



R PROGRAMMING LANGUAGE

VOJTĚCH VRÁNA
LUKÁŠ PEČINKA

Jazyk R

- programovací jazyk
- statistická analýza
- zpracování,
vyhodnocení
a vizualizace dat
- zdarma ke stažení
- velké množství balíčků



```
R Console

R version 4.2.2 (2022-10-31 ucrt) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```



R

není user-friendly (zobrazuje jen to, o co je požádán)

RStudio



VÝVOJOVÉ
PROSTŘEDÍ



INTERAKTIVNÍ



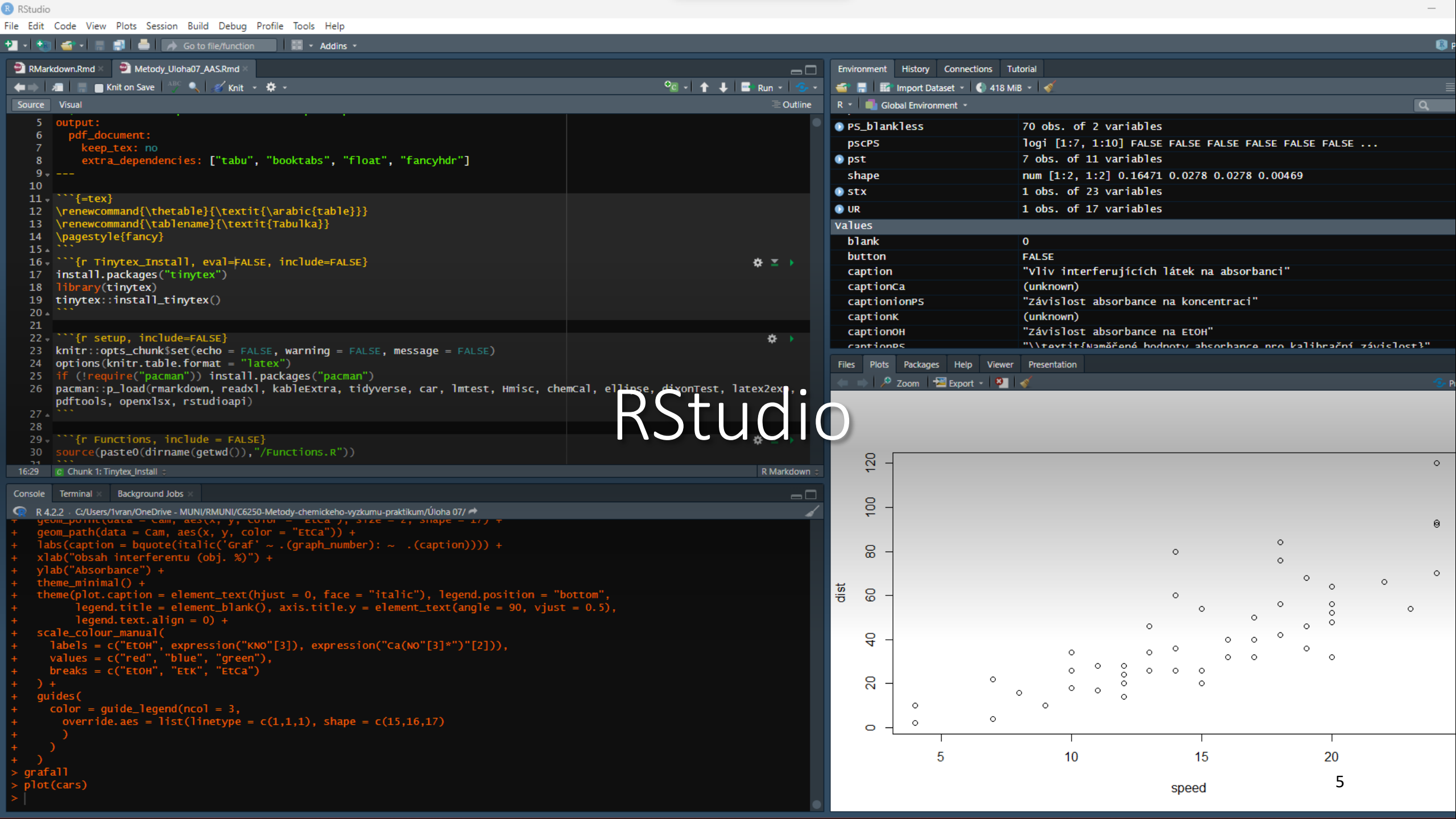
PŘEHLEDNĚJŠÍ



JEDNODUŠŠÍ



LEPŠÍ





Možnosti na MU

- Mgr. Vít Syrovátka, Ph.D.:
 - PŘF:Bi7560 Úvod do R (podzimní semestr)
 - PŘF:Bi8190 Vizualizace biologických dat (podzimní semestr)
 - PŘF:Bi7540 Zpracování dat v ekologii společenstev (jarní semestr)
 - PŘF:Bi7550 Analýza dat v ekologii společenstev v programu R (jarní semestr)
 - <https://vitsyrovatka.info/doku.php>
- Mgr. Eva Budinská, Ph.D.: PŘF:Bi7527 Analýza dat v R (jarní semestr)
- Mgr. Ing. Lubomír Prokeš, Ph.D.: PŘF:C2175 Využití programu R pro výpočty a zpracování dat v chemii (podzimní semestr)

Vyhodnocení

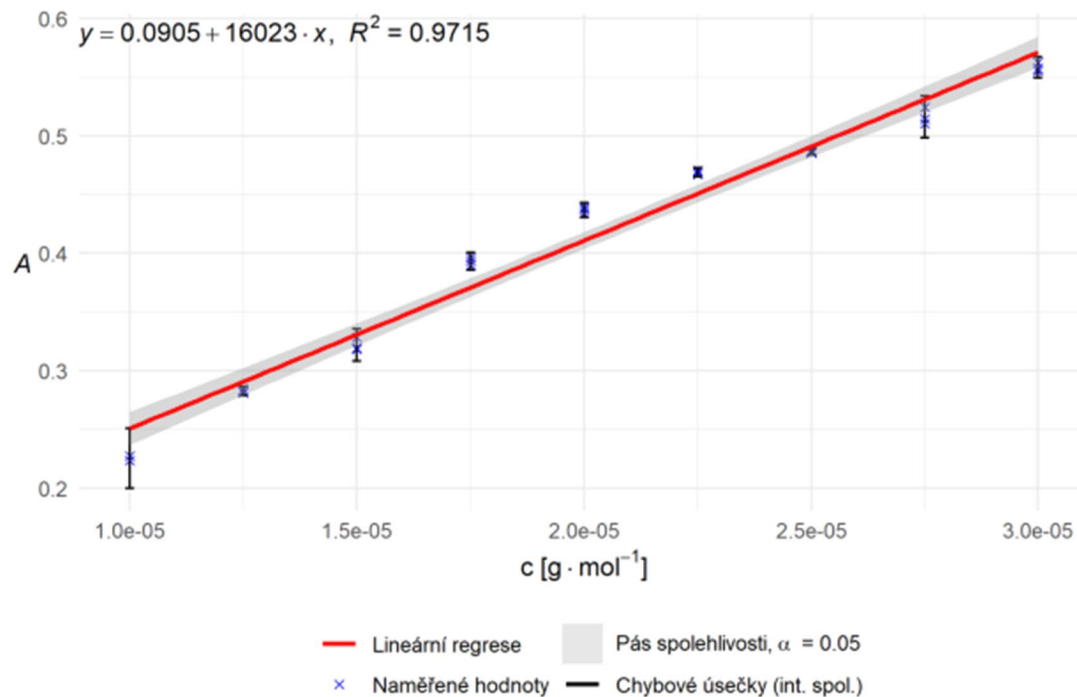
Kalibrační závislosti

Absorbance

Tabulka 2: Naměřené hodnoty absorbance pro kalibrační závislost

c [g·mol ⁻¹]	A		
1.00e-05	0.227	0.223	0.138
1.25e-05	0.284	0.281	0.282
1.50e-05	0.318	0.328	0.319
1.75e-05	0.390	0.396	0.393
2.00e-05	0.434	0.439	0.437
2.25e-05	0.467	0.469	0.470
2.50e-05	0.486	0.487	0.485
2.75e-05	0.524	0.510	0.514
3.00e-05	0.562	0.557	0.555

Červeně označené jsou hodnoty vyloučené ze souboru dat na základě Grubbsova testu.**



Graf 1: Závislost absorbance na koncentraci

Neznámý vzorek

Tabulka 5: Absorbance a vyhodnocení pro neznámý vzorek

Absorbance	Vypočtené hodnoty	
0.409	Směrodatná odchylka s_x	$6.25 \cdot 10^{-7} \text{ mol} \cdot \text{dm}^{-3}$
0.406		
0.404		
0.409	Interval spolehlivosti	$(1.98 \pm 0.17) \cdot 10^{-5} \text{ mol} \cdot \text{dm}^{-3}$
0.408		
0.404		

Soubor dat nevykazuje hodnoty, které by byly odlehle na základě Deanova-Dixonova Q-testu.*

Výpočet směrodatné odchylky s_x :

$$s_x = \frac{s_{y,x}}{b} \cdot \sqrt{\frac{1}{n} + \frac{1}{m} + \left(\frac{s_b}{b \cdot s_{y,x}}\right)^2 \cdot (\bar{y}_m - \bar{y})^2}$$

kde: $s_{y,x}$ je směrodatná odchylka bodů kolem regresní přímky,
 b je směrnice přímky,
 n je počet bodů v kalibrační přímce,
 m je počet paralelních měření,
 s_b je směrodatná odchylka směrnice,
 \bar{y}_m je průměrná hodnota signálu vzorku,
 \bar{y} je hodnota y v těžišti kalibrační přímky.

Závěr

Byly sestaveny kalibrační závislosti absorbance, fluorescence při 405 nm a fluorescence při 500 nm na koncentraci. Z rovnic regresních přímk byly vypočteny koncentrace fluoresceinu ve vzorku, směrodatná odchylka s_x a intervaly spolehlivosti.

Dále byly vypočteny meze detekce a meze stanovitelnosti měření. Pro neznámý vzorek také byla otestována shodnost výsledků stanovení pro F_{405} a F_{500} a bylo zjištěno, že výsledky se shodují.

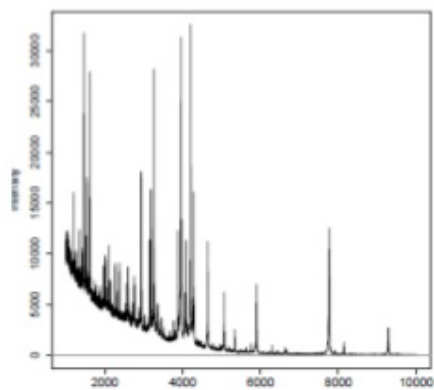
Tabulka 9: Výsledky stanovení obsahu fluoresceinu ve vzorku

	Absorbance	Fluorescence (405 nm)	Fluorescence (500 nm)
		mol · dm ⁻³	
Koncentrace	$(1.98 \pm 0.17) \cdot 10^{-5}$	$(2.04 \pm 0.25) \cdot 10^{-5}$	$(1.94 \pm 0.20) \cdot 10^{-5}$
Mez detekce	$2 \cdot 10^{-7}$	$1.4 \cdot 10^{-7}$	$8 \cdot 10^{-7}$
Mez stanovitelnosti	$7 \cdot 10^{-7}$	$5 \cdot 10^{-7}$	$3 \cdot 10^{-6}$

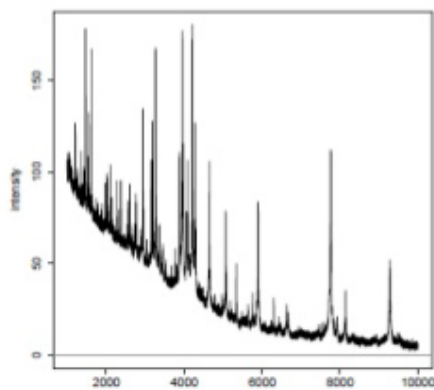
Závislosti nejsou dokonale lineární, což mohlo být způsobeno chybným pipetováním kalibračních roztoků.

(1) Further **stream-lining** of the **MALDI-TOF scoring of wheat proteins** we are following the **MALDIquant** process described by Sebastian Gibb (IMISE, University of Leipzig)

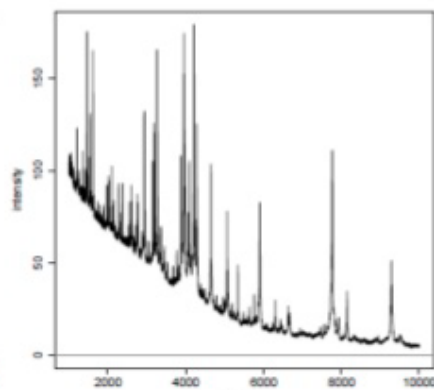
1: raw



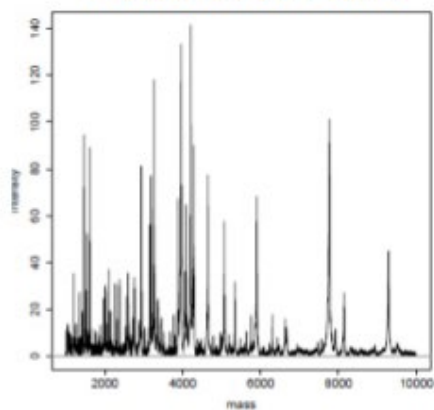
2: variance stabilization



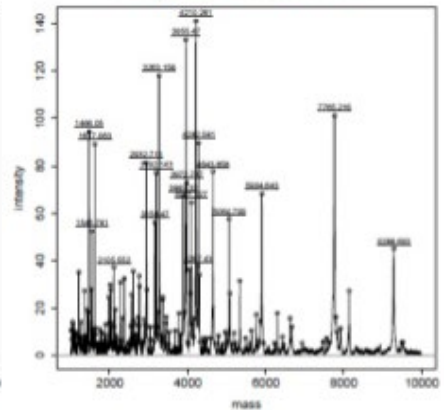
3: smoothing



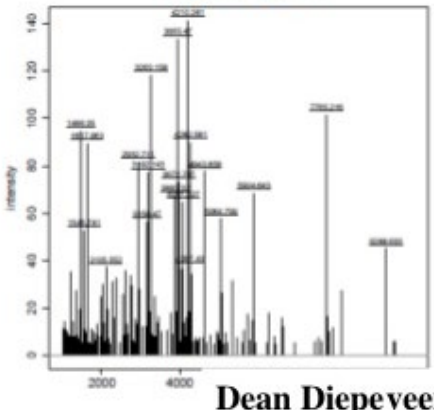
4: base line correction



5: peak detection

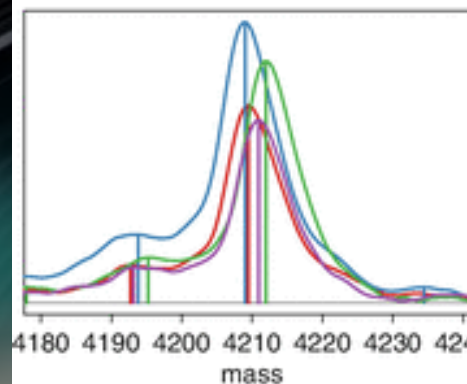


6: peak plot

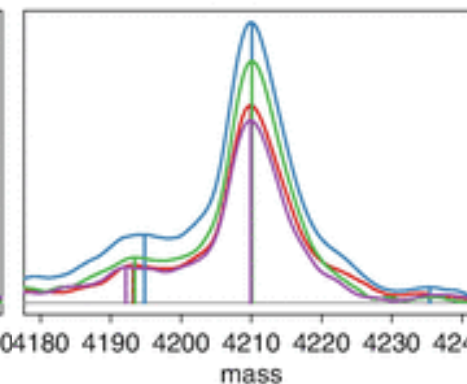


Dean Diepeveen

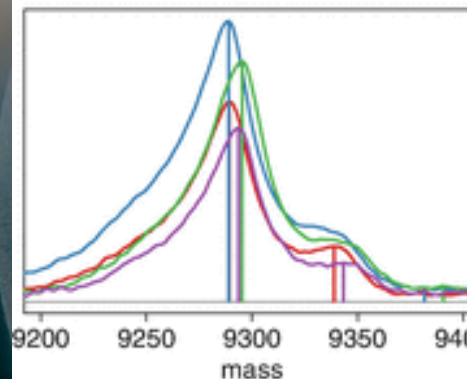
unwarped



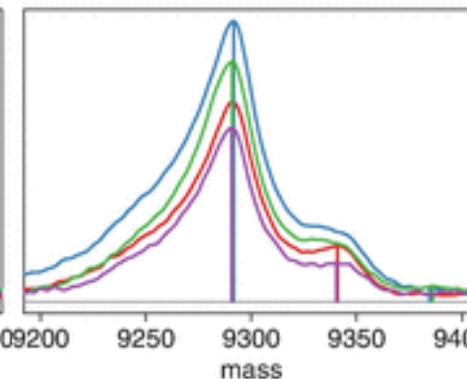
MALDIquant



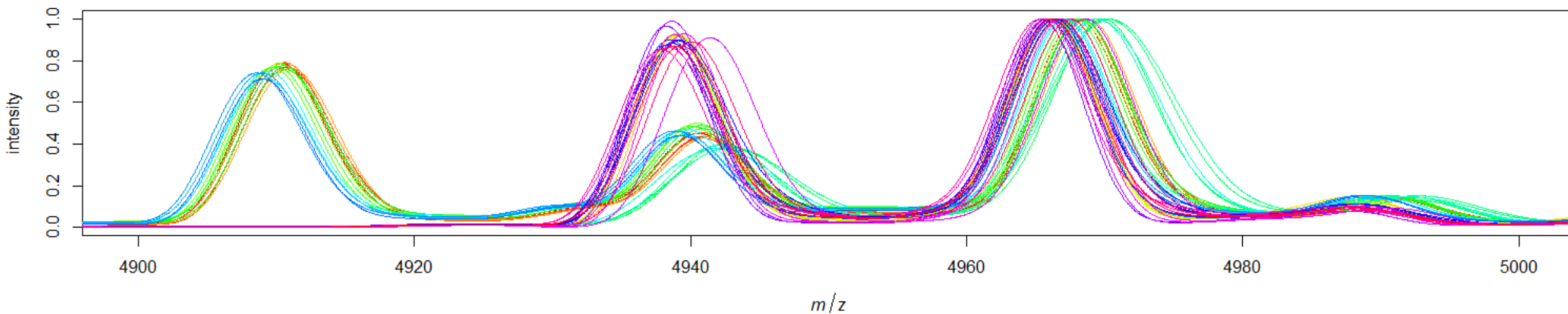
unwarped



MALDIquant

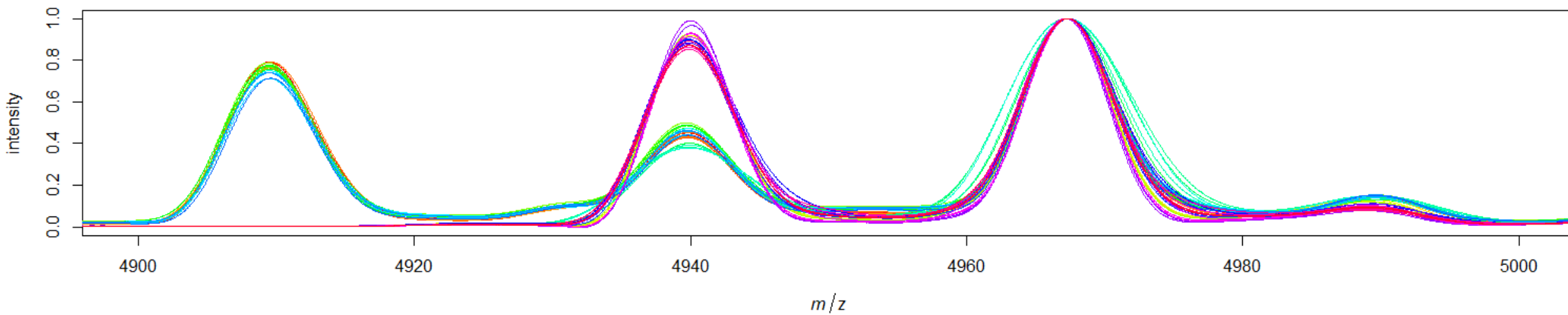


EM-10-015 unwarped spectra (mass 4900:5000 Da)

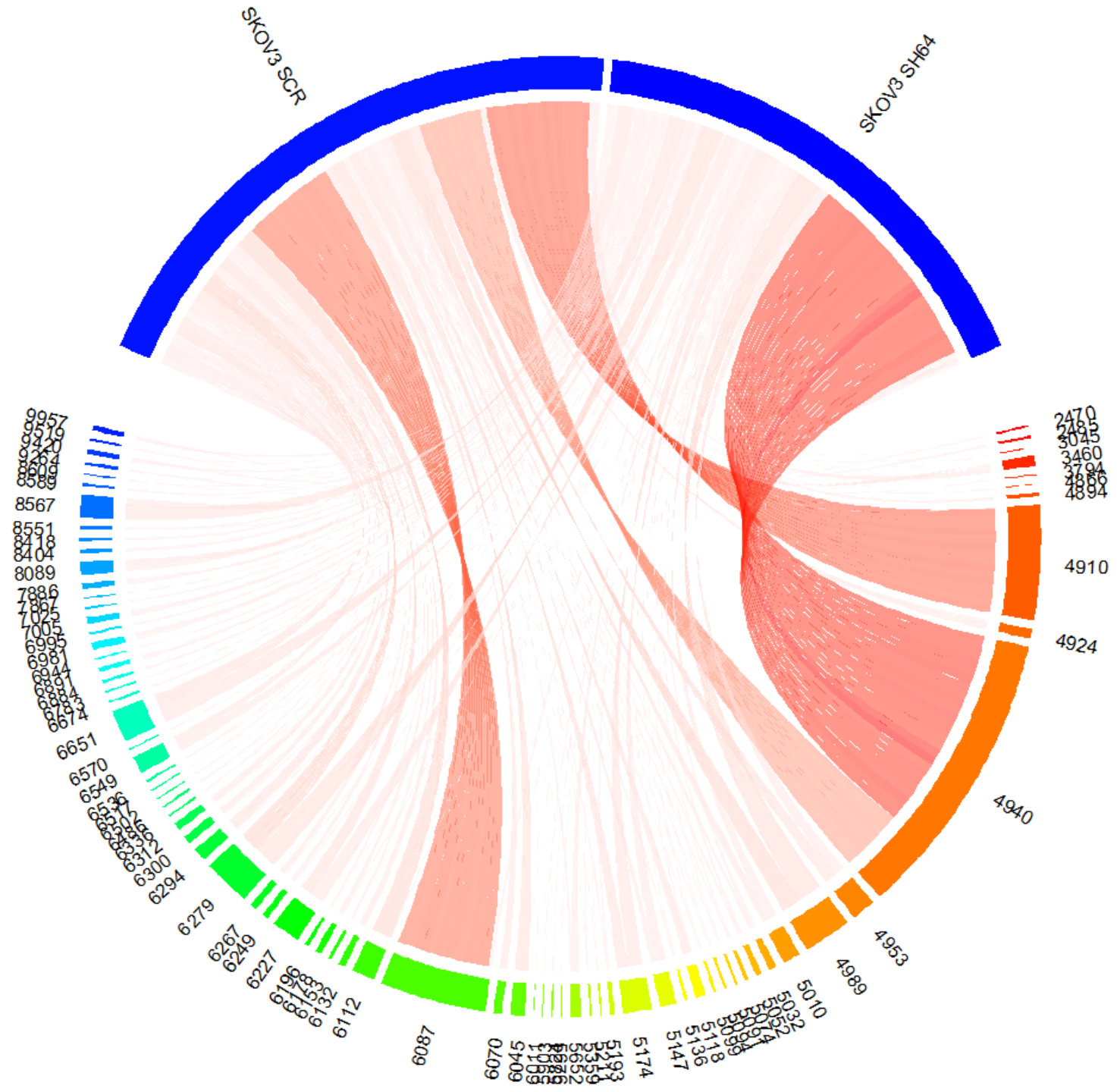


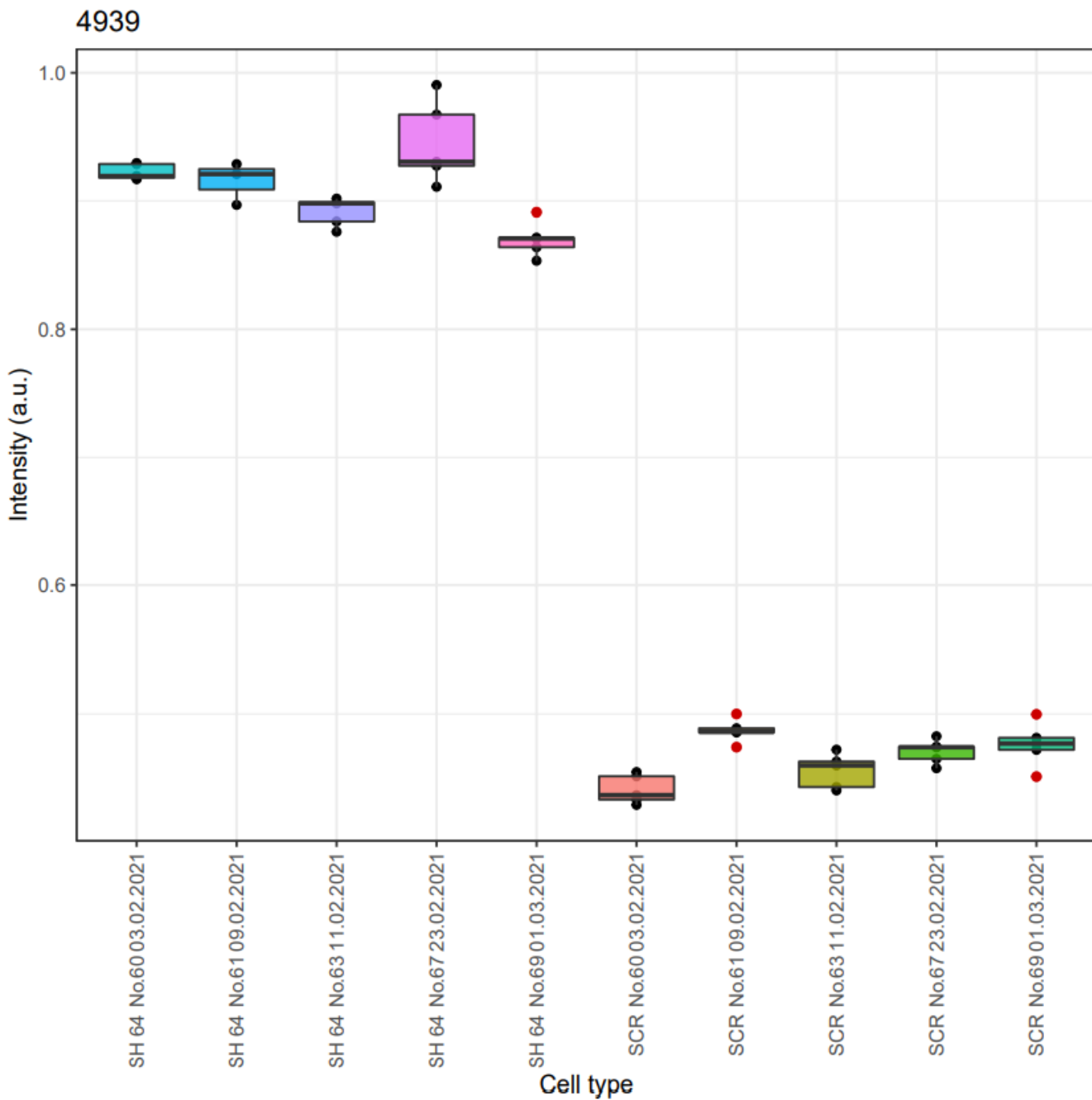
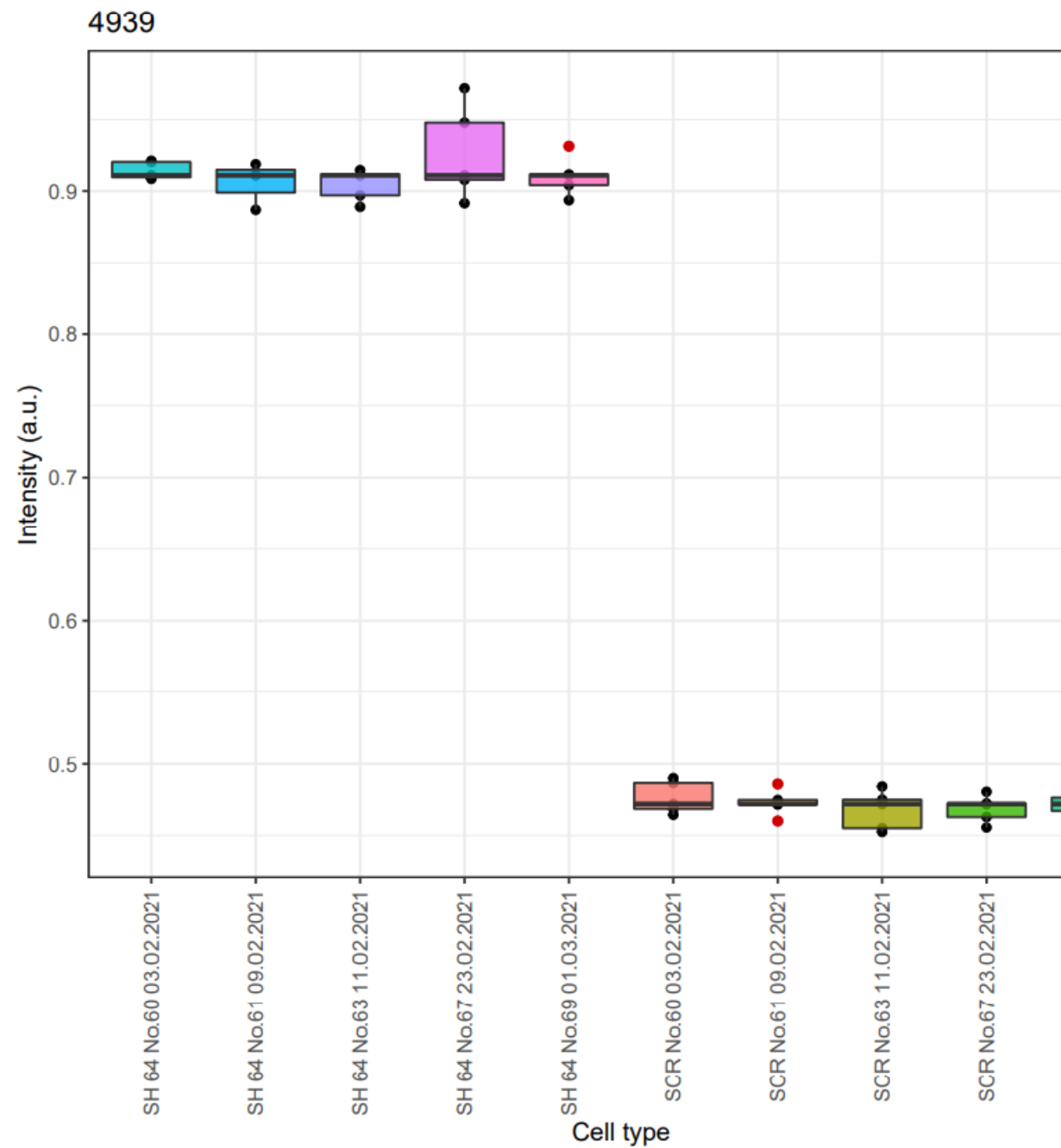
D:\ANALYZA\ANALYZA MS DAT ICIDATA\21-02-03_08-59_1A1.mzml

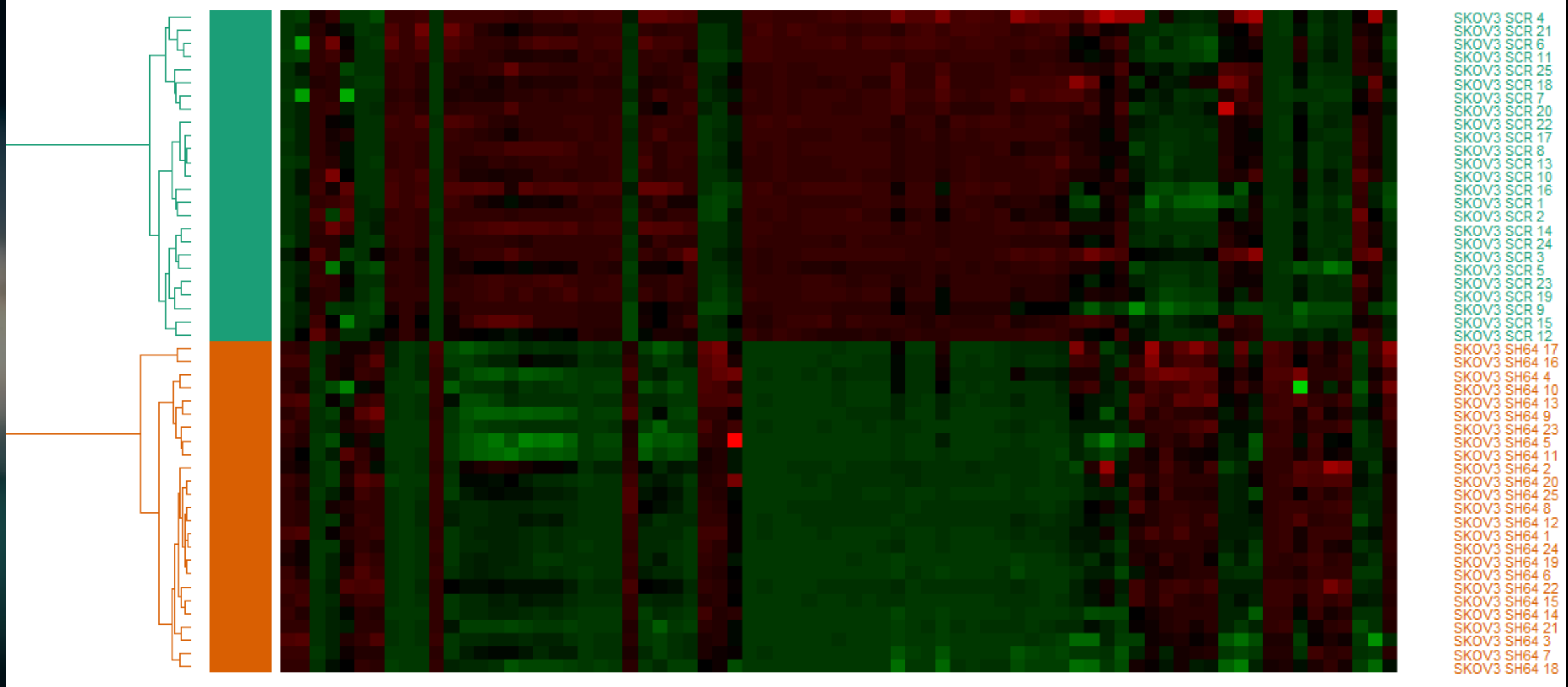
EM-10-015 warped spectra (mass 4900:5000 Da)



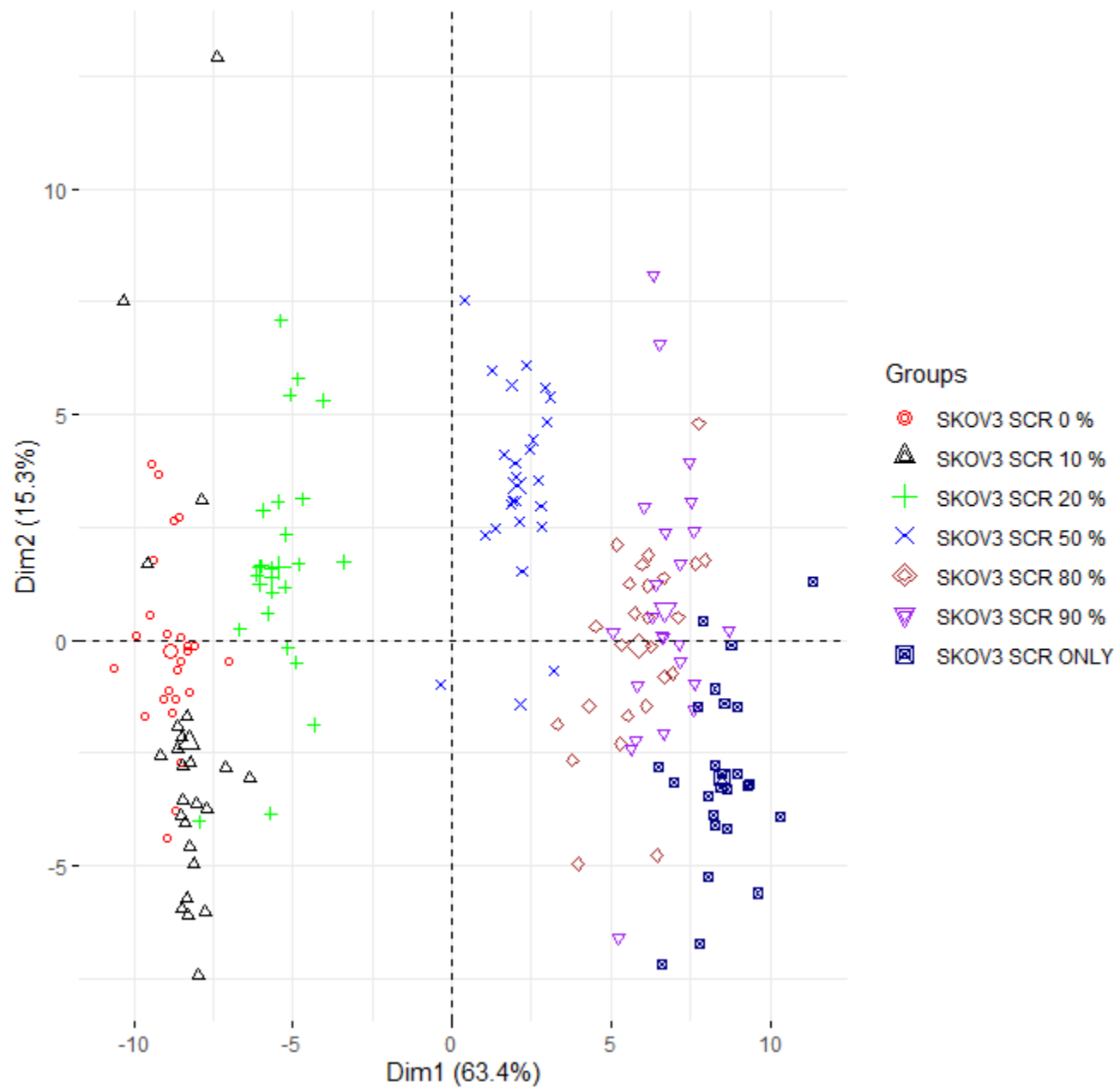
D:\ANALYZA\ANALYZA MS DAT ICIDATA\21-02-03_08-59_1A1.mzml

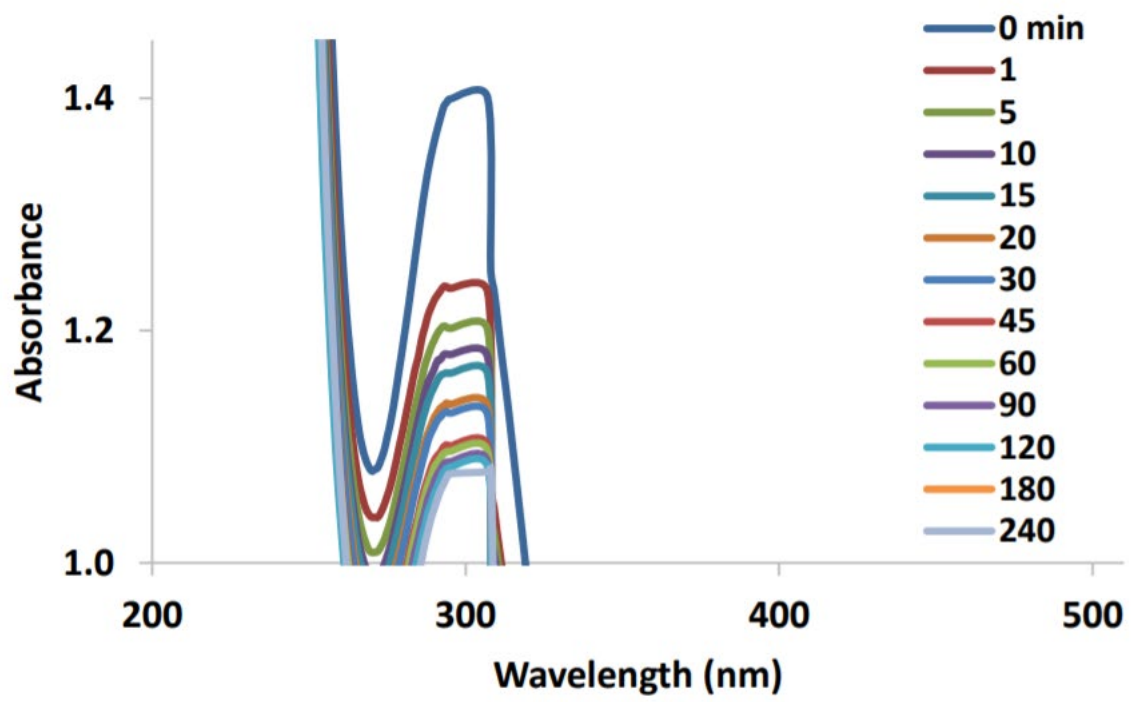


A) RAW DATA**B) MED CENTRED DATA**



Individuals - PCA





UV-Vis spectra concerning the adsorption of H_{AuCl}₄ on the Fe-BTC as a function of time

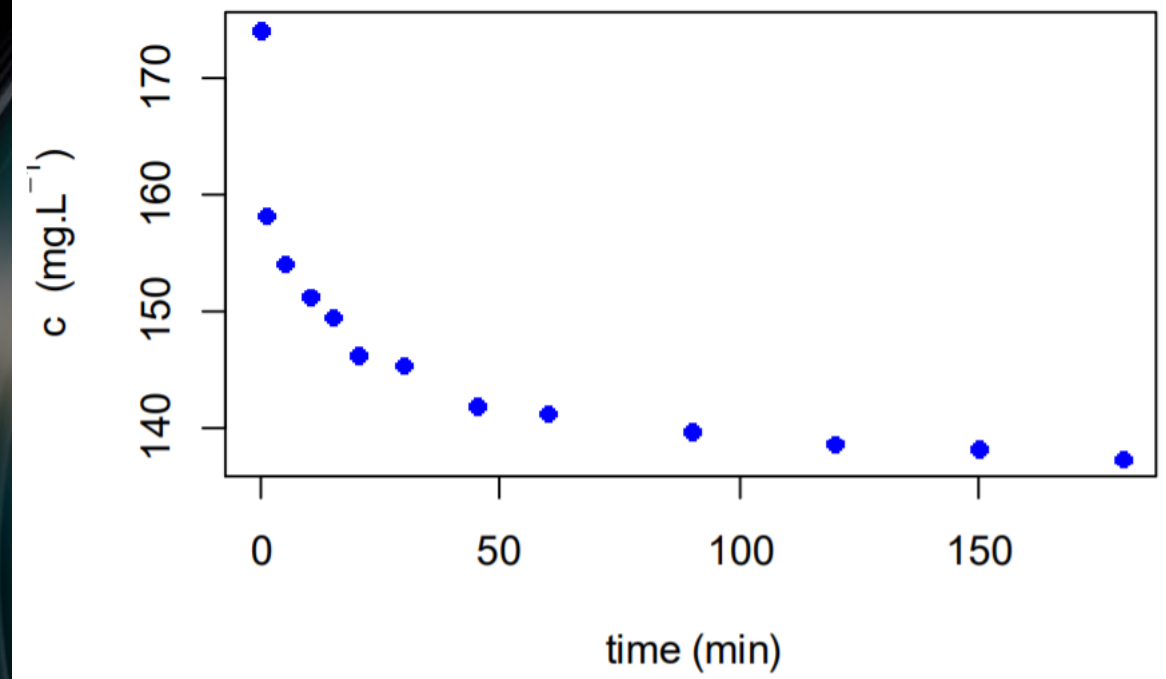


Fig. 22 Decrease of the H_{AuCl}₄ concentration in time

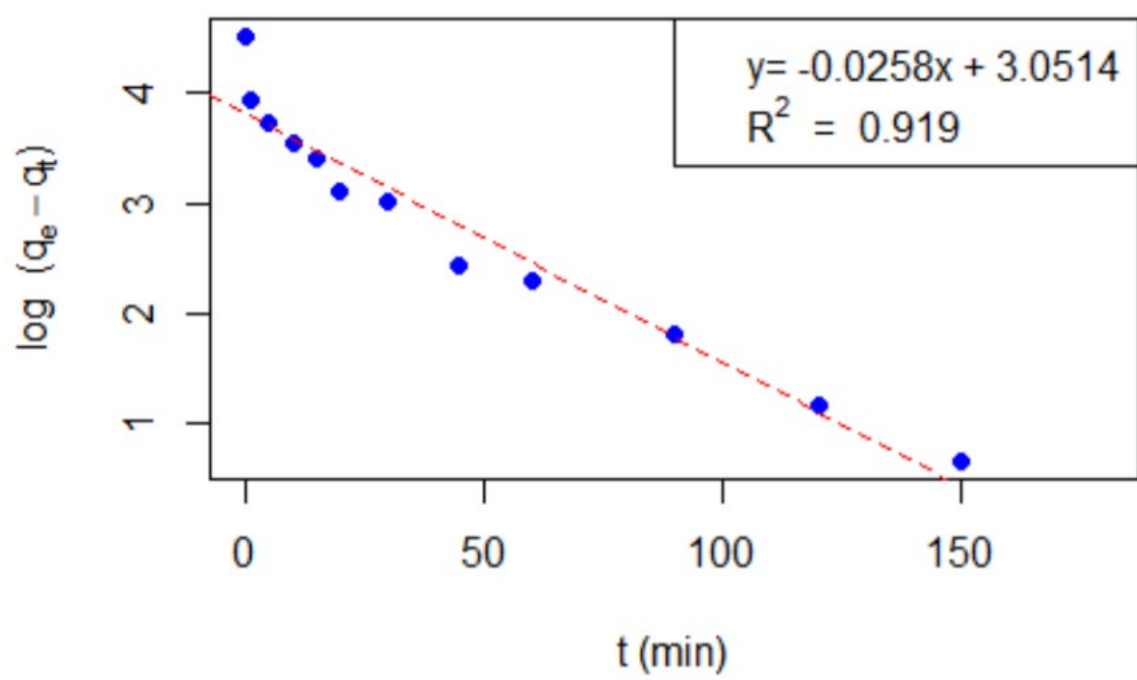


Fig 24 Pseudo-first-order kinetic model

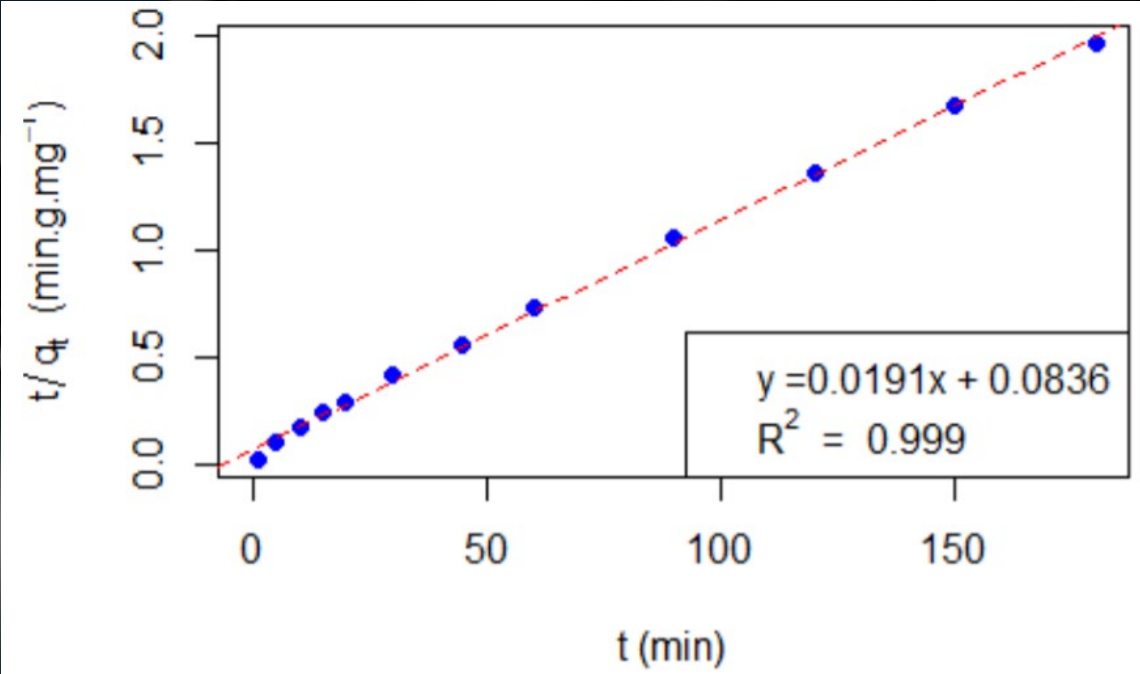
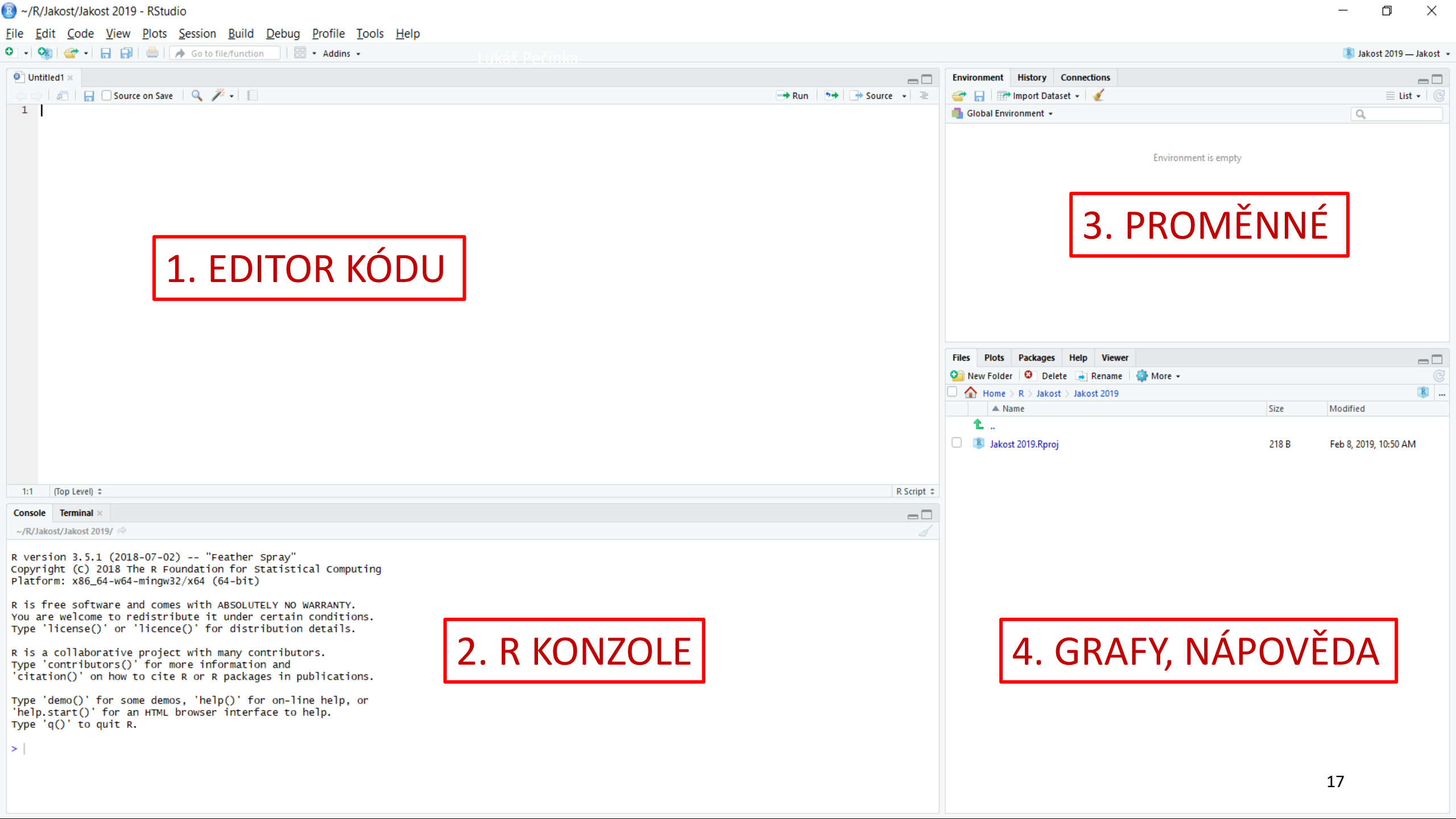


Fig. 25 Pseudo-second-order kinetic model

Table 2 Kinetic parameters for the adsorption of H₂AuCl₄ on the Fe-BTC

$q_{e,exp}$ (mg.g ⁻¹)	Pseudo-first-order			Pseudo-second-order		
	$q_{e,cal}$ (mg.g ⁻¹)	k_1 (min ⁻¹)	R^2	$q_{e,cal}$ (mg.g ⁻¹)	k_2 (g.mg ⁻¹ .min ⁻¹)	R^2
91.755	46.095	5.227×10^{-2}	0.945	93.673	1.644×10^{-3}	0.999



1. EDITOR KÓDU

3. PROMĚNNÉ

2. R KONZOLE

4. GRAFY, NÁPOVĚDA



Základní příkazy

- Přiřazení hodnoty do proměnné: `a <- 1`
- Přiřazení více hodnot: `a <- c(1,2,3)`
- Komentář: `a <- 1 # Můj milý deníčku,`
- Cesta k uloženému projektu: `getwd()`
- Soubory ve složce: `dir()`
- Ukázky funkcí: `demo()`
- Smazat proměnné: `rm(a)`
- Nápořěda k funkci: `?funkce`
- Ukončení: `q()`

Základní pravidla

- Rozlišuje velká a malá písmena
- Názvy proměnných bez mezer
- Matematické operátory +, -, *, /
- Logické hodnoty TRUE / FALSE
- Desetinná tečka



Statistika

- Posloupnost: `seq(from, to, by)`
- Aritmetický průměr: `mean()`
- Median: `median()`
- Rozpětí: `range()`
- Rozptyl: `var()`
- Směrodatná odchylka: `sd()`
- Extrémy: `max()`, `min()`
- Součet: `sum()`
- Seřazení: `sort()`
- Souhrnné info: `summary()`

$(a+b)^2 = a^2 + 2ab + b^2$
 $2\pi r = 22$
 $\sum_{n=0}^{\infty} \frac{x^n}{n!}$
 $(y+A) = \frac{2}{3}A$
 $6x + 3y = 12$
 $3x + y = 6$
 $t = \sqrt{\frac{a}{8}}$
 $y = \frac{\Delta x}{\Delta 3} a$
 $\int \sqrt{a^2 - x^2} dx = \frac{x}{2} + \frac{a^2}{2} \arcsin \frac{x}{a}$
 $d^2c = \pi v^2$
 $\alpha A = (1 + u^2 + v^2)$
 $a^2 = b^2 + c^2 \neq 2bc \cos$
 $V = 2RV$
 $\frac{df}{dx} = \frac{\partial f}{\partial y} x + \frac{\partial f}{\partial y_{xx}}$
 $\frac{1}{2} = \frac{1}{2}$
 $\frac{1}{3} + x^2 - 3x$

Matematika

- sčítání +
- odčítání -
- násobení *
- dělení /
- mocnina ^
- rovno ==
- není rovno !=
- větší, menší <, >, <=, >=



Číslíčka

- Logaritmus **log**(číslo, základ)
 - $\log(e) = 1$
 - $\log_{10}(10) = 1$
 - $\log(10, 10) = 1$
- Zaokrouhlení **signif**(x, digits = 2)
- Ludolfovo číslo **pi**
- Eulerovo číslo **exp**(1)
- Seznamy: **letters**, **month.name**



**POLICIE
V AKCI**

Příklady
POLICE
836412

Absorbance

$$\epsilon_{\lambda} = \frac{A_s}{b \cdot c_s}$$

- Vytvořte nový projekt „Absorbance“.
- Zjistěte, kde je umístěný a jaké soubory adresář obsahuje.
- Roztok manganistanu draselného ($M = 158,03$) o hmotnostní koncentraci $4,48 \text{ mg/l}$ vykazuje při vlnové délce 520 nm v 1 cm kyvetě absorbanci $0,510$. Jaká je hodnota molárního absorpčního koeficientu?
- Zaokrouhlete výsledek, aby byl bez desetinných čísel.
- Vygenerujte report.
- Postupujte tak, že nejprve vytvoříte objekty „A“, „b“ a „c“, které ponese hodnoty absorbance, tloušťky absorpční vrstvy a látkové koncentrace. Pak proveďte výpočet pomocí těchto objektů.



Absorbance

```
# Definovani promennych
```

```
A <- 0.510
```

```
b <- 1
```

```
c <- (4.48/1000)/158.03
```

```
# Vypocet koeficientu
```

```
E <- A/(b*c)
```

```
E
```

```
## [1] 17990.02
```

```
# Zaokrouhleni vysledku
```

```
E <- signif(E, digits = 5)
```

```
E
```

```
## [1] 17990
```

Tabulka

- `data.frame()` = tabulka
- `row.names = c(„název_řádku“,...)`
- `[řádek,sloupec]`

```
##   jmena vyska povolani  
## a  Ales   180  student  
## b  Tomas  175  student  
## c   Eva   160  student
```

Tabulka

```
# Zadani hodnot

jmena <- c("Ales", "Tomas", "Eva")
vyska <- c(180, 175, 160)
povolani <- rep("student", 3)

# Vytvoreni tabulky

seznam <- data.frame(jmena, vyska, povolani)
row.names(seznam) <- c("a", "b", "c")

# Nacteni tabulky

seznam
```

```
##   jmena vyska povolani
## a  Ales   180  student
## b  Tomas  175  student
## c   Eva   160  student
```

Koncentrace

řádky pojmenované: „první“, „druhý“, „třetí“

sloupce pojmenované: „koncentrace“ a „signál“

	koncentrace	signal
prvni	1.003	0.2
druhy	1.030	0.5
treti	1.123	1.2

Koncentrace

```
# Zadani hodnot

koncentrace <- c(1,2,5)
signal <- c(0.2,0.4,1)

# Vytvoreni tabulky

tabulka <- data.frame(koncentrace, signal)

# Nazvy radku

names <- c("prvni", "druhy", "treti")
row.names(tabulka) <- names

# Nacteni tabulky

tabulka
```

```
##      koncentrace signal
## prvni           1    0.2
## druhy           2    0.4
## tretí           5    1.0
```

Grafy

- Vykreslování os: **axes** = T/F
- Logaritmické měřítko os: **log** = „x,y“
- Název grafu: **main** = „název“
- Typ grafu: **type** = „p, l, b“
- Znak bodu: **pch** = 1-20
- Velikost bodu: **cex** = 69
- Barva: **col** = „duhová“
- Název osy x: **xlab** = „X“
- Hranice os: **xlim** = c(min, max)
- Font: **font** = 1-4
- Orámování grafu: **frame.plot** = T/F
- Vykreslení grafu: **plot(x)**, **plot(x,y)**
- Sloupcový graf: **barplot()**, **hist()**



```
## Warning: package 'readxl' was built under R version 3.5.3
```

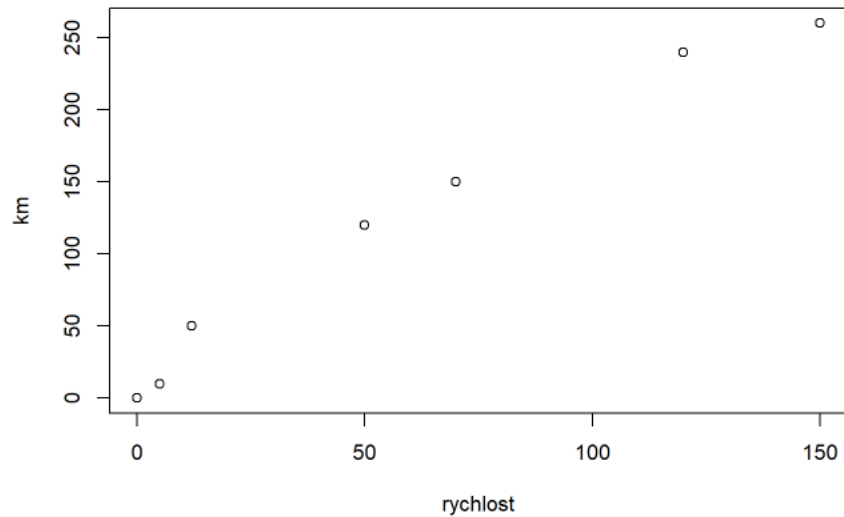
```
x<-read_excel("a.xlsx", sheet = "List1")  
getwd()
```

```
## [1] "C:/Users/Lukas/Documents/R/Jakost/Jakost 2019"
```

```
dir()
```

```
## [1] "a.xlsx" "aaaa.pdf"  
## [3] "Absorbance.html" "Absorbance.R"  
## [5] "Demograph.xlsx" "exponencialni funkce.R"  
## [7] "exponencialni_funkce.html" "Jakost 2019 tabulka.R"  
## [9] "Jakost 2019.Rproj" "Jakost_2019_tabulka.html"  
## [11] "Jakost_2019_tabulka_2.html" "Jakost_2019_tabulka_2.R"  
## [13] "kalibrace.xlsx" "lineární regrese.R"  
## [15] "lineární_regrese.html" "uvod-grafika.html"  
## [17] "uvod-grafika.R" "uvod-grafika.spin.R"  
## [19] "uvod-grafika.spin.Rmd" "uvod grafika.R"
```

```
plot(km~rychlost, x)
```



```
plot(x$rychlost, x$km, axes = T, main = "Závislost rychlosti na počtu ujetých km", type = "p", pch="$", col="bl  
ue", cex=2,  
xlab = "rychlost [km]", xlim = c(0,250), frame.plot = F)  
points(150,100, cex=5)  
model<-lm(km~rychlost, u)
```

```
points(150,100, cex=5)
model<-lm(km~rychlost, x)
abline(model, col="red", lty=2)
model
```

```
##
## Call:
## lm(formula = km ~ rychlost, data = x)
##
## Coefficients:
## (Intercept)      rychlost
##      16.031         1.764
```

```
summary(model)
```

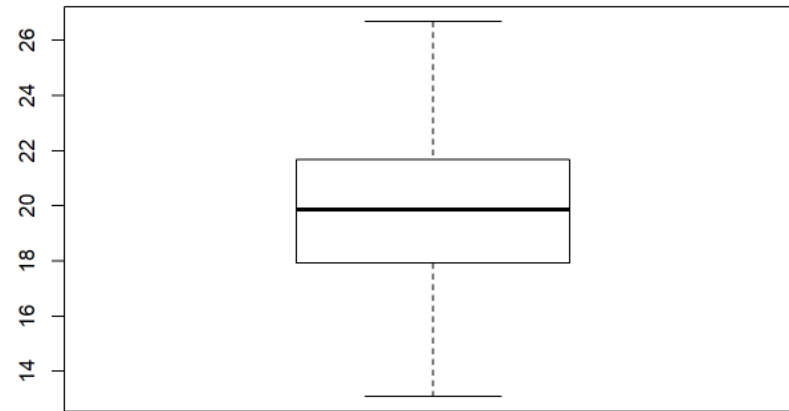
```
##
## Call:
## lm(formula = km ~ rychlost, data = x)
##
## Residuals:
##      1      2      3      4      5      6      7
## -16.03 -14.85  12.81  15.79  10.52  12.34 -20.57
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  16.0307     9.8128   1.634   0.163
## rychlost     1.7636     0.1231  14.325 2.99e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.76 on 5 degrees of freedom
## Multiple R-squared:  0.9762, Adjusted R-squared:  0.9715
## F-statistic: 205.2 on 1 and 5 DF,  p-value: 2.989e-05
```

```
predikce <- predict(model, interval = "confidence", level = 0.95)
predikce
```

```
##      fit      lwr      upr
## 1  16.03073 -9.1938487  41.25532
## 2  24.84873  0.7541575  48.94330
## 3  37.19392 14.5899836  59.79787
## 4 104.21069 86.7657777 121.65561
## 5 139.48268 121.8257956 157.13956
## 6 227.66263 201.5681560 253.75711
## 7 280.57061 246.7653359 314.37588
```

```
x<-as.matrix(x)
points(x[,1], predikce[,2], type = "l", col="blue")
points(x[,1], predikce[,3], type = "l", col="blue")
```

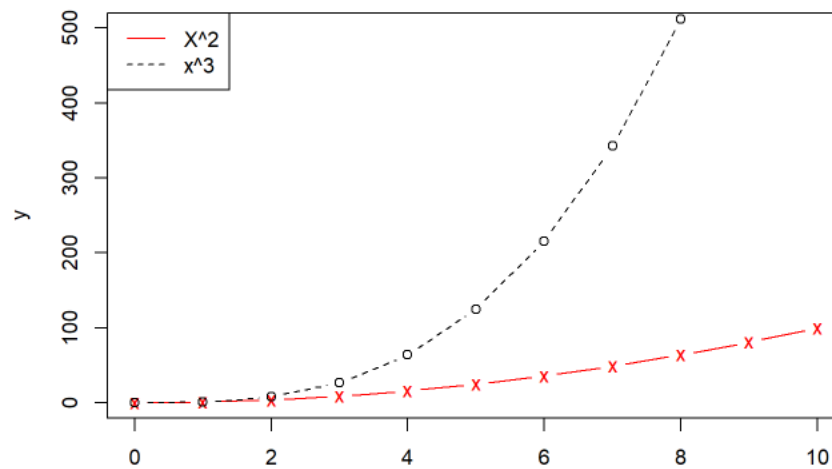
Závislost rychlosti na počtu ujetých km



```
x<-seq(0,10,1)
y<-x^2
y2<-x^3

plot(y~x, type = "b", pch = "x", col = "red", xlab = "hodnoty x", main = "Exponenciální funkce", ylim=c(0,500))
points(y2~x, type = "b", lty = 2)
legend("topleft", c("X^2", "x^3"), col = c("red", "black"), lty = 1:2)
```

Exponenciální funkce



exponencialni_funkce.R

Lukas

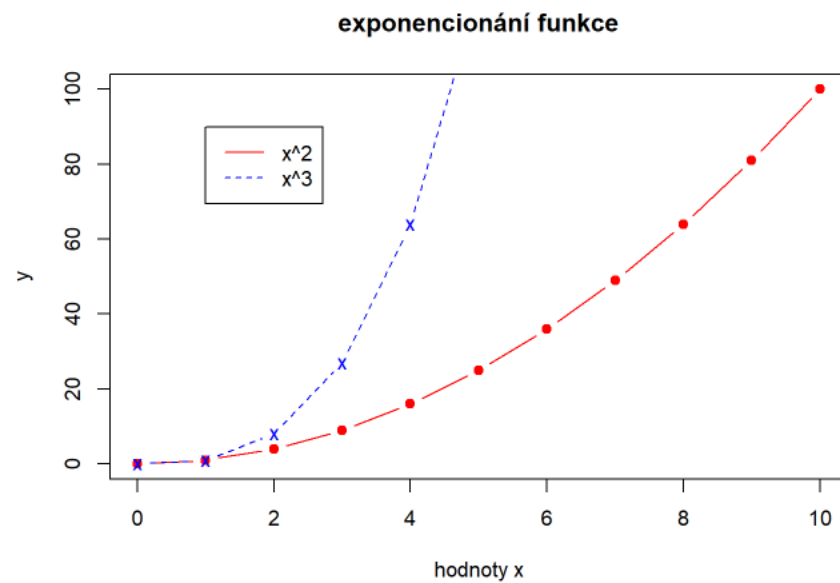
Wed Mar 06 10:47:17 2019

```
x <- seq(from = 0, to = 10, by = 1)
y <- x^2
y2 <- x^3

plot(y~x,
     type = "b",
     col = "red",
     pch = 19,
     xlab = "hodnoty x",
     main = "exponencionání funkce")

lines(y2~x,
      type = "b",
      col = "blue",
      pch = "x",
      lty = 2)

legend(1,90, c("x^2", "x^3"), col = c("red", "blue"), lty = 1:2)
```



Děkuji za pozornost!

