



# VYUŽITÍ PROGRAMU DATA MINING V ANALÝZE NÁKUPNÍHO CHOVÁNÍ

Petra Hloušková  
Stanislava Grosová



Petra Hloušková, Ústav ekonomiky a řízení, VŠCHT Praha, hlouskop@vscht.cz

# Definice funkčních potravin:



Funkční potraviny jsou potraviny, které se podobají běžným konvenčním potravinám a konzumují se jako součást obvyklé stravy, ale obsahují průkazně vyšší obsah některých fyziologicky významných složek anebo snižují riziko chronických chorob ve větší míře než běžné potraviny.

(Pokorný J., Dostálová J., 2002)



# Co má vliv na výběr potravin?



## Model výběru potravin: (podle Shepherd R., Engel Peter J.P., Olson J.C.)

### Atributy produktu:

- Kvalita
- Design obalu
- Značka
- Cena
- Přístupnost (prezentace v obchodech)
- Komunikace

### Charakteristika zákazníků, jejich kompetence a možnosti :

- Senzorické
- Psychologické
- Fyzické
- Ekonomické
- Osobní

### Akce zákazníka:

- Přijetí
- Preference
- Odmítnutí





# Hlavní cíle

- Určit segmenty
- Specifikovat spotřebitele
- Zjistiv vliv zdravotního stavu

# Využitá metodika

- Kvantitativní dotazníky
- Segmentace
- Shluková analýza



# Postup výzkumu typologie spotřebitele:



- Tvorba dotazníku
- Pracovní prostředí Witness Mineru
- Segmentace spotřebitele nakupujícího funkční potraviny
- Definice spotřebitele nakupujícího sojové produkty
- Vliv zdravotního stavu na nákupní chování spotřebitele



# Tvorba dotazníku a práce s daty



- Návaznost otázek
- Srozumitelnost otázek
- Využití škál
- Možnost kódování
- Přepis dat a export do Witnessu Mineru



# Witness Miner



- Nový nástroj pro získávání informací ze souborů dat
- Určen primárně jako samostatný analytický nástroj, možnost využití ve spojení se simulačním nástrojem WITNESS
- Identifikace skrytých závislostí mezi daty
- Silnou stránkou programu je schopnost nacházení 'důležitých info' v záplavě informací:
  - Finance: Odhalování projektů s vysokým pojišťovacím rizikem , odhalování finančních podvodů
  - Výroba: Výskyt poruch v rámci záruky, poruchy strojů, preventivní údržba
  - Zdravotnictví: Odhalování zdravotních rizik a závislostí, % úspěšné léčby
  - Obchod: Příležitosti k nabízení doplňkových produktů / služeb (X-selling)
  - Doprava: Chování zákazníků
  - Dodavatelské řetězce, obslužná centra: Výkonnostní faktory
  - Služby, obchod: Segmentace zákazníků, nabídka různých úrovní služeb

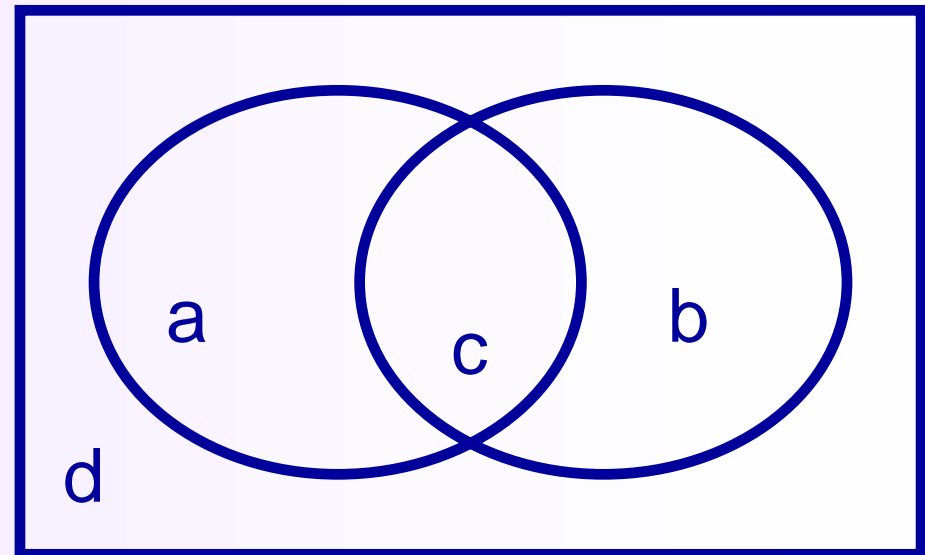


# Indukce pravidel



Pravidlo:

- IF *podmínka* THEN *třída*
  - Pakliže je splněna určitá podmínka, záznam je zařazen do určité třídy
  
- d – všechny záznamy v DB
- a – podmínka splněna
- b – záznamy patřící třídě
- c – správně klasifikované záznamy
  
- Stupeň pokrytí =  $a/d$
- Přesnost =  $c/a$



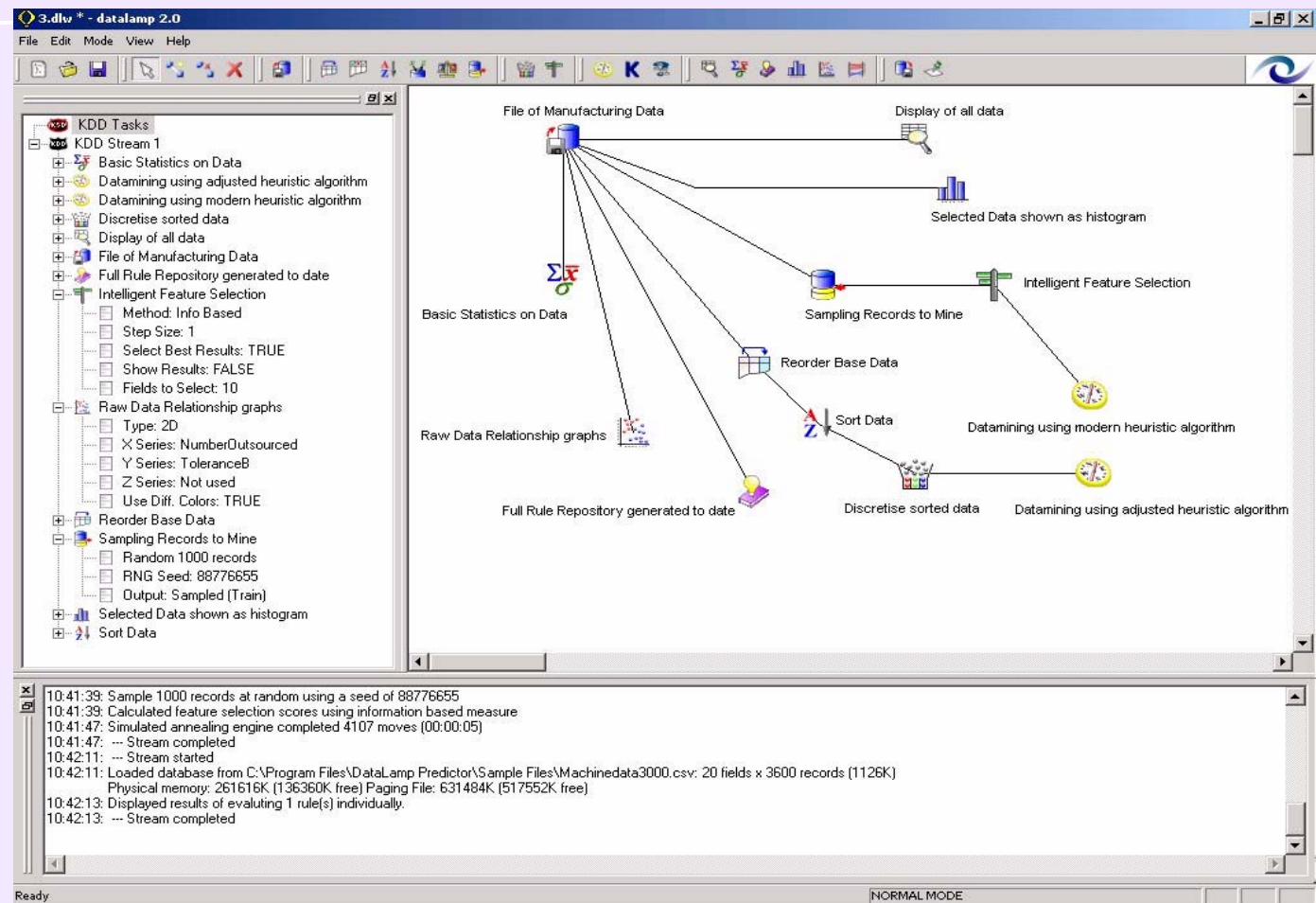


# Pracovní prostředí Witness Miner



Pracovní plocha  
obsahuje řádek s  
ikonami uzlů,  
model projektu

Uzly - operace  
jednoduše  
přetaženy na  
pracovní plochu a  
propojeny myší



# Witness Miner - segmentace



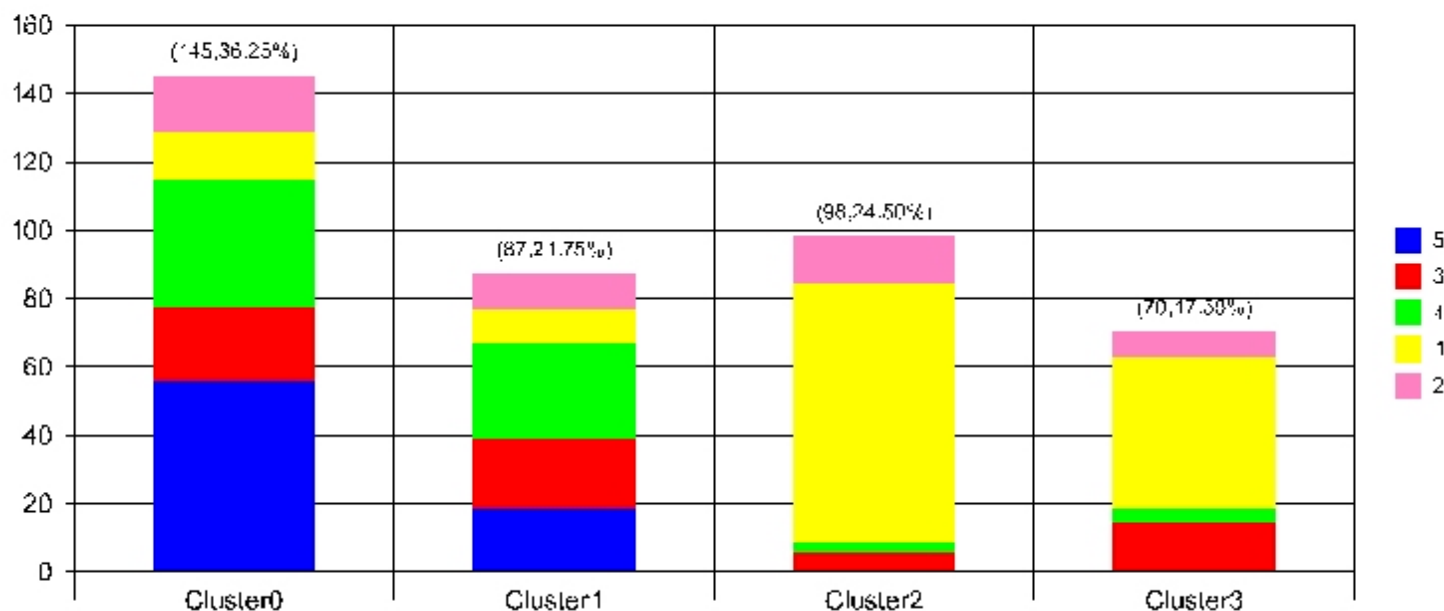
- Možnost volby hlavního kritéria
- Možnost volby počtu segmentů
- Výběr dat
- Generování hypotéz
- Ověření hypotéz
- Výstup v grafické podobě



# Rozdělení segmentů



Distribution of data in field "Cluster"



- 5: 2 – 3 x týdně □ 4: 1 x za týden □ 3: 2 – 3 x měsíčně,  
□ 2: 1 x měsíčně □ 1: nikdy



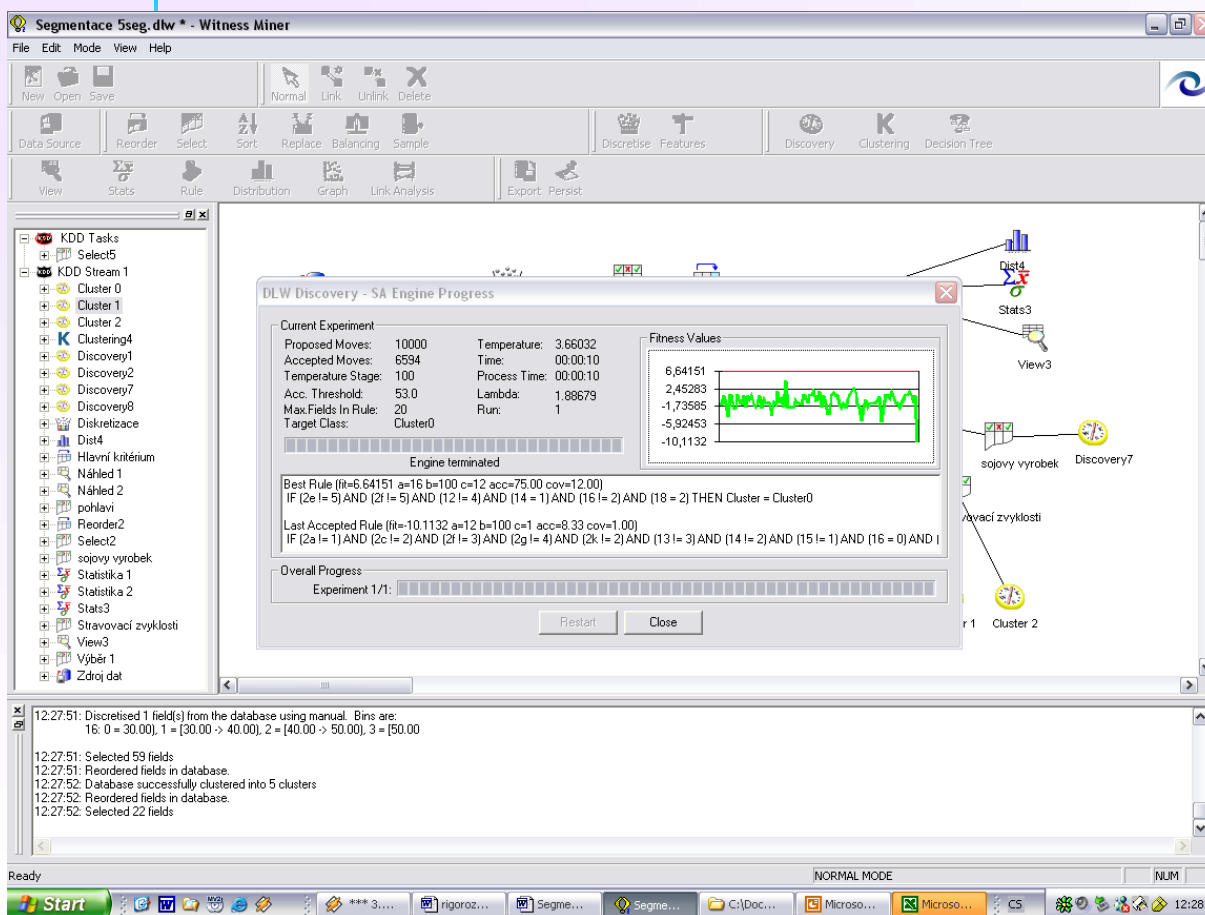
# Rozdělení segmentů



Segmenty/Návštevnost obchodu se zdravou výživou	2-3 x týdne	1 x za týden	2-3 x měsíčně	1 x měsíčně	nikdy
Cluster 0	56	22	37	14	16
Cluster 1	19	20	28	10	10
Cluster 2	0	6	3	76	13
Cluster 3	1	14	4	44	7



# Vzorové výstupy pro segmentaci



**Rule Report**  
 rule - evaluated  
 individually  
**IF** (2a = 4)  
**AND** (2e = 5)  
**AND** (2f = 4)  
**AND** (2i = 2)  
**AND** (2b = 1)  
**AND** (2h = 1)  
**AND** (2d = 2)  
**THEN** Cluster = Cluster0  
 a = 1, b = 145, c = 1  
 aM = 0, cM = 0, aAff = 0,  
 cAff = 0  
 Accuracy = 99.8,  
 Coverage = 0.69.





# Definice segmentů

## Cluster 0

- převážně v obchodech se zdravou výživou
- stravují se zdravě
- info z časopisů, internetu a prodejců
- vliv na zdraví, chuť, pohodlí, eko (bio) původ a cena
- výběr soj. výrobků je částečně vyhovující
- SŠ, pod 40 let, 1 dítě,
- bydliště do 20 000 obyvatel

## Cluster 1

- občas v obchodech se zdravou výživou
- stravují se zdravě
- info z časopisů, odborných článků
- chuť, vliv na zdraví, pohodlí, eko (bio) původ a cena
- výběr soj. výrobků je částečně vyhovující
- VŠ, pod 40 let, 2 děti,
- bydliště do 20 000 obyvatel



# Definice segmentů



## Cluster 2

- v obchodech se zdravou výživou nakupují sporadicky
- stravují se podle chuti
- info o zdravé výživě – internet nebo prodejci
- chuť a cena, méně důležitý je vliv na zdraví
- konzumace soj. výrobků jen kvůli prospěšným účinkům na zdraví
- SŠ, domácnosti, kde nejsou děti do 12 let
- bydliště nad 500 000 obyvatel

## Cluster 3

- v obchodech se zdravou výživou max. 1x měsíčně
- stravují se podle chuti
- nevyhledávají info o zdravé výživě, max. se poradí s přáteli
- chuť a cena, ostatní nedůležité
- konzumace soj. výrobků jen kvůli zdravotním omezením
- SŠ, domácnosti, kde nejsou děti do 12 let
- bydliště do 100 000 obyvatel



# Důsledky vyplývající z vyhodnocení jednotlivých segmentů

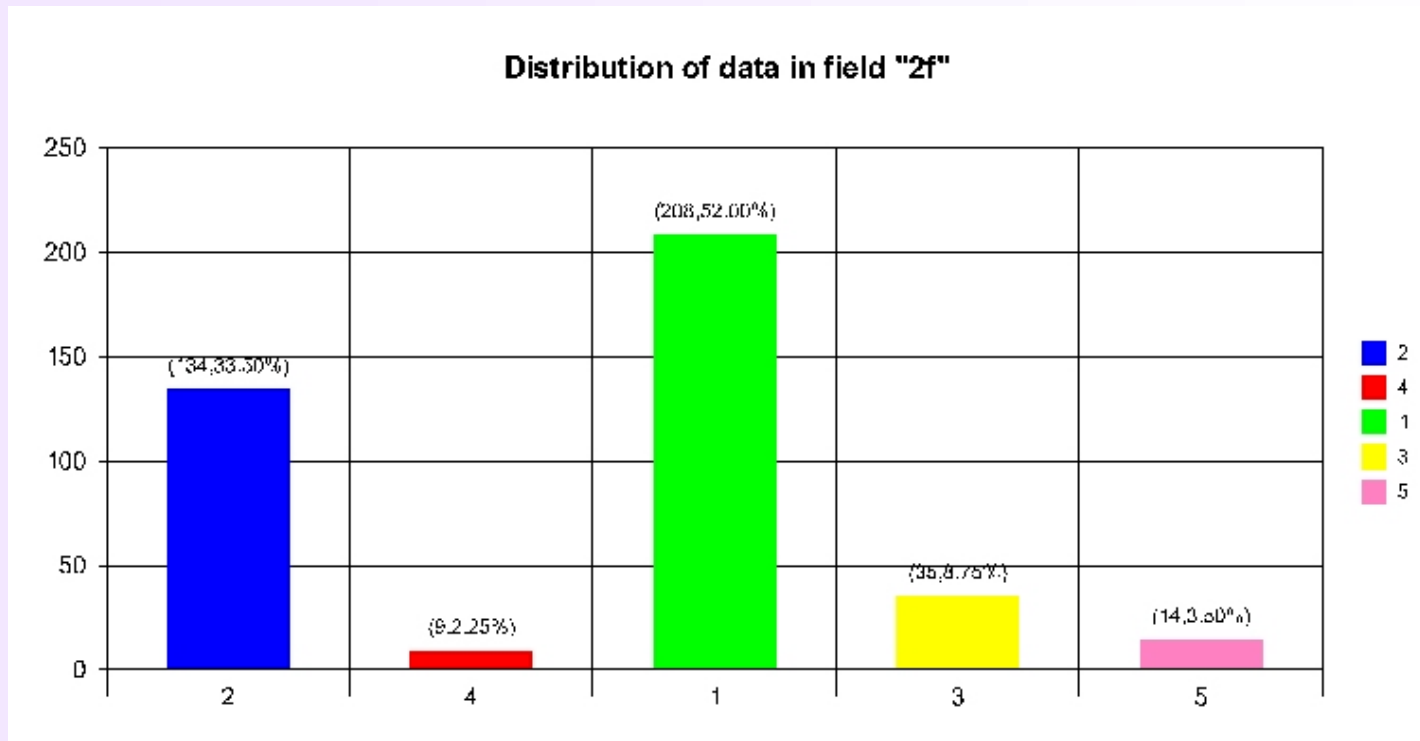


- Zaměření prodejců na segment Cluster 0 a Cluster 1
- Spotřebitelé nejsou citlivý na cenu
- Respondenti nakupují více ve zdravých výživách
- Potenciál je v Cluster 2 – ochota kupovat za přijatelnou cenu a srovnatelnou chuť – museli by se ovšem zdůraznit prospěšné účinky sóji - zaměření na hypermarkety
- Nevhodný je Cluster 3 – jsou ochotni tyto potraviny nakupovat pouze z vážných zdravotních důvodů





# Frekvence nákupů sojových výrobků



5: denně, 4: 4 – 5 x týdně, 3: 2 – 3 x týdně  
2: 1 x týdně, 1: nikdy



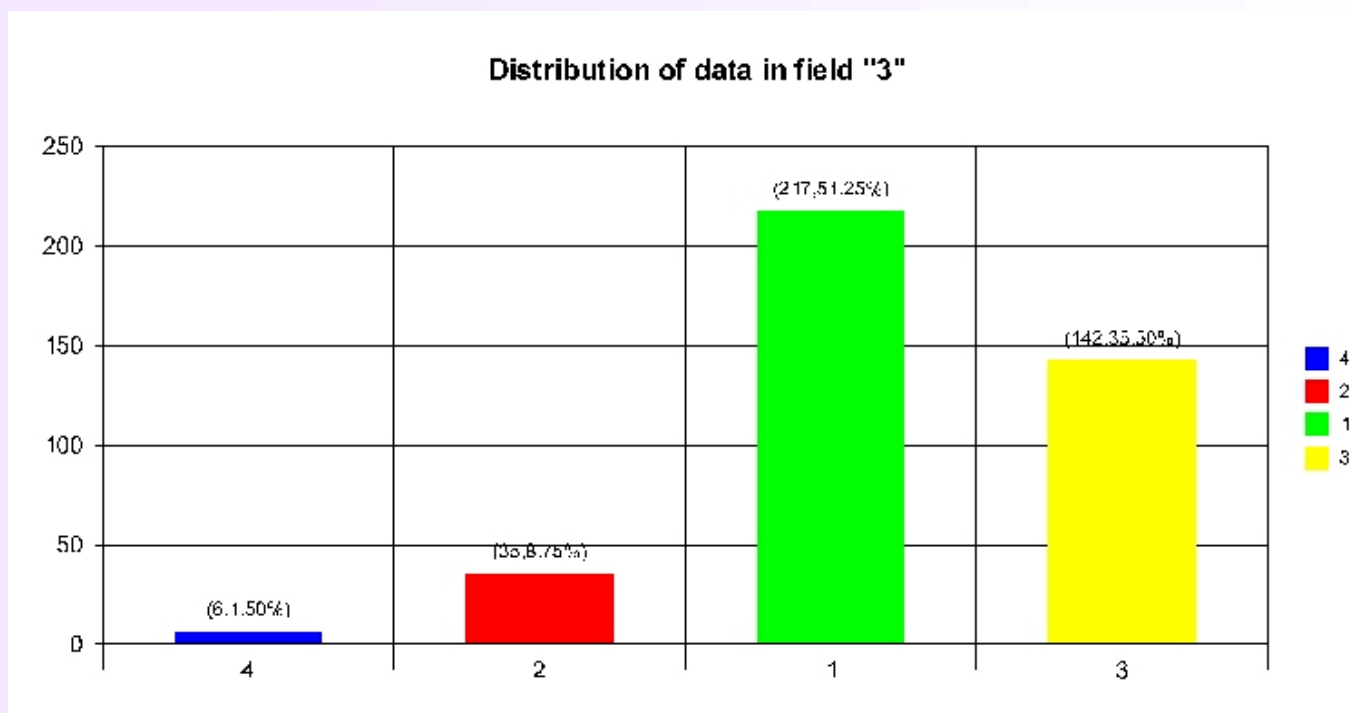
# Definice spotřebitele nakupujícího sojové výrobky



- 2 – 3 x měsíčně obchody se zdravou výživou a v průměru 1 x měsíčně hypermarkety
- denně ovoce a zeleninu, 2 – 3 x týdně cereální pečivo, jednou za týden kuřecí maso a luštěniny a nejedí vepřové maso
- info z časopisů, odborných článků a prodejců
- pohodlí při přípravě, důležitý vliv na zdraví a bio (eko) původ a nedůležité chuť i cena
- časopisy: Vlasta, Květy, Žena a život, Arnika, Regena a Sedmá generace
- zaměstnanci se středoškolským vzděláním mladší 40 let
- domácnosti bez dětí do 12 let
- průměrný počet obyvatel je do 500 000



# Hodnocení zdravotního stavu respondentů



- 4: vážné zdravotní potíže, 3: občas stonám  
2: opakující se obtíže, 1: zcela zdrav



## **Důsledky vyplívající z vyhodnocení vlivu zdravotního stavu spotřebitelů**



- respondenti se zdravotními problémy navštěvují prodejny zdravé výživy
- vitamíny, sojové výrobky a luštěniny
- info získávají od lékařů
- starší 30 let, žijí v domácnostech s dětmi
- průměrný počet obyvatel je nad 500 000
- zdravotní stav má vliv na spotřební chování





# **Děkuji Vám za pozornost**

**Petra Hloušková**

