

Česká zemědělská univerzita v Praze



Provozně ekonomická fakulta

Katedra statistiky

Bakalářská práce

Zpracování dotazníkových dat a jejich vytěžování v problematice zemědělského vzdělávání.

Petra Provazníková

Vedoucí diplomové práce: Doc. Ing. Rudolf Zeipelt, CSc.

Rapotín, 2011

© 2011

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra statistiky

Akademický rok 2010/2011

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Petra Provazníková

obor Veřejná správa a regionální rozvoj - Šumperk

Vedoucí katedry Vám ve smyslu Studijního a zkušebního řádu ČZU v Praze
čl. 16 určuje tuto bakalářskou práci.

Název práce: **Zpracování dotazníkových dat a jejich
vytěžování v problematice zemědělského
vzdělávání**

Osnova bakalářské práce:

1. Úvod
2. Cíl práce a metodika
3. Literární rešerše
4. Vlastní práce
5. Závěr
6. Seznam použitých zdrojů
7. Přílohy

Rozsah hlavní textové části: 30 - 40 stran

Doporučené zdroje:

1. Berka, P.: Dobývání znalostí z databází. Academia, Praha, 2001
2. Bjelka, M., Hanuš, O., Říha, J., Bezdíček, J., Pozdíšek, J., Suchá, S., Koza, M.: Dopad profesního vzdělávání v oboru chovu skotu, 6. část - šetrné čerpání přírodních zdrojů a údržba krajiny pomocí chovu krav bez tržní produkce mléka. Výzkum v chovu skotu/Cattle Research, 2009. ISSN 0139-7265
3. Hanuš, O., Vaculíková, J., Říha, J., Koza, M., Jedelská, R., Suchá, S.: Účinnost postgraduálního odborného vzdělávání v problematice chovu skotu a kvality jeho produktů, zdraví dojníc a kvalita produkce mléka. Výzkum v chovu skotu, 2006. ISSN 0139-7265
4. Hanuš O., Říha J., Suchá S., Kopecký J., Koza M., Jedelská R.: Dopad profesního vzdělávání v oboru chovu skotu – výrobní zemědělská praxe a potravinářské biotechnologické úpravy pro zvýraznění pozitivních zdravotních vlivů mléka a mléčných výrobků. Výzkum v chovu skotu/Cattle research, 2009. ISSN 0139-7265
5. Hendl, J.: Přehled statistických metod zpracování dat. Portál, Praha, 2004
6. Meloun, M., Militký, J.: Kompendium statistického zpracování dat. Praha, Academia, 2002

Další literatura bude doporučena v průběhu zpracování BP.

Vedoucí bakalářské práce: **doc. Ing. Rudolf Zeipelt, CSc.**

Termín odevzdání bakalářské práce: březen 2011



.....
Vedoucí katedry





.....
Děkan

V Praze dne: 7. 2. 2011

Prohlašuji, že jsem práci vypracoval samostatně za použití materiálů uvedených v Seznamu použitých zdrojů na konci práce.

V Rapotíně dne

Petra Provazníková

Na tomto místě bych ráda poděkovala Doc. Ing. Rudolfu Zeipeltovi, CSc. za cenné rady a odborné vedení při zpracování této práce.

Název práce česky

Zpracování dotazníkových dat a jejich vytěžování v problematice zemědělského vzdělávání.

Název práce anglicky

Processing of questionnaire data and their mining in the field of education in agriculture.

Souhrn

Práce se zabývá zpracováním dotazníkových dat s využitím metod statistiky a dataminingu ve dvou praktických aplikacích. Na základě sociologických dat respondentů je pomocí popisné statistiky a shlukovacích algoritmů v první části práce charakterizován typický účastník vzdělávacích seminářů v zemědělství za účelem zjištění této cílové skupiny. V druhé části práce jsou vyhodnocena dotazníková data týkající se preferencí konzumentů potravin živočišného původu a jejich zdrojů - mléka a masa. Byla provedena základní analýza dat pomocí popisné statistiky, formulovány hypotézy o vztazích mezi sociologickými údaji a odpověďmi respondentů ohledně preferencí při konzumaci mléka a mléčných výrobků, zdravotních benefitech mléka, preferencí druhů mas, četnosti jejich nákupu a preferencí úpravy hovězího masa. Následně byla data analyzována pomocí dvou algoritmů pro odvození asociačních pravidel. Na základě interpretace asociačních pravidel byla formulována doporučení pro producenty a prodejce dotčených výrobků. Byla ověřena možnost analýzy dotazníkových dat a jejich vytěžování pomocí metod dataminingu.

Klíčová slova

dotazník, datamining, statistické zpracování dotazníkových dat, vzdělávání v zemědělství, konzumenti potravin, preference konzumentů, mléko, maso, výrobky, asociační pravidla, shlukovací algoritmy.

Summary

This work is aimed to usage of statistical and datamining methods for questionnaire data processing. For description of target group purposes, we used descriptive statistics as well as clustering algorithms for task of characterization of typical educational agricultural seminaire participants. Questionnaire data about consument preferences in milk and meat as well as products were analyzed in the second part of work. Descriptive statistics were used for the first analysis of dataset. Hypothesis testing about relationships between social data and consumer preferences in milk and milk products consumption, health benefits in milk, preferences in meat, frequency of buying meat, preferences in beef preparation were performed. After that, association rules from dataset were derived using two algortihms and their interpretation were performed. On this base, we formulated recommandations for food producers and sellers. Usage of datamining methods for described questionnaire data processing and their mining process was proved in this work.

Keywords

questionnaire, datamining, statistical processing of questionare data, education in agriculture, foodstuff consumers, consumer preferences, milk, meat, products, as-sociation rules, clustering algorithms.

Obsah

1	Úvod	8
2	Cíl práce a metodika	9
2.1	Cíl práce	9
2.2	Materiál	9
2.2.1	Sociologická charakteristika účastníků vzdělávacích seminářů v zemědělství	9
2.2.2	Vyhodnocení dotazníků a využití metod dataminingu pro sledování preferencí konzumentů potravin živočišného původu	11
2.3	Použité metody	12
2.3.1	Kvantitativní a kvalitativní proměnné	13
2.3.2	Absolutní a relativní četnosti	13
2.3.3	Kontingenční tabulky	14
2.3.4	Průměr a medián	15
2.3.5	Směrodatná odchylka a rozptyl	15
2.3.6	Testování hypotéz pro kvantitativní proměnné	16
2.3.7	Datamining - vytěžování dat	17
2.3.8	Shlukovací algoritmy	18
2.3.9	Asociační pravidla	20
3	Literární rešerše	24
3.1	Sociologický výzkum a jeho úloha	24
3.1.1	Předmět, funkce sociologie	24
3.1.2	Sociální a sociálně psychické jevy	25
3.1.3	Sociologický výzkum a statistika	26
3.2	Dotazník	27
3.2.1	Dotazník – charakteristika a využití	27
3.2.2	Dotazníková šetření v rámci ČR	30
3.2.3	Metody sestavení dotazníku	31
3.3	Dotazníkové průzkumy u konzumentů potravin	34
4	Vlastní práce	37
4.1	Charakteristika souboru účastníků vzdělávacích seminářů v zemědělství	37
4.2	Charakteristika typického účastníka vzdělávacích seminářů v zemědělství pomocí shlukovacích algoritmů	42
4.3	Preference konzumentů potravin živočišného původu	47
4.3.1	Popis souboru respondentů dotazníkového šetření k preferencím u konzumace hovězího masa	47
4.3.2	Hypotézy o postojích respondentů ke konzumaci, preferované úpravě hovězího masa a oblíbené různých druhů mas	50
4.3.3	Asociační pravidla preferencí při konzumaci hovězího masa	53
4.3.4	Popis souboru respondentů dotazníkového šetření k preferencím u konzumace kravského mléka	56
4.3.5	Hypotézy o postojích respondentů ke konzumaci a zdravotní prospěšnosti mléka a mléčných výrobků	60
4.3.6	Asociační pravidla preferencí při konzumaci kravského mléka	66
5	Závěr	69
6	Seznam použitých zdrojů	73
7	Přílohy	75

1 Úvod

Práce se zabývá zpracováním dotazníkových dat jako efektivního nástroje sociologického a marketingového průzkumu s využitím metod statistiky a dataminingu. Z hlediska zaměření dotazníkových průzkumů lze práci rozdělit na dvě základní témata:

- Sociologický popis účastníků vzdělávacích seminářů v zemědělství.
- Průzkum preferencí, hypotéz a vztahů respondentů ke konzumaci mléka a masa.

K dosažení vytčených cílů práce je využito metod popisné statistiky a testování hypotéz, dále pak metod typických pro dynamicky se rozvíjející obor zpracování a vytěžování dat – tzv. datamining. K popisu sociologického profilu typických účastníků zemědělských seminářů je využito 2 algoritmů shlukové analýzy, k odvození vztahů v odpovědích konzumentů masa a mléka pak 2 algoritmů pro odvození asociačních pravidel.

Sociologický popis účastníků seminářů v zemědělství, které byly realizovány v rámci projektu zaměřeného na postgraduální vzdělávání v zemědělství může být využit jako zpětná vazba pro zkvalitnění vzdělávání v odvozených cílových skupinách a např. tematickému zaměření navazujících aktivit, podává však také zajímavé sociologické poznatky uvedené o vzdělávání v zemědělství, v dnešní době poněkud opomíjenému oboru z hlediska progresivity aplikace inovací, obecně.

Výsledky odvozené z analýzy preferencí, sociologických dat a názorů respondentů např. na zdravotní benefity konzumace potravin živočišného původu je možno využít jednak pro osvětu a vzdělávání v těchto oblastech, jednak pro cílené marketingové účely a analýzu situace na trhu s těmito produkty.

Předkládaná práce ukazuje, jak dosáhnout výše jmenovaných cílů za pomoci zpracování dotazníkových dat pomocí moderních metod jejich analýzy.

2 Cíl práce a metodika

2.1 Cíl práce

Předkládaná práce využívá aplikace statistické analýzy a metod dataminingu za účelem dosažení těchto dvou cílů:

1. Zjistit sociální strukturu respondentů dotazníků při vzdělávacích seminářích v problematice zemědělského výzkumu a potravinářství. Pomocí testování hypotéz nalézt významné rozdíly v hodnocení vzdělávací, odborné a technické úrovně seminářů, požadavcích na odborné zaměření podle jednotlivých sociálních efektů, které byly u respondentů sledovány.
2. Vyhodnocení odpovědí týkajících se zaměření konzumentů různého sociálního statutu vzhledem k jejich preferencím v konzumaci potravin a potravinových zdrojů (mléka a masa). Pomocí vícerozměrných metod dataminingu nalézt asociační pravidla, která popisují významné vztahy v dotazníkových datech. V diskuzi nalézt vhodné využití nalezených pravidel jako doporučení pro výrobce a producenty potravin.

2.2 Materiál

2.2.1 Sociologická charakteristika účastníků vzdělávacích seminářů v zemědělství

V rámci vzdělávacího projektu MŠMT 2E06033 „Výzkum možností růstu účinnosti vzdělávání a rozvoje lidských zdrojů zemědělsko-potravinářského oboru směrem k šetrnému hospodaření, k podpoře dobrého zdraví a welfare zvířat, kvality a bezpečnosti potravinového řetězce“ proběhlo v letech 2006-2008 celkem pět tematicky odlišných vzdělávacích seminářů pro odbornou zemědělskou veřejnost. Jednalo se o dva semináře tematicky zaměřené na problematiku produkce mléka:

- *Výrobní zemědělská praxe a potravinářské biotechnologické úpravy pro zvýraznění pozitivních zdravotních vlivů mléka a mléčných výrobků*, označeno **mléko1** – 81 respondentů,
- *Zdraví dojníc a kvalita produkce mléka*, označeno **mléko2** – 37 respondentů;

dva semináře tematicky zaměřené na problematiku produkce hovězího masa:

- *Šetrné čerpání přírodních zdrojů a údržba krajiny pomocí chovu krav bez tržní produkce mléka*, označeno **maso1** – 87 respondentů,
- *Údržba krajiny a kvalita produkce masa*, označeno **maso2** – 52 respondentů;

a jeden seminář týkající se problematiky výživy skotu:

- *Výživa dojníc a kvalita mléka - ekologické, zdravotní a hygienické faktory kvality a bezpečnosti mléka jako suroviny a potraviny*, označeno **výživa1** – 85 respondentů.

Účastníci seminářů byli kontaktováni pomocí inzerce jednotlivých akcí v odborných periodících, pomocí přímého poštovního nebo elektronického kontaktu a oznámení jednotlivých akcí na webových stránkách organizace.

Všem účastníkům jednotlivých vzdělávacích seminářů byl předložen k vyplnění dotazník. Dotazník sloužil k hodnocení dopadů podobných vzdělávacích akcí v sektoru zemědělské prvovýroby, jako hodnocení zajímavosti jednotlivých přednášených témat a ke zjištění cílových skupin postgraduálního vzdělávání v zemědělství.

Každý z jednotlivých seminářů byl pomocí zhodnocení dotazníkového průzkumu vyhodnocen a výsledky byly předmětem odborných publikací. [1, 2, 3, 4]. Dotazník ke zjištění zpětné vazby vzdělávacích akcí je strukturován do tří částí:

3. sociologická data respondenta,

- pohlaví (žena, muž),
- věk (do 20, 31-40, 41-50, nad 50 let),
- stupeň dosaženého vzdělání (základní, učební obor, středoškolské, vysokoškolské, vědecké),
- obor vzdělání (zemědělské, potravinářské, přírodovědné, veterinární, technické, humanitní, jiné),
- profese (student, řídicí technologický pracovník v živočišné výrobě, řídicí technologický pracovník v potravinářství, pracovník v zemědělské nebo potravinářské laboratoři, veterinární lékař, poradce v zemědělské výrobě, pracovník zemědělských služeb, učitel zemědělského oboru, pracovník státní správy v zemědělství, pracovník profesní zemědělské organizace, ostatní profese),
- region realizace profese.

4. hodnocení úrovně semináře,

- organizační úroveň (výborná-špatná),
- odborná úroveň (výborná-špatná),
- využitelnost poznatků (bezprostředně, později, vůbec),
- potřeba sborníku (nezbytný, potřebný pro využití poznatků, není potřebný),
- úroveň didaktických pomůcek (odpovídající modernímu trendu, průměrná úroveň, špatná úroveň).

5. hodnocení konkrétního tématu semináře.

Příklad dotazníku je uveden v Příloze 1.

Cílem této části práce je zjistit sociální strukturu účastníků vzdělávacích seminářů v zemědělství a jejich spokojenosti za účelem definice cílové skupiny vhodné pro další vzdělávací akce. K tomuto účelu jsme využili první dvě společné části struktury dotazníků – sociologická data respondenta a hodnocení úrovně semináře. V práci je uvažován tento soubor jako reprezentativní vzorek populace účastníků vzdělávacích akcí v zemědělství. Za účelem zpracování popisných charakteristik souboru a využití shlukovacích algoritmů pro odvození typických účastníků seminářů v zemědělství byl v práci využit software Statistica 8 [5].

2.2.2 Vyhodnocení dotazníků a využití metod dataminingu pro sledování preferencí konzumentů potravin živočišného původu

V rámci řešení projektu NPV II 2B08037 Biotechnologické metody pro inovace hodnocení zpracovatelské a spotřebitelské kvality hovězího masa jako potravinového zdroje živočišných proteinů proběhlo dotazníkové šetření týkající se postoje konzumentů k hovězímu masu. Šetření proběhlo v rámci výstavy TECHAGRO 2010. Celkem se dotazníkového průzkumu spojeného s degustací grilovaného hovězího masa zúčastnilo 132 respondentů. Dotazník (viz příloha 2) byl strukturován do několika částí:

1. sociologická data respondenta,
 - pohlaví (žena, muž),
 - věk respondenta (počet roků; student, produktivní, důchodce),
 - dosažené vzdělání (základní, středoškolské, vysokoškolské),
2. postoj a četnost nákupu masa,
 - jak často nakupujete výsekové maso (5x, 4x-2x, 1x za týden),
 - jak často nakupujete výsekové hovězí maso (1x za týden, 1x za měsíc, méně než 1x za měsíc, vůbec),
3. preference jednotlivých druhů masa,
 - jaké maso konzumujete nejčastěji (rybí, jehněčí, hovězí, vepřové, drůbeží, žádné),
 - jaké je vaše nejoblíbenější maso (rybí, jehněčí, hovězí, vepřové, drůbeží, žádné),
4. postoj ke konzumaci hovězího masa,
 - konzumovali jste někdy hovězí maso, na kterém jste si pochutnali (ano, ne),
 - kdy to bylo naposled (týden, měsíc, rok),
 - preferujete výsekové maso z tuzemské produkce (ano, ne, nezajímám se o původ, zajímám se o původ, ale nemohu jej získat),
5. preference úpravy masa,
 - jakou kuchyňskou úpravu hovězího masa preferujete (vaření, pečení, dušení, grilování),
 - jak propečený steaku preferujete (krvavý, méně krvavý, medium, téměř propečený, propečený),
 - grilovaný steak jsem doposud (měl, neměl; chutnal mi, nechutnal mi),
 - steak bych preferoval připravit (doma, v restauraci),
 - oslovila vás pozitivně úprava hovězího masa grilováním (ano, ne).

V práci se dále zabýváme vyhodnocením postojů respondentů dotazníkového průzkumu realizovaného v rámci výzkumného záměru MSM 2678846201, který se zaměřuje na sledování postoje konzumentů ke kravskému mléku a mléčným výrobkům. Dotazník (viz příloha 3) je strukturován do následujících částí:

1. sociologická data respondenta,
 - věk (počet let),
 - pohlaví (žena, muž),

- dosažené vzdělání (základní, vyučen, středoškolské, vysokoškolské, jiné),
- 2. stravovací návyky ve vztahu k mléku,
 - upřednostňuji (kravské mléko, mléčné výrobky – jogurty, sýry, tvaroh, máslo, jiné),
 - konzumuji (kravské mléko – denně, 2-3týdně, 1x týdně, jinak; mléčné výrobky – denně, 2-3týdně, 1x týdně, jinak),
- 3. názor na mléko a jeho konzumaci (ano, ne, nevím),
 - Je kravské mléko zdraví prospěšné?,
 - Je kravské mléko důležitým zdrojem základních živin?,
 - Má kravské mléko pozitivní vliv na zdraví člověka?,
 - Je kravské mléko hlavním zdrojem vápníku?,
 - Chrání kravské mléko před vznikem osteoporózy?,
 - Je kravské mléko dobře stravitelné?,
 - Obsahuje kravské mléko zvýšený podíl nezdravých tuků a cholesterol?,
 - Jsou zakysané mléčné výrobky zdravější než kravské mléko?,
 - Je podporována konzumace mléka ve Vaší rodině?,
 - Vedete Vaše děti ke konzumaci kravského mléka?,
 - Je kravské mléko propagováno dostatečně?,
 - Podpořil(a) by jste propagaci kravského mléka ve školách?,
 - Podpořil(a) by jste zavedení automatů na kravské mléko?,
 - Je vhodné konzumovat jiné mléko než kravské (sojové, kozí aj.)?.

Cílem této části práce je vyhodnocení odpovědí týkajících se zaměření konzumentů různého sociálního statutu vzhledem k jejich preferencím v konzumaci potravin a potravinových zdrojů (mléka a masa) a pomocí vícerozměrných metod dataminingu nalézt asociační pravidla, která popisují významné vztahy v dotazníkových datech. K tomuto účelu jsou v práci využity algoritmy strojového učení – asociační pravidla. Jedná se o algoritmy Apriori a Tertius implementované v software WEKA 3-6-1 [6]. Pro popis a testování hypotéz jsou použity metody implementované v software Statistica 8 [5] popsané v následující části práce.

2.3 Použité metody

Za účelem dosažení prvního z cílů bylo v práci využito metod popisné statistiky (popisné souhrné charakteristiky - tabulky s relativními a absolutními četnostmi, histogram) a dvou algoritmů shlukovací analýzy (k-means algoritmus, Expectation Maximalization algoritmus) k odvození typického účastníka vzdělávacích seminářů v oblasti zemědělství.

Zpracování analytické části druhého cíle bylo provedeno v následujících krocích:

1. popis a exploratorní analýza datových souborů pomocí metod popisné statistiky (relativní a absolutní četnosti odpovědí, kontingenční tabulky pro základní analýzu vícenásobných odpovědí, histogram, grafy interakcí vícenásobných odpovědí),

2. formulace a testování hypotéz, které vyplývají z exploratorní analýzy (Chí-kvadrát test, kontingenční tabulky, grafy interakcí, u spojitých proměnných souhrn metod analýzy variance ANOVA),
3. odvození asociačních pravidel - k tomuto účelu byly využity dva algoritmy z oblasti dataminingu - algoritmus Apriori a Tertius.

Metody popisné statistiky a shlukování byly využity pomocí softwarového rozhraní Statistica 8, metody odvození asociačních pravidel pomocí software WEKA 3-6-1 [6]. V další části textu je uveden použitý materiál a přehled metod využitých k výše popsanému způsobu zpracování dat.

2.3.1 Kvantitativní a kvalitativní proměnné

Při provádění statistických analýz a zpracování dat obecně dělíme datové proměnné - znaky na kvalitativní a kvantitativní, kvantitativní dále na diskrétní a spojitá [7, 8]. Kvalitativní proměnné jsou takové, jejichž hodnoty nevykazují možnost srovnání ve smyslu matematickém (je větší než, menší než, atd.) a vyžadují slovní popis. Typickým příkladem kvalitativní proměnné je např. barva, příčina poruchy, vyjádření pocitu atd.

Kvalitativní proměnné vyjadřují měřitelný znak. Dělíme je dále podle toho, zda nabývají diskrétních (konečný počet) nebo spojitých hodnot (např. reálná čísla). Spojitá proměnná může teoreticky nabývat libovolných hodnot z určitého intervalu reálných čísel [8]. Charakterizuje většinou určitou velikost, množství nebo objem statistické jednotky, údaje o nich se získávají přímým měřením, vážením atd. Příkladem takových znaků, které nazýváme spojité, jsou hektarový výnos, doba trvání pracovní operace, výkon motoru, hloubka orby, tučnost mléka, výška rostliny, dojivost, teplota výfukových plynů, měřená spotřeba paliva atd. [9].

Diskrétní proměnné neboli kategoriální proměnné nabývají naopak pouze konečného počtu hodnot [8], nabývají pouze jednotlivých izolovaných (diskrétních) hodnot – např. počet zrn v klasu, počet zmetků ve výrobní sérii, početní stav automobilového parku atd. [9], počty respondentů, kteří odpověděli kladně na určitou otázku - tzv. frekvenční data. Podle kontextu zpracování lze i s diskrétními údaji pracovat jako se spojitými, např. tepovou frekvenci můžeme považovat za spojitou proměnnou, pokud se setkáváme v analýze s velkými rozdíly tepových frekvencí. [8].

Podle typu dat používáme v práci k jejich vyhodnocení a analýze různé statistické metody, které jsou uvedeny v dalším přehledu.

2.3.2 Absolutní a relativní četnosti

Stručně uvedeme jednorozměrné popisné grafické a tabulační metody pro soubor kvantitativních měření jedné proměnné. V tomto případě si statistický soubor můžeme představit jako n -tici reálných čísel, v níž se jednotlivé prvky mohou opakovat, přičemž pořadí, jak byly prvky získány, nepřikládáme žádný význam.

Tabulka četností, relativních četností a kumulativních četností je základní numerické zobrazení, při kterém se v souboru přítomné hodnoty kvantitativní proměnné

setřídí a pro každou hodnotu se zjistí její absolutní i relativní četnost, dále absolutní a relativní kumulativní četnost [8]. Absolutní četnosti n_i udávají, kolikrát se která hodnota znaku v souboru vyskytuje, jejich součet je roven rozsahu souboru n [9].

$$n_1 + n_2 + n_3 + \dots + n_k = \sum n_i = n$$

Chceme-li provázat mezi sebou různá rozdělení četností, která se liší svým rozsahem, je vhodné převést absolutní četnosti na četnosti relativní f_i [9].

$$f_i = n_i / n, f_1 + f_2 + f_3 + \dots + f_k = \sum f_i = 1.$$

Četnosti se mohou zobrazit i graficky. Grafické zobrazení vytváří geometrický obraz dat. Přitom se využívají body, plochy, úsečky nebo různé další obrazce. Nejznámější způsob zobrazení jedné proměnné se nazývá histogram. V tomto případě osa X odpovídá hodnotám proměnné a osa Y absolutním nebo relativním četnostem [8].

2.3.3 Kontingenční tabulky

Kontingence je vztah dvou či více kvalitativních statistických znaků, z nichž alespoň jeden je znakem množným [7]. Zobrazení této závislosti umožňuje právě kontingenční tabulka. Názvy sloupců tvoří hodnoty, kterých nabývá jeden ze sledovaných znaků, názvy sloupců pak hodnoty druhého ze sledovaných znaků. V jednotlivých buňkách tabulky jsou pak uvedeny četnosti, které platí pro hodnoty znaků uvedených v odpovídajícím sloupci, resp. řádku. Toto zobrazení četností umožňuje měřit vzájemnou závislost znaků, tzv. kontingenci - těsnost závislosti, např. pomocí Pearsonova koeficientu kontingence [7].

	Znak B	b_1	b_2	...	b_j	b_m	Celkem
Znak A								
a_1		n_{11}	n_{12}	...	n_{1j}	...	n_{1m}	n_1
a_2		n_{21}	n_{22}	...	n_{2j}	...	n_{2m}	n_2
.								
.								
a_i				...	n_{ij}	...		n_i
.								
.								
a_k		n_{k1}	n_{k2}	n_{km}	n_k
Celkem		n_1	n_2	...	$n_{.j}$...	n_m	n

Tabulka 1. Příklad kontingenční tabulky.

2.3.4 Průměr a medián

V následujícím textu jsou popsány nejznámější statistické ukazatele kvalitativních proměnných - průměr a medián. Aritmetický průměr je definován jako součet všech naměřených údajů vydělený jejich počtem. Označujeme ho např. pomocí symbolu \bar{x} . Výpočet má tedy podobu:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i.$$

Tuto formu aritmetického průměru nazýváme prostý aritmetický průměr [9].

V případě, že pracujeme s údaji již utříděnými, použijeme vážený aritmetický průměr [9]:

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{\sum_{i=1}^k n_i}.$$

Podobně se počítá průměr pro data zadaná četnostním způsobem pomocí frekvenční tabulky, v níž pro hodnoty jsou ještě zadané jejich četnosti výskytu [8]:

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i}.$$

Vzhledem k citlivosti aritmetického průměru na extrémní hodnoty znaku je třeba mít k dispozici i jiné charakteristiky polohy, např. medián [9].

Medián \tilde{x} definujeme jako prostřední hodnotu řady pozorování, uspořádané podle velikosti na dvě stejně početné poloviny [8].

Při určování mediánu musíme rozlišit dva případy:

1. Rozsah n zkoumaného statistického souboru je vyjádřen lichým číslem. V tomto případě mediánem bude hodnota s pořadovým číslem $n+1$.
2. Rozsah n souboru je udán sudým číslem. Protože soubor se sudým počtem jednotek obsahuje dvě prostřední hodnoty, bývá zvykem volit za medián průměr těchto dvou prostředních hodnot [9].

Velkou výhodou mediánu je, že k jeho určení nepotřebujeme znát všechny hodnoty znaku v souboru, můžeme ho tedy určit i tehdy, jestliže získané údaje o hodnotách znaku v souboru jsou neúplné. Medián se velmi lehce určuje u souborů malého rozsahu [9].

2.3.5 Směrodatná odchylka a rozptyl

Dalšími ze základních popisných statistik kvantitativních proměnných jsou rozptyl znaku a jeho směrodatná odchylka. Charakteristiky variability měří rozptýlení hodnot příslušného souboru, tzn. určují rozmezí, v němž se výběrové údaje vyskytují [9]. Rozptyl a směrodatná odchylka spolu velmi úzce souvisejí. Oběma je společná vlastnost, že využívají při výpočtu všechny údaje a obě se vztahují k aritmetickému průměru – měří rozptýlenost dat kolem aritmetického průměru dat [8].

Rozptyl je definován jako průměrná kvadratická odchylka měření od aritmetického

průměru, přičemž při průměrování této odchylky dělíme číslem $(n-1)$ [8]:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

Směrodatná odchylka s je odmocnina z rozptylu a vrací míru rozptýlenosti do měřítka původních dat. Její výhodou oproti rozptylu je to, že je uvedena ve stejných měrných jednotkách jako zkoumaný statistický znak [9].

Výběrový rozptyl a výběrová směrodatná odchylka měří tzv. absolutní variabilitu. Pokud bychom chtěli porovnávat variabilitu statistického znaku u dvou nebo více souborů, které se výrazně liší úrovní znaků, nebo kdybychom chtěli porovnat variabilitu několika statistických znaků, vyjádřených v různých měrných jednotkách, užíváme obě popisné statistiky jako relativní charakteristiky - vyjádřené v procentech z průměru.

2.3.6 Testování hypotéz pro kvantitativní proměnné

U kvantitativních znaků nás obvykle zajímá, zda spolu navzájem souvisí, tj. zda se navzájem ovlivňují dva sledované jevy, které sledujeme, zda mezi nimi existuje závislost. K tomuto účelu se využívá teoretický aparát, který se nazývá testování hypotéz. Postup testování je v obecnosti takový, že nejprve formulujeme tzv. nulovou hypotézu, kterou následným testem můžeme vyvrátit. V případě, že nulová hypotéza je testem vyvrácena, platí tzv. hypotéza alternativní (logická negace nulové hypotézy). Formulování hypotéz a vyvozování závěrů z jejich testování je velmi častou příčinou disinterpretace statistických výsledků. V závislosti na typu testu a formulované hypotéze je využití aparátu testování hypotéz ve statistice velmi rozšířeno. V následujícím textu popisujeme testování pomocí Pearsonova Chí-kvadrát testu a Fisherova faktoriálního testu, které jsme využili pro zpracování výsledků v této práci.

Nulovou hypotézu formulujeme pro oba testy takto - H_0 : Mezi sledovanými znaky neexistuje závislost.

Použití zmíněných testů se řídí následujícími pravidly:

1. Je-li rozsah souboru větší než 40, použijeme Chí-kvadrát test pro nezávislost.
2. Je-li rozsah souboru menší než 20, použijeme Fisherův faktoriální test.
3. Je-li $20 < n < 40$, potom je nutno vyjádřit očekávané četnosti a_0, b_0, c_0, d_0 .

Dále pak platí:

1. Pokud jsou všechny očekávané četnosti větší než 5, potom použijeme Chí-kvadrát test pro nezávislost.
2. Jestliže alespoň jedna z očekávaných četností je menší než 5, je nutno užít Fisherův faktoriální test [7].

Chi-kvadrát test nezávislosti

H_0 : Mezi sledovanými znaky neexistuje závislost.

A/B	Ano	Ne	Součet
Ano	a	b	a+b
Ne	c	d	c+d
Součet	a+c	b+d	n

Tabulka 2. Tabulka k výpočtu Chí-kvadrát testu nezávislosti pro dva znaky.

Nulovou hypotézu o nezávislosti testujeme pomocí testového kritéria Chí-kvadrát:

$$\chi^2 = \frac{n(ad - bc)^2}{(a+b)(a+c)(b+d)(c+d)} .$$

V tabulkách Chí-kvadrát rozdělení poté nalezneme kritické hodnoty χ^2_α a porovnáme s vypočtenou hodnotou testového kritéria. Je-li $\chi^2 > \chi^2_\alpha$, nulovou hypotézu o nezávislosti zamítáme [7].

Fisherův faktoriálový test

H_0 : mezi sledovanými znaky neexistuje závislost.

V případě, že musíme užít Fisherův faktoriálový test, postupujeme následovně:

- Vyhledáme nejmenší skutečnou sdruženou četnost
- Tuto četnost postupně v pomocných tabulkách 2×2 zmenšujeme po jedné až po nulu, při zachování okrajových četností. Pro každou tabulku vypočítáváme pravděpodobnost p_i , a to pomocí faktoriálů:

$$p_i = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

- Součet všech p_i je hodnotou testového kritéria a je porovnáván s hladinou významnosti alfa. Pokud $\sum p_i < \alpha$, pak nulovou hypotézu zamítáme [7].

2.3.7 Datamining - vytěžování dat

Datamining, česky vytěžování dat je nástrojem, který umožňuje odhalit skryté vztahy v datech, které na první pohled, ani při použití statistické abstrakce pomocí různých charakteristik, nejsou v datech patrné. Tyto vztahy reprezentuje pomocí tzv. modelů, které umožňují reprezentovat znalosti v podobě čitelné pro člověka. Datamining využívá metody, které jsou schopny zpracovat díky výpočetní technice obrovské množství dat (z hlediska množství záznamů i sledovaných proměnných).

Datamining je proces porozumění datům. Týká se:

- plánu experimentu,
- sběru a reprezentace dat,
- metod jejich vytěžování a
- metod ověření vybudovaných modelů.

Metody vytěžování dat jsou převzaty z klasické statistiky (regrese, Bayesovský klasifikátor), matematického modelování (PCA, LCA, ICA), ale také z matematického aparátu umělé inteligence (neuronové sítě) a strojového učení (regresní stromy, rozhodovací stromy, algoritmy založené na učení z instancí atd.) [10].

Typické úkoly dataminingu z hlediska modelování jsou:

- hledání podobných případů v rámci datového souboru a jejich reprezentace - shluková analýza,
- hledání souvislostí - tzv. analýza nákupního košíku - asociační pravidla,
- predikce - rozhodovací stromy a pravidla, neuronové sítě, Bayesovská klasifikace.

Datamining nalézá v současné době uplatnění především v marketingu, ke zpracování dat o zákaznících za účelem cílených marketingových kampaní, identifikace klíčových skupin zákazníků, ale také např. k odhadu vytížení internetového připojení, předpovědích počasí, plánování dopravy apod. Své uplatnění má však také v řadě výzkumných a softwarových aplikací [10].

2.3.8 Shlukovací algoritmy

Shluková analýza (clusterizace) využívá k rozdělení datového prostoru případů do shluků, které lze popsat jednotnými vlastnostmi. Jedná se o nalezení skupin případů, které jsou si navzájem podobné, podle různých kritérií. V kontextu dataminingu se jedná o nalezení struktur, které popisují datový prostor případů a umožňují jednoduše popsat podobné případy. Shlukovací algoritmy se používají bez znalosti jakékoli třídy klasifikace (identifikace), jedná se o učení bez učitele [10].

Mezi shlukovací algoritmy řadíme např. metodu shlukování pomocí k -středů (k -means), Expectation-Maximization (EM) algoritmus, které popisujeme dále, mezi další řadíme např. random sampling algoritmy založené na Markovových řetzcích nebo hierarchické shlukování.

k-means algoritmus

1. rozděl příklady náhodně do k shluků,
 2. nalezní centroidy (střední případy) pro všechny shluky v aktuálním rozkladu,
 3. pro každý příklad v datovém souboru proved':
 - a. určí vzdálenost od centroidu každého shluku podle kritéria (Euklidická vzdálenost, Čebyševova, Hammingova),
 - b. přesuň příklad do shluku, k jehož centroidu má příklad nejbliže,
 - c. byl-li přesunut nějaký případ, jdi na bod 2, jinak konec.
-

EM algoritmus

Expectation maximalization algoritmus je iterativní shlukovací algoritmus založený na pravděpodobnostním počtu. EM algoritmus upravuje v každém průchodu podle distribuce (rozložení) jednotlivých parametrů pravděpodobnostní příslušnost k jednotlivým shlukům u každé instance - datové položky v souboru. Výsledným stavem je dosažení maximální spolehlivosti odhadů distribuce jednotlivých parametrů. Může být využit i ke shlukování s předem nespécifikovaným počtem výsledných shluků - dokáže je odvodit např. formou křížové validace [10].

Inicializace

1. Zvol náhodně distribuci jednotlivých parametrů.

Iteration

1. Spočítej hodnoty distribuce jednotlivých parametrů.
2. Na aktualizovaných hodnotách distribuce proved' odhad jejich maximální spolehlivosti.
3. Jestliže se změnila distribuce parametrů, jdi na 1, jinak konec.

K úloze charakteristiky typického účastníka vzdělávacích seminářů v zemědělském sektoru bylo využito shlukové analýzy pomocí algoritmů k-means a EM implementované v software Statistica 8 [5]. Vstupem do obou shlukovacích algoritmů byla data popisující sledovaný sociologický profil účastníků semináře – sociologická data respondenta.

Nastavení obou algoritmů bylo následující:

Algorithm: k-Means

Distance method: Euclidean distances
Initial centers: Maximize initial distance
MD casewise deletion: Yes
Cross-validation: No
Testing sample: 0
Training cases: 274
Training error: 1.626441
Number of clusters: 2

Algorithm: EM

Continuous distribution: Normal
MD casewise deletion: Yes
Cross-validation: No
Testing sample: 0
Training cases: 274
Training log-likelihood: -8.743955
Number of clusters: 2

2.3.9 Asociační pravidla

Asociační pravidla reprezentují odvozené znalosti pomocí aparátu, který je velmi blízký lidskému pochopení, proto jsou velmi široce využívána především v marketingu - pro úlohu analýzy nákupního košíku. Asociační pravidla slouží k vyhledávání souvislostí mezi hodnotami jednotlivých proměnných a tyto souvislosti reprezentují pomocí logického aparátu (většinou implikace) [10]:

Předpoklad \Rightarrow Závěr,

kde *Předpoklad* je logickou kombinací (většinou za pomoci operátoru AND) hodnot parametrů. *Závěr* pravidla bývá tvořen hodnotou jednoho parametru, možná je však i jejich logická kombinace. Před použitím asociačních pravidel je nutné převést numerické parametry např. do binárních hodnot nebo do intervalů hodnot proměnné. Asociační pravidla se odvozují pomocí kontingenčních tabulek.

Při interpretaci výsledků asociačních pravidel je nutné sledovat jejich charakteristiky, které zohledňují využitelnost pravidla a využití - např.: podpora (počet případů, které splňují předpoklad i závěr), spolehlivost (pravděpodobnost splnění závěru za platnosti předpokladu), dále ukazatele jako např. pokrytí, kvalita, zajímavost, zjištění kauzální spolehlivosti. Všechny z vyjmenovaných parametrů je možno pro pravidla kalkulovat z kontingenční tabulky, která obsahuje počty případů splňujících pro dané pravidlo předpoklad, závěr a jejich negace.

Algoritmy pro hledání asociačních pravidel prohledávají prostor kombinací hodnot parametrů, které jsou pro odvození pravidel použity. Prohledávání prostoru kombinací (počet kombinací roste exponenciálně s počtem hodnot a počtem parametrů) je výpočetně náročné.

Znáмым algoritmem pro generování asociačních pravidel je algoritmus apriori. Algoritmus hledá logické kombinace hodnot parametrů, které dosahují požadované četnosti [10, 11].

Algoritmus apriori

1. Do množiny pravidel L_1 přiřaď všechny hodnoty parametrů, které dosahují alespoň požadované četnosti
 2. $k = 2$
 3. Dokud L_{k-1} není prázdná
 - a. Vygeneruj na základě L_{k-1} množinu kandidátů C_k
 - b. Do L_k zařaď kombinace z C_k , které dosáhly alespoň požadovanou četnost
 - c. $k = k + 1$
-

Algoritmus Tertius

Algoritmus Tertius odvozuje asociační pravidla na základě kritéria deskriptivního potvrzení (rozdíl pravěpodobnosti platnosti předpokladu a zároveň závěru a platnosti předpokladu a zároveň negace závěru pravidel). Tertius odvozuje asociační pravidla, u nichž je závěr pravidla také tvořen logickou kombinací hodnot atributů. Algoritmus u každého pravidla ve výstupu vypisuje hodnotu deskriptivního potvrzení a míru správně určených hodnot, které pokrývá pravidlo [12].

Za účelem využití asociačních pravidel k analýze datových souborů popsaných výše byly využity algoritmy Apriori a Tertius implementované v software WEKA 3-6-1 [6].

Datový set pro zjištění významných asociačních pravidel z dotazníkového průzkumu preferencí při konzumaci hovězího masa a sociologických údajů respondentů byl vytvořen úpravou datového souboru popsaného v části 2.2.2. Byly z něj vyřazeny proměnné týkající se preferencí s četností možných odpovědí nižší než 10 – monotematické odpovědi (věková kategorie, otázky na nejčastěji konzumované a nejoblíbenější maso, otázka, zda jste si někdy pochutnali na hovězím, preference tuzemské produkce, otázka, zda vám steak chutnal, zda vás oslovilo grilování hovězího masa), aby se zabránilo generování zřejmých či zavádějících pravidel. Spojitá proměnná věku je reprezentována pomocí 10 stejně dlouhých intervalů se stejnými frekvencemi (weka.filters.unsupervised.attribute.discretize -F -B 10 -M -1.0 -R first).

Vstupem do algoritmu Apriori byl tak následující datový set a nastavení:

=== Run information ===

Scheme: weka.associations.Apriori -N 30 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0
Relation: data_konzument_clear1.txt-weka.filters.unsupervised.attribute.Remove-R8-weka.filters.unsupervised.attribute.Remove-R16-weka.filters.unsupervised.attribute.Remove-R17-weka.filters.unsupervised.attribute.Remove-R23-weka.filters.unsupervised.attribute.Remove-R24-weka.filters.unsupervised.attribute.Remove-R7-15-weka.filters.unsupervised.attribute.Remove-R3-weka.filters.unsupervised.attribute.Discretize-F-B10-M-1.0-Rfirst-last

Instances: 132

Attributes: 13

- pohlaví
- věk
- vzdělání
- nakup výsekového masa
- nakup výsekového hovězího
- kdy naposled
- X10vaření
- X10pečení
- X10dušení
- X10grilování
- propečenost
- steak
- doma.restaurace

=== Associator model (full training set) ===

Pro algoritmus Tertius bylo využito následující nastavení a vstupní set:

=== **Run information** ===

Scheme: weka.associations.Tertius -K 10 -F 0.0 -N 1.0 -L 4 -G 0 -c 0 -I 0 -P 0
Relation: data_konzument_clear1.txt-weka.filters.unsupervised.attribute.Remove-R8-weka.filters.unsupervised.attribute.Remove-R16-weka.filters.unsupervised.attribute.Remove-R17-weka.filters.unsupervised.attribute.Remove-R23-weka.filters.unsupervised.attribute.Remove-R24-weka.filters.unsupervised.attribute.Remove-R7-15-weka.filters.unsupervised.attribute.Remove-R3-weka.filters.unsupervised.attribute.Discretize-F-B10-M-1.0-Rfirst-last
Instances: 132
Attributes: 13

Popis datového setu a nastavení algoritmu Apriori pro generování 30 nejlepších rozhodovacích pravidel z dotazníkového průzkumu pro preference při konzumaci mléčných výrobků a informovanosti respondentů je uveden v následujícím výpisu.

=== **Run information** ===

Scheme: weka.associations.Apriori -N 30 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Relation: dotaznik.txt-weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.NumericToBinary
Instances: 200
Attributes: 26
 věk
 pohlaví
 vzdělání
 upřednostňuji
 jogurty_binarized
 sýry_binarized
 tvaroh_binarized
 máslo_binarized
 jiné_binarized
 konzumuji mléko
 konzumuji výrobky
 je mléko prospěšné zdraví?
 zdroj živin?
 pozitivní vliv?
 zdroj vápníku?
 ochrana před osteoporézou?
 dobře stravitelné?
 cholesterol a tuky?
 zakysané zdravější?
 podpora v rodině?
 děti?
 dostatečná propagace?

propagace ve školách?
automaty?
je vhodné jiné mléko?
negativní vliv na zdraví?

=== Associator model (full training set) ===

Nastavení a datový set pro odvození rozhodovacích pravidel pomocí algoritmu Tertius je uvedeno v následujícím výstupu.

=== **Run information** ===

Scheme: weka.associations.Tertius -K 10 -F 0.0 -N 1.0 -L 4 -G 0 -c 0 -I 0 -P 0
Relation: dotaznik.txt-weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.NumericToBinary-weka.filters.unsupervised.attribute.Remove-R1-3-weka.filters.unsupervised.attribute.Remove-R6-weka.filters.unsupervised.attribute.Remove-R9-13-weka.filters.unsupervised.attribute.Remove-R11-15-weka.filters.unsupervised.attribute.Remove-R1

Instances: 200

Attributes: 11

jogurty_binarized
sýry_binarized
tvaroh_binarized
máslo_binarized
konzumuji mléko
konzumuji výrobky
je mléko prospěšné zdraví?
cholesterol a tuky?
zakysané zdravější?
je vhodné jiné mléko?
negativní vliv na zdraví?

=== Associator model (full training set) ===

3 Literární rešerše

3.1 Sociologický výzkum a jeho úloha

3.1.1 Předmět, funkce sociologie

Podle svého zakladatele francouzského filozofa a tvůrce názvu této nové disciplíny Augusta Comta, měla být sociologie hlavním reprezentantem všech společenských věd, tedy oborem poměrně univerzálním a vlastně zastřešujícím ostatní společenské vědy. Sociologie spolu s dalšími společenskými vědami studuje společnost jakožto určitý historický proces, v němž se uplatňují nejrůznější formy (od původních, kmenových pospolitostí až po moderní národní státy i nadnárodní společenství) [13]. Abychom pochopili podstatu „sociologické perspektivy“, je vhodné připomenout, že sociologie je velice úzce spjata zejména s následujícími vědními disciplínami:

- filozofií: sociologie je s filozofií svázána historicky, vyvinula se z tzv. sociální filozofie, filozofie dějin a morální filozofie
- historií: všechny sociální instituce (rodina, vzdělávací systémy apod.) se vyvinuly a jejich lepší pochopení je možné jen na bázi historického vědění, např. přechod od tradiční k moderní společnosti a od moderní k postindustriální společnosti bez historické vědy není možné
- psychologií, zejména se sociální psychologií: sociologie se snaží vysvětlit sociální život psychickými pohnutkami, bez psychického vědění je však neúplná
- ekonomii: hospodářský život je podstatnou složkou sociálního života, ekonomické chování může být dokonce modelem sociálního chování
- antropologií: sociální a kulturní antropologie se soustřeďuje na studium malých, písma neznalých společenství a jejich životní způsob, v poslední době však vznikají antropologické studie věnované moderním společnostem, jde tedy o disciplínu sociologii velmi blízkou
- politickou vědou: spojena se sociologií při zkoumání témat jako je volební chování, neformální mocenské struktury apod.

V poslední době se sociologie silně ekologizuje a ekologie se sociologizuje – ukazuje se, že ekologický problém není jen problémem přírodovědeckým, ale také sociálním, politickým a ekonomickým. Sociologie má tedy poměrně úzký vztah ve smyslu využití vstupních dat, metod, ale také prolínání poznatků a předmětů zkoumání prakticky ke všem sociálním vědám a také k některým vědám exaktním a přírodním (k matematice, logice a statistice, k demografii, biologii apod.) [14].

Její záběr zasahuje od globálních problémů celého lidstva až po zkoumání jedince jako konkrétního aktéra, k těmto hranicím se však přibližuje jen zřídka. Převažuje stádium útvarů, které tvoří celky (sociální systémy), založené na vztazích mezi částmi i mezi systémy navzájem. Soustřeďuje se na sociální jednání, jehož základní rysy jsou (podle Maxe Webera) zaměřenost na druhé lidi, jeho smysl a význam, jemuž je třeba porozumět. Sociologie se zabývá společenskými vztahy a principy jejich vzniku v nejrůznějších oblastech či sférách života.

Na jedné straně tedy sociologové studují společnost jako celek a zákonitosti jejího vývoje [13], sociologie se snaží podat celkový obraz společnosti, formulovat obecně platné pravidelnosti společenského vývoje, tedy vypovídat o společnosti jako celku [14], na druhé straně pak probíhají výzkumy jednotlivých jevů nebo sociálních problémů [13], dílčí sociální jevy sociologie zkoumá ve vazbě na jiné sociální jevy, tedy nikoli samostatně [14]. Tyto výzkumy samozřejmě předpokládají využití určitého pojmového aparátu, teoretických předpokladů (hypotéz). Právě jejich prostřednictvím jsou získávána jednotlivá fakta, která mohou být východiskem zobecnění, formulace zákonitostí či doporučení [13].

Z tohoto pohledu lze sociologii definovat jako teoreticko-empirickou vědu, která má dvě základní poznávací funkce, jež ji potom fakticky vnitřně strukturují:

- společenské jevy a procesy popsat pomocí vlastního metodologického instrumentária (funkce nebo dimenze empirická),
- podle možností, výskyt a průběh určitých jevů vysvětlit (funkce teoretická) [14].

Sociologická abstrakce hraje významnou úlohu v aplikaci zkoumaných jevů i v našem životě. Aktéři v sociálním prostředí si sice v každodenním životě obvykle nekladou nárok systematicky vysvětlovat vše, co kolem nich probíhá, potřebují však rozumět jednání ostatních, chápat jeho smysl a záměr. Sami usilujeme o dosažení určitých hodnot či statků, a k tomu potřebujeme umět volit správnou cestu, efektivní prostředky. Tím nesledujeme nějaký abstraktní poznávací zájem, nýbrž většinou zcela konkrétní, praktický cíl [13], k jehož dosažení nám napomáhá sociologická metodologie a obecně známé výsledky zkoumání konkrétních jevů i jejich principů.

3.1.2 Sociální a sociálně psychické jevy

Z obecně metodologických důvodů a také z logiky sociologického a sociálně psychologického výzkumu je třeba nejdříve zodpovědět otázku co je předmětem výzkumu. Metodologie každého výzkumu musí vycházet z povahy zkoumaných jevů a nástroje pro získávání informací ve zkoumané realitě musí této povaze odpovídat. Proto i sociologický a sociálně psychologický výzkum by měl ve své metodologii vycházet z povahy předmětných jevů, z jejich projevu i utváření. Předmětem sociologického a sociálně psychologického výzkumu jsou sociální a sociálně psychické jevy. Sociální a sociálně psychické jevy vznikají v aktivitách jedinců orientovaných na jiné jedince. Obsahově jsou charakterizovány jako:

- percepce (vnímání, přijímání informací),
- komunikace (výměna informací a proces sdělování),
- interakce (výměna činností s efektem nastolení nové situace, podmínek pro zúčastněné jedince).

Sociální a sociálně psychické jevy se formují ve vazbách (vztazích) vytvořených v průběhu těchto aktivit.

Charakter těchto vazeb je sice určen vlastnostmi osobností, které do nich vstupují, ale popis jednotlivých individuí a jejich chování nevystihuje charakter propojení mezi nimi ve všech zásadních momentech. Druhou důležitou částí charakteru těchto vazeb je způsob jejich formování obecnějšími faktory, např. jaké jsou zvyklosti, jaké jsou obvyklé způsoby komunikace a interakce, jaká pravidla jednání platí,

jaké jsou společné hodnoty a cíle atd. Souhrnně se tyto skutečnosti označují jako sociální instituce. Každé propojení, které je součástí sociálního a sociálně psychického jevu, je institucionální. Je tím vyjádřeno právě určité podřízení v chování jedinců zmíněným faktorům.

Sociální instituce korigující chování jedinců se sice v konečném důsledku projevují v chování a v jeho výsledcích, ale jejich utváření, působení a poté poznání není vždy na úrovni každého jedince možné. Jsou vnímány ve svém souhrnu a celkovém významu jako kultura daného společenství, skupiny. Sociální vazby nejsou tedy určeny jen přáním osob, mezi kterými existují nějaké vztahy, ale i kulturou společenství, skupiny, která uchovává, spoluutváří a zvyrazňuje sociální instituce [15].

3.1.3 Sociologický výzkum a statistika

Výzkum ve společenských vědách stojí před základním úkolem řešit určitý problém (ať již má teoretickou nebo praktickou povahu) na základě shromážděných faktů, tzn. spolehlivých, empiricky podložených informací o jevu. Proto také předpokladem úspěchu je správné vymezení základní otázky a nalezení objektivní metody získání a vyhodnocení informací. Většina výzkumů pracuje s teoreticky podloženými výchozími předpoklady v podobě hypotéz, jež musejí být formulovány tak, aby umožňovaly sběr dat a jejich analýzu. Záleží však nejen na zvolené metodě (kvalitativní nebo kvantitativní přístup), ale i na její další konkretizaci v tzv. procedurách (monografická, statistická, typologická, experimentální), které určují způsob výběru objektů, volbu technik sběru dat i zaměření analýzy [13].

V souvislosti s určením obsahu a základních postupů sociologického výzkumu je třeba se zmínit o rozdílu mezi ním a statistikou, myšleno především statistikou ekonomickou. Mnohdy je totiž sociologický výzkum chápán jako statistika v oblasti sociálních jevů. Je to zřejmě způsobeno využíváním matematického a statistického aparátu v tzv. kvantitativním výzkumu, který doposud převládal nebo převládá v empirickém výzkumu.

Statistika zachycuje v soustavě ukazatelů stav a proměny sledovaných skutečností. Ukazatele jsou definovány s ohledem na co nejužitečnější a nejpreciznější charakteristiku širší skutečnosti (ekonomiky, školství, životní úroveň atd.). Ukazatel je buď homogenní, nebo má stálou a ve značné míře známou dopočitatelnou (kvantifikovatelnou) strukturu. Je maximálně stabilizovaný a standardizovaný. Každý ukazatel má informační hodnotu sám za sebe, vypovídá bezprostředně o sledované skutečnosti [13, 15].

Sociologický výzkum má samozřejmě také takové cíle, jako je zachytit stav a proměnu sledovaných skutečností [15]. Termínem výzkum označujeme nepřilíh rozsáhlé výzkumné akce v sociologii s cílem vzájemné závislosti a podmíněnosti sledovaných jevů. Specifikem sociologického výzkumu je, že předmět poznání je vždy velmi složitý, a proto obtížně definovatelný v jednoznačné podobě nebo stavu [13].

Tato skutečnost se projevuje i v hledání a konstrukci ukazatelů zkoumaného jevu - parametrů sledování. V sociologickém výzkumu jsou utvářeny nebo hledány takové ukazatele, které by především označovaly širší sledovanou skutečnost a popsaly předmět výzkumu. Izolované jednotlivé ukazatele nemusí nést relevantní informaci z hlediska cílů výzkumu. Informace, které jsou relevantní ve vztahu k sociální realitě, vznikají kombinací více ukazatelů, tj. ve vztazích mezi jednotlivými

ukazateli. Z tohoto pohledu se jeví poněkud opomíjeným oborem zpracování sociologických dat tzv. datamining - hledání vícerozměrných vztahů v datech a sumarizace těchto vztahů, které nejsou na první pohled v datech patrné ve srozumitelné formě - tzv. modelu, který umožňuje snadnou interpretaci vztahů, stejně jako výpočet kvalitativních parametrů (vypovídací schopnosti) modelu.

Nejjednodušším příkladem rozdílu mezi statistickým ukazatelem a ukazatelem sociologického výzkumu je otázka v dotazníku. Sama odpověď na otázku v dotazníku má mnohdy velmi malou informační hodnotu. Lidé nemusí věci, kterých se otázka týká znát, mohou vědomě odpovídat jinak, než odpovídá skutečnost apod. Teprve vzájemné porovnání odpovědí na další otázky, i když budou stejně nepřesné, umožní odhalit sběr a tendenci v odpovědích a tato tendence, popř. směr, charakterizuje odpovídajícího. Naopak statistický výkaz obsahuje ukazatele, jejichž hodnoty jsou získávány dotazy, ale ty musí být, bez ohledu na subjektivitu výpovědi objektivní. K překonání tohoto paradoxu u některých ukazatelů si statistika vytváří příslušné nástroje - např. křížové otázky, omezení časového limitu pro odpověď atd.

Bez ohledu na rozdíl mezi statistickým (ve smyslu exaktního měření, i když s určitou chybou) a sociologickým zjišťováním je využití statistiky a matematicko-statistického aparátu v sociologických výzkumech značné, protože právě statistika disponuje základním aparátem dokazování (vyvrácení) výzkumných hypotéz obecně a metodami pro sumarizaci a vyjádření naměřených faktů na vyšší úrovni abstrakce. Podobně jako v jiných oborech, které statistiku využívají především jako nástroj, je samotné vložení smyslu a nalezení významu statistických zjišťování a jejich zařazení do sociologického poznání právě tím, čím se obory odlišují [15].

3.2 Dotazník

3.2.1 Dotazník – charakteristika a využití

Dotazníková metoda se nejčastěji charakterizuje jako výzkumná metoda, pomocí níž se zkoumají názory lidí na jednotlivé jevy. Sám dotazník se chápe jako listina předem připravených a formulovaných otázek o jistém problému, který je předmětem zkoumání; na tyto otázky zkoumaná osoba - tzv. respondent - odpovídá. Dotazník je tedy metodou hromadného zjišťování jistého jevu [16] a jako forma zpětné vazby náleží ke zcela specifickým metodám, používaným ve společenských vědách. Je to metoda, která shromažďování dat zakládá na dotazování osob. Charakterizována je tím, že je určena pro hromadné získávání údajů. Např. A. Ničkovič charakterizuje dotazník jako měrný prostředek, jehož pomocí se zkoumá mínění lidí o jednotlivých jevech. O možnostech dotazníku jako vědecké výzkumné metody se rozvinuly velké spory. Počátky užívání této metody jsou spojeny se jménem amerického psychologa a pedagoga G. S. Halla. Jeho zásluhou se používání dotazníkové metody značně rozšířilo. Stala se navzdory pochybnostem módní metodou výzkumu [17].

Dotazník je asi nejčastější výzkumný nástroj při sociologických výzkumech. Je to kontrolní aparát paměti, který chrání před zapomenutím výsledků. Písemné otázky v něm se mají stavět tak, aby byly získány pokud možno vyčerpávající od-

povědi na jednotlivé otázky, ale také s maximálním ohledem na popis daného jevu jako celku. Dotazník musí nabádat k odpovědím. Otázky musí být pro respondenta srozumitelné a formulovány tak, aby odpovědi mohly být jednoznačné. Respondent musí být schopen odpovídat bez nadměrné myšlenkové námahy. Je vyloučeno žádat odpovědi, které vyžadují hledat v dokumentech. Dotazník by měl být krátký a obsažný. V rozesílaném dotazníku se předpokládají jednodušší způsoby odpovědí než v dotazníku rozdávaném [18].

Dotazník je nenahraditelný v případech, kdy je třeba sebráný materiál podrobit kvantitativní analýze. Kvantitativní výzkum je zaměřen na níže uvedené charakteristiky jevů:

- rozsah výskytu, zastoupení. Rozumí se tím četnost nebo okruh sociálních subjektů, u kterých se vyskytuje, nebo oblast, ve které se vyskytuje (např. jaká část národa, pracovníků podniku nebo vysokoškoláků apod.),
- frekvence, tj. projev složek sociálního jevu v čase, např. opakování jeho výskytu, opakování výkon činnosti (např. frekvence hodnocení pracovníků, častost četby beletrie, opakování nákupů zboží rychlé spotřeby),
- intenzita jako mohutnost nebo síla jednotlivých složek sociálního jevu (např. síla zájmu, výše příjmu, míra spokojenosti) [18].

Tyto charakteristiky se převádějí do souboru stabilních ukazatelů, které jsou zajišťovány na velkých souborech jednotek (lidských jedinců).

Základním úkolem kvantitativního výzkumu je měření výše uvedených charakteristik a jejich souvislostí [15]. Ve výzkumech veřejného mínění, jako ve všech společenskovědních výzkumech, není možné pracovat s úplným popisem reality, ale vždy je třeba redukovat získaná data a informace [19]. Kvalitu zjištěných výsledků podmiňují nejvíce dva faktory: správně položené otázky (které jsou také správně interpretovány) a sestavený soubor respondentů, tedy těch, kteří na tyto otázky odpovídají [13]. Je prokázáno, že k tomu, aby byl zjištěn názor všech občanů na nějaký problém, není třeba dotazovat každého jedince. Stačí vytvořit podle určitých pravidel zmenšený reprezentativní vzorek zkoumané populace, který zaručuje, že názory v něm zjištěné, je možné vztáhnout na celou zkoumanou populaci. Zmenšený vzorek celé populace, v odborné terminologii nazývaný výběrový soubor, nemusí být extrémně velký, empiricky je dokázáno, že plně postačí dotázat se vzorku o velikosti cca 700 - 800 lidí. Podmínkou ovšem je, aby byl tento počet lidí správně vybrán. K tomu mají výzkumy veřejného mínění dobře vypracované a ověřené postupy. Jsou v zásadě dva. První je založen na tzv. kvótním výběru, což znamená, že se předem stanoví, které charakteristiky členů zkoumané populace (věk, pohlaví, vzdělání, velikost místa bydliště atd.) mohou zásadním způsobem ovlivnit rozdílnost názorů. Podle Gallupa, spočívá kvótní výběr v tom, že na základě osobního úsudku vybereme ze základního souboru skupinu jednotek v takovém složení, aby byl vytvořen zmenšený reprezentativní model základního souboru [20]. Rozhodující pak je zachovat ve výběrovém souboru stejné poměrné zastoupení všech vybraných charakteristik, jako je v celé populaci. Druhý postup je založen na náhodném výběru, který zaručuje, že každý člen zkoumané populace má stejnou šanci dostat se do vybraného vzorku dotázaných. Při tomto postupu jsou lidé pro výzkum vybíráni podle určitých kroků z existujících seznamů, které zahrnují všechny členy zkoumané populace [19].

Výzkum kvalitativní je, pokud jde o jeho zaměření, doplňkem výzkumu kvantitativního (nikoli ve smyslu podřízené části). Tím má být vyjádřeno, že sledovaná problematika má dvě vzájemně se doplňující části, tj. obsah a jejich jednotlivé kvantifikovatelné charakteristiky.

Úkolem kvalitativního výzkumu v sociologii je odhalovat neznámé skutečnosti o sociálních jevech, a to především:

- existenci těchto jevů a jejich strukturu,
- jejich vlastnosti a funkce,
- faktory, které sociální jevy ovlivňují nebo s nimi jinak souvisejí.

Kvalitativní výzkum v sociologii se orientuje na pochopení smyslu jednajících sociálních subjektů. Cílem kvalitativního výzkumu je vytváření adekvátního popisu nebo logické konstrukce celku sociálního jevu. Zachycuje vlastní dynamiku sociálních jevů a její příčiny. Novost, proměnlivost, individuální specifičnost těchto jevů v kvalitativním výzkumu způsobují, že standardizace postupů zkoumání není možná.

Okruhy metod a technik kvalitativního a kvantitativního výzkumu není možné přesně rozlišit, protože mnohé se, po případné lehké modifikaci, využívají jak v jednom, tak i druhém typu výzkumu. Pozorování je běžně považováno za součást kvalitativního výzkumu, ale při dobré formulaci pozorovaných objektů a při dostatečném množství pozorovacích situací lze výsledky statisticky zpracovat. Analýza dokumentů má rovněž charakter techniky kvalitativního výzkumu, ale podle charakteru dokumentů je tato technika použitelná i v kvantitativním výzkumu. Hloubkové nestandardizované rozhovory jsou uplatňovány vždy ve výzkumu kvalitativním a standardizované ve výzkumu kvantitativním. Anketa je rovněž součástí jak kvalitativního, tak kvantitativního výzkumu.

Kvalitativní a kvantitativní výzkum je také možné chápat jako dvě etapy empirického výzkumu. V první etapě je nutné rozkrýt sledovaný problém, má výzkum kvalitativní charakter. Poté, po dostatečném popisu problému je možné měřit jeho jednotlivé stránky a výzkum nabývá charakteru kvantitativního [15].

V souvislosti s problematikou dotazníkového průzkumu jako exaktního nástroje sociologie sloužícího pro účely kvantitativního výzkumu je nutné zmínit také existenci jevů, které jsou nepochybně sociologické povahy, zároveň je však nelze statisticky zkoumat. Jedná se především o jevy, jejichž částí jsou emoce (neexistuje objektivní ani správně stanovená stupnice, která by reflektovala objektivní stav jejich vnímání), subjektivní pocity (vnímání krásna, ale také např. účelnosti) atd. Přesto, že jsme schopni získat výsledky kvantitativního výzkumu jako v ostatních případech, musíme mít na zřeteli, že pracujeme s daty, které nelze objektivně interpretovat, neboť z principu (řeč, neidentifikovatelné skupinové jevy) nelze přesně stanovit měřící škálu těchto jevů. Z hlediska pohledu na sociologické jevy jako celek je pak zřejmé, že nelze objektivně zkoumat jevy jako např. historické změny, nálady ve společnosti atd. bez velmi pečlivě a měřitelně připravených otázek, které mají však omezenou vypovídací schopnost, což může být velmi snadno zneužito při účelových interpretacích výsledků. Popsané skutečnosti však nejsou pouze doménou sociologie, k disinterpretaci statistiky může dojít i v mnoha jiných, daleko exaktněji zaměřených oborech.

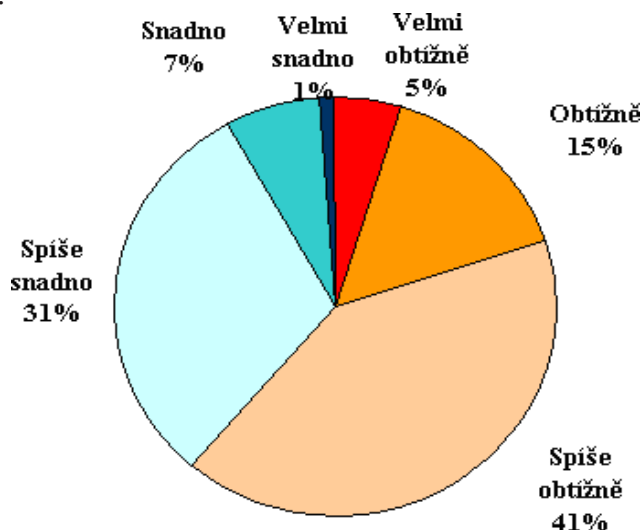
3.2.2 Dotazníková šetření v rámci ČR

Organizací, která v podmínkách ČR využívá dotazníkové šetření prováděné na velkém počtu respondentů jako jeden ze základních podkladů sociologicky zaměřených průzkumů je Český statistický úřad (ČSÚ). ČSÚ je ústředním orgánem státní správy České republiky. Byl zřízen dne 8. ledna 1969 zákonem č. 2/1969 Sb., o zřízení ministerstev a jiných ústředních orgánů státní správy. V souladu se zákonem o postavení ČSÚ, je jeho činnost, organizace, hospodaření a vztah k veřejnosti, k ostatním orgánům státní správy a územní samosprávy upraven Statutem Českého statistického úřadu. Činnost ČSÚ přináší velké množství cenných údajů o ČR (demografii, sociologické struktuře, hospodářství, ekonomice), které nelze jiným způsobem efektivně zjistit [22]. Typicky jsou získávané údaje periodicky zpracovávány (týdně, měsíčně, ročně, atd.) a řada z nich zveřejňována v podobě statistických ročenek a jiných materiálů.

Např. "Sčítání lidu, domů a bytů" patří k nejrozsáhlejšími statistickým zjišťováním v rámci ČR vůbec. Přípravu, organizaci, samotné provedení sčítání, zpracování a zpřístupnění jeho výsledků zajišťuje na základě zákona č. 296/2009 Sb., o sčítání lidu, domů a bytů v roce 2011 právě ČSÚ [21]. Následující sčítání je připravováno na březen 2011.

Dalším mediálně významným subjektem využívajícím dotazníkových šetření jako jeden ze základních nástrojů je např. agentura STEM s.r.o. (<http://www.stem.cz/>) – Středisko empirických výzkumů se zabývá celou řadou výzkumů jako např. stranické preference, popularita politiků, instituce, ekonomika, my a svět, vysvětlení naší společnosti a také sociální politika. Příkladem činnosti společnosti, která realizuje řadu privátních i veřejných zakázek týkajících se průzkumu veřejného mínění, může být např. dotazníkový výzkum v oblasti Sociální politiky, který byl zaměřen na skutečnosti spojené se sociální úrovní populace "Jak vycházíte se současným příjmem, který máte v domácnosti?" [19]. Průzkum ukázal, že výše příjmů obyvatelstva se odráží mimo jiné také ve výběru a konzumaci potravin.

Výpovědi o tom, jak domácnosti vycházejí se svými příjmy, se příliš nemění. V současné době 61 % lidí uvádí, že jejich domácnost vychází s příjmy obtížně, spořit se daří necelé polovině obyvatel (46 %). Snadněji vycházejí s rodinným rozpočtem lidé s maturitou a především vysokoškoláci nebo podnikatelé. O problematickém hospodaření s rodinnými příjmy hovoří nejčastěji lidé se základním vzděláním a nezaměstnaní [19].



Obrázek 1. Pramen: STEM, Trendy 3/2007, 1272 respondentů starších 18 let. [19]

Další agentury, které se zabývají výzkumem veřejného mínění jsou např. TNS Factum a CVVM. CVVM používá při svých šetřeních reprezentativní vzorek populace starší 15 let, zpravidla se jedná o soubor 1 000 – 1 200 respondentů, kteří jsou do výzkumu vybráni na základě metody kvótního výběru. CVVM provádí také výzkumy s větším souborem respondentů, k čemuž používá také pravděpodobnostních výběrových metod, jejichž volba je vždy závislá na potřebách konkrétního výzkumu. Tazatelskou síť CVVM tvoří cca 800 školených externích spolupracovníků [23].

3.2.3 Metody sestavení dotazníku

Přesná formulace konkrétního cíle a úkolů dotazníku ve vztahu ke zvolenému problému jsou základní podmínkou účelného koncipování dotazníku. Proto je zvláště důležité si uvědomit, že objektivní výsledek dotazníkové metody závisí v první řadě na teorii, z níž vědecký pracovník vychází, na promyšlené vědecké hypotéze. Z hlediska dané hypotézy stylizuje jednotlivé otázky a zaměřuje je na podstatné stránky zkoumaných jevů a procesů [17]

Dotazník je tedy metodou hromadného zjišťování jistého jevu [16]. Přípravu a hodnocení dotazníkového šetření lze popsat pomocí několika etap:

1. Stanovení výzkumného problému (vytyčení hypotézy a určení cíle) [1].

Tato první etapa začíná stanovením tzv. programové otázky, která přesně definuje cíl celého výzkumu. Jednoduše ji lze charakterizovat dotazy: Co chceme zkoumat a proč [7]? Po stanovení výzkumného problému a následného určení konkrétního cíle formulujeme hypotézy možného řešení problému. Hypotézy naznačují možné souvislosti mezi dvěma nebo více jevy, tedy proměnnými. Smyslem hypotéz je jejich potvrzení nebo vyvrácení. Na základě dalších souvisejících analýz (např. studium odborné literatury, využití předchozích výzkumů z dané oblasti konzultace s odborníky) jsou některé hypotézy vyloučeny.

2. Vybírání reprezentativního vzorku pro výzkum, stanovení rozsahu a způsobu výběru.

Počet respondentů má být 2x až 3x větší, než vyžaduje cíl výzkumu. Je to proto, že návrat vyplněných dotazníků tvoří pouze 30-40% z celkového počtu zaslaných, zejména pokud jde o dotazníky zaslané poštou [1].

Přesné zjištění počtu zkoumaných osob vyžaduje stanovení nutného rozsahu výběru. Při správném postupu není porušena reprezentativnost výběru, zejména je potom možné využít všech dostupných technik zpracování [7]. V praxi se obvykle setkáváme se třemi odlišnými přístupy:

- a. Slepý odhad – zde se odhad provádí subjektivně na základě dosavadních zkušeností. [7] Čím je zkušenost výzkumníka větší, tím je lepší přesnost a spolehlivost takového výzkumu.
- b. Nákladový přístup – zde se vychází z kalkulace nákladů na šetření a zpracování.[7]
- c. Statistický přístup - velikost vzorku je zde stanovena na základě statistických metod s ohledem na požadovanou míru přesnosti a spolehlivosti.

Nutný rozsah lze vyjádřit ze vztahu:

$$n = \frac{u_a^2 \cdot p \cdot q}{\Delta^2}, \text{ kde}$$

α = hladina významnosti (obvykle volíme 0,05 či 0,01),

u_a = tabelovaná hodnota normovaného normálního rozdělení,

p = podíl počtu respondentů znalých problematiky, příklánějící se k jedné variantě,

q = podíl počtu respondentů znalých problematiky, příklánějících se k druhé variantě. Obvykle tyto počty dopředu neznáme a v takovém případě volíme $p=q=0,5$ [7]

3. Vlastní rozpracování dotazníku [7].

Vlastní sestavení dotazníku má kromě výše uvedených požadavků vyhovovat i určité taktice v pořadí otázek. První otázky se mají týkat zkoumaných osob a teprve další problému, který chceme zkoumat [1]. Každá otázka by měla mít vztah ke zkoumanému problému. Otázky musí vytvářet uzavřený konsekventní celek [18]. Pořadí otázek sestavujeme tak, že zpočátku uvádíme otázky, které vzbuzují zájem respondentů o zkoumanou problematiku a nevyžadují mnoho přemýšlení [1]. Je třeba začínat s jednoduchými otázkami (ne otázkami typu „proč“) [18]. Pokud možno postupujeme od lehčího k těžšímu a od obecnějšího k specifickému [1].

Otázky dělíme podle způsobů odpovědi na uzavřené a otevřené. Odpovídá-li zkoumaná osoba na otázky, které jsou v dotazníku uvedeny podtrháním, kroužkováním a jiným označením odpovědi, hovoříme o uzavřené formě dotazníkové otázky [1]. Uzavřené otázky mají předem stanoveny varianty odpovědí. Tyto varianty by měly být správně připraveny, aby nenutily respondenta přiklonit se k odpovědi, která mu ne zcela vyhovuje. To lze zajistit při důkladném předvýzkumu. Uzavřené otázky se používají pro svou jednoduchost, kvalitativní stejnorodost odpovědí a snadnost jejich zpracování [15].

Má-li zkoumaná osoba možnost si sama zformulovat odpověď na otázku podle vlastního uvážení, hovoříme o otevřené formě dotazníkové otázky [1]. Otevřené otázky dávají respondentovi velkou svobodu k vyjádření, ale jsou pro něj velmi obtížné, stejně, jako jsou obtížné k interpretaci. Variabilita odpovědí na tyto otázky je velmi široká. Často jsou odpovědi na různé úrovni obecnosti a nespolehlivé a zdlouhavě se zpracovávají [15]. Z hlediska statistického zpracování dotazníku je první forma, tedy uzavřená forma otázek dotazníku výhodnější.

Podle formulace otázky dělíme na přímé a nepřímé. Přímé otázky jsou formulovány tak, že je respondentovi zřejmý smysl dotazu. Podle tohoto zjevného významu nemá otázka žádný jiný skrytý význam. Otázky přímé formou jsou např. psychotaktické otázky. Umožňují respondentovi odpovědět pravdivě bez zvyšování vnitřního napětí i na otázky, které ohrožují prestiž respondenta. Skutečný význam otázky je skryt zdánlivým významem, nebo je vztažen k jiným osobám.

Pokud se má např. ve výzkumu zjistit jaký je podíl všech kuřáků, včetně příležitostných, volí se místo přímé otázky: „Jste kuřák?“ otázka: „Kdybyste měl(a) možnost si vybrat pro sebe cigaretu ze všech značek, které znáte, kterou značku byste si vybral(a)?“ Na tuto druhou otázku skuteční nekuřáci odpovědí jednoznačně: „Žádnou.“ I výjimeční kuřáci jednu značku s největší pravděpodobností uvedou.

Druhým typem psychosomatického dotazu je otázka zjišťovací. Například při odpovědích na otázku „Myslíte si, že většina vašich spolupracovníků jsou lidé upřímní?“

je ověřeno, že s vysokou mírou pravděpodobnosti lidé, kteří jsou sami upřímní, odpovědí na otázku: „Ano“. Lidé, kteří upřímní nejsou, odpovědí: „Ne“. Tyto psychotaktické otázky pomohou překlenout problém zábran respondenta odpovídat pravdivě.

Nepřímé otázky jsou opakem otázek přímých. Tyto otázky mají pro respondenta zjevný a samozřejmý význam, ale za ním se skrývá více nebo méně skrytý význam další. Takové otázky se používají především v sociálně psychologických výzkumech [15].

Důležité je stanovení počtu otázek. Příliš málo nebo naopak příliš velké množství otázek může ohrozit kvalitu následných analýz. V zásadě jde o to, aby byl počet otázek takový, aby bylo možné dobře analyzovat daný problém a přitom dotazník a jeho vyplnění neunavoval respondenta [7].

Podle R. Ničkoviče [16] musí dotazník splňovat tyto podmínky:

- Význam zkoumaného tématu odpovídá vynaloženým prostředkům, času a lidské síle.
- Význam tématu a cíl je třeba vysvětlit v návodu.
- Vyžaduje jen nevyhnutelné informace, které není možné získat jiným způsobem.
- Zabezpečuje vysoký stupeň motivace zkoumaného.
- Není příliš dlouhý a jeho vyplňování nevyžaduje mnoho času.
- Svou vnější formou je uspořádaný a přitažlivý.
- Nezasahuje do intimních citů a života zkoumaného.
- Zabezpečuje anonymitu.
- Má jasný důvod.
- Umožňuje lehké a rychlé uspořádání a tabulaci.
- Otázky v dotazníku mají být formulovány jasně, krátce, jednoznačně, srozumitelně.
- Nemají se používat zbytečná slova a nezvyklé termíny.
- Otázky nesmějí vsugerovat odpovědi podle přání výzkumníka.
- Otázky mají být uspořádány podle logického pořadí a logických stupňů [16].

4. **Ověření dotazníku**, tzv. předvýzkum

Po ukončení všech předchozích stádií dotazníkového průzkumu nastává etapa ověření dotazníku [7], předvýzkum je jakousi „generální zkouškou“ vlastního výzkumu [15], tzn. vyzkoušení dotazníku na předběžném vzorku osob [1]. Podle V. Příhody [16] má být tento vzorek tvořen alespoň 30 osobami. Osoby se nemají ničím lišit od těch, které chceme zkoumat [16]. Předvýzkum obvykle ukáže, nakolik jsou navrhované techniky sběru dat použitelné a srozumitelné. Předvýzkum poskytuje první empirická data, která lze využít při přípravě kategorizace a při vytváření schématu statistického zpracování. Výsledky předvýzkumu se do vlastního terénního sběru dat nezahrnují především z důvodu změn, které mohou být po předvýzkumu provedeny [15]. V případě zjištěných nedostatků na malém vzorku osob je možné udělat jednoduše a rychle nápravu, aby vlastní šetření proběhlo správně a bylo z něj možno činit odpovídající závěry [7].

Dotazník kromě předností, které z něj činí jednu z nejpoužívanějších technik výzkumu (získání značného množství údajů týkajících se určitého problému, obsáhne velký počet respondentů a na poměrně velkém prostoru, nevyžaduje mnoho času na provedení, není spojen s velkými finančními náklady, dává čas na odpověď,

snadno se zpracovává apod.) [18] má i své záporné stránky. Patří k nim zejména těžkosti při získávání otevřených a pravdivých odpovědí na tiskovém formuláři [1], malá návratnost dotazníku, technika je nepružná vzhledem k nepředvídaným situacím [18].

5. Vlastní shromažďování materiálu [7].

6. Analýza získaného materiálu a jeho zobecnění [7].

Po získání údajů v terénu a po jejich kontrole a přípravě na další zpracování je nutné přistoupit ke skutečné analýze dat a k jejich interpretaci. Data postupně procházejí dalším zpracováním (surové, primární údaje v dotazníku postupně procházejí kódováním a dalšími transformacemi), mění se na sekundární data: počítače dnes zpracovávají různé statistiky, nabízejí tabulky, údaje o souvislostech mezi proměnnými, umožňují i další transformace jednotlivých znaků. Po následné analýze se stávají podkladem formulace empirických tvrzení (faktů) [14].

3.3 Dotazníkové průzkumy u konzumentů potravin

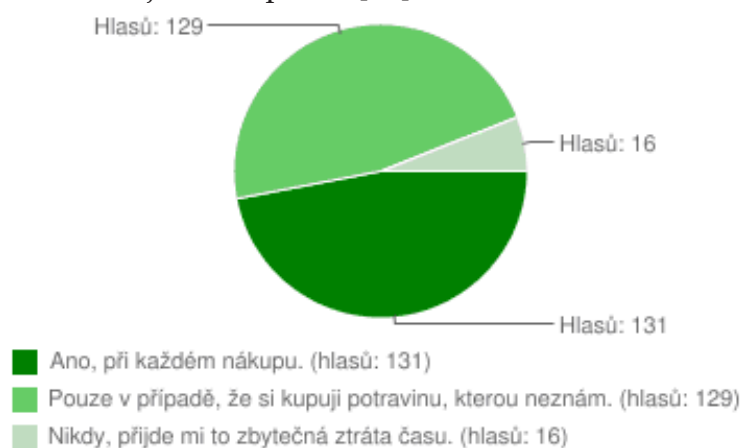
Dotazníkové průzkumy jsou jednou z velmi efektivních metod marketingu. Cílem marketingu je směnou - transakcí - uspokojící potřeby zákazníka, zvítězit nad silou konkurentů a reakcemi na podněty vytvářet trvalý vztah se zákazníkem. Různé druhy prostředí, například globální, lokální, marketingové, logistické nebo procesní, profilují aktivní konkurenty. Marketing konkuruje nemarketingovým přístupům v okamžiku přípravy a vlastního prodeje nebo nákupu (směny). Marketing je společenský a řídicí proces, kterým jednotlivci a skupiny získávají to, co potřebují, prostřednictvím tvorby, nabídky a směny hodnotných produktů s ostatními [24]. Potraviny a jejich zdroje tvoří stále velmi významnou část ekonomiky ČR i jiných zemí. S nadsázkou se dá říci, že potraviny zůstanou stabilním sektorem prodeje v jakékoli, i krizové situaci. Z tohoto důvodu existuje v sektoru také vysoká nasycenost trhu a zejména v sociálně problematickém období vysoké zaměření konzumenta na cenu produktu. V souvislosti s ekologizací a zdravotní osvětou společnosti však také přibývají parametry, na jejichž základě zákazník potraviny či jejich zdroje vybírá. Proto je marketing v potravinářské produkci nezastupitelným nástrojem ekonomického profitu.

V současné době v souvislosti s výše uvedenými skutečnostmi na trhu s potravinami probíhají různé marketingové průzkumy zabývající se např. zdravím konzumentů v souvislosti s potravinami, potravinami a produkty prospěšnými zdraví (ať už s jasně či méně jasně definovatelnými benefity) a také vzdělávání v oblasti šetrného či účelného využívání potravinových zdrojů. Tyto průzkumy jsou obvykle součástí koncepce obchodní politiky firmy, které zahrnuje průzkum trhu, propagaci výrobků a služeb, vývoj nových výrobků, přičemž cílem je vždy dosáhnout maximálního ekonomického efektu.

V této souvislosti můžeme zmínit např. projekt Škola spotřebitele, který vznikl ve spolupráci Státní zemědělské a potravinářské inspekce a České televize [25]. Tento projekt vznikl jako reakce SZPI na výsledky výzkumu veřejného mínění "Bezpečnost potravin a český spotřebitel", který inspekce realizovala koncem roku 2005.

Česká televize odvysílala jednotlivé krátké informační šoty v pořadu Sama doma. Cílem projektu Škola spotřebitele je poskytnout občanům základní rady a informace týkající se nakupování potravin a nakládání s nimi. Součástí projektu jsou zpřístupněná videa na internetových stránkách s tématy např. Zdravé potraviny, rizikové potraviny, mražené potraviny, údaje, kterých by si spotřebitel měl u potravin všimnout, informace na obalech a další.

Státní zemědělská a potravinářská inspekce provádí na svých internetových stránkách různá pravidelná anketní šetření, např. "Uvítali byste uvádění nutričních hodnot u většiny druhů potravin?", "Podle jakých kritérií si vybíráte doplňky stravy?", "Jak často si kupujete instantní potraviny?", "Sledujete informace na obalech potravin před tím, než si je zakoupíte?" [26].



Obrázek 2. Vyhodnocení ankety: Sledujete informace na obalech potravin před tím, než si je zakoupíte? Celkem hlasů: 276 Hlasováno dne 1.2.2011. Zdroj: SZPI [26].

Dotazníková šetření z oblasti potravin jsou častým tématem marketingových studií jako součást závěrečných prací studentů vysokých škol (zdroj <http://www.vyplnto.cz>), jak dokládají následující témata [27]:

- Nakupování v domácnostech.
- Nákupní zvyklosti.
- Jak nakupujete potraviny? - malé prodejny vs. hypermarkety".
- Situace na trhu potravin v ČR.
- Prodej ze dvora v ČR.
- Stav výživy a zdraví lidí v ČR.
- Vliv reklamy na spotřebitele.
- Ochrana spotřebitelů-označování potravin.

Dalším příkladem marketingového průzkumu v potravinářství zacíleného na screening trhu, nalezení segmentů trhu a podkladů pro výhled na ekonomický profit, je například následující průzkum sektoru trhu s biopotraviny. Trh BIO potravin v ČR je zatím poměrně málo rozvinut, má ovšem předpokládaný velký potenciál růstu, jako přirozený následek rostoucí oblíbenosti těchto výrobků po celém světě.

Např. časopis Foodservice (<http://www.foodservice.cz>) uvádí, že biopotraviny byly i v letošním roce v obchodních řetězcích dražší průměrně o 100 % [28]. Jak zjistil průzkum občanského sdružení PRO-BIO Liga, realizovaný v roce 2010 ve čtyřech českých městech, nejmenší příplatek za biokvalitu zaznamenáváme u mléčných výrobků (44 %), alkoholických nápojů (96 %) a masa a masných výrobků (109 %), nejvyšší naopak u olejů (279 %) [29].

Prezentovaná data jsou výsledkem sběru cen biopotravin ve čtyřech největších městech České republiky a sběru cen konvenčních potravin v prodejnách obchodních řetězců. V každém městě byl sběr proveden v jedné prodejně každého z osmi nejvýznamnějších řetězců s potravinami a ve třech největších specializovaných prodejnách biopotravin v daném městě. Nejedná se o reprezentativní průzkum průměrných cen konvenčních potravin s cenami biopotravin, ale víceméně o kvalitativní průzkum cen a nabídky biopotravin a s nimi porovnatelných potravin z konvenčního zemědělství. Ceny byly sbírány v červnu a listopadu 2010 v Praze, Brně, Olomouci a Plzni.

Biopotraviny bývají v průměru o 100 % dražší než srovnatelné konvenční potraviny. Nicméně jednotlivé komodity mohou být dražší i o několik stovek procent, ale i o pouhé jednotky procent. Nejvyšší cenový rozdíl byl zaznamenán u olejů a sušenek (až 300 %) naopak nejnižší tradičně u třtinového cukru a mléčných výrobků (2 % a 20 %). Z kvalitativního hlediska je nutné připomenout, že významný cenový rozdíl u olejů a sušenek je dán skutečností, že výrobky v biokvalitě bývají nabízeny pouze ve specifické velmi vysoké kvalitě (např. rostlinné oleje Emile Noël, slepované sušenky Molenaartje), tudíž kvantitativní srovnání cen pohled může zkreslovat.

Pokud zákazník nakupuje biopotraviny v obchodních řetězcích, ušetří průměrně 19 % z ceny oproti nákupu biopotravin ve zkoumaných specializovaných obchodech. Nicméně bioobchody nabízejí vyšší standard informací, atp. o biopotravinách a ekologickém zemědělství.

Z poptávaného košíku 100 položek nabízela prodejní místa průměrně 33 potravin. Nelze však říci, že by některé položky nebyly dostupné vůbec a jiné všude, spíše v jednotlivých místech byly dostupné různé druhy položek. Podstatné rozdíly byly zaznamenány mezi jednotlivými městy. Nejvyšší nabídka byla zaznamenána v Praze nejnižší naopak v Plzni.

Pro milovníky lokálních potravin průzkum přinesl zjištění, že tuzemských či alespoň v tuzemsku zabalených či místním dodavatelem distribuovaných biopotravin je v nabídce průměrně 38 % [29].

4 Vlastní práce

4.1 Charakteristika souboru účastníků vzdělávacích seminářů v zemědělství

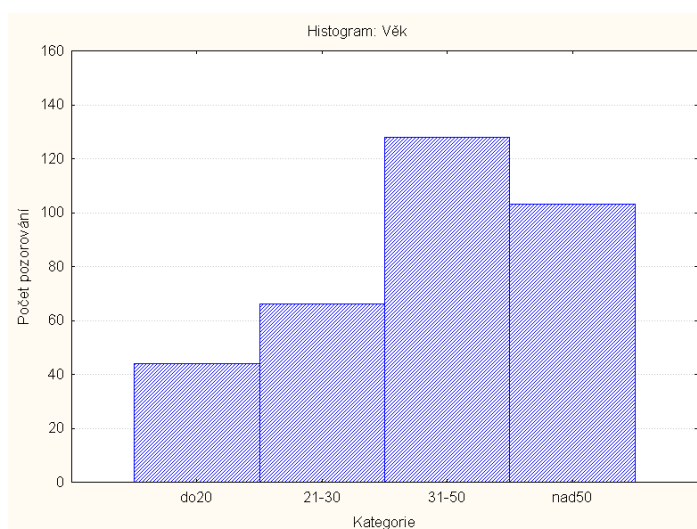
Datový soubor obsahuje celkem 342 vyplněných dotazníků. Počet respondentů souhrnně zobrazuje tabulka 3.

	Count	Percent
maso1	87	25.43860
maso2	52	15.20468
mléko1	81	23.68421
mléko2	37	10.81871
výživa1	85	24.85380

Tabulka 3. Absolutní a relativní počty respondentů na jednotlivých seminářích.

Z celkového počtu 342 respondentů se vzdělávacích akcí zúčastnilo 191 mužů a 114 žen, 37 respondentů nevedlo do dotazníku své pohlaví. Z tohoto hlediska je počet mužů a žen vzhledem ke struktuře pracovníků v zemědělství [30] vyrovnaný.

Graf 1. ukazuje věkovou strukturu účastníků. Nejvíce zastoupena je věková skupina 31-50 let (128 respondentů), dále skupina nad 50 let (103 respondentů), 66 respondentů ve skupině 21-30 let a 44 respondentů ve věku menším nebo rovném 20 rokům.



Graf 1. Věková struktura účastníků vzdělávacích akcí v zemědělství.

Při průzkumu úrovně vzdělání jednotlivých respondentů byly zjištěny počty uvedené v tabulce 4.

	Count	Cumulative - Count	Percent	Cumulative - Percent
základní	30	30	8.77193	8.7719
učební obor	9	39	2.63158	11.4035
středoškolské	102	141	29.82456	41.2281
vysokoškolské	161	302	47.07602	88.3041
vědecké	40	342	11.69591	100.0000
Missing	0	342	0.00000	100.0000

Tabulka 4. Úroveň vzdělání respondentů.

Lze konstatovat, že vzdělávacích akcí se nejvíce zúčastňovali respondenti s vysokoškolským a středoškolským vzděláním, překvapivým závěrem vypovídajícím o vysoké odborné úrovni seminářů je účast 40 respondentů s vědeckým vzděláním.

Obory vzdělání respondentů shrnuje tabulka 5. Naprostou většinou je zastoupeno vzdělání v zemědělských oborech (287 respondentů), minoritně některé blízké obory – potravinářské (17 respondentů) a veterinární vzdělání (13 respondentů). Všechny semináře navštívilo pouze 24 respondentů s jiným oborem vzdělání (přírodovědné, humanitní, technické, jiné). Jeden respondent neuvedl obor vzdělání. Popsané skutečnosti ukazují, že semináře byly cíleny na skupinu účastníků ze zemědělského sektoru, poněkud překvapivá je vzhledem k tematické náplni seminářů směrem ke zpracování živočišné produkce poměrně nízká účast respondentů s potravinářským či veterinárním vzděláním.

	Count	Cumulative - Count	Percent	Cumulative - Percent
zemědělské	287	287	83.91813	83.9181
přírodovědné	3	290	0.87719	84.7953
potravinářské	17	307	4.97076	89.7661
veterinární	13	320	3.80117	93.5673
humanitní	3	323	0.87719	94.4444
technické	9	332	2.63158	97.0760
jiné	9	341	2.63158	99.7076
Missing	1	342	0.29240	100.0000

Tabulka 5. Zastoupení oborů vzdělání u respondentů.

V tabulce 6 a grafu 2 je uvedena profesní skladba respondentů.

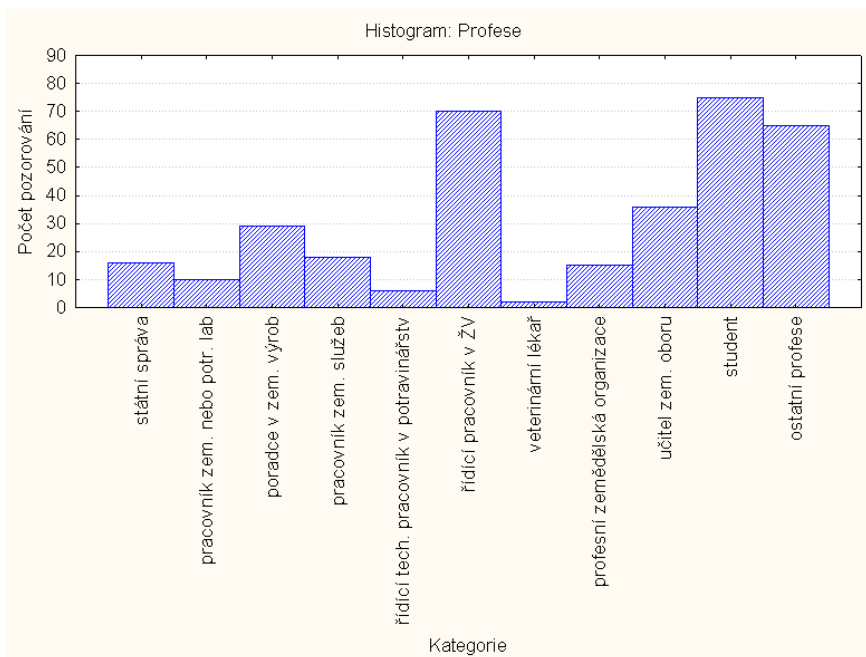
	Count	Cumulative - Count	Percent	Cumulative - Percent
státní správa	16	16	4.67836	4.6784
pracovník zem. nebo potr. lab	10	26	2.92398	7.6023
poradce v zem. výrobě	29	55	8.47953	16.0819

pracovník zem. služeb	18	73	5.26316	21.3450
řídící tech. pracovník v potravinářství	6	79	1.75439	23.0994
řídící pracovník v ŽV	70	149	20.46784	43.5673
veterinární lékař	2	151	0.58480	44.1520
profesní zemědělská organizace	15	166	4.38596	48.5380
učitel zem. oboru	36	202	10.52632	59.0643
student	75	277	21.92982	80.9942
ostatní profese	65	342	19.00585	100.0000
Missing	0	342	0.00000	100.0000

Tabulka 6. Profesní skladba respondentů – účastníků seminářů.

Z tabulky 6 je patrné, že vzdělávacích seminářů se z celkového počtu respondentů zúčastnilo 70 řídicích pracovníků v živočišné výrobě, což ukazuje na vysoký zájem dalšího vzdělávání vedoucích pracovníků v provozních podnicích. Podobně lze registrovat poměrně vysokou snahu o další vzdělávání u poradců v zemědělské výrobě – 29 respondentů. Pozitivně lze hodnotit taktéž účast pedagogů středních a vysokých škol – 36 respondentů. Celkem se pěti vzdělávacích seminářů zúčastnilo 75 studentů středních a vysokých škol, bez dalších údajů o respondentech však není možné z tohoto počtu usuzovat na generalizované výsledky. Počet respondentů, jejichž profese je realizována v rámci státní správy byl v rámci všech vzdělávacích akcí 16. Zastoupen byl také sektor zemědělských služeb, který má vysoký potenciál přenosu recentních poznatků do praxe, počtem 18 respondentů. Ostatní profese jsou zastoupeny počtem 65 respondentů.

Ze sociologického hlediska je zajímavé srovnání úrovně vzdělání, oboru vzdělání a profesního zaměření respondentů. Pomocí tabulky vícenásobných odpovědí jsme dospěli k následujícím výsledkům (diskutovány jsou počty nad 10 respondentů). 29 respondentů se základním vzděláním v zemědělském oboru uvádí jako profesi student. Je tedy zřejmé, že absolvovali vzdělávací semináře v rámci výuky. 34 respondentů se středoškolským vzděláním (celkový počet 102 respondentů) v zemědělském oboru pracuje na pozicích řídicích pracovníků v živočišné výrobě. 29 těchto respondentů uvedlo jako profesi student – jedná se tedy o studenty vysokých škol. Z celkového počtu 161 respondentů s vysokoškolským vzděláním 134 respondentů uvedlo obor vzdělání zemědělský. Z vysokoškolsky vzdělaných respondentů pracuje 11 ve státní správě, 21 jako poradci v zemědělské výrobě, 12 jako pracovníci zemědělských služeb, 25 jako řídicí pracovníci v živočišné výrobě, 22 jako učitelé zemědělských oborů a 22 v ostatních, nevyjmenovaných profesích. Signifikantní počty účastníků s vědeckým vzděláním pocházely ze sektoru vysokoškolských pedagogů (učitel zemědělských oborů – 11) a ostatních profesí (14 respondentů) z celkového počtu 40 účastníků s vědeckým vzděláním.



Graf 2. Profesní skladba respondentů.

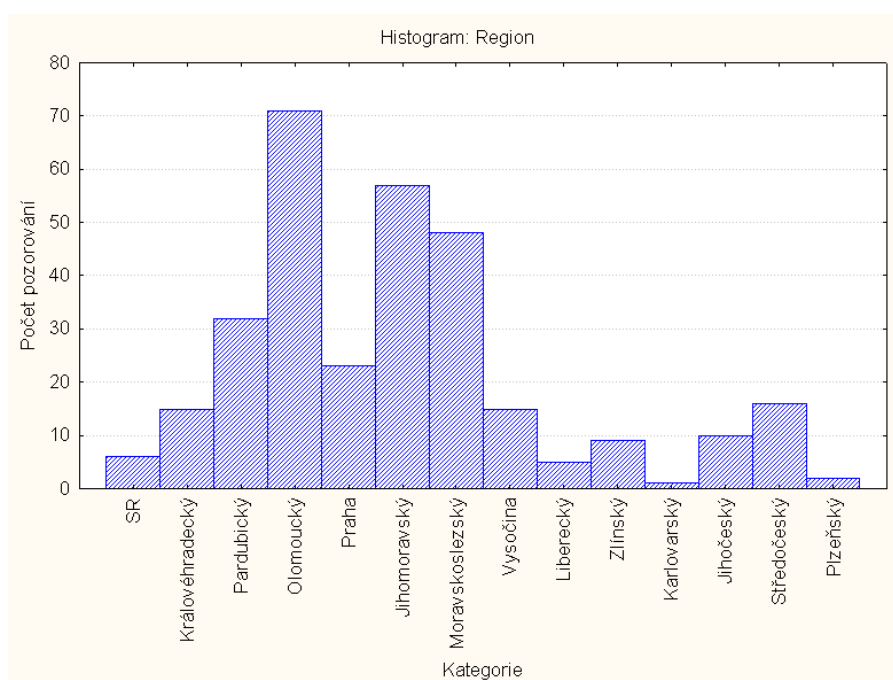
Tabulka 7 a graf 3 uvádí počty účastníků vzdělávacích seminářů z jednotlivých krajů ČR a SR. Semináře zaměřené na produkci mléka a hovězího masa se konaly v Rapotíně u Šumperka, seminář zaměřený na výživu skotu v Pohořelicích. Těmto skutečnostem odpovídá i regionální skladba respondentů – spádová oblast Olomouckého (71 respondentů), Jihomoravského kraje (57 respondentů). Významně je zastoupen také Moravskoslezský kraj (48 respondentů) a Pardubický kraj (32 respondentů). 6 účastníků navštívilo semináře ze Slovenské republiky. Z kraje Praha bylo přítomno 23 respondentů, ze Středočeského kraje 16, což ukazuje na vysoký zájem o další vzdělávání i v regionech, kde zemědělská produkce není primární.

Lze konstatovat, že krajově jsou zastoupeny především tradičně zemědělské regiony s většími výrobními podniky (což by ukazovalo na vyšší zájem managementu o další vzdělávání v oboru). Tradičně zemědělský region Vysočina je však zastoupen pouze 15 účastníky.

	Count	Cumulative - Count	Percent	Cumulative - Percent
SR	6	6	1.75439	1.7544
Královéhradecký	15	21	4.38596	6.1404
Pardubický	32	53	9.35673	15.4971
Olomoucký	71	124	20.76023	36.2573
Praha	23	147	6.72515	42.9825
Jihomoravský	57	204	16.66667	59.6491
Moravskoslezský	48	252	14.03509	73.6842

Vysočina	15	267	4.38596	78.0702
Liberecký	5	272	1.46199	79.5322
Zlínský	9	281	2.63158	82.1637
Karlovarský	1	282	0.29240	82.4561
Jihočeský	10	292	2.92398	85.3801
Středočeský	16	308	4.67836	90.0585
Plzeňský	2	310	0.58480	90.6433
Missing	32	342	9.35673	100.0000

Tabulka 7. Regionální zastoupení účastníků seminářů.



Graf 3. Struktura souboru respondentů z hlediska regionálního zařazení.

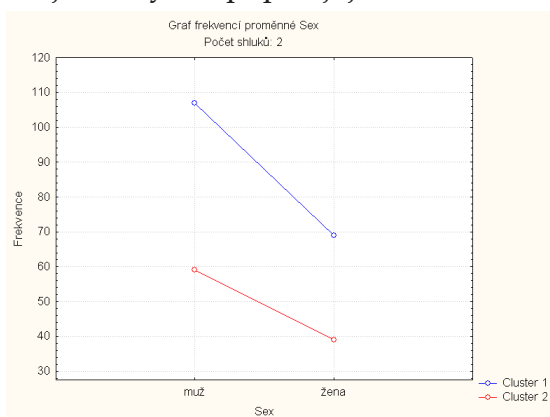
4.2 Charakteristika typického účastníka vzdělávacích seminářů v zemědělství pomocí shlukovacích algoritmů

Tabulka 8. zobrazuje výsledky shlukové analýzy pomocí K-means algoritmu.

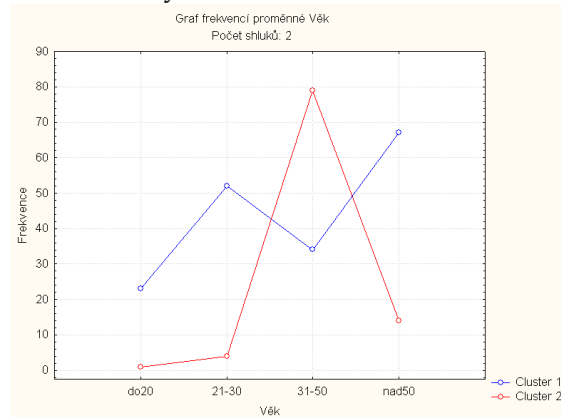
Cluster	Sex	Age	Education	Field of Education	Profession	Region	Number of cases	Percentage(%)
1	muž	nad50	VŠ	zemědělské	student	Moravskoslezský	176	64.23358
2	muž	31-50	SŠ	zemědělské	řídící prac. ŽV	Olomoucký	98	35.76642

Tabulka 8. Výsledky shlukové analýzy pomocí k-means algoritmu.

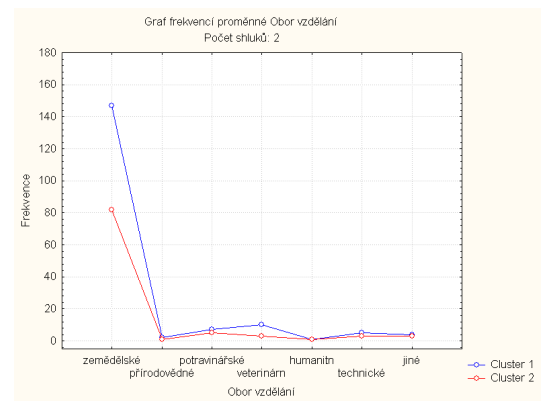
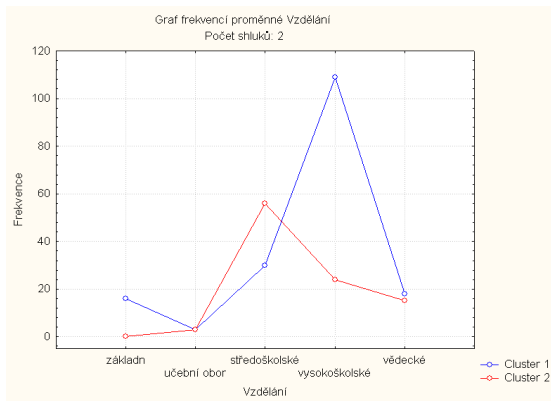
Typickými účastníky odvozenými podle tohoto algoritmu na datové sadě reprezentovanými centroidy 2 odvozených shluků jsou: muž ve věku nad 50 let se zemědělským vysokoškolským vzděláním a profesí studenta (interpretací nesmyslný centroidní reprezentací výsledků) z Moravskoslezského kraje – tuto podmínku, která zahrnuje profesní zařazení, by splňovalo 64,23 % respondentů. Další odvozenou skupinou reprezentovanou centroidem shluku, která by teoreticky zahrnovala 35,77 % respondentů jsou muži se středoškolským vzděláním se zemědělským zaměřením, kteří pracují jako řídící pracovníci v živočišné výrobě v Olomouckém kraji. Grafy 4-9 popisují jasně strukturu obou odvozených shluků.



Graf 4. Struktura odvozených shluků pomocí k-means algoritmu z hlediska proměnné Pohlaví.

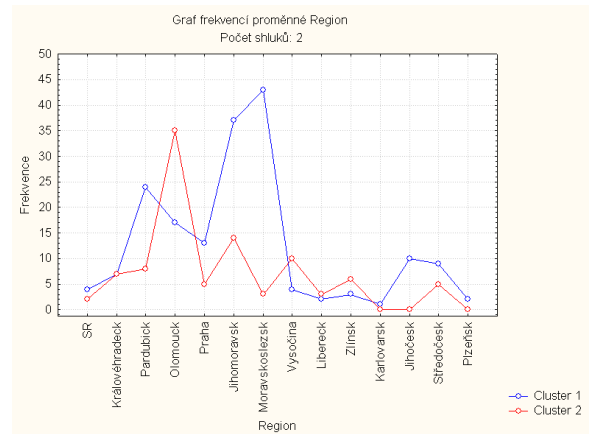
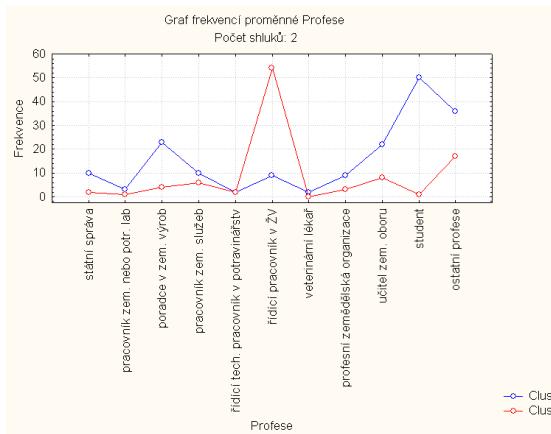


Graf 5. Struktura odvozených shluků pomocí k-means algoritmu z hlediska proměnné Věk.



Graf 6. Struktura odvozených shluků pomocí k-means algoritmu z hlediska proměnné Vzdělání.

Graf 7. Struktura odvozených shluků pomocí k-means algoritmu z hlediska proměnné Obor vzdělání.



Graf 8. Struktura odvozených shluků pomocí k-means algoritmu z hlediska proměnné Profese.

Graf 9. Struktura odvozených shluků pomocí k-means algoritmu z hlediska proměnné Region.

	df	Chi-square	p value
Sex	1	0.0092	0.923508
Age	3	99.7914	0.000000
Education	4	61.2127	0.000000
Field of Education	6	1.4409	0.963322
Profession	10	104.5367	0.000000
Region	13	64.5490	0.000000

Tabulka 9. Chí-koadrát test mezi odvozenými shluky v jednotlivých sledovaných znacích pomocí k-means algoritmu.

Následující grafy ukazují výsledky shlukování pomocí EM algoritmu, který pokrývá celý datový prostor, na kterém demonstruje jeho ideální dělení. Pomocí grafu 10 je zobrazena struktura obou odvozených shluků typického účastníka vzdělávacího semináře v zemědělském sektoru. První z odvozených shluků obsahuje 13,69 % všech případů, druhý potom 86,31 % všech případů. Shluky tedy dělí úplně datový prostor případů.

Oba shluky jsou reprezentovány následujícími frekvenčními tabulkami.

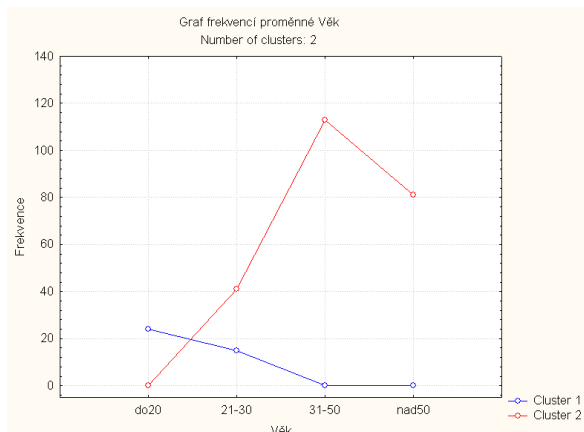
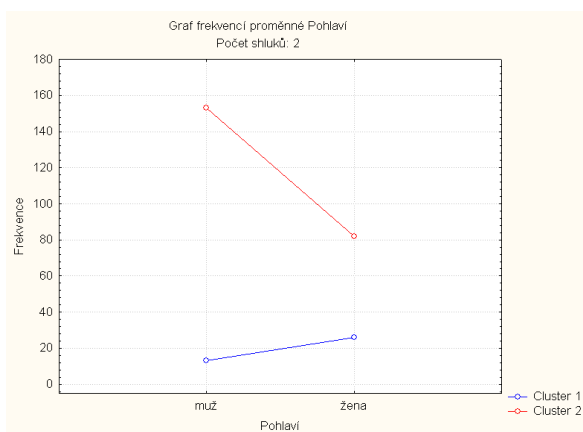
Cluster	Sex	Age	Education	Field of Education	Profession	Region	Number of cases	Percentage(%)
1	muž	do20	středoškolské	zemědělské	student	Moravsko-slezský	39	13.69
2	muž	31-50	vysokoškolské	zemědělské	řídící pracovník v ŽV	Olomoucký	235	86.31

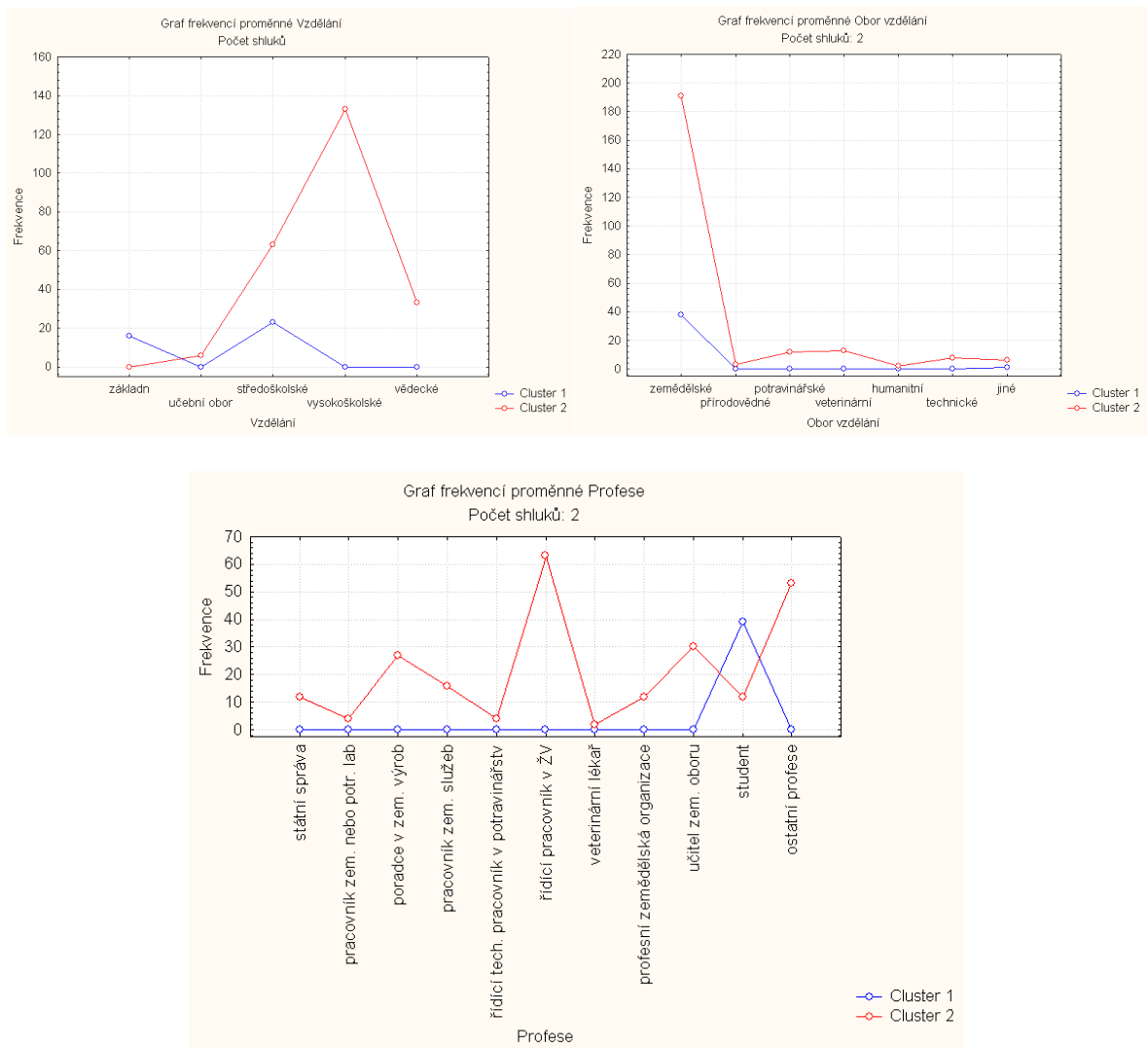
Tabulka 10. Výsledky shlukové analýzy pomocí EM algoritmu.

	Cluster 1	Cluster 2	Total
muž	13	153	166
žena	26	82	108
do20	24	0	24
21-30	15	41	56
31-50	0	113	113
nad50	0	81	81
základní	16	0	16
učební obor	0	6	6
středoškolské	23	63	86
vysokoškolské	0	133	133
vědecké	0	33	33
základní	16	0	16
učební obor	0	6	6
středoškolské	23	63	86
vysokoškolské	0	133	133
vědecké	0	33	33
státní správa	0	12	12
pracovník zem. nebo potr. lab	0	4	4
poradce v zem. výrobě	0	27	27
pracovník zem. služeb	0	16	16

řídící tech. pracovník v potravinářství	0	4	4
řídící pracovník v ŽV	0	63	63
veterinární lékař	0	2	2
profesní zemědělská organizace	0	12	12
učitel zem. oboru	0	30	30
student	39	12	51
ostatní profese	0	53	53
SR	0	6	6
Královéhradecký	0	14	14
Pardubický	11	21	32
Olomoucký	1	51	52
Praha	0	18	18
Jihomoravský	0	51	51
Moravskoslezský	27	19	46
Vysočina	0	14	14
Liberecký	0	5	5
Zlínský	0	9	9
Karlovarský	0	1	1
Jihočeský	0	10	10
Středočeský	0	14	14
Plzeňský	0	2	2

Tabulka 11. Struktura shluků odvozených pomocí EM algoritmu.





Graf 10. Interpretace dvou shluků odvozených pomocí EM algoritmu (proměnné Pohlaví, Věk, Vzdělání, Obor vzdělání a Profese).

Z jednotlivých grafů a tabulky 12 je jasně patrné, že algoritmus dělil datový prostor především podle regionu respondenta, jeho věku a úrovně dosaženého vzdělání.

Test nezávislosti pro kategoriální proměnné Počet shluků: 2 Počet případů: 274			
	df	Chi-square	p value
Sex	1	14.1406	0.000170
Age	3	184.0387	0.000000
Education	4	135.9811	0.000000
Field of Education	6	7.3515	0.289563
Profession	10	198.8300	0.000000
Region	13	115.4787	0.000000

Tabulka 12. Tabulka Chí-kvadrát test mezi odvozenými shluky v jednotlivých sledovaných znacích pomocí EM algoritmu.

Z uvedených výsledků lze odvodit, že oba algoritmy jsou schopny dospět k překvapivě konzistentním závěrům týkajících se typického účastníka vzdělávacích seminářů v zemědělství, jak je uvedeno v předchozích výsledcích.

4.3 Preference konzumentů potravin živočišného původu

4.3.1 Popis souboru respondentů dotazníkového šetření k preferencím u konzumace hovězího masa

Dotazníkového průzkumu se zúčastnilo 132 respondentů, z toho 87 mužů, 35 žen, 10 respondentů své pohlaví neuvedlo v dotazníku. 95 respondentů uvedlo svůj věk jako produktivní, 28 respondentů bylo studenty a 7 respondentů v důchodovém věku. V kategorii vzdělání uvedlo základní vzdělání 14 respondentů, středoškolské 41 a vysokoškolské 74 respondentů.

Tabulky 13 a 14 shrnují výsledky průzkumu k postoji a četnosti nákupu masa.

	Count	Cumulative - Count	Percent	Cumulative - Percent
1x za týden	77	77	58.33333	58.3333
4x-2x	21	98	15.90909	74.2424
vlastní zdroj	11	109	8.33333	82.5758
5x a více za týden	1	110	0.75758	83.3333
Missing	22	132	16.66667	100.0000
Missing	7	132	5.30303	100.0000

Tabulka 13. Tabulka četností nákupů výsekového masa u respondentů.

	Count	Cumulative - Count	Percent	Cumulative - Percent
1x za měsíc	43	43	32.57576	32.5758
méně než 1x za měsíc	29	72	21.96970	54.5455
1x za týden	20	92	15.15152	69.6970
vlastní zdroj	13	105	9.84848	79.5455
vůbec	20	125	15.15152	94.6970
Missing	7	132	5.30303	100.0000

Tabulka 14. Tabulka četností nákupů výsekového hovězího masa u respondentů.

U dotazu na frekvenci nákupu výsekového masa uvedlo 58,33 % respondentů jeho nákup alespoň 1x za týden. Více než jednou týdně nakupuje výsekové maso 16,66 % respondentů. 8,33 % respondentů překvapivě uvedlo, že využívá výsekové maso z vlastních zdrojů.

Při dotazu na nákup výsekového hovězího masa uvedlo 32,57 % respondentů jeho nákup alespoň jedenkrát do měsíce a 15,15 % respondentů alespoň 1x za týden. Vzhledem k poměrně nízké spotřebě hovězího masa v ČR [31] se jedná o poměrně vysoké zastoupení četnosti nákupů, dané pravděpodobně místem dotazníkového šetření a vzorkem respondentů (odborná veřejnost na zemědělské výstavě), což koresponduje také s 9,85 % respondentů, kteří využívají hovězí maso z vlastních zdrojů. V protikladu k tomuto zjištění je však 15,15 % respondentů, kteří hovězí výsekové maso nekupují vůbec.

Na otázku jaké maso konzumujete nejčastěji uvedla většina respondentů více než jednu odpověď. Jednotlivé odpovědi lze tedy interpretovat samostatně jako preferenci jednotlivých druhů mas. 16,66 % respondentů zvolilo z možných odpovědí rybí maso, pouze 3,78 % zvolilo jehněčí maso, 31,06 % hovězí maso, vepřové maso zvolilo 64,39 % a drůbeží 43,93 % respondentů.

V dotazu na oblíbenost masa opět řada respondentů určila více než jednu odpověď, a proto je interpretujeme jako samostatné otázky. Rybí maso určilo jako oblíbené 22,27 % respondentů, jehněčí 10,61, hovězí 41,67, vepřové 41,67 a drůbeží 26,51 % respondentů. Výsledky týkající se této preferenční otázky ve vztahu k hovězímu masu shrnuje tabulka 15.

rybí	jehněčí	vepřové	drůbeží	hovězí - ano	hovězí - ne	Počet odpovědí
ne	ne	ne	ne	30	3	33
ne	ne	ne	ano	1	14	15
Total				31	17	48
ne	ne	ano	ne	6	26	32
ne	ne	ano	ano	5	6	11
Total				11	32	43
ne	ano	ne	ne	2	7	9
ne	ano	ne	ano	0	0	0
Total				2	7	9
ne	ano	ano	ne	0	1	1
ne	ano	ano	ano	0	0	0
Total				0	1	1
ano	ne	ne	ne	3	10	13
ano	ne	ne	ano	1	2	3
Total				4	12	16
ano	ne	ano	ne	3	3	6
ano	ne	ano	ano	2	2	4
Total				5	5	10

ano	ano	ne	ne	1	1	2
ano	ano	ne	ano	0	1	1
Total				1	2	3
ano	ano	ano	ne	0	0	0
ano	ano	ano	ano	1	0	1
Total				1	0	1
Column Total				55	76	131

Tabulka 15. Souhrnné preferenční odpovědi oblíbenosti jednotlivých druhů mas ve vztahu k hovězímu masu.

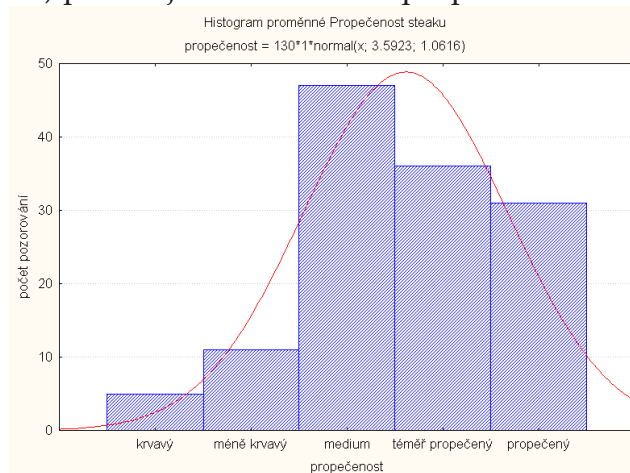
Z tabulky 15 jsou patrné některé zajímavé skutečnosti. 26 respondentů uvedlo, že preferují vepřové maso, nepreferují žádný jiný druh mas. 14 respondentů, kteří preferují drůbeží maso taktéž nepreferují žádný jiný druh. 30 respondentů jednoznačně preferuje hovězí maso na úkor jiných druhů mas.

Při vyhodnocení části dotazníku týkající se postoje ke konzumaci hovězího masa uvedlo 95,45 % respondentů, že si někdy pochutnali na hovězím masu. Z tohoto počtu uvedlo 41,67 %, že se tak stalo před méně než týdnem, 28,03 % před méně než měsícem a 22,73 % méně než před rokem.

Při dotazu na preferenci tuzemské produkce hovězího masa odpovědělo 73,48 % respondentů, že tuzemskou produkci preferuje, 9,85 % uvedlo, že se o původ masa nezajímá, 6,82 % respondentů uvedlo, že není schopno získat informace o původu masa a 6,06 % nepreferuje tuzemskou produkci.

Kuchyňskou úpravu hovězího masa vařením má v oblibě pouze 15,15 %, pečením 35,61 %, dušením 27,27 %, grilováním 43,18 % respondentů. Z tohoto hlediska lze výsledky interpretovat, jako posun českého konzumenta hovězího masa od tradičních způsobů úpravy (vaření, dušení) k modernímu grilování, ale také pečení (roastbeef, rolády atd.).

Na grafu 11 a v tabulce 16 jsou uvedeny výsledky preference respondentů, která se týká způsobu úpravy hovězího steaku. Z grafu i tabulky je patrné, že většina respondentů (86,36 %) preferuje středně a více propečené steaky.



Graf 11. Preference úpravy hovězího steaku.

	Count	Cumulative - Count	Percent	Cumulative - Percent
krvavý	5	5	3.78788	3.7879
méně krvavý	11	16	8.33333	12.1212
medium	47	63	35.60606	47.7273
téměř propečený	36	99	27.27273	75.0000
propečený	31	130	23.48485	98.4848
Missing	2	132	1.51515	100.0000

Tabulka 16. Preference úpravy steaku.

Při dotazu, zda respondent ochutnal grilovaný hovězí steak odpovědělo 78,03 % kladně, 17,42 % záporně, 6 respondentů odpověď nevedlo, z tohoto počtu 71,21 % respondentů uvedlo, že grilovaný steak jim chutnal, 2,27 % uvedlo zápornou odpověď, 35 respondentů odpověď nevedlo. Doma by si grilovaný hovězí steak raději připravilo 45,45 % respondentů, v restauraci by si jej nechalo připravit 21,97 %, odpověď nevedlo 43 respondentů. Ze 132 dotázaných bylo kladně osloveno grilováním hovězího masa 93,18 %, 3,03 % na otázku odpověděla záporně, 5 respondentů odpověď nevedlo.

4.3.2 Hypotézy o postojích respondentů ke konzumaci, preferované úpravě hovězího masa a oblíbenosti různých druhů mas

Pro zjištění cílových skupin konzumentů hovězího masa jsme formulovali následující tři hypotézy (a odpovídající nulové):

1. souvisí pohlaví respondenta s preferencí hovězího masa jako nejčastěji konzumovaného masa,
2. souvisí pohlaví respondenta s hovězím masem jako nejoblíbenějším masem
3. existuje závislost mezi věkovou kategorií a preferovanou kuchyňskou úpravou hovězího masa,
4. existuje závislost mezi pohlavím a preferencí úpravy grilovaného hovězího steaku.

Tabulka 17 ukazuje výsledky testování první hypotézy. Z 85 respondentů, kteří nevedli hovězí maso jako nejčastěji konzumované bylo 60 mužů a 25 žen. Z 37 respondentů, kteří uvedli hovězí maso jako nejčastěji konzumované bylo 27 mužů a 10 žen. Pro tyto hodnoty a očekávané frekvence byla vypočtena hodnota Pearsonova Chi-square koeficientu 0,072, což při jednom stupni volnosti odpovídá hodnotě $p=0,79$ a nulová hypotéza tedy není vyvrácena a mezi znaky nebyla prokázána závislost.

	hovězí - ne	hovězí - ano	Row - Totals
M	60	27	87
Ž	25	10	35
Totals	85	37	122

Tabulka 17. Kontingenční tabulka pozorovaných hodnot pro znaky pohlaví a preferenci hovězího masa jako nejčastěji konzumovaného.

Tabulka 18 zobrazuje výsledky testování druhé hypotézy. Ze 70 respondentů, kteří neuvedli hovězí maso jako nejoblíbenější bylo 46 mužů a 24 žen. Z 52 respondentů, kteří uvedli hovězí maso jako nejoblíbenější bylo 41 mužů a 11 žen. Pro tyto hodnoty a očekávané frekvence byla vypočtena hodnota Pearsonova Chi-square koeficientu 2,515, což při jednom stupni volnosti odpovídá hodnotě $p=0,113$, odpovídající nulová hypotéza tedy není vyvrácena a mezi znaky nebyla prokázána závislost.

	hovězí - ano	hovězí - ne	Row - Totals
M	41	46	87
Ž	11	24	35
Totals	52	70	122

Tabulka 18. Kontingenční tabulka pozorovaných hodnot pro znaky pohlaví a preferenci hovězího masa jako nejoblíbenějšího.

Tabulka 19 shrnuje výsledky testování třetí hypotézy o souvislosti věkové skupiny a preferovanému způsobu kuchyňské úpravy hovězího masa.

	vaření - ne	vaření - ano	Row - Totals
PV	80	14	94
S	24	4	28
D	7	0	7
Totals	111	18	129

Pearson Chi-square 1.206906, $df=2$, $p=.54692$

	dušení - ne	dušení - ano	Row - Totals
PV	69	25	94
S	23	5	28
D	2	5	7
Totals	94	35	129

Pearson Chi-square 8.179369, $df=2$, $p=.01675$

	pečení - ne	pečení - ano	Row - Totals
PV	61	33	94
S	16	12	28
D	6	1	7
Totals	83	46	129

Pearson Chi-square 2.038587, $df=2$, $p=.36085$

	grilování - ano	grilování - ne	Row - Totals
PV	42	52	94
S	12	16	28
D	2	5	7
Totals	56	73	129

Pearson Chi-square .6926924 $df=2$ $p=.70727$

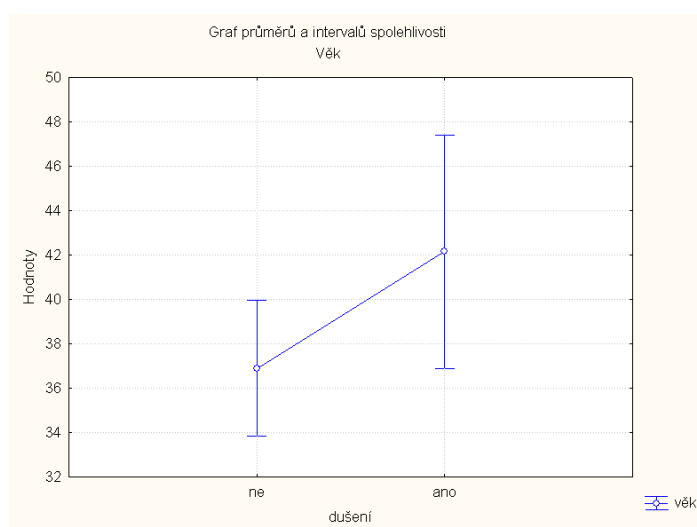
Tabulka 19. Výsledky testování hypotéz o vztahu mezi věkovou skupinou a způsobem kuchyňské úpravy hovězího masa.

Z výsledků je patrné, že nulová hypotéza je vyvrácena ($p=0,01675$) pouze v případě vztahu věkové skupiny k preferenci dušení jako způsobu úpravy hovězího masa.

K ověření tohoto výsledku byla následně použita analýza ANOVA s efektem odpovědi na preferenci způsobu úpravy pomocí dušení a závislou spojitou proměnou věku udaného respondentem. Tabulka 20 a graf 12 zobrazuje výsledky analýzy.

	SS - Effect	df - Effect	MS - Effect	SS - Error	df - Error	MS - Error	F	p
věk	666.189	1	666.189	23558.29	115	204.854	3.252	0.0739

Tabulka 20. Výsledky analýzy variance pro efekt věku a preferenci/zamítání dušení jako preferované úpravy hovězího masa.



Graf 12. Variance věku ve skupinách preferujících, zamítajících dušení jako preferovanou úpravu hovězího masa.

Z výsledků je patrné, že hodnota hypotézy o shodných rozptylech v rámci obou skupin nebyla vyvrácena ($p=0,074$), lze ji však nazvat statistickým trendem a interpretovat pomocí tvrzení, že starší lidé (průměrný věk 42,15 let) spíše nezavrhují/ preferují dušení hovězího masa v porovnání s mladšími lidmi (průměrný věk 36,89 let).

Tabulka 21 zobrazuje výsledky testování čtvrté hypotézy. Pro tyto hodnoty a očekávané frekvence byla vypočtena hodnota Pearsonova Chi-square koeficientu 5,932, což při čtyřech stupních volnosti odpovídá hodnotě $p=0,204$, odpovídající nulové hypotéze tedy není vyvrácena a mezi znaky nebyla prokázána závislost.

Propečnost	krvavý	méně krvavý	medium	téměř propečený	propečený	Row - Totals
M	4	6	35	25	15	85
Ž	1	2	9	10	13	35
Totals	5	8	44	35	28	120

Tabulka 21. Kontingenční tabulka pozorovaných hodnot pro znaky pohlaví a preferenci úpravy steaku.

4.3.3 Asociační pravidla preferencí při konzumaci hovězího masa

Výstupem algoritmu Apriori je následujících 30 nejlepších nalezených asociačních pravidel.

Apriori

Best rules found:

1. X10grilovani=ano doma.restaurace=doma 29 ==> steak=měl 29 acc:(0.99421)
2. pohlaví=M X10dušení=ne propečenost=medium 28 ==> steak=měl 28 acc:(0.99416)
3. X10pečení=ano X10grilování=ne steak=měl 27 ==> X10dušení=ne 27 acc:(0.9941)
4. pohlaví=M X10pečení=ano X10grilování=ne 25 ==> X10dušení=ne 25 acc:(0.99396)
5. pohlaví=M X10vaření=ne X10pečení=ano steak=měl 25 ==> X10dušení=ne 25 acc:(0.99396)
6. X10dušení=ne X10grilování=ano propečenost=medium 24 ==> steak=měl 24 acc:(0.99387)
7. pohlaví=M vzdělání=VŠ X10pečení=ne X10dušení=ne 24 ==> steak=měl 24 acc:(0.99387)
8. vzdělání=VŠ kdy naposled=týden X10pečení=ne 22 ==> steak=měl 22 acc:(0.99367)
9. X10pečení=ne X10grilování=ano propečenost=medium 22 ==> steak=měl 22 acc:(0.99367)
10. pohlaví=M vzdělání=VŠ X10pečení=ne X10grilování=ano 21 ==> steak=měl 21 acc:(0.99355)
11. pohlaví=M nákup výsekového masa=1x za týden kdy naposled=měsíc 19 ==> X10vareni=ne 19 acc:(0.99324)
12. vzdělání=VŠ X10pečení=ano steak=měl 19 ==> X10dušení=ne 19 acc:(0.99324)
13. nákup výsekového masa=1x za týden X10pečení=ano X10grilování=ne 19 ==> X10dušení=ne 19 acc:(0.99324)
14. kdy naposled=týden X10pečení=ne doma.restaurace=doma 17 ==> steak=měl 17 acc:(0.99283)
15. X10vaření=ne X10dušení=ne doma.restaurace=restaurace 17 ==> nákup výsekového masa=1x za týden 17 acc:(0.99283)
16. pohlaví=M propečenost=medium doma.restaurace=doma 16 ==> steak=měl 16 acc:(0.99256)
17. nákup výsekového hovězího=1x za měsíc X10vaření=ne X10pečení=ano 16 ==> X10dušení=ne 16 acc:(0.99256)
18. pohlaví=M vzdělání=VŠ X10pečení=ne propečenost=medium 16 ==> steak=měl 16 acc:(0.99256)
19. pohlaví=M X10vaření=ne X10pečení=ano doma.restaurace=doma 16 ==> X10dušení=ne 16 acc:(0.99256)
20. nákup výsekového masa=1x za týden steak=neměl 15 ==> X10vaření=ne 15 acc:(0.99225)
21. nákup výsekového hovězího=1x za týden X10dušení=ne 15 ==> steak=měl 15 acc:(0.99225)

22. pohlaví=M X10pečení=ano propečenost=téměř propečený 15 ==> X10dušení=ne 15 acc:(0.99225)
23. X10dušení=ne propečenost=téměř propečený doma.restaurace=doma 15 ==> steak=měl 15 acc:(0.99225)
24. nákup výsekového masa=4x-2x X10dušení=ne 14 ==> steak=měl 14 acc:(0.99186)
25. pohlaví=M X10vaření=ne X10dušení=ano X10grilování=ne 14 ==> X10pečení=ne 14 acc:(0.99186)
26. nákup výsekového hovězího=1x za měsíc doma.restaurace=restaurace 13 ==> nákup výsekového masa=1x za týden 13 acc:(0.99138)
27. nákup výsekového hovězího=méně než 1x za měsíc steak=neměl 13 ==> X10vaření=ne 13 acc:(0.99138)
28. X10dušení=ano steak=neměl 13 ==> X10vaření=ne 13 acc:(0.99138)
29. věk=(23.5-28.5]' 12 ==> vzdělání=VŠ 12 acc:(0.99078)
30. věk=(57-inf)' 12 ==> X10vaření=ne 12 acc:(0.99078)

Z výsledků asociačních pravidel je zřejmé, že algoritmus Apriori upřednostňuje nejčetnější odpovědi. Konverguje tak především k pravidlům obsahujícím odpovědi týkající se preferované úpravy hovězího masa a pozitivní odpovědi na otázku, zda respondent měl v minulosti hovězí steak. Ze zajímavých pravidel lze zmínit například ta, jejichž interpretační význam má smysl pro majitele restaurací při plánování skladby jídelních lístků - pokud je respondent vysokoškolsky vzdělaný a preferuje úpravu masa pečením, pak nevyžaduje dušenou úpravu masa, odpovědělo 19 respondentů. Jestliže muži upřednostňují pečení hovězího masa s propečeností téměř propečený, pak nemají rádi dušené maso, odpovědělo 15 mužů. Naopak jestliže 14 mužů, kteří mají v oblibě dušenou úpravu masa nepreferují vaření a grilování, pak nepreferují ani úpravu masa pečením.

Pro prodejce výsekového masa je zajímavým pravidlem např. - jestliže respondent nakupuje výsekové hovězí maso méně než 1x za měsíc a neměl steak, odpovědělo 13 dotázaných, pak nemá rád ani vařené maso. Není tedy vhodné této skupině zákazníkům nabízet steakové maso zároveň s masem vhodným k vaření.

Z marketingově hodnotných výsledků lze vybrat například pravidlo popisující skupinu, která se aktivně (průzkum byl spojen s ochutnávkou) zajímá o úpravu hovězího masa grilováním - jestliže jsou dotazované osoby ve věku 23-28 roků, pak jsou s vysokoškolským vzděláním, odpovědělo 12 respondentů. Na tuto skupinu by měla být cílena reklamní kampaň.

Z pravidla 30 vyplývá, že lidé ve věku 57 a více nepreferují úpravu masa vařením; odpovědělo 12 respondentů. Interpretaci tohoto pravidla lze zdůvodnit tím, že starší lidé mají problém s konzumací vařeného hovězího masa, které díky kolísavé kvalitě této suroviny nelze vždy ideálně vařením upravit. Nabízí se zde prostor pro oslovení této skupiny respondentů např. s jiným, resp. jistě kvalitním hovězím masem vhodným pro úpravu vařením, neboť z tabulky 19 je patrné, že lidé v produktivním věku jsou nejsilnější skupinou respondentů, kteří preferují vařená masa, v důsledku přibývajících věku se však tato skupina potenciálních zákazníků z výše uvedených důvodů ztrácí.

Výstupem algoritmu Tertius je soubor následujících nejlepších nalezených asociačních pravidel:

Tertius

1. /* 0,651707 0,022727 */ X10vaření = ne and X10pečení = ne ==> X10dušení = ano or X10grilování = ano
2. /* 0,547635 0,068182 */ X10pečení = ne and X10dušení = ne ==> X10grilování = ano or propečenost = méně krvavý
3. /* 0,531737 0,068182 */ X10pečení = ne and X10dušení = ne and steak = měl ==> X10grilování = ano
4. /* 0,507755 0,098485 */ X10pečení = ne and X10dušení = ne ==> X10grilování = ano
5. /* 0,506506 0,090909 */ X10pečení = ne ==> X10dušení = ano or X10grilování = ano or propečenost = krvavý
6. /* 0,486672 0,083333 */ X10grilování = ne ==> nákup výsekového masa = 5x a více za týden or X10pečení = ano or X10dušení = ano
7. /* 0,484946 0,075758 */ X10grilování = ne ==> X10pečení = ano or X10dušení = ano or propečenost = propečený
8. /* 0,475170 0,098485 */ X10vaření = ne and X10grilování = ne ==> X10pečení = ano or steak = neměl
9. /* 0,473784 0,053030 */ X10pečení = ne and X10grilování = ne ==> nákup výsekového hovězího = méně než 1x za měsíc or X10dušení = ano
10. /* 0,472406 0,090909 */ X10grilování = ne ==> vzdělání = Z or X10pečení = ano or X10dušení = ano

Number of hypotheses considered: 267326

Number of hypotheses explored: 99537

Na základě vyhodnocení algoritmu Tertius jsou patrné následující skutečnosti: Pokud dotazovaní nepreferují vaření a pečení hovězího masa, pak dávají přednost dušení nebo grilování. Toto pravidlo je prakticky využitelné například pro restaurace, které by měly nabízet na jídelním lístku jídla obou úprav, ale také pro prodejce výsekového masa. Zákazník bude v prodejně volit maso vhodné pro jeden nebo druhý druh úpravy, prodejce by z hlediska objemu prodeje měl nabízet oba dva.

Z hlediska restaurací rozvíjí předchozí komentář následující pravidla. Pokud konzumenti nepreferují pečení a dušení masa, pak si rádi pochutnají na grilovaném masu s propečeností méně krvavý. Jestliže dotázaní nevyžadují pečení, pak dávají přednost úpravě masa dušením nebo grilováním, propečenost krvavý.

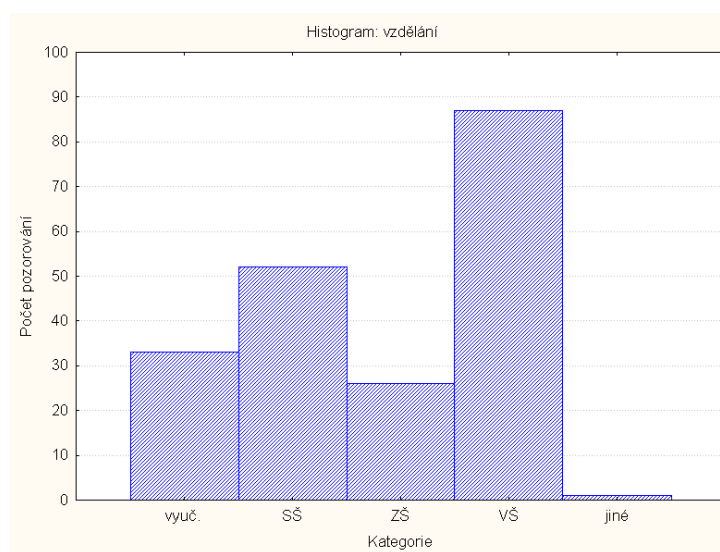
Z pohledu prodejce výsekového masa, případně jeho producenta jsou zajímavá např. tato pravidla: jestliže respondenti negrilují, pak kupují výsekové maso 5x a více za týden, upřednostňují úpravu masa pečením nebo dušením. Když dotázaní neradi grilují, pak mají dokončené pouze základní vzdělání a preferují pečenou nebo dušenou úpravu masa. Pokud respondenti nepreferují pečení masa a zároveň jeho grilování, nakupují výsekové maso méně než jednou za měsíc a upřednostňují jeho dušení. Toto pravidlo může v praxi znamenat doporučení pro prodejce, aby výsekové maso vhodné k dušení udržoval pouze mražené.

4.3.4 Popis souboru respondentů dotazníkového šetření k preferencím u konzumace kravského mléka

Dotazníkového průzkumu týkajícího se preferencí u konzumentů kravského mléka a mléčných výrobků se zúčastnilo 200 respondentů, 88 žen, 111 mužů, 1 respondent pohlaví v dotazníku neuvedl. Tabulka 22 ukazuje věkové rozložení respondentů, které je více vyrovnané v rámci jednotlivých skupin, než u dotazníkového průzkumu popsaneho v předchozí části práce.

	Count	Cumulative - Count	Percent	Cumulative - Percent
nad 50	69	69	34.50000	34.5000
21-30	40	109	20.00000	54.5000
do 20	31	140	15.50000	70.0000
31-50	60	200	30.00000	100.0000
Missing	0	200	0.00000	100.0000

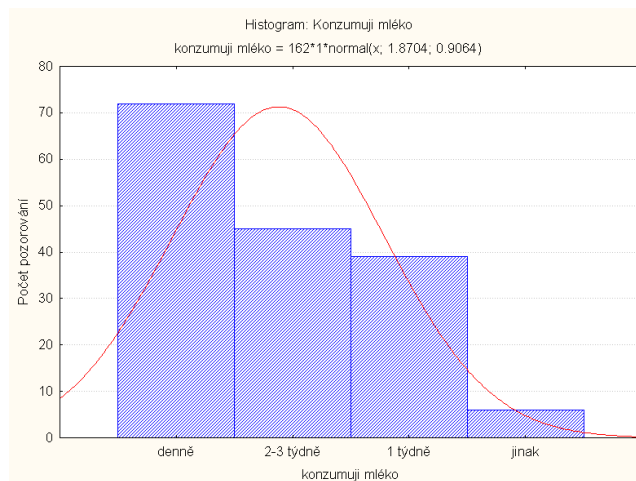
Tabulka 22. Věková struktura respondentů dotazníkového průzkumu o konzumaci mléka a mléčných výrobků.



Graf 13. Struktura vzdělání respondentů dotazníkového průzkumu preferencí při konzumaci mléka a mléčných výrobků.

Graf 13 zobrazuje strukturu vzdělání respondentů dotazníku. Procentně vyjádřené počty ukazují 13 % respondentů se základním vzděláním, 16,5 % vyučených, 26 % středoškolsky vzdělaných, 43,5 % vysokoškolsky vzdělaných respondentů. 1 z respondentů uvedl jiné vzdělání, 1 své vzdělání neuvedl.

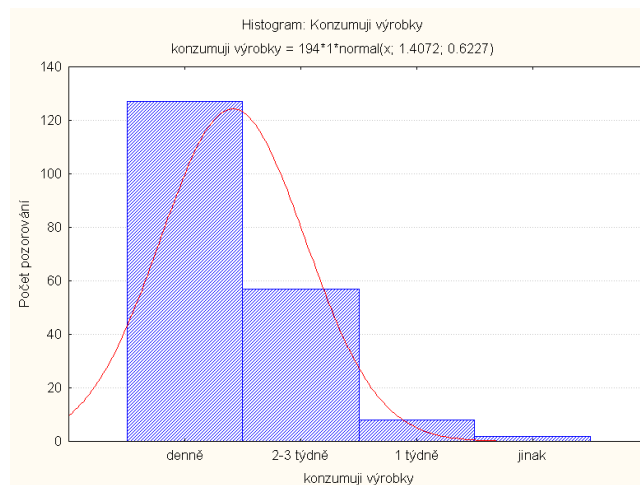
Při odpovědích na stravovací návyky v souvislosti s konzumací mléka a mléčných výrobků uvedlo 28,5 % respondentů, že preferuje jak konzumaci mléka, tak mléčných výrobků, 49 % respondentů preferuje pouze mléčné výrobky a 22,5 % preferuje konzumaci mléka. Jako oblíbený výrobek konzumace uvedlo jogurty 76,5 %, sýry 83 %, tvaroh 44 %, máslo 55,5 % a jiné mléčné výrobky 4 % respondentů.



Graf 14. Frekvence konzumace mléka u respondentů.

Graf 14 zobrazuje jednotlivé podíly frekvencí konzumace mléka u respondentů dotazníku. 36 % dotázaných konzumuje mléko denně, 22,5 % 2-3x týdně, 19,5 % 1x týdně 3 % uvedlo odpověď jinak, 38 respondentů odpověď nevedlo.

Mléčné výrobky konzumuje denně 63,5 % respondentů, 28,5 % 2-3x týdně, 4 % jednou týdně, 1 % uvedlo odpověď jinak a 6 respondentů na otázku neodpovědělo. Z uvedených údajů vyplývá, že mezi respondenty je oblíbenější konzumace mléčných výrobků, než samotného mléka.



Graf 15. Frekvence konzumace mléčných výrobků u respondentů.

Při vyhodnocení části dotazníku týkající se názorů na konzumaci a zdravotní prospěšnost mléka jsme dospěli k výsledkům uvedeným v tabulce 23.

je mléko prospěšné zdraví?	Count	Percent	zdroj živin?	Count	Percent
ano	184	92	ano	184	92
nevím	11	5.5	nevím	13	6.5
ne	4	2	ne	3	1.5
Missing	1	0.5	Missing	0	0

pozitivní vliv?	Count	Percent	zdroj vápníku?	Count	Percent
ano	178	89	ano	150	75
nevím	18	9	nevím	31	15.5
ne	4	2	ne	17	8.5
Missing	0	0	Missing	2	1
ochrana před osteoporózou?	Count	Percent	dobře stravitelné?	Count	Percent
ano	128	64	nevím	36	18
nevím	60	30	ano	146	73
ne	8	4	ne	12	6
Missing	4	2	Missing	6	3
cholesterol a tuky?	Count	Percent	zakysané zdravější?	Count	Percent
nevím	71	35.5	nevím	60	30
ano	27	13.5	ano	113	56.5
ne	99	49.5	ne	25	12.5
Missing	3	1.5	Missing	2	1
podpora v rodině?	Count	Percent	děti?	Count	Percent
ano	171	85.5	ano	141	70.5
ne	24	12	ne	44	22
nevím	3	1.5	nevím	13	6.5
Missing	2	1	Missing	2	1
dostatečná propagace?	Count	Percent	propagace ve školách?	Count	Percent
ano	38	19	ano	184	92
nevím	44	22	nevím	8	4
ne	118	59	ne	8	4
Missing	0	0	Missing	0	0
automaty?	Count	Percent	je vhodné jiné mléko?	Count	Percent
ano	153	76.5	ano	117	58.5
nevím	28	14	nevím	51	25.5
ne	19	9.5	ne	32	16
Missing	0	0	Missing	0	0

negativní vliv na zdraví?	Count	Percent	-	-	-
ne	79	39.5	-	-	-
jiné	45	22.5	-	-	-
ano	8	4	-	-	-
nevím	6	3	-	-	-
Missing	62	31	-	-	-

Tabulka 23. Výsledky shrnující názory na konzumaci a zdravotní nezávadnost mléka.

Z respondentů dotazníku považuje mléko za zdraví prospěšné 92 %, 2 % odpověděla, že mléko není zdraví prospěšné, 5,5 % uvedla, že neví, zda mléko je či není prospěšné zdraví. Za zdroj živin považuje mléko 92 % respondentů, jisto si není 6,5 % a 1,5 % nepovažuje mléko za zdroj živin. Pozitivní vliv mléka na zdraví kladně ohodnotilo 89 % respondentů, 9 % uvedlo odpověď nevím a 2 % nepovažuje mléko za pozitivně působící na zdraví člověka. Překvapivé odpovědi v souvislosti s předchozími výsledky jsou uvedeny při poměrně všeobecně známé otázce, zda je mléko považováno za zdroj vápníku. V této otázce uvedlo 75 % dotázaných odpověď ano, 8,5 % odpověď ne a celých 15,5 % odpověď nevím. Ještě více překvapivé výsledky jsou v souvislosti s neinformovaností veřejnosti o benefitech konzumace mléka uvedeny u odpovědi na otázku, zda je konzumace mléka ochranou před onemocněním osteoporózou. Zde uvedlo pouze 64 % dotázaných odpověď ano, celých 30 % odpověď nevím a 4 % respondentů tvrdila, že mléko není zdrojem ochrany před tímto onemocněním.

Při dotazu na poměrně diskutovatelné téma stravitelnosti mléka naopak 73 % respondentů uvádí, že mléko je snadno stravitelné, 18 % neví a 6 % uvádí, že mléko není lehce stravitelné.

Vysokou neinformovanost prokázali respondenti také v otázce, zda kravské mléko je zdrojem nezdravých tuků a cholesterolu. 35,5 % zvolilo odpověď nevím, 49,5 % odpověď ne (což je při rozsáhlé kritice živočišných produktů v médiích překvapivé) a 13,5 % se domnívalo, že mléko je zdrojem nezdravých tuků a cholesterolu.

Zakysané mléčné výrobky považuje za zdravější 56,5 % dotázaných, jisto si není 30 % a za zdravější nepovažuje tyto výrobky 12,5 % respondentů. Konzumace mléka v rodině je podporována v 85,5 % případů, není podporována ve 12,5 % a na otázku neodpovědělo, nebo si nebylo jisto 2,5 % respondentů. Své děti vede ke konzumaci kravského mléka 70,5 % dotázaných, 22 % děti v konzumaci nepodporuje a 7,5 % se k otázce nebylo schopno vyjádřit. Poněkud v protikladu jsou v této souvislosti výsledky odpovědí na otázku, zda by respondent podpořil propagaci mléka na školách. Plných 92 % respondentů odpovědělo kladně, 4 % záporně a 4 % zvolilo odpověď nevím. Jako dostatečnou hodnotí propagaci mléka 19 % respondentů, 59 % hodnotí propagaci jako nedostatečnou a 22 zvolilo odpověď nevím. Mléčná automaty podporuje 76,5 % dotázaných, nepodporuje 9,5 % a proti je 9,5 % respondentů.

Překvapivé výsledky přineslo také vyhodnocení otázky, zda je vhodné konzumovat jiné mléko, než kravské (sójové, kozí atd.). Celkem 58,5 % odpovídajících je pře-

svědčeno, že je vhodné konzumovat jiné druhy mléka, 16 % je proti jejich konzumaci a celých 25,5 % zvolilo odpověď nevím.

Možné negativní vlivy na zdraví člověka při konzumaci kravského mléka připouští 4 % dotázaných, žádné negativní vlivy si nepřipouští 39,5 %, neví 3 %, a jiné negativní vlivy než zdravotní spatřuje v konzumaci mléka 22,5 % dotázaných. 31 % na otázku neodpovědělo.

Celkově výsledky vypovídají o nepříliš vysoké úrovni informovanosti populace v otázkách zdravotních benefitů konzumace kravského mléka a celkové problematiky; ukazují spíše na povrchní informace spotřebitelů.

4.3.5 Hypotézy o postojích respondentů ke konzumaci a zdravotní prospěšnosti mléka a mléčných výrobků

Na základě výsledků diskutovaných v předchozí části jsme formulovali následující hypotézy a provedli jejich otestování:

1. Souvisí věk a pohlaví s konzumací mléka a mléčných výrobků?
2. Ovlivňuje věk a vzdělání pohled na zdravotní prospěšnost mléka?
 - a. Souvisí pohlaví, věk a vzdělání s pohledem na zdravotní prospěšnost mléka?
 - b. Souvisí věk a vzdělání s názorem, že mléko je zdrojem vápníku?
 - c. Souvisí věk a vzdělání s názorem, že mléko je ochranou před osteoporézou?
 - d. Existuje souvislost mezi názorem na zdravotní prospěšnost mléka a názorem, že mléko ochranou před osteoporézou?
3. Považují konzumenti podle věku a vzdělání zakysané mléčné výrobky za zdravější?

Tabulky 24-25 shrnují výsledky pozorování pro testování první hypotézy. Ani v jednom případě se závislost (věk, pohlaví) na preferenci konzumace mléka či mléčných výrobků neprokázala. Pro věk je udána hodnota Pearsonova koeficientu 8,263, což při 6 stupních volnosti odpovídá hodnotě $p=0,215$. Pro pohlaví vychází hodnota Pearsonova Chi-kvadrát koeficientu 3,038, což při 3 stupních volnosti odpovídá hodnotě $p=0,219$. Mezi sledovanými znaky nebyla prokázána souvislost, můžeme tedy říci, že respondenti neupřednostňují konzumaci mléka, mléčných výrobků podle věku či pohlaví.

	upřednostňuji - obojí	upřednostňuji - výrobky	upřednostňuji - mléko	Row - Totals
nad 50	19	28	22	69
21-30	13	21	6	40
do 20	6	17	8	31
31-50	19	32	9	60
Totals	57	98	45	200

	upřednostňuji - obojí	upřednostňuji - výrobky	upřednostňuji - mléko	Row - Totals
ž	30	42	16	88
m	27	55	29	111
Totals	57	97	45	199

Tabulka 24. Kontingenční tabulka pozorovaných hodnot pro znaky vzdělání, pohlaví a preferenci mléka či mléčných výrobků.

Tabulka 25 shrnuje výsledky pozorování pro znaky pohlaví v závislosti na názoru o zdravotním benefitu mléka. Nulová hypotéza není ani v tomto případě vyvrácena (Pearson Chi-Square=0,606, df=2, p=0,739).

	je mléko prospěšné zdraví? - ano	je mléko prospěšné zdraví? - nevím	je mléko prospěšné zdraví? - ne	Row - Totals
ž	79	6	2	87
m	104	5	2	111
Totals	183	11	4	198

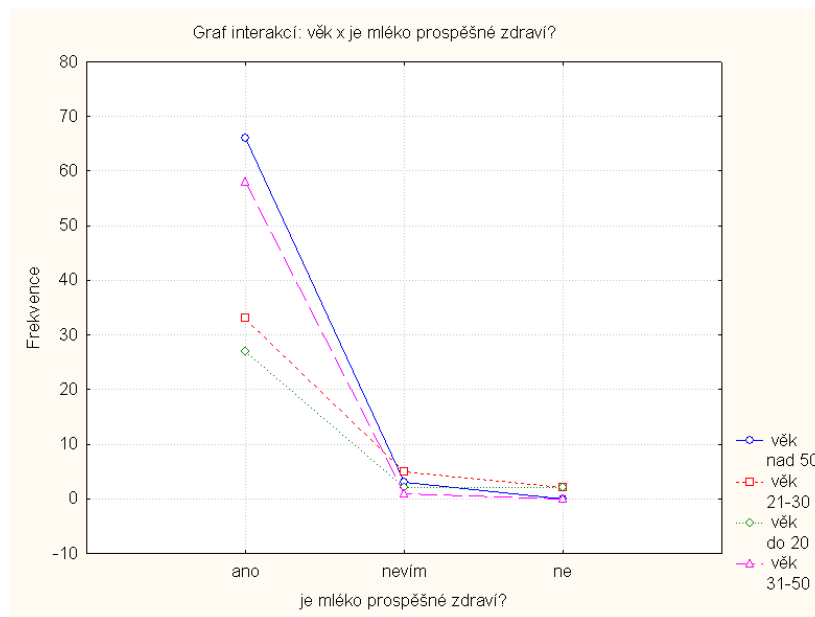
Tabulka 25. Kontingenční tabulka pohlaví a názoru na zdravotní prospěšnost mléka.

Výsledky potřebné pro testování hypotézy o vlivu věku na názor o zdravotní prospěšnosti mléka jsou shrnuty v tabulce 26.

	je mléko prospěšné zdraví? - ano	je mléko prospěšné zdraví? - nevím	je mléko prospěšné zdraví? - ne	Row - Totals
nad 50	66	3	0	69
21-30	33	5	2	40
do 20	27	2	2	31
31-50	58	1	0	59
Totals	184	11	4	199

Tabulka 26. Kontingenční tabulka pro proměnné zdravotní prospěšnost mléka a věk respondenta.

Nulová hypotéza o nezávislosti obou jevů je v tomto případě vyvrácena (Pearson Chi-Square=13,52; df=6; p=0,0354). Z grafu 16 je patrné, že především starší lidé hodnotí zdravotní přínosy mléka kladně (skupiny 31-50 let a nad 50 let).



Graf 16. Hodnocení zdravotní prospěšnosti mléka v závislosti na věkové skupině respondentů.

Hypotéza o vlivu úrovně vzdělání na názor o zdravotní prospěšnosti mléka nebyla prokázána (resp. nulová hypotéza o neexistenci závislosti nebyla vyvrácena) (Pearson Chi-Square=6,537; df=8; p=0,587).

Při podrobnějším rozboru problematiky informovanosti respondentů o zdravotních účincích mléka jsme testovali hypotézy o vlivu věku a vzdělání na názory o mléku jako zdroje vápníku a mléku jako ochraně před osteoporózou. Tabulka 27 a graf 17 shrnují tyto výsledky.

	zdroj vápníku? - ano	zdroj vápníku? - nevím	zdroj vápníku? - ne	Row - Totals
nad 50	57	6	5	68
21-30	31	5	4	40
do 20	17	9	5	31
31-50	45	11	3	59
Totals	150	31	17	198

Pearson Chi-square 11.66417, df=6, p=.06991.

	zdroj vápníku? - ano	zdroj vápníku? - nevím	zdroj vápníku? - ne	Row - Totals
vyuč.	26	7	0	33
SŠ	40	6	5	51
ZŠ	14	8	4	26
VŠ	68	10	8	86
jiné	1	0	0	1
Totals	149	31	17	197

Pearson Chi-square 12.33004; df=8; p=.13709.

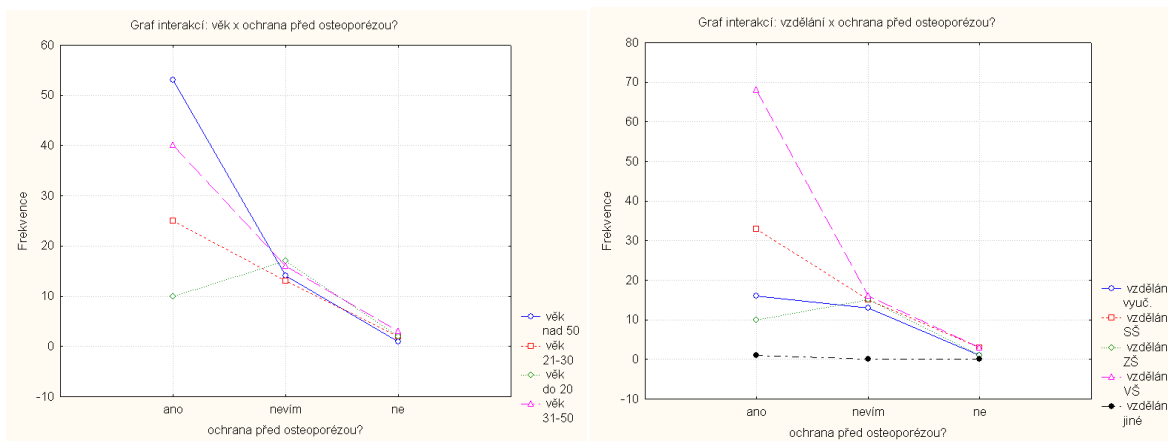
	ochrana před osteoporézou? - ano	ochrana před osteoporézou? - nevím	ochrana před osteoporézou? - ne	Row - Totals
nad 50	53	14	1	68
21-30	25	13	2	40
do 20	10	17	2	29
31-50	40	16	3	59
Totals	128	60	8	196

Pearson Chi-square 17.85779; df=6; p=.00660.

	ochrana před osteoporézou? - ano	ochrana před osteoporézou? - nevím	ochrana před osteoporézou? - ne	Row - Totals
vyuč.	16	13	1	30
SŠ	33	15	3	51
ZŠ	10	15	1	26
VŠ	68	16	3	87
jiné	1	0	0	1
Totals	128	59	8	195

Pearson Chi-square 18.98244; df=8; p=.01496.

Tabulka 27. Kontingenční tabulky pro věk, vzdělání respondentů v závislosti na jejich názoru, zda mléko je ochranou před osteoporézou a zdrojem vápníku.



Graf 17. Závislost mezi věkem a vzděláním respondentů a jejich názorem, zda mléko je ochranou před osteoporézou.

Z výsledků je patrné, že existuje závislost mezi věkem respondentů, úrovní jejich vzdělání a jejich názorem na mléko jako ochranu před osteoporézou. Z grafů 17 vyplývá, že starší (31-50, nad 50 let) a vzdělanější lidé (SŠ, VŠ) jsou lépe informovaní o tomto zdravotním benefitu mléka.

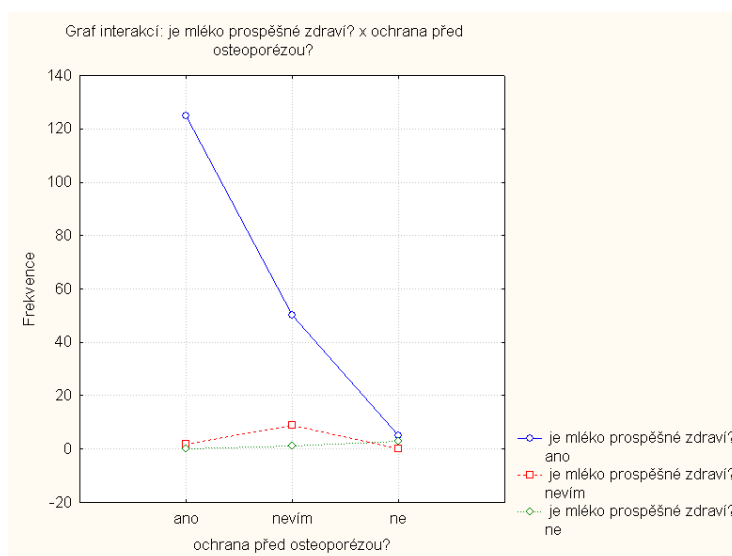
Tyto závěry implikují testování další hypotézy – zda existuje závislost v pohledu na zdravotní prospěšnost mléka a názorem, že mléko je ochranou před osteoporézou. Tato hypotéza je výsledky potvrzena. Lidé, kteří se domnívají, že mléko je

zdraví prospěšné, mají přehled o tom, že je ochranou před osteoporézou. Výsledky shrnuje tabulka 28 a graf 18.

	ochrana před osteoporézou? - ano	ochrana před osteoporézou? - nevím	ochrana před osteoporézou? - ne	Row - Totals
ano	125	50	5	180
nevím	2	9	0	11
ne	0	1	3	4
Totals	127	60	8	195

Pearson Chi-square 66.95483; df=4; p=.00000.

Tabulka 28. Souvislost mezi pohledem na mléko jako zdraví prospěšné a jeho hodnocením jako ochrany před osteoporézou.



Graf 18. Závislost mezi odpověďmi na zdravotní prospěšnost mléka a jeho hodnocením jako ochrany před osteoporézou.

Poslední testovanou hypotézou je vztah mezi věkem a vzděláním respondentů a jejich hodnocením mléčných zakysaných výrobků jako zdravějších. Výsledky testování shrnují kontingenční tabulky pozorovaných hodnot (tabulka 29), grafy 19 a výsledky statistických testů.

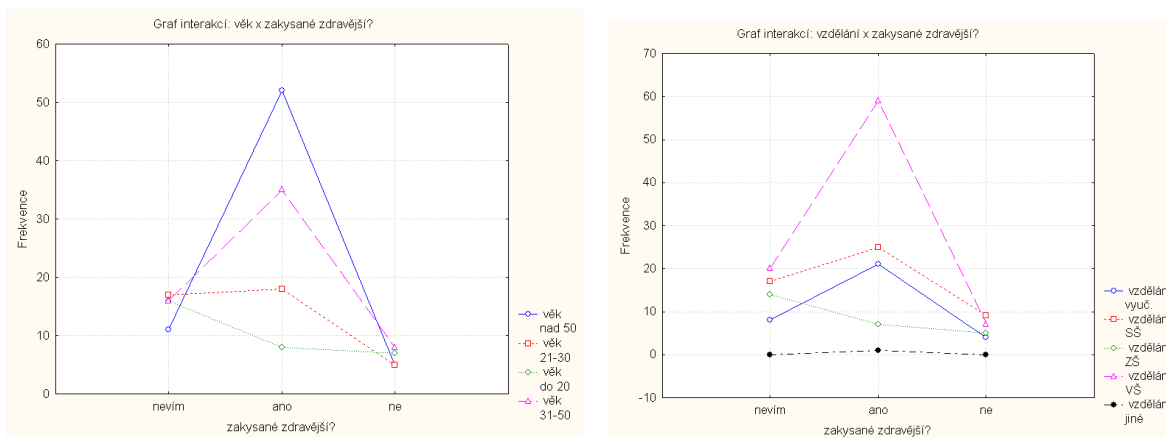
	zakysané zdravější? - nevím	zakysané zdravější? - ano	zakysané zdravější? - ne	Row - Totals
nad 50	11	52	5	68
21-30	17	18	5	40
do 20	16	8	7	31
31-50	16	35	8	59
Totals	60	113	25	198

Pearson Chi-square 26.12376; df=6; p=.00021.

	zakysané zdravější? - nevím	zakysané zdravější? - ano	zakysané zdravější? - ne	Row - Totals
vyuč.	8	21	4	33
SŠ	17	25	9	51
ZŠ	14	7	5	26
VŠ	20	59	7	86
jiné	0	1	0	1
Totals	59	113	25	197

Pearson Chi-square 17.75704; df=8; p=.02313.

Tabulka 29. Tabulky závislosti věku a vzdělání ve vztahu k názoru na zdravotní benefity zakysaných mléčných produktů.



Graf 19. Interakce mezi vzděláním, věkem respondentů a jejich názorem na zdravotní prospěšnost zakysaných mléčných výrobků.

Obě nulové hypotézy o nezávislosti sledovaných jevů byly vyvráceny. Respondenti hodnotí zdravotní prospěšnost zakysaných mléčných výrobků v závislosti na vzdělání i věku. Doporučením pro výrobce by mohlo být cílení reklamy na skupinu respondentů ve věku 21-30 let, která nevykazuje příliš vysoké hodnocení oproti starším věkovým skupinám. Obdobně respondenti se základním vzděláním vykazují poměrně nízkou frekvenci hodnocení zdravotní prospěšnosti zakysaných mléčných výrobků, doporučením je proto výrazně zvýšit informovanost mladších věkových skupin o benefitech zakysaných mléčných výrobků, případně připravit produkt přímo pro tyto cílové skupiny.

4.3.6 Asociační pravidla preferencí při konzumaci kravského mléka

Výsledky algoritmu Apriori – 30 nejlepších nalezených asociačních pravidel shrnuje následující výstup:

Apriori

Best rules found:

1. zdroj živin?=ano pozitivní vliv?=ano děti?=ano 133 ==> je mléko prospěšné zdraví?=ano 133 acc:(0.99162)
2. zdroj živin?=ano děti?=ano propagace ve školách?=ano 131 ==> je mléko prospěšné zdraví?=ano 131 acc:(0.99152)
3. sýry_binarized=1 zdroj živin?=ano podpora v rodině?=ano propagace ve školách?=ano 126 ==> je mléko prospěšné zdraví?=ano 126 acc:(0.99127)
4. sýry_binarized=1 zdroj živin?=ano pozitivní vliv?=ano podpora v rodině?=ano 124 ==> je mléko prospěšné zdraví?=ano 124 acc:(0.99117)
5. zdroj živin?=ano pozitivní vliv?=ano podpora v rodině?=ano automaty?=ano 123 ==> je mléko prospěšné zdraví?=ano 123 acc:(0.99111)
6. zdroj živin?=ano pozitivní vliv?=ano zdroj vápníku?=ano podpora v rodině?=ano 122 ==> je mléko prospěšné zdraví?=ano 122 acc:(0.99106)
7. dobře stravitelné?=ano děti?=ano 115 ==> je mléko prospěšné zdraví?=ano 115 acc:(0.99064)
8. ochrana před osteoporézou?=ano podpora v rodině?=ano 113 ==> je mléko prospěšné zdraví?=ano 113 acc:(0.99051)
9. zdroj živin?=ano zdroj vápníku?=ano děti?=ano 112 ==> je mléko prospěšné zdraví?=ano 112 acc:(0.99044)
10. zdroj živin?=ano děti?=ano automaty?=ano 112 ==> je mléko prospěšné zdraví?=ano 112 acc:(0.99044)
11. zdroj živin?=ano zdroj vápníku?=ano dobře stravitelné?=ano 109 ==> je mléko prospěšné zdraví?=ano 109 acc:(0.99023)
12. dobře stravitelné?=ano podpora v rodině?=ano automaty?=ano 107 ==> je mléko prospěšné zdraví?=ano 107 acc:(0.99008)
13. ochrana před osteoporézou?=ano děti?=ano 106 ==> je mléko prospěšné zdraví?=ano 106 acc:(0.99001)
14. zdroj živin?=ano pozitivní vliv?=ano zdroj vápníku?=ano ochrana před osteoporézou?=ano 106 ==> je mléko prospěšné zdraví?=ano 106 acc:(0.99001)
15. ochrana před osteoporézou?=ano dobře stravitelné?=ano 104 ==> propagace ve školách?=ano 104 acc:(0.98985)
16. ochrana před osteoporézou?=ano automaty?=ano 104 ==> propagace ve školách?=ano 104 acc:(0.98985)
17. zdroj živin?=ano ochrana před osteoporézou?=ano dobře stravitelné?=ano 100 ==> je mléko prospěšné zdraví?=ano 100 acc:(0.98952)
18. jogurty_binarized=1 ochrana před osteoporézou?=ano 99 ==> je mléko prospěšné zdraví?=ano pozitivní vliv?=ano 99 acc:(0.98943)
19. jogurty_binarized=1 zdroj vápníku?=ano dobře stravitelné?=ano 93 ==> je mléko prospěšné zdraví?=ano 93 acc:(0.98886)
20. zdroj živin?=ano pozitivní vliv?=ano zdroj vápníku?=ano dostatečná propaga-

- ce?=ne 88 ==> je mléko prospěšné zdraví?=ano 88 acc:(0.98833)
21. sýry_binarized=1 zdroj živin?=ano pozitivní vliv?=ano dostatečná propagace?=-ne 87 ==> je mléko prospěšné zdraví?=ano 87 acc:(0.98821)
22. ochrana před osteoporézou?=ano dostatečná propagace?=-ne 86 ==> propagace ve školách?=ano 86 acc:(0.98809)
23. zdroj živin?=ano zdroj vápníku?=ano podpora v rodině?=ano dostatečná propagace?=-ne 83 ==> je mléko prospěšné zdraví?=ano 83 acc:(0.98772)
24. zdroj živin?=ano zdroj vápníku?=ano cholesterol a tuky?=-ne 76 ==> je mléko prospěšné zdraví?=ano 76 acc:(0.98673)
25. zdroj živin?=ano děti?=ano 137 ==> je mléko prospěšné zdraví?=ano 136 acc:(0.98652)
26. pozitivní vliv?=ano děti?=ano 137 ==> je mléko prospěšné zdraví?=ano 136 acc:(0.98652)
27. ochrana před osteoporézou?=ano dobře stravitelné?=ano dostatečná propagace?=-ne 74 ==> je mléko prospěšné zdraví?=ano 74 acc:(0.98642)
28. děti?=ano propagace ve školách?=ano 135 ==> je mléko prospěšné zdraví?=ano 134 acc:(0.98631)
29. ochrana před osteoporézou?=ano cholesterol a tuky?=-ne 70 ==> propagace ve školách?=ano 70 acc:(0.98572)
30. máslo_binarized=1 děti?=ano 69 ==> podpora v rodině?=ano 69 acc:(0.98554)

Z uvedeného přehledu jsou z hlediska interpretace nejzajímavější následující pravidla. Celkově algoritmus konvergoval (díky vysoké četnosti odpovědí o zdravotní prospěšnosti mléka) k výběru pravidel se závěrem o zdravotní prospěšnosti mléka. Např. jestliže respondent vnímá mléko jako zdroj živin s pozitivním vlivem na vývoj dětí, pak považuje mléko také za prospěšné zdraví, odpovědělo 133 respondentů. Pokud považují respondenti mléko za důležitý zdroj živin pro děti a podporují jeho propagaci ve školách, pak se také domnívají, že je mléko zdraví prospěšné (pravidlo pokrývá 131 respondentů).

Překvapivým pravidlem je např. následující pravidlo s četností 126 respondentů: jestliže respondenti upřednostňují konzumaci sýrů a podporují konzumaci mléka v rodině a jeho propagaci ve školách, pak považují mléko za prospěšné zdraví. První část pravidla tak formuje skupinu s pozitivní odpovědí ohledně zdravotní prospěšnosti mléka jako konzumenty, kteří preferují sýry a zároveň podporují propagaci mléka. Pravidla týkající se názorů na zdravotní prospěšnost mléka nepřinesla žádná překvapivá zjištění - např. jestliže respondent vnímá konzumaci mléka jako ochranu před osteoporézou a považuje mléko za dobře stravitelné, pak podporuje propagaci mléka ve školách, odpovědělo 104 dotázaných. Pokud dotázaný upřednostňuje konzumaci jogurtů a považuje mléko za ochranu před osteoporézou, pak považuje mléko za zdraví prospěšné s pozitivním vlivem na organismus, odpovědělo 93 dotázaných. Jestliže respondenti považují mléko za zdroj živin a vápníku a myslí si, že mléko neobsahuje cholesterol a tuky, pak považují mléko za prospěšné zdraví, odpovědělo 76 dotázaných.

Pravidlo 30 - pokud respondent upřednostňuje konzumaci másla a vede děti ke konzumaci mléka, podporuje konzumaci mléka v rodině - umožňuje odvodit, že při případné propagaci např. nových mléčných výrobků by bylo vhodné se zaměřit na rodiče s dětmi, kteří nakupují máslo. Odvozené pravidlo umožňuje aktivně zacílit reklamní kampaň na tuto skupinu zákazníků. Výsledek také naznačuje, že není možno spoléhat na jednoduché třídění výstupu algoritmu Apriori, které upřednostňuje nejčetnější, ale málo zajímavá pravidla.

Výstup algoritmu Tertius s 10 nejlepšími nalezenými asociačními pravidly.

Tertius

1. /* 0,261551 0,005000 */ cholesterol a tuky? = nevím and je vhodné jiné mléko? = nevím ==> zakysané zdravější? = nevím or negativní vliv na zdraví? = nevím
2. /* 0,261372 0,030000 */ sýry_binarized = 1 and konzumuji mléko = denně and jogurty_binarized = 1 ==> konzumuji výrobky = denně
3. /* 0,258493 0,015000 */ máslo_binarized = 1 and je vhodné jiné mléko? = nevím ==> zakysané zdravější? = nevím or negativní vliv na zdraví? = nevím
4. /* 0,256330 0,045000 */ sýry_binarized = 1 and konzumuji mléko = denně ==> konzumuji výrobky = denně
5. /* 0,249841 0,015000 */ máslo_binarized = 1 and je vhodné jiné mléko? = nevím ==> zakysané zdravější? = nevím or negativní vliv na zdraví? = ano
6. /* 0,247149 0,045000 */ je vhodné jiné mléko? = nevím ==> cholesterol a tuky? = ano or zakysané zdravější? = nevím or negativní vliv na zdraví? = nevím
7. /* 0,246653 0,045000 */ cholesterol a tuky? = ne and jogurty_binarized = 1 ==> konzumuji mléko = 1 týdně or je vhodné jiné mléko? = ano
8. /* 0,241461 0,060000 */ tvaroh_binarized = 1 and jogurty_binarized = 1 ==> konzumuji výrobky = denně or cholesterol a tuky? = ano
9. /* 0,236400 0,045000 */ konzumuji výrobky = 2-3 týdně ==> tvaroh_binarized = 0 or zakysané zdravější? = ne or jogurty_binarized = 0
10. /* 0,235139 0,140000 */ sýry_binarized = 1 and jogurty_binarized = 1 ==> konzumuji výrobky = denně or cholesterol a tuky? = ano

Number of hypotheses considered: 120563

Number of hypotheses explored: 71197

Výstupem algoritmu Tertius byla nalezena následující interpretačně zajímavá asociační pravidla. Pokud dotázaný upřednostňuje konzumaci sýrů a konzumuje mléko denně, pak konzumuje denně také mléčné výrobky.

Jestliže respondenti neznají odpověď na otázku zda je vhodné konzumovat jiné mléko než kravské, pak si myslí, že mléko obsahuje cholesterol a tuky nebo neumí posoudit, zda je zakysané mléko zdravější nebo jestli má mléko negativní vliv na zdraví. Toto pravidlo odráží vysokou neinformovanost respondentů o mléce obecně. Vysoký počet respondentů v předpokladu pravidla ukazuje na potřebu kampaně o složení a zdravotních účincích mléka, která by podle tohoto pravidla měla dopad na velkou cílovou skupinu, která si neuvědomuje přednosti zakysaných mléčných výrobků.

Pokud dotázaný upřednostňuje konzumaci tvarohu a jogurtů, pak konzumuje mléko denně nebo si myslí, že mléko obsahuje cholesterol a tuky. V tomto odvozeném pravidle se odráží úspěšná marketingová kampaň výrobců jogurtů, kteří masivní reklamou dokázali přesvědčit část populace o zdravotních benefitech svých výrobků, přestože tato skupina respondentů považuje mléko za nezdravé (zdroj tuků a cholesterolu).

Celkově algoritmus konvergoval k odvození pravidel obsahujících v předpokladu kombinaci preferencí konzumovaných mléčných výrobků v kombinaci s názory na zdravotní prospěšnost mléka. Názory na zdravotní prospěšnost mléka se pak nejčastěji objevují v závěru odvozených asociačních pravidel.

5 Závěr

Práce se zabývá zpracováním dotazníkových dat pro dvě stanovené úlohy:

- charakterizaci typického účastníka vzdělávacích seminářů v zemědělství
- zjištění preferencí respondentů při konzumaci potravin živočišného původu a jejich zdrojů – mléka a masa.

K analýze zvolených úloh byly v práci využity metody popisné statistiky a dataminingu. V rámci první úlohy bylo analyzováno 342 vyplněných dotazníků získaných na 5 vzdělávacích seminářích se zemědělskou tematikou (produkce mléka, hovězího masa, problematika výživy skotu). Pomocí metod popisné statistiky (tabulky charakterizující absolutní a relativní počty respondentů) byl soubor účastníků popsán z hlediska jejich sociologických údajů (věk, pohlaví, profese, stupeň a obor vzdělání, region). Byla zjištěna vyrovnaná účast mužů a žen na seminářích se zemědělskou tematikou – 191 mužů, 114 žen. Seminářů se profesně zúčastnil největší počet respondentů (70), kteří zastávají místa vedoucích pracovníků v živočišné výrobě, což ukazuje na vysoký zájem o další vzdělávání u této profese. Taktéž pedagogové středních a vysokých zemědělských škol projevují zájem o postgraduální vzdělání v oboru (36 účastníků). Ze sociologického hlediska přinesla zajímavé výsledky analýza úrovně vzdělání, oboru vzdělání a profesního zaměření respondentů. 34 ze 102 respondentů se středoškolským vzděláním pracuje na pozici řídicích pracovníků v živočišné výrobě, nepokračovalo tedy ve studiu na vysoké škole. Seminářů se zemědělskou problematikou se z vysokoškolsky vzdělaných respondentů účastní především lidé se zemědělsky zaměřeným vzděláním (134 respondentů ze 161 vysokoškolsky vzdělaných). Vysokoškolsky vzdělaní respondenti pracují v zemědělství zejména jako pracovníci státní správy (11), poradci (21), řídicí pracovníci (25), učitelé (22).

Regionální příslušnost účastníků seminářů reflektuje místo jejich konání, stejně jako tradiční zemědělské regiony s většími výrobními podniky, což lze připsat většímu zájmu jejich managementu o další vzdělávání a aktuální poznatky v oboru.

K popisu typického účastníka semináře byly kromě zmíněných metod zvoleny dva algoritmy shlukové analýzy dat. Jednalo se o k-means a EM algoritmus. Oba algoritmy odvodily z popsání datového souboru typické účastníky seminářů následujícím způsobem. K-means algoritmus reprezentuje dva typické účastníky jako muže ve věku nad 50 let se zemědělským vysokoškolským vzděláním z Moravskoslezského kraje – tyto podmínky splňuje 64,23 % respondentů. Druhý shluk je pak reprezentován 35,77 % respondentů, kteří jsou muži se středoškolským vzděláním se zemědělským zaměřením, kteří pracují jako řídicí pracovníci v živočišné výrobě v Olomouckém kraji.

EM algoritmus zkonvergoval řešení, jehož dva odvozené shluky reprezentují typické účastníky jako muž ve věku do 20 let s ukončeným středoškolským vzděláním a profesí student z Moravskoslezského kraje, a muž ve věku 31-50 let s vysokoškolským vzděláním pracující na pozici řídicího pracovníka v živočišné výrobě v Olomouckém kraji.

Oba algoritmy efektivně popisují typického účastníka vzdělávacích seminářů v zemědělství a umožňují cílit navazující vzdělávací aktivity na tyto skupiny. Mezi shluky vždy existují statisticky významné rozdíly ve většině sledovaných proměnných, oba algoritmy lze tedy k podobným úlohám doporučit. Díky lepší interpretaci výsledků se kloníme spíše k využití sofistikovanějšího EM algoritmu.

Dotazníková data týkající se preferencí konzumentů potravin živočišného původu a jejich zdrojů byla zpracována pomocí metod popisné statistiky, byla u nich provedena exploratorní analýza a formulovány hypotézy o zajímavých vztazích mezi preferencemi konzumentů a jejich sociologickými údaji (věk, pohlaví, vzdělání atd.). Další aplikovanou analýzou bylo použití dvou algoritmů pro odvození asociačních pravidel, která by odhalila souvislosti mezi analyzovanými parametry týkající se preferencí konzumentů. V rámci popsání analýz byla otestována data pocházející ze dvou dotazníkových průzkumů zaměřených na: preference konzumentů a názor na zdravotní benefity konzumace mléka a mléčných výrobků (200 respondentů); preference při nákupu, konzumaci a úpravě hovězího masa (132 respondentů).

V rámci popisu souboru respondentů dotazníkového průzkumu týkajícího se hovězího masa, provedeného mezi odbornou veřejností na zemědělské výstavě, byly mimo jiné zjištěny následující závěry. 58,33 % respondentů nakupuje výsekové maso alespoň jedenkrát za týden, 8,33 % využívá výsekové maso z vlastních zdrojů. Pouze 32,57 % respondentů nakupuje výsekové hovězí maso alespoň 1x za měsíc, 15,15 % respondentů vůbec nekupuje hovězí maso. V otázce nejčastěji konzumovaného druhu masa 16,66 % zvolilo rybí maso, pouze 3,78 % jehněčí, 31,06 % hovězí, 64,39 % vepřové a 43,93 % drůbeží maso. Jako oblíbené maso označilo ze všech respondentů drůbeží pouze 26,51 %, vepřové 41,67 %, hovězí 41,67 %, jehněčí 10,61 a rybí 22,27 % respondentů. Z těchto výsledků je patrné, že jeho častější konzumaci některých druhů mas brání jiné vlivy, než jejich oblíbenost (např. cena, dostupnost, kolísavá kvalita).

Kuchyňskou úpravu hovězího masa vařením má v oblibě pouze 15,15 %, pečením 35,61 %, dušením 27,27 %, grilováním 43,18 % respondentů. Z tohoto hlediska lze výsledky interpretovat, jako posun českého konzumenta hovězího masa od tradičních způsobů úpravy (vaření, dušení) k modernímu grilování, ale také pečení (roastbeef, rolády atd.).

Při testování formulovaných hypotéz nebyl zjištěn vztah mezi pohlavím respondentů a volbou hovězího jako nejčastěji konzumovaného ani nejoblíbenějšího masa. Při testování hypotéz o preferované kuchyňské úpravě hovězího masa a věkovou kategorií byl nalezen vztah mezi věkovou kategorií respondentů a jejich vztahu k dušení jako kuchyňské úpravě. Následně byla hypotéza otestována pomocí ANOVA pro věk jako spojitou veličinu a efekt preference dušení. Výsledky odpovídají statistickému trendu, který lze vyjádřit tvrzením, že starší lidé (průměrný věk 42,15 let) spíše nezavrhují/preferují dušení hovězího masa v porovnání s mladšími lidmi (průměrný věk 36,89 let). Mezi proměnnými pohlaví a preferovaný způsob úpravy steaku nebyl v rámci testování této hypotézy prokázán vztah.

Popsaný datový set byl použit jako vstup pro odvození asociačních pravidel pomocí algoritmu Apriori a Tertius. Na základě odvozených asociačních pravidel bylo formulováno několik doporučení pro producenty a prodejce hovězího masa. Jestliže například respondent nakupuje výsekové hovězí maso méně než 1x za měsíc a neměl nikdy steak, pak nemá rád ani vařené maso. Toto pravidlo pokrývá skupinu 13 respondentů. Není tedy vhodné této skupině zákazníků nabízet steakové maso zároveň s masem vhodným k vaření. Podobným příkladem je pravidlo, které popisuje, že pokud dotazovaní nepreferují vaření a pečení hovězího masa, pak dávají přednost dušení nebo grilování. Toto pravidlo je prakticky využitelné například pro restaurace, které by měly nabízet na jídelním lístku jídla obou úprav, ale také pro

prodejce výsekového masa. Zákazník bude v prodejně volit maso vhodné pro jeden nebo druhý druh úpravy, prodejce by z hlediska objemu prodeje měl nabízet oba dva. Další z pravidel odvozených pomocí algoritmu Tertius říká, že pokud respondenti nepreferují pečení masa a zároveň jeho grilování, nakupují výsekové maso méně než jednou za měsíc a upřednostňují jeho dušení. Toto pravidlo může v praxi znamenat doporučení pro prodejce, aby výsekové maso vhodné k dušení udržoval pouze mražené. Z marketingově hodnotných výsledků lze vybrat například pravidlo popisující skupinu, která se aktivně (průzkum byl spojen s ochutnávkou) zajímá o úpravu hovězího masa grilováním - jestliže jsou dotazované osoby ve věku 23-28 roků, pak jsou s vysokoškolským vzděláním. Toto pravidlo pokrývá množinu 12 respondentů. Na popsanou skupinu by měla být cílena reklamní kampaň.

Z popisných statistik respondentů dotazníkového průzkumu týkajícího se konzumace mléka, mléčných výrobků a názorů na jejich zdravotní přínosy vyplývá např., že 28,5 % respondentů preferuje jak konzumaci mléka, tak mléčných výrobků, 49 % respondentů preferuje pouze mléčné výrobky a 22,5 % preferuje pouze konzumaci mléka. Jako oblíbený výrobek uvedlo jogurty 76,5 %, sýry 83 %, tvaroh 44 %, máslo 55,5 % a jiné mléčné výrobky 4 % z celkového počtu respondentů. Pokud se týká názorů na zdravotní prospěšnost konzumace mléka, považuje mléko za zdraví prospěšné 92 % respondentů, 2 % odpověděla, že mléko není zdraví prospěšné, 5,5 % uvedla, že neví, zda mléko je či není prospěšné zdraví. Vysokou neinformovanost prokázali respondenti při hodnocení zdravotních benefitů mléka v otázkách, zda je mléko zdrojem vápníku, ochranou před osteoporézou, a také zda je zdrojem tuků a cholesterolu (ano odpovědělo 13,5 %, neví 35,5 % respondentů). Celkově výsledky vypovídají o nepříliš vysoké úrovni informovanosti populace v otázkách zdravotních benefitů konzumace kravského mléka a celkové problematiky; ukazují spíše na povrchní informace spotřebitelů.

V rámci analýzy datového souboru byly otestovány hypotézy a o vztahu věku a pohlaví k preferenci konzumace mléka nebo mléčných výrobků. Ani v jednom případě se závislost testovaných parametrů neprokázala, lze tvrdit, že respondenti neupřednostňují konzumaci mléka, mléčných výrobků podle věku či pohlaví.

Naopak testování hypotéz o vztahu věku a úrovně vzdělání vzhledem k názoru na zdravotní prospěšnost mléka přinesly následující výsledky. Především starší lidé hodnotí zdravotní přínosy mléka kladně (skupiny 31-50 let a nad 50 let) a považují jej za prospěšné zdraví. Byla prokázána také závislost mezi věkem respondentů, úrovní jejich vzdělání a jejich názorem na mléko jako ochranu před osteoporézou. Starší (31-50, nad 50 let) a vzdělanější lidé (SŠ, VŠ) jsou lépe informovaní o tomto zdravotním benefitu mléka. Poslední testovanou hypotézou byl vztah mezi věkem a vzděláním respondentů a jejich hodnocením mléčných zakysaných výrobků jako zdravějších. Obě nulové hypotézy o nezávislosti sledovaných jevů byly vyvráceny. Respondenti hodnotí zdravotní prospěšnost zakysaných mléčných výrobků v závislosti na vzdělání i věku. Doporučením pro výrobce by mohlo být cílení reklamy na skupinu respondentů ve věku 21-30 let, která nevykazuje příliš vysoké hodnocení oproti starším věkovým skupinám. Obdobně respondenti se základním vzděláním vykazují poměrně nízkou frekvenci hodnocení zdravotní prospěšnosti zakysaných mléčných výrobků, doporučením je proto výrazně zvýšit informovanost mladších věkových skupin o benefitech zakysaných mléčných výrobků, případně připravit produkt přímo pro tyto cílové skupiny.

Při odvození asociačních pravidel pomocí algoritmu Apriori byly zjištěny např.

tyto pravidla: „pokud respondent upřednostňuje konzumaci másla a vede děti ke konzumaci mléka, podporuje konzumaci mléka v rodině“ - umožňuje odvodit, že při případné propagaci např. nových mléčných výrobků by bylo vhodné se zaměřit na rodiče s dětmi, kteří nakupují máslo. Odvozené pravidlo umožňuje aktivně zacílit reklamní kampaň na tuto skupinu zákazníků. Při plánování obchodní strategie může být užitečné také následující pravidlo odvozené algoritmem Tertius: „Pokud dotázaný upřednostňuje konzumaci sýrů a konzumuje mléko denně, pak konzumuje denně také mléčné výrobky“. Pravidlo: „Pokud dotázaný upřednostňuje konzumaci tvarohu a jogurtů, pak konzumuje mléko denně nebo si myslí, že mléko obsahuje cholesterol a tuky“ odráží např. úspěšnou marketingovou kampaň výrobců jogurtů, kteří dokázali přesvědčit část populace o zdravotních benefitech svých výrobků, přestože tato skupina respondentů považuje mléko za nezdravé (zdroj tuků a cholesterolu).

Dá se konstatovat, že oba algoritmy jsou vhodné pro stanovenou úlohu – hledání souvislostí v dotazníkových datech a mohou být vhodným nástrojem pro jejich analýzu s komerční aplikací. Algoritmus Apriori má sklon odvozovat více obecná pravidla, neboť pracuje na základě kritéria četnosti jejich literálů. Algoritmus Tertius obecně produkuje interpretačně zajímavější pravidla založená na kritériu spolehlivosti (potvrzení) pravidel, která v závěru pravidel obsahují logické kombinace vytvořené pomocí operátorů AND a OR. Přes stejnou vypovídací schopnost aparátu je interpretace asociačních pravidel odvozených pomocí algoritmu Tertius vhodnější pro popsanou úlohu.

6 Seznam použitých zdrojů

- [1] HANUŠ, O. a kol. *Dopad profesního vzdělávání v oboru chovu skotu – výrobní zemědělská praxe a potravinářské biotechnologické úpravy pro zvýraznění pozitivních zdravotních vlivů mléka a mléčných výrobků*. Výzkum v chovu skotu. ISSN 0139-7265. sv. 185, ročník LI, číslo 1. S 49-63.
- [2] HANUŠ, O. a kol. *Účinnost postgraduálního odborného vzdělávání v problematice chovu skotu a kvality jeho produktů – zdraví dojníc a kvalita produkce mléka*. Výzkum v chovu skotu. 2006. ISSN 0139-7265.
- [3] HANUŠ, O. a kol. *Dopad postgraduálního vzdělávání v oboru chovu skotu – ekologické, zdravotní a hygienické faktory kvality a bezpečnosti mléka jako suroviny a potraviny*. Výzkum v chovu skotu. 2007. ISSN 0139-7265.
- [4] HANUŠ, O. a kol. *Dopad profesního vzdělávání v oboru chovu skotu - 6. část – šetrné čerpání přírodních zdrojů a údržba krajiny pomocí chovu krav bez tržní produkce mléka*. Výzkum v chovu skotu/Cattle research. 2009. ISSN 0139-7265.
- [5] StatSoft Inc. (2008). Statistica 8. Tulsa, OK.
- [6] Waikato Environment for Knowledge Analysis - Weka 3.6.1. The University of Waikato, Hamilton, New Zealand, URL: <http://www.cs.waikato.ac.nz/~ml/weka/>.
- [7] SVATOŠOVÁ, Libuše, KÁBA, Bohumil. *Statistické metody II*. 1. vyd., dotisk. Praha: Reprografické studio PEF ČZU v Praze, 2008. 105s. ISBN 978-80-213-1736-9.
- [8] HENDL, Jan. *Přehled statistických metod zpracování dat*. 1. vyd. Praha: nakladatelství Portál, 2004. 584s. ISBN 80-7178-820-1.
- [9] SVATOŠOVÁ, Libuše, KÁBA, Bohumil. *Statistické metody I*. 1. vyd., 2. dotisk. Praha: Reprografické studio PEF ČZU v Praze, 2007. 132s. ISBN 978-80-213-1672-0.
- [10] BERKA, Petr. *Dobývání znalostí z databází*. 1. vyd. Praha: Academia, 2003. 366s. ISBN 80-200-1062-9.
- [11] HOFMANN, T: *Introduction to Machine Learning*. [studijní materiál], Brown University, Rhode Island, USA, 2003.
- [12] FLACH, P. A., LACHICHE, N. (1999). *Confirmation-Guided Discovery of first-order rules with Tertius*. Sec. 42:61-95. Machine Learning.
- [13] BURIÁNEK, Jiří. *Sociologie*, 1. vyd. Praha: Nakladatelství Fortuna, 1996. 128 s. ISBN 80-7168-304-3.
- [14] PETRUSEK, Miroslav a kol. *Sociologie*, 2. vyd. Praha: Státní pedagogické nakladatelství, 1994. 118s. ISBN 80-04-26588-X.
- [15] SURYNEK, Alois, KOMÁRKOVÁ, Růžena, KAŠPAROVÁ, Eva. *Základy sociologického výzkumu*, 1. vyd. Praha: Nakladatelství Management Press, 2001.

160s. ISBN 80-7261-038-4.

- [16] ĎURIČ, L. *Úvod do pedagogické psychologie*, Praha: SPN, 1979. 286s.
- [17] SKALKOVÁ, J. *Úvod do metodologie a metod pedagogického výzkumu*, Praha: SPN, 1985, 209s.
- [18] HOLÝ, Ivan. *Úvod do sociologie*, 1. vyd. Brno – Kraví hora: Vydavatelství MU, 1994. 65s. ISBN 80-210-0985-3.
- [19] Středisko empirických výzkumů: <http://www.stem.cz/pages/faq.php#faq15>, [20011-01-18].
- [20] GALLUP, Georgie. *Průvodce po výzkumu veřejného mínění*, 1. vyd. Praha : 1948.
- [21] Český statistický úřad: http://www.scitani.cz/slodb2011/redakce.nsf/i/o_scitani, [2011-01-21].
- [22] Český statistický úřad: http://www.czso.cz/csu/redakce.nsf/i/o_ceskem_statistickem_uradu, [2011-01-18].
- [23] Výzkumy veřejného mínění: <http://regionalka.wz.cz/reg%20info/vyzkumy%20VM.htm>, [2011-01-21].
- [24] KOTLER, P.: *Marketing Management*. Prentice-Hall International Editions, Englewood Cliffs, N.J. 1991.
- [25] Státní zemědělská a potravinářská inspekce: <http://www.szpi.gov.cz/lstDoc.aspx?nid=11328> [2011-01-21].
- [26] Státní zemědělská a potravinářská inspekce: <http://www.szpi.gov.cz/default.aspx> [2011-01-21].
- [27] Vyplň to.cz, Ing. Marek Demčák: <http://www.vyplnto.cz/realizovane-pruzkumy/> [2011-01-21].
- [28] České a slovenské odborné nakladatelství: <http://www.foodservice.cz/e-zpravodaj/ceny-biopotravin-zustavaji-stabilni.htm>, [2011-02-27].
- [29] PRO-BIO LIGA: http://www.pro-bio.cz/bioliga/kdo_jsme.htm [2011-01-21].
- [30] Veřejná databáze ČSÚ: <http://vdb.czso.cz/vdbvo/tabparam.jsp?voa=tabulka&cislatab>, [2011-03-13].
- [31] STEINHAUSER, Ladislav. *Produkce masa*. Tišnov : Last, 2000. 464 s. ISBN 80-900-2607-9.

7 Přílohy

Příloha 1. Dotazník použitý pro úlohu charakterizace typického účastníka vzdělávacích seminářů v zemědělství.

**Exploratorní dotazník
pro hodnocení názorů respondentů jako účastníků
vzdělávacího semináře**

Datum konání semináře:

Téma semináře:

*Vyplnění dotazníku je dobrovolné a striktně anonymní!
Hoditci se zaškrtněte, nebo vyplňte požadovaný údaj.*

Část I A

(týkající se obecných dotazů o charakteristice respondenta)

Respondent:

žena	
muž	

Pohl. respondent:

do 20 roků	
21 až 30 roků	
31 až 50 roků	
nad 50 roků	

Stupeň dosaženého vzdělání respondentem:

základní	
všechni obor bez maturity	
střední s maturitou	
vyšší odborné (univerzita)	
akademický stupeň (vědecké)	

Obor vzdělání respondentem:

zemědělské	
potravinářské	
přírodovědné	
veterinární	
technické	
humanitní	
jiné	

Profese respondentem:

student	
řídící technologický pracovník v živočišné výrobě	
řídící technologický pracovník v potravinářství	
pracovník v zemědělské nebo potravinářské laboratoři	
veterinární lékař	
poradce v zemědělské výrobě	
pracovník zemědělských služeb	
učitel zemědělského oboru	
pracovník sítní správy v zemědělství	
pracovník profesní zemědělské organizace	
ostatní profese	

Revize (třeba okres) popřípadě místo realizace profese (zaměstnání) respondentem:

Část I B

(týkající se obecných dotazů k seminářům)

Hodnocení organizační úrovně semináře:

výborná	
velmi dobrá	
dobrá	
špatná	

Hodnocení odborné úrovně semináře:

výborná	
velmi dobrá	
dobrá	
špatná	

Možná využitelnost poznatků:

bezprostředně	
později	
vůbec	

Potřeba sborníků:

je nezbytný	
je potřebný pro pozdější využití poznatků	
není potřebný	

Použití didaktické pomůcky při prezentaci:

odpovídající modernímu trendu (PC a dataprojektor ap.)	
průměrná úroveň	
špatná úroveň	

Úroveň zařčení pohostění jako doprovodného servisu:

výborná	
velmi dobrá	
dobrá	
špatná	

Hodnocení - vynecháváte-li nejlépe pedagogicky a technicky provedené prezentace od nejlépeší:
(vysluhování lektorů, přeches a jeho srozumitelnost, dodržení času)

1.	
2.	
3.	

Návrh pro příští zlepšení organizace vzdělávacího semináře:


.....

.....

Příloha 2. Dotazník použitý pro zjištění preferencí konzumentů masa.

<p style="text-align: center;">Dotazník</p> <p><input type="checkbox"/> žena <input type="checkbox"/> muž</p> <p>1. Jaký je Váš věk? <input type="text"/> student <input type="text"/> produktivní věk <input type="text"/> důchodce</p> <p>2. Jaké je Vaše dosažené vzdělání? <input type="text"/> základní <input type="text"/> střední <input type="text"/> střední odborné <input type="text"/> vysoké odborné</p> <p>3. Jak často nakupujete výsekové maso? <input type="text"/> 5x a více za týden <input type="text"/> 4x-2x <input type="text"/> 1x za týden</p> <p>4. Jak často nakupujete výsekové hovězí maso? <input type="text"/> 1x za týden <input type="text"/> 1x za měsíc <input type="text"/> méně než 1x za měsíc <input type="text"/> vůbec</p> <p>5. Jaké maso konzumujete nejčastěji? <input type="text"/> rybí <input type="text"/> jehněčí <input type="text"/> hovězí <input type="text"/> vepřové <input type="text"/> drůbeží <input type="text"/> žádné <input type="text"/> jiné</p> <p>6. Vaše nejoblíbenější maso? <input type="text"/> rybí <input type="text"/> jehněčí <input type="text"/> hovězí <input type="text"/> vepřové <input type="text"/> drůbeží <input type="text"/> žádné <input type="text"/> jiné</p> <p>7. Konzumovali jste někdy hovězí maso, na kterém jste si pochutnali? <input type="text"/> ano <input type="text"/> ne</p> <p>8. Kdy to bylo naposledy? minulé <input type="text"/> týden <input type="text"/> měsíc <input type="text"/> rok</p> <p>9. Preferujete výsekové maso z tuzemské produkce? <input type="text"/> ano <input type="text"/> ne <input type="text"/> nezajímám se o původ <input type="text"/> zajímám se o původ nemohu ho získat</p> <p>10. Jakou kuchyňskou úpravu hovězího masa preferujete? <input type="text"/> vaření <input type="text"/> pečení <input type="text"/> dušení <input type="text"/> grilování</p> <p>11. Jakou propéčenost steaku preferujete? <input type="text"/> krvavý <input type="text"/> méně krvavý <input type="text"/> medium <input type="text"/> téměř propéčený <input type="text"/> propéčený</p>	<p style="text-align: center;">Dotazník</p> <p>Grilovaný steak jsem doposud <input type="text"/> měl <input type="text"/> neměl <input type="text"/> nechutnal mi <input type="text"/> nechutnal mi</p> <p>Steak, odpovídající kvalitou vzorku č. <input type="text"/> bych si chtěl <input type="text"/> objednat v restauraci <input type="text"/> připravit doma <input type="text"/> objednat v restauraci</p> <p>Oslovila Vás pozitivně úprava hovězího masa grilováním? <input type="text"/> ano <input type="text"/> ne <input type="text"/></p> <p style="text-align: center;">Děkujeme</p>
---	--

Příloha 3. Dotazník použitý pro zjištění preferencí konzumentů mléka a jejich názoru na jeho zdravotní prospěšnost.



VÝZKUMNÝ ÚSTAV PRO CHOV SKOTU, S.R.O., RAPOTÍN
ODDĚLENÍ VÝŽIVY ZVÍŘAT A KVALITY ŽIVOTNÍCH PRODUKTŮ
PRACOVIŠTĚ POHOŘELICE, VÍDEŇSKÁ 699, 691 23 POHOŘELICE

DOTAZNÍK – MLÉKO A MLÉČNÉ VÝROBKY

Dobrý den, rádi bychom Vám předem poděkovali za inováci Vašich odpovědí v následujícím anonymním dotazníku, které budou zpracovány v rámci připravovaného projektu zaměřeného na výzkum, vývoj a inovaci v oblasti konzumace kravského mléka a mléčných výrobků. Prosíme Vás tímto o pečlivé vyplnění a předem velice děkujeme.

1/ Osobní údaje

- Věk: let
- Pohlaví: žena muž
- Dosažené vzdělání: základní
 vyučení(á)
 středškolské s maturitou
 vysokškolské
jiné

2/ Stravovací návyky

- Upřednostňuji: kravské mléko
 mléčné výrobky jogurty
 sýry
 tvaroh
 máslo
jiné
- Konzumuji: kravské mléko
 mléčné výrobky denně
 2-3krát týdně
 1krát týdně
jiné

3/ Váš názor

Je kravské mléko zdravoti prospěšné? ano ne nevím

Je kravské mléko důležitým zdrojem základních živin? ano ne nevím

Má kravské mléko pozitivní vliv na zdraví člověka? ano ne nevím

Je kravské mléko hlavním zdrojem vápníku? ano ne nevím

Chrání kravské mléko před vznikem osteoporózy? ano ne nevím

Je kravské mléko dobře stravitelné? ano ne nevím

Obsahuje kravské mléko zvýšený podíl nezdravých tuků a cholesterol? ano ne nevím

Jsou zakysané mléčné výrobky zdravější než kravské mléko? ano ne nevím

Je podporována konzumace mléka ve Vaší rodině? ano ne nevím

Vedete Vaše děti ke konzumaci kravského mléka? ano ne nevím

Je kravské mléko propagováno dostatečně? ano ne nevím

Podpoří(a) byste propagaci kravského mléka ve školách? ano ne nevím

Podpoří(a) byste zavedení automatů na kravské mléko? ano ne nevím

Je vhodné konzumovat jiné mléko než kravské (sojové, kozí aj.)? ano ne nevím

3/ Myslete si, že kravské mléko může mít negativní vliv na zdraví člověka, a proč?

.....

.....

.....

Kontakt: Ing. Jarmila Svobodová, tel: 519 426 002, fax: 519 424 548, e-mail: Jarmila@emil.cz