7032
```

Následující dvouvýběrový t-test testuje hypotézu, že střední hodnota nových případů na milión v ČR je rovna střední hodnotě nových případů na milión v Rakousku. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů na milión v Rakousku se nerovná střední hodnotě nových případů na milión v ČR.

Two Sample t-test

```
data: new_cases_per_million_czechia and new_cases_per_million_austria
t = 7.2242, df = 1150, p-value = 9.157e-13
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 95.01377 165.86690
sample estimates:
mean of x mean of y
 274.1910 143.7507
```

4.3 Wilcox test

Následující Wilcoxonův testuje hypotézu, že střední hodnota nových případů na milion v ČR je rovna střední hodnotě nových případů na milion na Slovensku. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů na milion na Slovensku se nerovná střední hodnotě nových případů na milion v ČR. Vzhledem k zešíkmení dat poskytuje tento test přesnější výsledky oproti dvouvýběrovému t-testu.

Wilcoxon rank sum test with continuity correction

```
data:  new_cases_per_million_czechia and new_cases_per_million_slovakia
W = 205293, p-value = 2.97e-12
alternative hypothesis: true location shift is not equal to 0
```

Následující Wilcoxonův testuje hypotézu, že střední hodnota nových případů na milion v ČR je rovna střední hodnotě nových případů na milion v Německu. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní. Při této hladině významnosti tudíž můžeme tvrdit, že střední hodnota nových případů na milion v Německu se nerovná střední hodnotě nových případů na milion v ČR. Vzhledem k zešíkmení dat poskytuje tento test přesnější výsledky oproti dvouvýběrovému t-testu.

Wilcoxon rank sum test with continuity correction

```
data:  new_cases_per_million_czechia and new_cases_per_million_germany
W = 188720, p-value = 5.261e-05
alternative hypothesis: true location shift is not equal to 0
```

4.4 Fisherův test

Následující Fisherův test zkoumá zda jsou rozptyly hodnot nových případů na milión v ČR a na Slovensku stejné. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní, tudíž že jsou rozptyly těchto dat různé.

```
F test to compare two variances
```

```
data:  new_cases_per_million_czechia and new_cases_per_million_slovakia
```

```
F = 4.5141, num df = 575, denom df = 575, p-value < 2.2e-16
```

```
alternative hypothesis: true ratio of variances is not equal to 1
```

```
95 percent confidence interval:
```

```
3.832723 5.316656
```

```
sample estimates:
```

```
ratio of variances
```

```
4.514119
```

4.5 Shapiro Wilk test

Následující Shapiro Wilk test testuje zda je veličina nových případů v ČR nabývá normálního rozdělení. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní, tudíž že tato veličina nenabývá normálního rozdělení.

```
Shapiro-Wilk normality test
```

```
data:  new_cases_czechia
```

```
W = 0.72003, p-value < 2.2e-16
```


Následující Shapiro Wilk test testuje zda je veličina nových testů v ČR nabývá normálního rozdělení. Vzhledem ke skutečnosti, že p-value je menší než hladina významnosti ($\alpha = 0,05$), zamítáme tuto hypotézu ve prospěch alternativní, tudíž že tato veličina nenabývá normálního rozdělení.

Shapiro-Wilk normality test

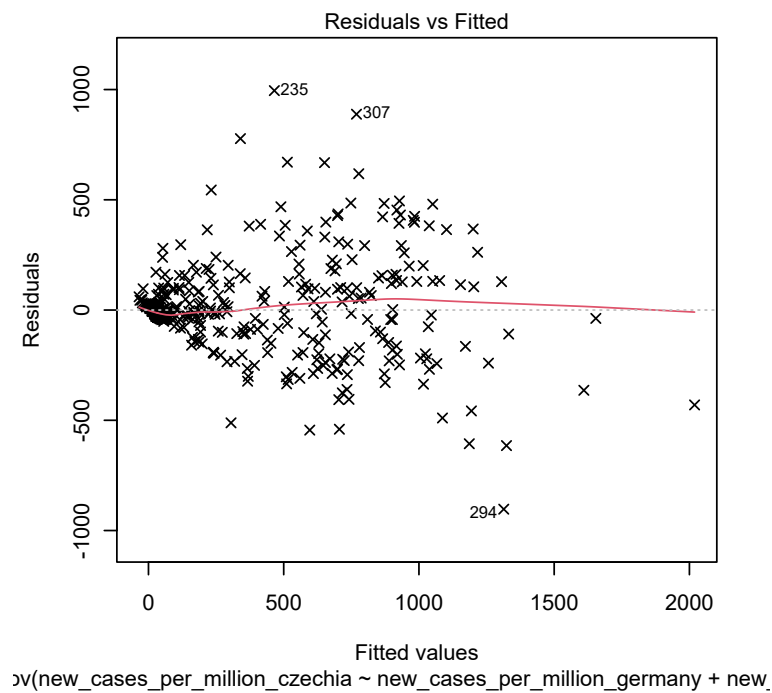
data: new_tests_czechia

W = 0.82915, p-value < 2.2e-16

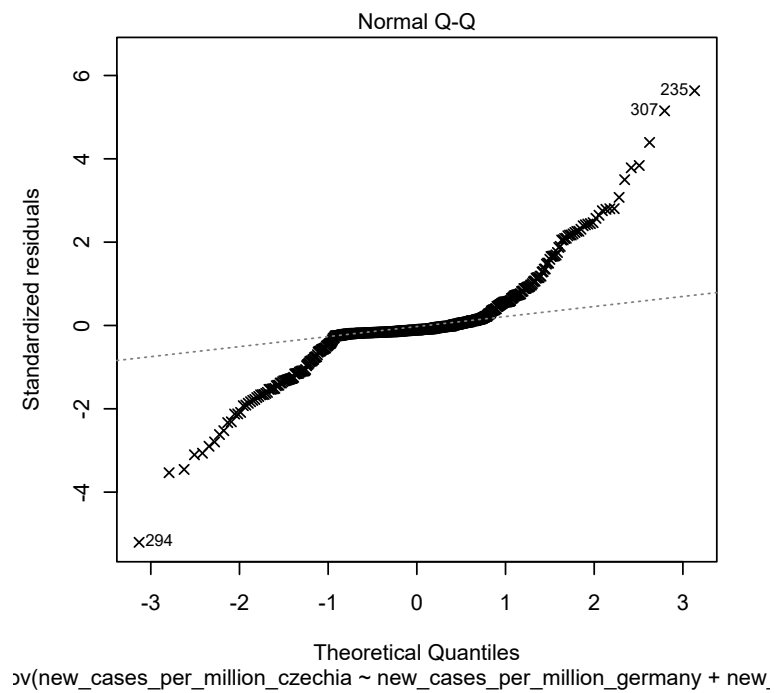
5 ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
new_cases_per_million_germany	1	32058325	32058325	1020.43	< 2e-16 ***
new_cases_per_million_slovakia	1	37842394	37842394	1204.54	< 2e-16 ***
new_cases_per_million_poland	1	2637470	2637470	83.95	< 2e-16 ***
new_cases_per_million_austria	1	978148	978148	31.14	3.72e-08 ***
Residuals	571	17938795	31416		

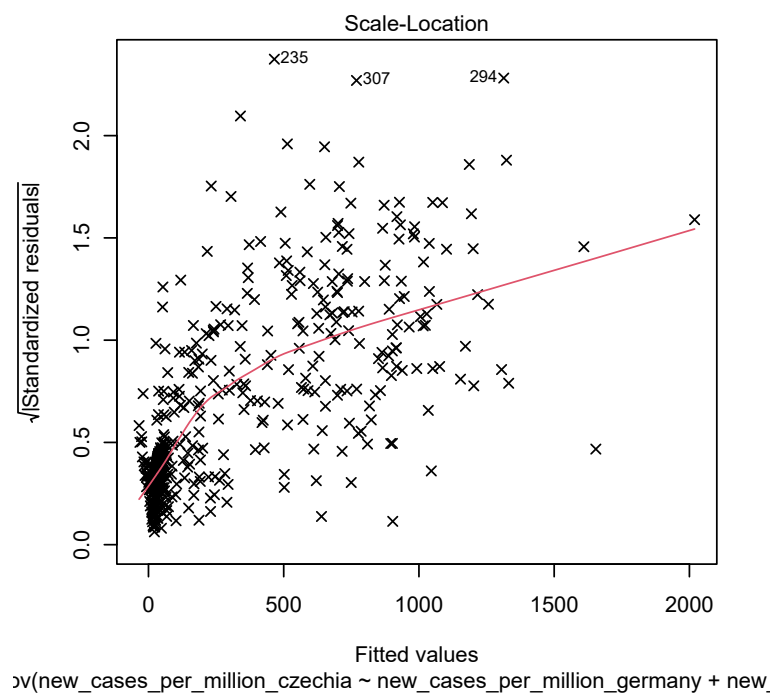
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



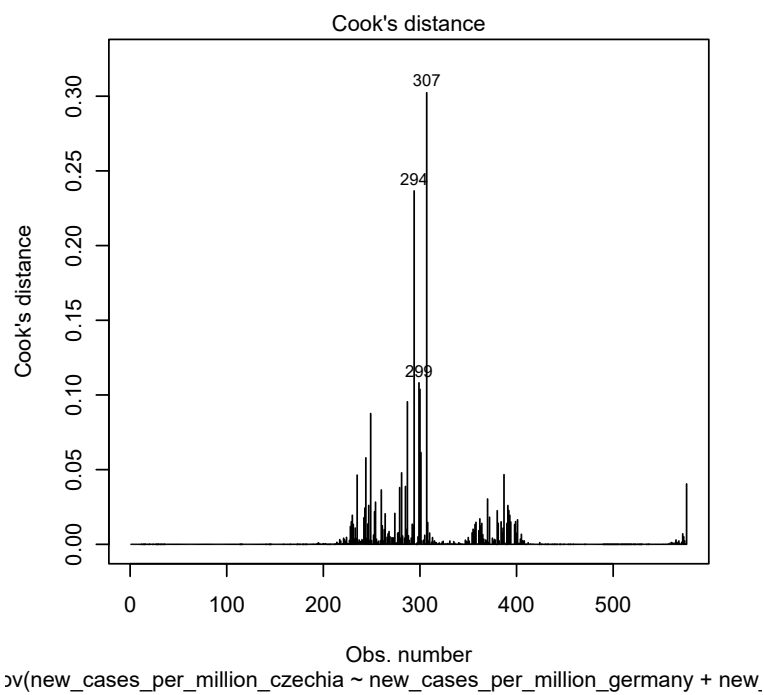
Obrázek 31: Anova graf nových testů, případů a úmrtí



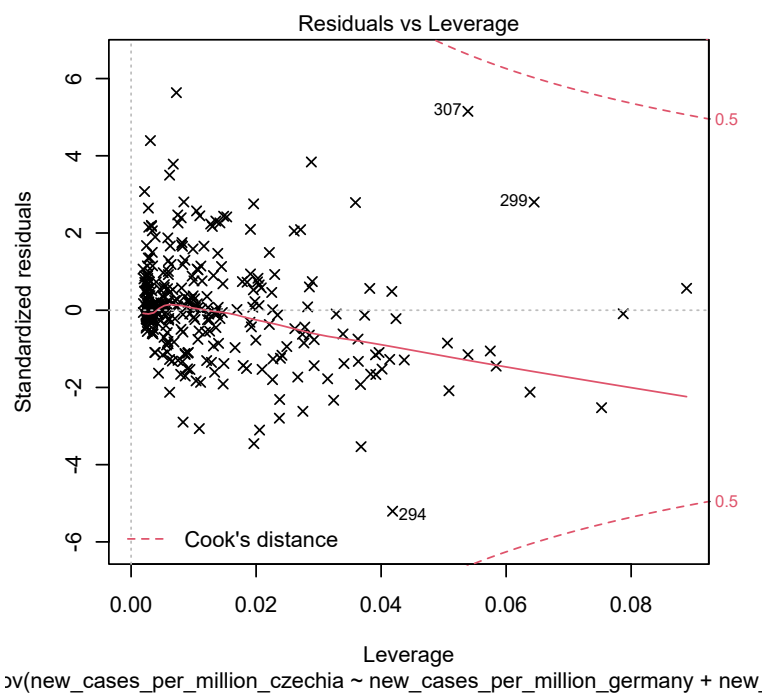
Obrázek 32: Anova graf nových testů, případů a úmrtí



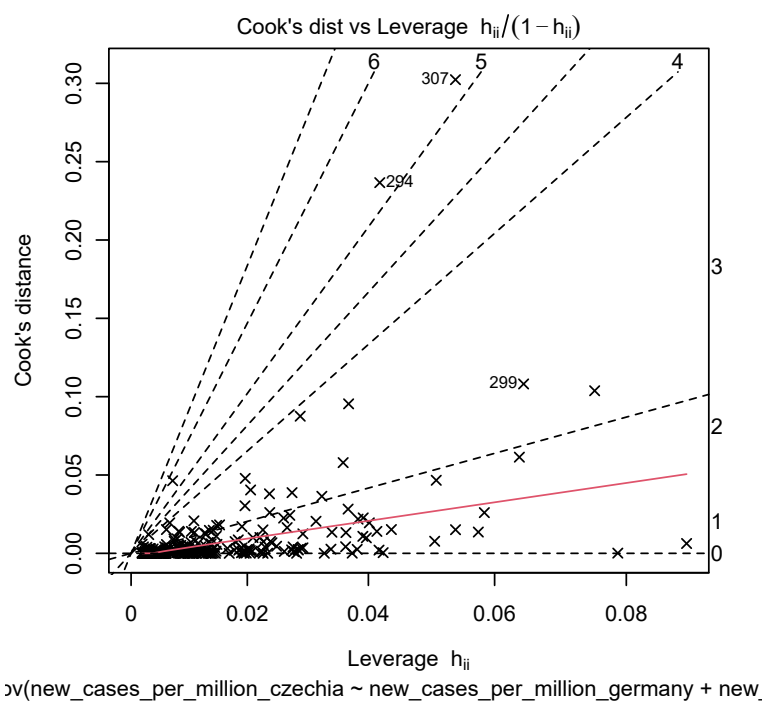
Obrázek 33: Anova graf nových testů, případů a úmrtí



Obrázek 34: Anova graf nových testů, případů a úmrtí



Obrázek 35: Anova graf nových testů, případů a úmrtí



Obrázek 36: Anova graf nových testů, případů a úmrtí

6 VARIANCE

Níže jsou popsány střední hodnoty kvadrátů odchylek od střední hodnoty nových testů a nových případů v ČR.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-61749419	-61749419	-61749419	-61749419	-61749419	-61749419

7 KORELACE

7.1 Korelační matice

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	1.00000000	-0.54948723	-0.06330228	-0.253692480	0.47034101	0.494350433
[2,]	-0.54948723	1.00000000	0.10390139	0.304800900	-0.39360394	-0.432884095
[3,]	-0.06330228	0.10390139	1.00000000	0.275607246	-0.36319740	-0.356972846
[4,]	-0.25369248	0.30480090	0.27560725	1.000000000	-0.13231174	-0.093132516
[5,]	0.47034101	-0.39360394	-0.36319740	-0.132311738	1.00000000	0.614784606
[6,]	0.49435043	-0.43288409	-0.35697285	-0.093132516	0.61478461	1.000000000
[7,]	0.13896782	-0.24782103	-0.24161919	0.353519423	0.11078661	0.122226465
[8,]	0.91233626	-0.51795164	0.02909308	-0.148550657	0.46939039	0.628649280
[9,]	0.38448342	-0.21217598	-0.47549035	0.171646550	0.39143921	0.502760435
[10,]	0.68502263	-0.63069457	0.17535665	-0.330560209	0.33462476	0.148609873
[11,]	-0.64882654	0.67676868	-0.09582765	0.134138762	-0.38301108	-0.369705597
[12,]	0.48112726	-0.28106527	-0.03481675	0.371208006	0.27535388	0.142940292
[13,]	0.36370252	-0.10131812	-0.12177299	-0.227428260	0.25138402	0.567271627
[14,]	-0.25061897	0.19107363	-0.16457797	0.300565332	-0.02590209	-0.298372912
[15,]	0.65224771	-0.47181602	0.39361671	-0.248687122	0.14513369	0.160025578
[16,]	-0.60115609	0.36763122	-0.40667922	0.035650419	-0.11878008	-0.136299917
[17,]	-0.27929283	0.19938197	0.55497164	0.232351502	-0.31316661	-0.537716357
[18,]	-0.58263640	0.46392759	-0.24461296	-0.056007261	-0.20473288	-0.075240976
[19,]	-0.57197373	0.42554401	0.47913549	0.406242617	-0.44004508	-0.589522947
[20,]	-0.34878234	0.44526385	0.08116204	-0.007928781	-0.41037971	0.050826337
[21,]	0.29677449	-0.21285419	-0.58429260	-0.171106614	0.25634765	0.001672992
[22,]	0.24980279	-0.06935493	0.43528786	-0.104363550	-0.11317344	0.065020929
[23,]	0.27047145	-0.25948603	-0.41088871	0.102449282	0.37728356	0.356624344
[24,]	0.64438294	-0.40150453	0.23617268	0.067762686	0.16029020	0.040603940
	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]
[1,]	0.13896782	0.91233626	0.38448342	0.68502263	-0.648826545	0.48112726
[2,]	-0.24782103	-0.51795164	-0.21217598	-0.63069457	0.676768676	-0.28106527

[3,]	-0.24161919	0.02909308	-0.47549035	0.17535665	-0.095827647	-0.03481675
[4,]	0.35351942	-0.14855066	0.17164655	-0.33056021	0.134138762	0.37120801
[5,]	0.11078661	0.46939039	0.39143921	0.33462476	-0.383011082	0.27535388
[6,]	0.12222647	0.62864928	0.50276043	0.14860987	-0.369705597	0.14294029
[7,]	1.00000000	0.10588364	0.47183815	-0.12094563	-0.396003117	0.70243376
[8,]	0.10588364	1.00000000	0.40770185	0.56084862	-0.613745392	0.37776913
[9,]	0.47183815	0.40770185	1.00000000	-0.14958123	-0.130837957	0.42596685
[10,]	-0.12094563	0.56084862	-0.14958123	1.00000000	-0.703393692	0.27306666
[11,]	-0.39600312	-0.61374539	-0.13083796	-0.70339369	1.00000000	-0.52613516
[12,]	0.70243376	0.37776913	0.42596685	0.27306666	-0.526135161	1.00000000
[13,]	-0.45255954	0.45063651	0.12427722	0.11789103	-0.023479719	-0.42757786
[14,]	0.46682606	-0.26358165	0.40503744	-0.39469955	0.202854241	0.27722660
[15,]	-0.35783885	0.69041056	-0.20157288	0.81320174	-0.506722870	0.06291983
[16,]	-0.21519710	-0.60441403	0.03139535	-0.54507460	0.739181725	-0.49050054
[17,]	-0.02167874	-0.35129619	-0.64287605	0.14882860	-0.025697540	0.20143055
[18,]	-0.29559534	-0.56474419	-0.09736262	-0.72478193	0.744146397	-0.59415141
[19,]	0.17203370	-0.58181795	-0.46322626	-0.38445043	0.285458484	0.05858271
[20,]	-0.37702642	-0.18381080	-0.18632785	-0.51760988	0.518994291	-0.52891745
[21,]	0.28854548	0.15213681	0.42044943	0.14524456	-0.008954537	0.22913874
[22,]	-0.65417123	0.32802492	-0.41109552	0.53362687	-0.128194939	-0.28596395
[23,]	0.30795289	0.22526131	0.82623540	-0.02252771	0.002852469	0.35803125
[24,]	0.22804531	0.52552584	0.26305110	0.62902881	-0.558655403	0.61550857

	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]
[1,]	0.363702516	-0.25061897	0.65224771	-0.601156092	-0.27929283	-0.58263640
[2,]	-0.101318122	0.19107363	-0.47181602	0.367631222	0.19938197	0.46392759
[3,]	-0.121772986	-0.16457797	0.39361671	-0.406679217	0.55497164	-0.24461296
[4,]	-0.227428260	0.30056533	-0.24868712	0.035650419	0.23235150	-0.05600726
[5,]	0.251384018	-0.02590209	0.14513369	-0.118780078	-0.31316661	-0.20473288
[6,]	0.567271627	-0.29837291	0.16002558	-0.136299917	-0.53771636	-0.07524098
[7,]	-0.452559538	0.46682606	-0.35783885	-0.215197096	-0.02167874	-0.29559534
[8,]	0.450636511	-0.26358165	0.69041056	-0.604414029	-0.35129619	-0.56474419
[9,]	0.124277219	0.40503744	-0.20157288	0.031395351	-0.64287605	-0.09736262

[10,]	0.117891027	-0.39469955	0.81320174	-0.545074603	0.14882860	-0.72478193
[11,]	-0.023479719	0.20285424	-0.50672287	0.739181725	-0.02569754	0.74414640
[12,]	-0.427577855	0.27722660	0.06291983	-0.490500543	0.20143055	-0.59415141
[13,]	1.000000000	-0.48370608	0.37235746	0.007854382	-0.59342008	0.15076511
[14,]	-0.483706078	1.000000000	-0.43772261	0.227934789	-0.09355538	-0.11651835
[15,]	0.372357465	-0.43772261	1.000000000	-0.551198459	0.08898115	-0.64927987
[16,]	0.007854382	0.22793479	-0.55119846	1.000000000	-0.17511202	0.64418732
[17,]	-0.593420081	-0.09355538	0.08898115	-0.175112017	1.000000000	-0.15175859
[18,]	0.150765110	-0.11651835	-0.64927987	0.644187321	-0.15175859	1.000000000
[19,]	-0.697742234	0.24346443	-0.41535576	0.058533021	0.68958292	0.17216858
[20,]	0.299396639	-0.20106417	-0.16859112	0.136032790	-0.08775309	0.52510429
[21,]	-0.016238034	0.40625378	-0.06460999	0.049283183	-0.47724091	-0.15852356
[22,]	0.409076559	-0.57036169	0.76873691	-0.251268110	0.16070128	-0.27278028
[23,]	-0.052452643	0.40725390	-0.15049820	0.289472148	-0.39521422	-0.09733461
[24,]	-0.204324511	-0.13841903	0.49181120	-0.501437272	0.22031395	-0.61874308

	[,19]	[,20]	[,21]	[,22]	[,23]
[1,]	-0.57197373	-0.348782340	0.296774495	0.24980279	0.270471448
[2,]	0.42554401	0.445263852	-0.212854193	-0.06935493	-0.259486029
[3,]	0.47913549	0.081162036	-0.584292598	0.43528786	-0.410888707
[4,]	0.40624262	-0.007928781	-0.171106614	-0.10436355	0.102449282
[5,]	-0.44004508	-0.410379707	0.256347654	-0.11317344	0.377283561
[6,]	-0.58952295	0.050826337	0.001672992	0.06502093	0.356624344
[7,]	0.17203370	-0.377026416	0.288545480	-0.65417123	0.307952892
[8,]	-0.58181795	-0.183810804	0.152136808	0.32802492	0.225261313
[9,]	-0.46322626	-0.186327852	0.420449433	-0.41109552	0.826235405
[10,]	-0.38445043	-0.517609879	0.145244559	0.53362687	-0.022527712
[11,]	0.28545848	0.518994291	-0.008954537	-0.12819494	0.002852469
[12,]	0.05858271	-0.528917453	0.229138740	-0.28596395	0.358031250
[13,]	-0.69774223	0.299396639	-0.016238034	0.40907656	-0.052452643
[14,]	0.24346443	-0.201064170	0.406253782	-0.57036169	0.407253903
[15,]	-0.41535576	-0.168591119	-0.064609988	0.76873691	-0.150498202
[16,]	0.05853302	0.136032790	0.049283183	-0.25126811	0.289472148

```

[17,]  0.68958292 -0.087753087 -0.477240905  0.16070128 -0.395214219
[18,]  0.17216858  0.525104295 -0.158523556 -0.27278028 -0.097334613
[19,]  1.00000000  0.118347789 -0.397322928 -0.25397306 -0.325774330
[20,]  0.11834779  1.000000000 -0.387687664  0.24459112 -0.381138177
[21,] -0.39732293 -0.387687664  1.000000000 -0.32992909  0.385696405
[22,] -0.25397306  0.244591125 -0.329929089  1.00000000 -0.333322819
[23,] -0.32577433 -0.381138177  0.385696405 -0.33332282  1.000000000
[24,] -0.16097037 -0.538536252  0.033197305  0.22275144  0.359999086

```

[,24]

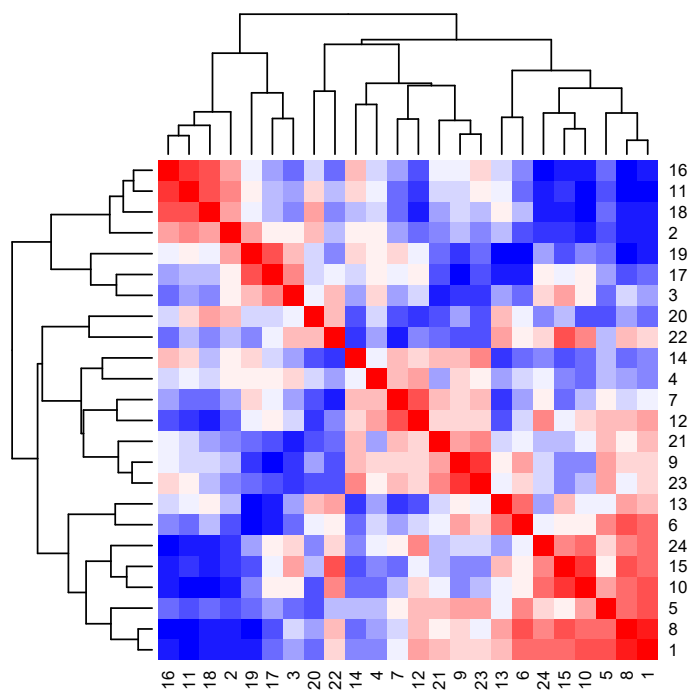
```

[1,]  0.64438294
[2,] -0.40150453
[3,]  0.23617268
[4,]  0.06776269
[5,]  0.16029020
[6,]  0.04060394
[7,]  0.22804531
[8,]  0.52552584
[9,]  0.26305110
[10,] 0.62902881
[11,] -0.55865540
[12,]  0.61550857
[13,] -0.20432451
[14,] -0.13841903
[15,]  0.49181120
[16,] -0.50143727
[17,]  0.22031395
[18,] -0.61874308
[19,] -0.16097037
[20,] -0.53853625
[21,]  0.03319730
[22,]  0.22275144

```

[23,] 0.35999909

[24,] 1.00000000



Obrázek 37: Heatmap graf korelační matice

8 KOVARIANCE

8.1 Kovarianční matice

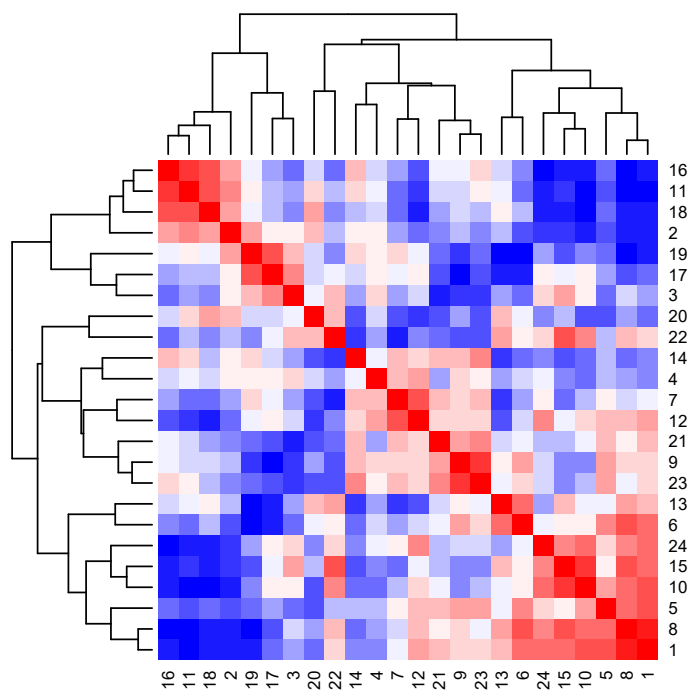
	[,1]	[,2]	[,3]	[,4]	[,5]					
[1,]	11655.5199	-5581.3007	-157.7681	-507.15036	3681.1304					
[2,]	-5581.3007	8851.6449	225.6667	530.99638	-2684.5652					
[3,]	-157.7681	225.6667	532.9275	117.81159	-607.8261					
[4,]	-507.1504	530.9964	117.8116	342.86775	-177.6087					
[5,]	3681.1304	-2684.5652	-607.8261	-177.60870	5255.3913					
[6,]	3620.6649	-2762.9384	-559.0580	-116.99094	3023.5217					
[7,]	1366.3623	-2123.4203	-507.9855	596.15942	731.4348					
[8,]	41641.5344	-20601.8949	283.9420	-1162.90399	14386.0870					
[9,]	49127.1286	-23625.7790	-12991.3333	3761.63225	33584.9565					
[10,]	271050.3170	-217475.4601	14836.6377	-22433.29529	88907.7826					
[11,]	-275592.5308	250510.1051	-8703.5797	9772.16123	-109241.3043					
[12,]	119305.1105	-60736.8732	-1846.1014	15787.51993	45848.6957					
[13,]	205291.0054	-49837.5761	-14697.4783	-22017.37500	95279.1739					
[14,]	-68113.3080	45254.8333	-9564.4058	14010.52536	-4727.0435					
[15,]	269795.3587	-170075.1087	34814.7826	-17643.01087	40311.3043					
[16,]	-285503.4964	152153.7754	-41299.4638	2903.93116	-37879.5217					
[17,]	-70292.1884	43729.9855	29866.6377	10029.75362	-52924.8261					
[18,]	-58929.0399	40891.0362	-5290.2899	-971.56884	-13904.5217					
[19,]	-21161.2754	13720.0725	3790.4638	2577.79710	-10932.0000					
[20,]	-3419.7101	3804.5072	170.1594	-13.33333	-2701.8261					
[21,]	2275.2772	-1422.1196	-957.8696	-224.99457	1319.6957					
[22,]	1561.2409	-377.7428	581.7246	-111.87138	-474.9565					
[23,]	3170.6467	-2650.8587	-1029.9565	205.98370	2969.8261					
[24,]	12932.9565	-7022.4783	1013.5652	233.26087	2160.2174					
	[,6]	[,7]	[,8]	[,9]	[,10]					
[1,]	3620.664855	1366.3623	41641.534	49127.129	271050.317					
[2,]	-2762.938406	-2123.4203	-20601.895	-23625.779	-217475.460					

[3,]	-559.057971	-507.9855	283.942	-12991.333	14836.638
[4,]	-116.990942	596.1594	-1162.904	3761.632	-22433.295
[5,]	3023.521739	731.4348	14386.087	33584.957	88907.783
[6,]	4602.302536	755.1594	18030.259	40367.013	36949.998
[7,]	755.159420	8294.1449	4076.812	50857.710	-40369.667
[8,]	18030.259058	4076.8116	178735.955	203998.317	869022.042
[9,]	40367.012681	50857.7101	203998.317	1400734.650	-648834.900
[10,]	36949.998188	-40369.6667	869022.042	-648834.900	13432546.563
[11,]	-98677.153986	-141891.8406	-1020862.415	-609234.792	-10142631.676
[12,]	22272.835145	146934.7246	366830.705	1157942.502	2298693.705
[13,]	201203.918478	-215486.1304	996070.614	769001.614	2259007.745
[14,]	-50956.452899	107026.9420	-280526.149	1206772.257	-3641648.453
[15,]	41594.228261	-124861.7391	1118328.707	-914041.946	11419144.315
[16,]	-40676.307971	-86214.5507	-1124084.221	163456.286	-8788069.395
[17,]	-85039.811594	-4602.5797	-346226.942	-1773726.449	1271589.797
[18,]	-4781.981884	-25220.2464	-223678.199	-107953.344	-2488584.895
[19,]	-13705.289855	5369.0725	-84293.290	-187875.884	-482858.029
[20,]	313.144928	-3118.3623	-7057.420	-20027.406	-172286.290
[21,]	8.059783	1866.1304	4567.538	35337.321	37802.538
[22,]	255.356884	-3448.9275	8028.226	-28166.150	113220.139
[23,]	2626.994565	3045.3043	10340.777	106179.908	-8965.136
[24,]	512.086957	3860.9565	41303.609	57877.000	428585.696
	[,11]	[,12]	[,13]	[,14]	[,15]
[1,]	-275592.531	119305.111	205291.005	-68113.308	269795.36
[2,]	250510.105	-60736.873	-49837.576	45254.833	-170075.11
[3,]	-8703.580	-1846.101	-14697.478	-9564.406	34814.78
[4,]	9772.161	15787.520	-22017.375	14010.525	-17643.01
[5,]	-109241.304	45848.696	95279.174	-4727.043	40311.30
[6,]	-98677.154	22272.835	201203.918	-50956.453	41594.23
[7,]	-141891.841	146934.725	-215486.130	107026.942	-124861.74
[8,]	-1020862.415	366830.705	996070.614	-280526.149	1118328.71
[9,]	-609234.792	1157942.502	769001.614	1206772.257	-914041.95

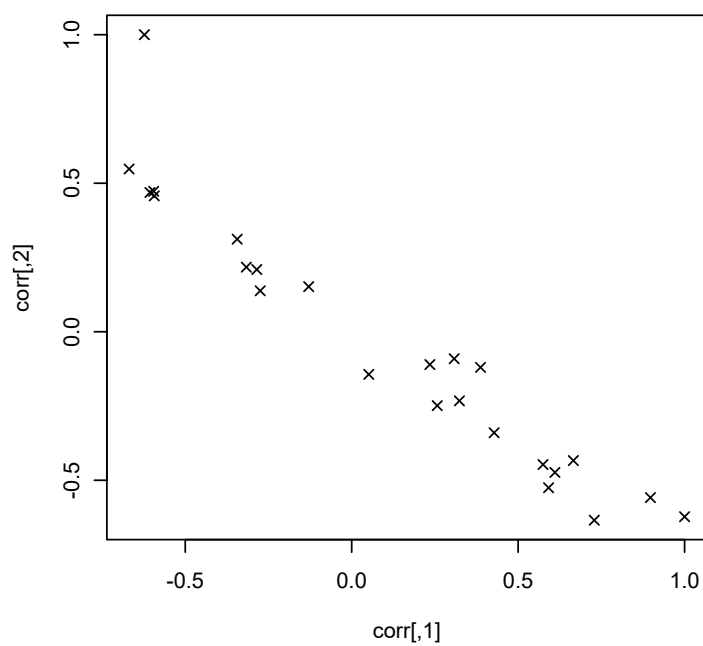
[10,]	-10142631.676	2298693.705	2259007.745	-3641648.453	11419144.32
[11,]	15479111.042	-4754486.187	-482973.929	2009135.895	-7638350.77
[12,]	-4754486.187	5275535.042	-5134595.679	1602951.438	553702.99
[13,]	-482973.929	-5134595.679	27334779.853	-6366365.207	7458885.49
[14,]	2009135.895	1602951.438	-6366365.207	6337302.341	-4221893.59
[15,]	-7638350.772	553702.989	7458885.489	-4221893.587	14679530.63
[16,]	12793299.953	-4955997.764	180645.924	2524185.007	-9290143.72
[17,]	-235692.638	1078549.449	-7232717.261	-549038.029	794759.17
[18,]	2742820.062	-1278487.004	738455.902	-274797.297	-2330525.98
[19,]	384871.623	46110.797	-1250122.957	210033.058	-545351.17
[20,]	185440.493	-110329.116	142158.783	-45968.029	-58662.39
[21,]	-2501.832	37374.353	-6028.832	72625.902	-17579.14
[22,]	-29197.839	-38023.437	123813.842	-83120.736	170506.38
[23,]	1218.582	89292.462	-29777.332	111321.250	-62610.62
[24,]	-408606.217	262818.130	-198594.043	-64779.261	350301.87
	[,16]	[,17]	[,18]	[,19]	[,20]
[1,]	-285503.496	-70292.19	-58929.0399	-21161.275	-3419.71014
[2,]	152153.775	43729.99	40891.0362	13720.072	3804.50725
[3,]	-41299.464	29866.64	-5290.2899	3790.464	170.15942
[4,]	2903.931	10029.75	-971.5688	2577.797	-13.33333
[5,]	-37879.522	-52924.83	-13904.5217	-10932.000	-2701.82609
[6,]	-40676.308	-85039.81	-4781.9819	-13705.290	313.14493
[7,]	-86214.551	-4602.58	-25220.2464	5369.072	-3118.36232
[8,]	-1124084.221	-346226.94	-223678.1993	-84293.290	-7057.42029
[9,]	163456.286	-1773726.45	-107953.3442	-187875.884	-20027.40580
[10,]	-8788069.395	1271589.80	-2488584.8949	-482858.029	-172286.28986
[11,]	12793299.953	-235692.64	2742820.0616	384871.623	185440.49275
[12,]	-4955997.764	1078549.45	-1278487.0036	46110.797	-110329.11594
[13,]	180645.924	-7232717.26	738455.9022	-1250122.957	142158.78261
[14,]	2524185.007	-549038.03	-274797.2971	210033.058	-45968.02899
[15,]	-9290143.717	794759.17	-2330525.9783	-545351.174	-58662.39130
[16,]	19351592.341	-1795790.88	2654828.3841	88238.725	54346.42029

[17,]	-1795790.884	5434541.62	-331436.7536	550893.797	-18578.59420
[18,]	2654828.384	-331436.75	877671.2101	55273.855	44676.63768
[19,]	88238.725	550893.80	55273.8551	117435.623	3683.23188
[20,]	54346.420	-18578.59	44676.6377	3683.232	8247.79710
[21,]	15395.685	-79006.17	-10546.3370	-9669.087	-2500.30435
[22,]	-63988.591	21687.41	-14794.0036	-5038.420	1285.92754
[23,]	138269.293	-100040.22	-9901.3370	-12122.087	-3758.47826
[24,]	-410074.652	95479.78	-107761.6522	-10254.957	-9092.26087
	[,21]	[,22]	[,23]	[,24]	
[1,]	2275.277174	1561.2409	3170.6467	12932.9565	
[2,]	-1422.119565	-377.7428	-2650.8587	-7022.4783	
[3,]	-957.869565	581.7246	-1029.9565	1013.5652	
[4,]	-224.994565	-111.8714	205.9837	233.2609	
[5,]	1319.695652	-474.9565	2969.8261	2160.2174	
[6,]	8.059783	255.3569	2626.9946	512.0870	
[7,]	1866.130435	-3448.9275	3045.3043	3860.9565	
[8,]	4567.538043	8028.2264	10340.7772	41303.6087	
[9,]	35337.320652	-28166.1504	106179.9076	57877.0000	
[10,]	37802.538043	113220.1395	-8965.1359	428585.6957	
[11,]	-2501.831522	-29197.8388	1218.5815	-408606.2174	
[12,]	37374.353261	-38023.4366	89292.4620	262818.1304	
[13,]	-6028.831522	123813.8424	-29777.3315	-198594.0435	
[14,]	72625.902174	-83120.7355	111321.2500	-64779.2609	
[15,]	-17579.141304	170506.3804	-62610.6196	350301.8696	
[16,]	15395.684783	-63988.5906	138269.2935	-410074.6522	
[17,]	-79006.173913	21687.4058	-100040.2174	95479.7826	
[18,]	-10546.336957	-14794.0036	-9901.3370	-107761.6522	
[19,]	-9669.086957	-5038.4203	-12122.0870	-10254.9565	
[20,]	-2500.304348	1285.9275	-3758.4783	-9092.2609	
[21,]	5042.940217	-1356.3424	2974.0489	438.2609	
[22,]	-1356.342391	3351.3025	-2095.2337	2397.2609	

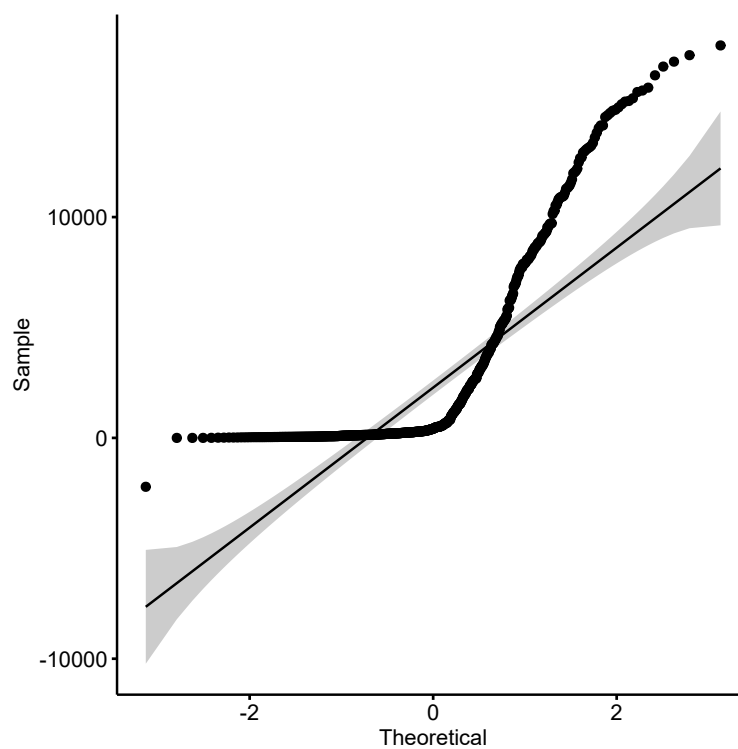
[23,]	2974.048913	-2095.2337	11790.2011	7266.9130
[24,]	438.260870	2397.2609	7266.9130	34560.1739



Obrázek 38: Heatmap graf kovarianční matice



Obrázek 39: Graf kovarianční matice



Obrázek 40: GGQQPlot graf korelační matice

9 TESTOVÁNÍ V KONTINGENČNÍCH TABULKÁCH

9.1 Pearsonův Chí-kvadrát test

Následující kontingenční tabulka zobrazuje četnost výskytu hodnot nových případů na milion pro ČR, Slovensko, Polsko, Německo a Rakousko v rámci skupin. Pomocí chí-kvadrát testu, který poté následuje zamítáme hypotézu, že se tato veličina napříč zmíněnými zeměmi řídí stejným rozdělením pravděpodobnosti.

	x<10	10-100	100>x
Czech	104	232	240
Slovakia	219	149	208
Poland	198	172	206
Germany	122	272	181
Austria	104	205	267

Tabulka 5: Kontingenční tabulka nových případů na milion

```
Pearson's Chi-squared test
```

```
data: table
```

```
X-squared = 146.95, df = 8, p-value < 2.2e-16
```

Následující chí-kvadrát test zkoumá zda má veličina nových případů v ČR stejné rozdělení jako veličina nových případů v ČR. Vzhledem ke skutečnosti, že p-value je vyšší než hladina významnosti ($\alpha = 0,05$), tuto hypotézu nemůžeme zamítnout.

```
Pearson's Chi-squared test
```

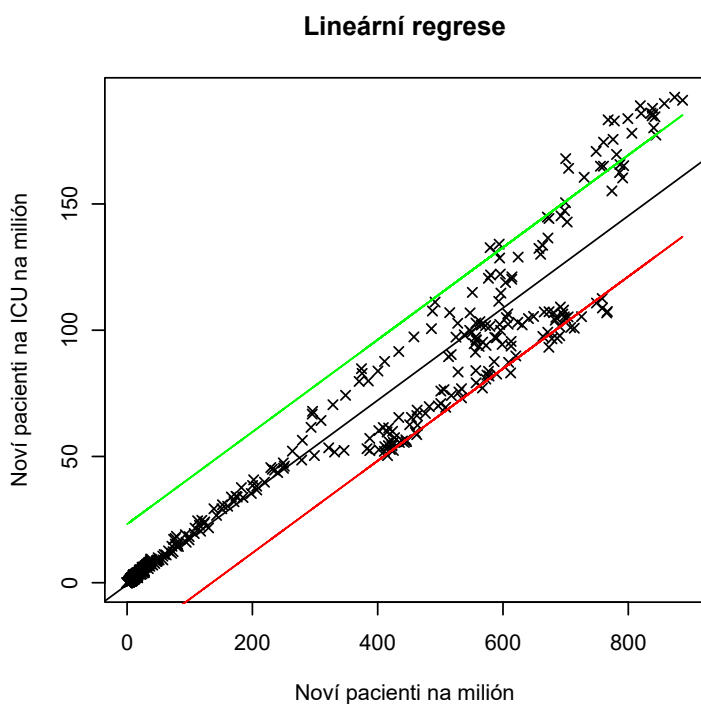
```
data: new_tests_czechia and new_cases_czechia
```

```
X-squared = 147356, df = 146982, p-value = 0.245
```

10 REGRESE

10.1 Lineární regrese

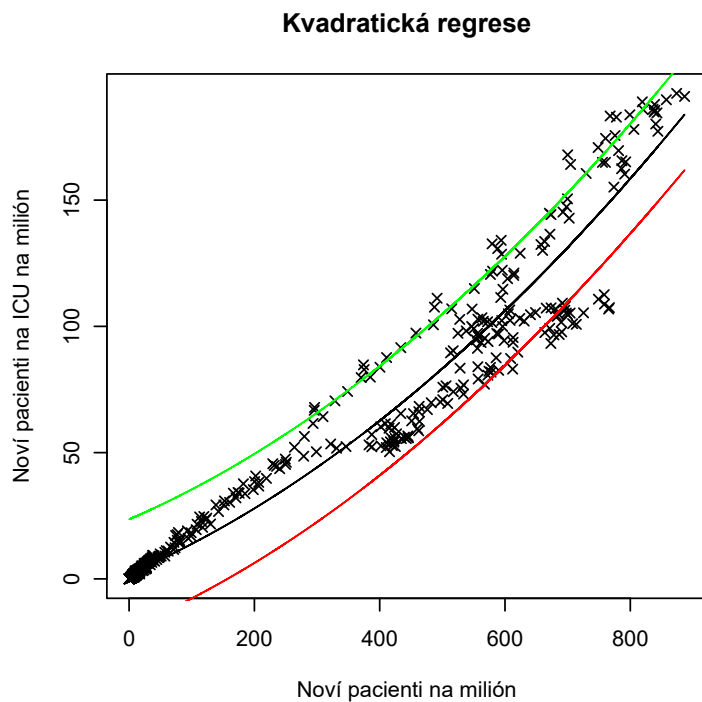
Následující graf zobrazuje jakých hodnot bude s 95% pravděpodobností nabývat hodnota pacientů na ICU na milión v ČR v závislosti na počtu nově hospitalizovaných pacientů na milión v ČR. Tato závislost je zde vyjádřena jako lineární funkce $y = -0.7746 + 0.1826x$.



Obrázek 41: Graf lineární regrese

10.2 Kvadratická regrese

Následující graf zobrazuje jakých hodnot bude s 95% pravděpodobností nabývat hodnota pacientů na ICU na milión v ČR v závislosti na počtu nově hospitalizovaných pacientů na milión v ČR. Tato závislost je zde vyjádřena jako kvadratická funkce $y = 1.9982244 + 0.1074610x + 0.0001102x^2$.



Obrázek 42: Graf kvadratické regrese

ZÁVĚR

V této semestrální práci byl analyzován vývoj epidemie nemoci Covid-19 v ČR. Jak je zřejmé z grafů, zkoumaná data se neřídí dle normálního rozdělení pravděpodobnosti a zpravidla jsou výrazně zešikmena vlevo. Pomocí grafů byly porovnány sedmidenní klouzavé průměry nových případů na milion v Rakousku a České republice. Pomocí grafů byly také vizualizovány další veličiny jako například počet nových hospitalizací, počet nových testů, reprodukční číslo, počet pacientů na ICU či pozitivita testů. Pomocí testů bylo otestováno například zda se střední hodnota nových případů rovná konkrétní hodnotě či zda se střední hodnota nových případů výrazněji v průběhu času změnila. Tyto analýzy byly protě vzhledem k sešikmení dat provedeny kromě t-testu také pomocí Wilcox testu, jelikož by jeho výsledky zešikmení dat nemělo případně tolik ovlivnit. Pro řádné srovnání nových případů na milion je poté na data aplikován test ANOVA, který tuto veličinu porovnává mezi ČR, Německem, Slovenskem, Polskem a Rakouskem. Nakonec je pomocí regrese navržena lineární a kvadratická funkce popisující možnou závislost přírůstku nových pacientů na ICU na milion na přírůstku nově hospitalizovaných pacientů na milion. Těmito operacemi práce jistě poskytuje bližší pohled na vývoj současné epidemie v naší zemi jakož i nebezpečí, které tento virus představuje.

POUŽITÁ LITERATURA

- [1] Our World in Data *Data on COVID-19 (coronavirus)* [online]. 2021 [cit. 2021-11-18]. Dostupné z: <https://github.com/owid/covid-19-data/tree/master/public/data>

SEZNAM PŘÍLOH

Příloha A	??
-----------------	----

PŘÍLOHA A

Příloha A zahrnuje ZIP soubor, který obsahuje:

- Zdrojové kódy
- Zdrojová data použita v práci