

**Univerzita Pardubice
Fakulta ekonomicko-správní**

**Návrh modelu pro plánování preventivních kontrol bezpečnosti
práce**

Bc. Kateřina Štěpánková

**Diplomová práce
2009**

Univerzita Pardubice
Fakulta ekonomicko-správní
Ústav systémového inženýrství a informatiky
Akademický rok: 2008/2009

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Kateřina ŠTĚPÁNKOVÁ**
Studijní program: **N6209 Systémové inženýrství a informatika**
Studijní obor: **Pojistné inženýrství**

Název tématu: **Návrh modelu pro plánování preventivních kontrol bezpečnosti práce**

Z á s a d y p r o v y p r a c o v á n í :

Základní pojmy z oblasti bezpečnosti práce a pojištnictví.
Návrh modelu pro plánování kontrol BOZP u jednotlivých subjektů.
Ověření vytvořeného modelu a návrhy na zlepšení.

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

1. **DUCHÁČKOVÁ, Eva.** Principy pojištění v pojištnictví. 2. aktualiz. vyd. Praha : Ekopress, 2005. 178 s. ISBN 80-86119-92-0.
2. **BERKA, Petr.** Dobývání znalostí z databází. 1. vyd. Praha : [s.n.], 2003. 366 s. ISBN 80-200-1062-9
3. **BERRY, Michael J. A., LINOFF, Gordon.** Data mining techniques : For marketing, sales, and customer support. 1st edition. [s.l.] : [s.n.], 1997. 454 s. ISBN 0-471-17980-9
4. **PARR RUD, Olivia.** Data mining : Praktický průvodce dolováním dat pro efektivní prodej, cílený marketing a podporu zákazníků (CRM). 1. vyd. [s.l.] : [s.n.], 2001. 329 s. ISBN 80-7226-577-6

Vedoucí diplomové práce:

doc. Ing. Pavel Petr, Ph.D.

Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce:

6. října 2008

Termín odevzdání diplomové práce:

1. května 2009

doc. Ing. Renáta Myšková, Ph.D.

děkanka

L.S.

doc. Ing. Jiří Křupka, Ph.D.

vedoucí ústavu

V Pardubicích dne 6. října 2008

Prohlašuji:

Tuto práci jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně Univerzity Pardubice.

V Pardubicích dne 17.8.2009

Kateřina Štěpánková

Na tomto místě bych velmi ráda poděkovala svému vedoucímu práce, kterým byl doc. Ing. Pavel Petr, Ph.D., za pomoc, rady, cenné připomínky a inspirující vedení při tvorbě a kompletování této práce do finální podoby.

ANOTACE

Cílem práce je navrhnout model pro plánování preventivních kontrol dodržování bezpečnosti práce a tím také charakterizovat ekonomický subjekt, kde se stává nejvíce smrtelných pracovních úrazů. Jsou zde vymezeny základní právní předpisy týkající se bezpečnosti a ochrany zdraví při práci a statistika smrtelných pracovních úrazů za rok 2008. Dále jsou zde popsány jednotlivé etapy metodiky CRISP-DM, podle které je celá práce sestavena. Modelování dat proběhlo v software Clementine.

KLÍČOVÁ SLOVA

smrtelný pracovní úraz; data mining, CRISP-DM

TITLE

Scheme for plan steering control of work safety

ANNOTATION

The aim of this labour is to propose a scheme for plan steering control observantion of safety at work and by this to describe economic entity, where are happened the most of the infustrial fatalitys. Basic legislations of safety and protection of health in the work and statistics of industrial fatalitys in the 2008 are defined here. The separate phases of methodics CRISP-DM are described here. The labour is compiled by this methodics. The modeling of data slipped in software Clementine.

KEYWORDS

industrial fatality; data mining; CRISP-DM

OBSAH:

1	ÚVOD	8
2	Legislativa upravující pracovní úrazy.....	9
3	Pracovní a smrtelné úrazy ve vztahu k pojišťovnictví.....	10
3.1	Komerční pojištění	10
3.2	Sociální pojištění	13
3.3	Shrnutí kapitoly	14
4	Statistika smrtelných pracovních úrazů	14
4.1	Smrtelné pracovní úrazy v ČR v roce 2008.....	14
4.2	Smrtelné pracovní úrazy v Evropě	19
4.3	Shrnutí kapitoly	21
5	Data mining a metodologie CRISP-DM	21
5.1	Data mining	22
5.2	CRISP-DM	24
5.3	Shrnutí kapitoly	59
6	Závěr	60
7	Použitá literatura	62
	Seznam tabulek	64
	Seznam obrázků	65
	Seznam příloh.....	66

1 ÚVOD

V dnešní době, kdy čas je velmi cenný, se v oblasti zaměstnanosti klade čím dál větší důraz na dodržování bezpečnosti při práci. Ze zákona je jsou předepsaná každoroční školení, v některých odvětvích i vícekrát do roka. Každý nový zaměstnanec musí být řádně proškolen. Tato opatření se ovšem neprovádějí pouze v pracovně-právním vztahu, ale také ve školách, při nástupu do nemocnice jako pacient, v případě placeného a organizovaného adrenalinového sportu. Po proškolení dostane každý ochranné pomůcky, pokud jsou pro výkon jeho práce nutné, a tyto bezpečně dodržovat. Zdá se, že těmito postupy by se mělo pracovním, ale také smrtelným pracovním, úrazům zamezit. Realita je přitom odlišná. K úrazům dochází stále, i když s klesající tendencí.

Pro dodržování bezpečnosti ochrany zdraví při práci jsou zde specializované instituce. Tyto instituce provádějí kontroly v ekonomických subjektech po celé České republice. Především se zaměřují na záznamy ze školení o bezpečnosti práce a na používání předepsaných ochranných pomůcek. Po dokončení kontroly je zpracován zápis o průběhu kontroly. V lepším případě je celý pozitivní, totiž bezpečnost práce by měla být opravdu bezpečná. V opačném případě je subjekt na nedostatky upozorněn, vyzván k nápravě a samozřejmě předepsána pokuta.

Ovšem je nutné říci, že pracovníků provádějících tyto kontroly není mnoho, proto je velice nutné efektivně vybírat rizikové subjekty. V rámci projektu jsou zkoumána dostupná data o smrtelné pracovní úrazovosti v letech 2002 – 2007.

Cílem práce je vytvoření návrhu a ověření modelu pro plánování kontrol dodržování bezpečnosti práce. Je potřeba nalézt takovou organizaci, na základě průběhu smrtelné pracovní úrazovosti, která bude nejrizikovější. V takovém případě se kontroly bezpečnosti práce velmi zkvalitní a zefektivní.

V práci jsou využity prvky data miningu a metodiky CRISP-DM, která je rozdělena do šesti částí. Od fáze porozumění problému, přes modelování až po samotné využití výsledků v praxi. V našem případě pro provedení kontroly v předem identifikovaném ekonomickém subjektu.

V případě, že byl projekt správný, bude daná kontrola pozitivní. Tudíž můžeme říci, že kontrola byla efektivní a projekt správný.

2 Legislativa upravující pracovní úrazy

Zákonů a nařízení vlády je v oblasti pracovních úrazů celá řada. Zabývají se především bezpečností ochrany zdraví při práci, dále jen BOZP, evidencí, hlášení a řešení pracovních úrazů, právy a povinnostmi zaměstnance a zaměstnavatele, pracovními pomůckami, oděvem. Mezi nejdůležitější legislativu v této oblasti patří [11], [13], [23], [18], [21]:

1. Zákon č. 262/2006 Sb., zákoník práce, kde část V. je nazvaná Bezpečnost a ochrana zdraví při práci. Tato část obsahuje tři hlavy. Hlava 1 hovoří o předcházení ohrožení života a zdraví při práci. Obsahuje dva §, § 101 a § 102. Hlava 2, V. části zákoníku práce, pojednává o povinnostech zaměstnavatele, práva a povinnosti zaměstnance. Tato hlava obsahuje čtyři § (§ 103 - § 106). V § 103 jsou vymezeny základní povinnosti zaměstnavatele a zaměstnance. Následující § 104 upravuje otázky týkající se ochranných pracovních prostředků, pracovních oděvů a obuvi, mycích, čistících a dezinfekčních prostředků a ochranných nápojů. Další § 105 se týká povinností zaměstnavatele při pracovních úrazech a nemocech z povolání. Poslední § druhé hlavy, tj. § 106, hovoří o právech a povinnostech zaměstnance. Třetí, poslední, hlava páté části zákoníku práce obsahuje dva § (§ 107 a § 108), kde se hovoří o účasti zaměstnanců na řešení otázek BOZP [11].
2. Zákon č. 309/2006 Sb., kterým se upravují další požadavky BOZP v pracovněprávních vztazích a o zajištění bezpečnosti a ochrany zdraví při činnosti nebo poskytování služeb mimo pracovněprávní vztahy (zákon o zajištění dalších podmínek BOZP) [13].
3. Zákon č. 251/2005 Sb. o inspekci práce [23].
4. Nařízení vlády č. 494/2001 Sb., způsob evidence, hlášení a zasílání záznamu o úrazu [18].
5. Nařízení vlády č. 361/2007 Sb., kterým se stanoví podmínky ochrany zdraví při práci [21].

3 Pracovní a smrtelné úrazy ve vztahu k pojišťovnictví

Pojištění lze členit do několika skupin podle [7]:

- a) závaznosti:
 - zákonné (zdravotní, sociální),
 - smluvní – povinné a dobrovolné,
- b) druhu:
 - životní,
 - neživotní,
- c) míry rizikovosti:
 - rizikové,
 - rezervotvorné (např. pojištění na dožití).

Z důvodu pojistné ochrany pracovníků budeme dále hovořit pouze o komerčním pojištění a sociálním pojištění. Komerční pojištění patří mezi smluvní dobrovolné. Sociální pojištění patří mezi zákonná pojištění.

3.1 Komerční pojištění

Pojištění je nástroj finančního krytí negativních důsledků nahodilosti. Přenáší se jím riziko na specializovanou instituci, pojistitele. Pojištění je finanční služba a právní vztah, při kterém pojistitel přebírá závazek pojištěnému poskytnout pojistné plnění, nastane-li pojistná událost.

Neživotní pojištění zahrnuje krytí velké škály rizik neživotního charakteru. V rámci neživotních pojištění jsou kryta rizika různého charakteru. Jsou to rizika ohrožující zdraví a životy osob, rizika vyvolávající přímé věcné škody a rizika vyvolávající finanční ztráty.

Neživotní pojištění má pro případ pracovního úrazu hned několik produktů. Patří mezi ně úrazové pojištění a pojištění nemoci.

V případě úrazového pojištění se jedná o úrazové pojištění [7]:

- s jednorázovým plněním,
- splněním povahy náhrady škody,
- s kombinovaným plněním,
- cestujících (např. jako doplněk k povinnému ručení).

Úrazové pojištění zahrnuje výplatu pojistného plnění v případě, že v důsledku úrazu dojde k přechodnému nebo trvalému tělesnému poškození nebo smrti pojištěného. Přitom úraz bývá v pojišťovnictví charakterizován jako náhlé a neočekávané působení zevních sil nebo vlastní síly, neočekávané a nepřetržité působení vysokých teplot, par, plynů, záření a jedů, které měly za následek tělesné poškození nebo smrt [7].

Základem úrazového pojištění je krytí trvalých následků úrazu a smrti následkem úrazu.

V rámci úrazového pojištění může být též sjednáno pojistné plnění pro řadu dalších případů, např. [7]:

- plnění za dobu nezbytného léčení,
- denní odškodné za dobu pracovní neschopnosti,
- úhrada nákladů spojená s úrazem,
- pojištění drobných úrazů.

Pro případ pojištění nemoci [7]:

- s jednorázovým plněním,
- s plněním povahy náhrady škody,
- s kombinovaným plněním,
- smluvní zdravotní pojištění.

Toto pojištění, nemocenské pojištění, které je provozované komerčními pojišťovnami se uplatňuje jako doplněk povinného všeobecného zdravotního pojištění a povinného sociálního nemocenského pojištění, a to pro případ, kdy nejsou dané potřeby povinným pojištěním kryty.

Do této oblasti lze zařadit pojistné produkty zaměřené na krytí důsledků nemoci, které se projevují ve ztrátě nebo snížení příjmů v souvislosti s pracovní neschopností. Jde především o tyto produkty [7]:

- pojištění na denní dávku při pracovní neschopnosti, tedy nemocenské pojištění pro osoby, které nejsou povinně zahrnuty do sociálního nemocenského pojištění (osoby samostatně výdělečně činné), nebo jako doplňkové pojištění k povinnému nemocenskému pojištění (především pro osoby s vyššími příjmy),
- pojištění denní podpory při pobytu v nemocnici, tedy nemocenské pojištění v případě pobytu v nemocnici.

V některých profesích hrozí zaměstnanci určitá rizika spojená se zvýšenými nároky pracovní pozice na jeho zdraví. Pracovní úraz nebo nemoc z povolání může potkat ale každého zaměstnance. Nesouvisí totiž pouze s fyzicky náročnou nebo nebezpečnou prací. Zákon proto stanoví v rámci ochrany zaměstnance poměrně široké následky odpovědnosti zaměstnavatele za pracovní úrazy a nemoci z povolání zaměstnanců [7].

Potřeby v souvislosti s pracovními úrazy a nemocemi z povolání jsou kryty v rámci zákonného odpovědnostního pojištění, které provozují komerční pojišťovny.

V praxi to probíhá tak, že zaměstnavatel uzavře s komerční pojišťovnou smlouvu za své zaměstnance. Pojistné se vypočítává ze základu stanovené pro určení vyměřovacího základu na sociální zabezpečení. Sazby pojistného jsou stanoveny v právním předpise a jsou diferencované podle odvětví činnosti zaměstnavatele. Například v odvětví dobývání černého uhlí činí sazba 0,0504 %, v odvětví stavebnictví činí sazba 0,0098 %, v odvětví školství, cestovního ruchu a zdravotnictví činí sazba 0,0042 %, vyšší sazba je pro činnosti, kde se pracuje s výbušninami, radioaktivními látkami, jedy, činnosti ve velkých výškách, hloubkách a tato sazba zde činí 0,0105 [7].

Na základě existence pojištění odpovědnosti zaměstnavatele za škodu při pracovním úrazu a nemoci z povolání má zaměstnavatel právo, aby pojišťovna za něho nahradila škodu, která vznikla zaměstnanci při pracovním úrazu nebo nemoci z povolání v rozsahu, v jakém za škodu odpovídá zaměstnavatel podle zákoníku práce. Pojistnou událostí je zde vznik povinnosti zaměstnavatele hradit škodu.

Následující tabulka (tabulka 1) nastiňuje přijaté pojistné a vyplacené pojistné plnění v rámci zákonného pojištění odpovědnosti zaměstnavatele za škodu způsobenou při pracovním úrazu a nemoci z povolání.

Tabulka 1 - Přijaté pojistné a vyplacené poj. plnění v rámci odpovědnostního pojištění, upraveno s využitím [7]

Rok	1995	1996	1997	1998	1999	2000	2001
Přijaté pojistné (v mil. Kč)	1656,5	1922,2	2151,3	2281,9	2356,0	2402,2	2618,7
Vyplacené poj. plnění (v mil. Kč)	589,7	885,8	1180,0	1460,3	1650,8	1639,6	1849,1

Životní pojištění může vést také ke snížením následků vzniklých pracovními úrazy, ale převážně smrtelnými pracovními úrazy. Jedná se o finanční pomoc pozůstalým zemědělo.

Životní pojištění zahrnuje krytí rizik ohrožujících životy lidí a zahrnuje krytí pouze dvou rizik, a to riziko úmrtí a riziko dožití.

Výše pojistného plnění u životního pojištění není dána výší a velikostí škody, to by bylo jen velmi těžko odhadnutelné, ale je dána velikostní pojistné částky, kterou pojistník sjednal. Vlastní pojistné plnění, vedle krytí rizika, plní úsporovou funkci.

3.2 Sociální pojištění

Sociální pojištění patří mezi zákonné povinné pojištění, kdy se pojistná smlouva nesjednává, povinnost platit pojistné vyplývá ze zákona.

V rámci tohoto pojištění jsou kryta rizika, která mají sociální charakter, a z nich vyplývající potřeby. Jedná se o potřeby [7]:

- v souvislosti s dlouhodobou pracovní neschopností (sociální důchodové pojištění),
- v souvislosti s krátkodobou pracovní neschopností (sociální nemocenské pojištění),
- v souvislosti s nezaměstnaností (pojištění nezaměstnanosti).

Pro sociální pojištění je typické, že velikost pojistného není stanovena v závislosti na riziku, ale je pro všechny účastníky tohoto pojištění stanovena stejným způsobem, obvykle určitým procentem z příjmů. Další typickou vlastností sociálního pojištění je princip solidarity, na rozdíl od komerčního pojištění kde se uplatňuje princip ekvivalence. Princip solidarity je založen na tom, že veškeré příspěvky od pojištěnců jsou soustředovány do státního rozpočtu a následně vypláceny pouze těm zúčastněným, kteří byli postiženi daným sociálním rizikem. Mezi další vlastnost tohoto pojištění lze zařadit subjekty financování. Jsou jimi jednotliví pojištěnci (zaměstnanci), ale také zaměstnavatelé [7].

Výše pojistného na sociální pojištění v současné době činí [10]:

- a) placené zaměstnancem – 6,5 % z vyměřovacího základu,
- b) placené zaměstnavatelem – 25 % z úhrnu vyměřovacích základů,
- c) placené osobou samostatně výdělečně činnou – 29,2 % z vyměřovacího základu.

3.3 Shrnutí kapitoly

Pojištění jako takové lze členit do několika skupin podle závaznosti, druhu a míry rizikovosti. V těchto případech najdeme jak dobrovolná tak povinná pojištění. Dobrovolné (komerční) pojištění v sobě skrývá řadu rizik, proti kterým se lze pojistit. Pro případ pracovního úrazu je to úrazové pojištění a pojištění nemoci. Zákonným povinným pojištěním je sociální pojištění, které vyplývá ze zákona, neuplatňuje se zde princip ekvivalence, nýbrž princip solidarity.

4 Statistika smrtelných pracovních úrazů

Každoročně se vedou statistiky týkající se smrtelných úrazů, nemocí z povolání, pracovní neschopnosti delší jak 3 dny. Tyto statistické informace zpracovává Český statistický úřad, dále jen ČSÚ. ČSÚ provádí statistiky týkající se inflace, nezaměstnanosti, makroekonomických údajů (např. HDP, měnové ukazatele) a otázky týkající se obyvatelstva.

4.1 Smrtelné pracovní úrazy v ČR v roce 2008

Zatím posledními statistikami v oblasti smrtelných úrazů jsou údaje za rok 2008. Statistiky jsou rozděleny do několika kategorií a to podle krajů a ČR celkem, podle velikosti podniku, podle odvětvové klasifikace ekonomických činností, dále jen OKEČ, dělených podle krajů a ČR celkem. Další statistikou jsou smrtelné úrazy osob samostatně výdělečně činných, dále jen OSVČ.

Jednotkami pro účely evidence a statistického vykazování pracovní neschopnosti a smrtelných pracovních úrazů jsou všechny nefinanční podniky, finanční instituce, vládní instituce včetně příspěvkových, neziskové instituce sloužící domácnostem atd., pokud tyto jednotky samostatně provádějí agendu nemocenského pojištění.

Statistiky o smrtelných pracovních úrazech jsou sledovány na ekonomických jednotkách, které mají více jak 25 zaměstnanců, tudíž se nejedná o malé organizace.

V následující tabulce (tabulka 2) je úhrn smrtelných úrazů v absolutních číslech podle krajů jednotlivě a za celou ČR. Je patrné, že nejvíce smrtelných úrazů v roce 2008 bylo v Ústeckém kraji a to 27, dále v Praze a Moravskoslezském kraji po 26 případech.

Dalším statistickým výstupem v oblasti smrtelných pracovních úrazů je rozdělení podle velikosti podniku. V další tabulce (tabulka 3) je tato situace vyčíslena, opět se zde

hovoří o absolutních číslech. V prvním sloupci je velikost organizace rozdělena ČSÚ podle počtu zaměstnanců.

Tabulka 2 - Úhrn smrtelných úrazů v krajích a ČR celkem v roce 2008, upraveno s využitím [5]

Kraj	Smrtelné úrazy	
	celkem	ženy
Hl. m. Praha	26	2
Středočeský	14	0
Jihočeský	15	0
Plzeňský	8	0
Karlovarský	2	0
Ústecký	27	0
Liberecký	2	0
Královéhradecký	6	0
Pardubický	13	0
Vysočina	7	0
Jihomoravský	18	2
Olomoucký	7	0
Zlínský	3	0
Moravskoslezský	26	4
ČR celkem	174	8

Tabulka 3 - Smrtelné úrazy dle velikosti organizace v roce 2008, upraveno s využitím [5]

Velikost organizace (dle počtu zaměstnanců)	Smrtelné úrazy	
	celkem	ženy
1 – 49	73	0
50 – 99	20	1
100 – 249	20	0
250 – 499	14	0
500 +	40	7
Celkem	167	8

V tabulce 3 nejsou započítány OSVČ. Smrtelných případů v roce 2008 OSVČ bylo 7.

Další částí statistik je rozdělení smrtelných úrazů podle OKEČ pro jednotlivé kraje a za celou ČR celkově. OKEČ má několik sekcí, subsekcí a oddílů. Sekce OKEČ jsou následující [4]:

- A zemědělství, myslivost, lesnictví,
- B rybolov a chov ryb,
- C těžba nerostných surovin,
- D zpracovatelský průmysl,
- E výroba a rozvod elektřiny, plynu a vody,
- F stavebnictví,
- G obchod; opravy motorových vozidel a výrobků pro osobní potřebu a převážně pro domácnost,
- H ubytování a stravování,
- I doprava, skladování a spoje,
- J finanční zprostředkování,
- K činnosti v oblasti nemovitostí a pronájmu; podnikatelské činnosti,
- L veřejná správa a obrana; povinné sociální zabezpečení,
- M vzdělávání,
- N zdravotní a sociální péče; veterinární činnosti,
- O ostatní veřejné, sociální a osobní služby,
- P činnosti domácností,
- Q exterritoriální organizace a instituce.

Nejvíce subsekcí má sekce D zpracovatelský průmysl a to celkem 14. Každá ze subsekcí má ještě své oddíly.

Následující tabulka (tabulka 4) stanovuje počty smrtelných úrazů v absolutních číslech za celou ČR podle OKEČ.

Z této tabulky je patrné, že nejvíce smrtelných úrazů se stalo ve zpracovatelském průmyslu (sekce D) a následně ve stavebnictví (sekce F).

Tabulka 4 - Smrtelné pracovní úrazy dle OKEČ za celou ČR v roce 2008, upraveno s využitím [5]

OKEČ	Smrtelné úrazy	
	celkem	ženy
A	15	
B	0	
C	5	
D	49	
E	1	
F	46	1
G	13	
H	0	
I	25	5
J	3	2
K	6	
L	4	
M	0	
N	0	
O	7	
Celkem	174	8

Další tabulka (tabulka 5) vypovídá o počtu smrtelných úrazů dle OKEČ a Ústeckého kraje v absolutních číslech, kde podle tabulky 2 bylo v roce 2008 nejvíce smrtelných pracovních úrazů.

Tabulka 5 - Smrtelné pracovní úrazy dle OKEČ za kraj Ústecký v roce 2008, upraveno s využitím [5]

OKEČ	Smrtelné úrazy	
	celkem	ženy
A	1	
B	0	
C	0	
D	15	
E	0	
F	5	
G	0	
H	0	
I	4	
J	0	
K	0	
L	1	
M	0	
N	0	
O	1	
Celkem	27	0

Z tabulky 5 je jasné, že v Ústeckém kraji bylo nejvíce smrtelných pracovních úrazů v sekci D, která představuje zpracovatelský průmysl.

V dalším kraji, Moravskoslezském, kde je dále nejvíce smrtelných úrazů nastává nejvyšší smrtelnost také v sekci D, zpracovatelském průmyslu. Také v Pardubickém kraji zahynulo nejvíce lidí v odvětví zpracovatelského průmyslu a se stejným počtem případů také v sekci stavebnictví. V každém odvětví po čtyřech smrtelných úrazech.

Nejvyšší smrtelná pracovní úrazovost byla v roce 2008 v odvětví stavebnictví (F), a to celkem v 5 krajích (Praha, Středočeský, Jihočeský, Plzeňský, Vysočina, Olomoucký). Dalším odvětvím OKEČ (s důrazem na kraje), nastávají smrtelné pracovní úrazy ve zpracovatelském průmyslu (D), v těžbě nerostných surovin (C), v obchodě a opravárenství motorových vozidel (G).

Následující tabulka (tabulka 6) představuje kraje České republiky a sekci, ve které bylo nejvíce smrtelných pracovních úrazů v daném kraji.

Tabulka 6 - Největší úmrtnost v OKEČ za kraje v roce 2008, upraveno s využitím [5]

Kraj	Smrtelné úrazy	
	sekce	počet případů v sekci
Hl. m. Praha	F	6
Středočeský	D; F	5; 5
Jihočeský	F	6
Plzeňský	F	3
Karlovarský	F; G	1; 1
Ústecký	C	15
Liberecký	G; K	1; 1
Královéhradecký	A; D; F; G; K	1; 1; 1; 1; 1
Pardubický	D; F	4; 4
Vysočina	F	3
Jihomoravský	D	5
Olomoucký	F	4
Zlínský	F; G; I	1; 1; 1
Moravskoslezský	D	8

Další statistikou týkající se problematiky smrtelných pracovních úrazů je úhrn úmrtí podle okresů v krajích ČR. V Pardubickém kraji zahynulo nejvíce lidí v okrese Ústí nad Orlicí a to v počtu 6, dále v okrese Svitavy a Pardubice v počtu 3, následuje okres Chrudim s 1 případem.

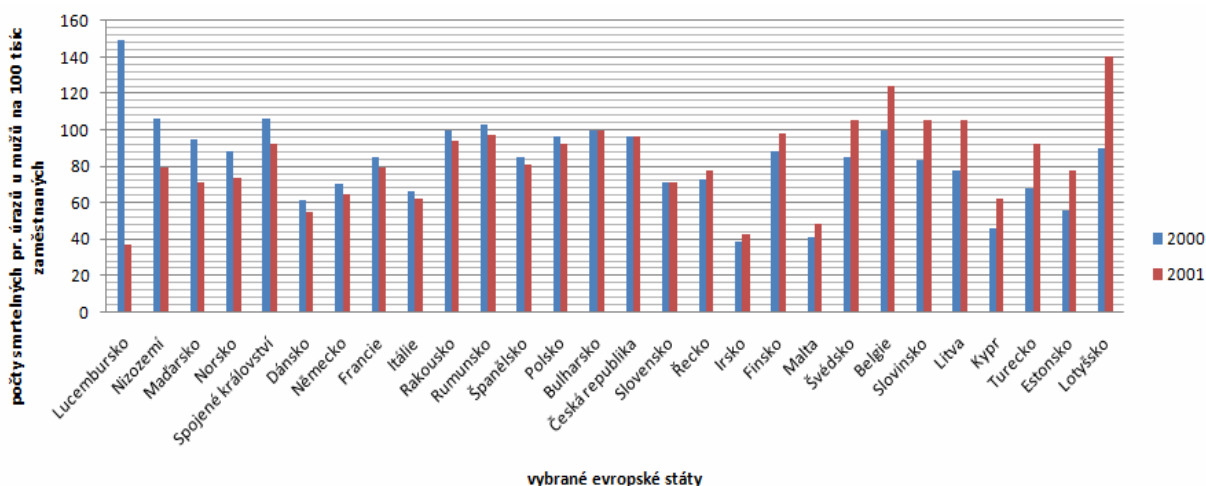
Nejvíce smrtelných pracovních úrazů v ČR (podle okresů) se stalo v okrese Ostrava – město, jednalo se o 11 případů.

4.2 Smrtné pracovní úrazy v Evropě

Smrtné pracovní úrazy trápí všechny země na světě. Bezpečnosti při práci se začíná klást mnohem větší důraz než dříve, tento důraz rok od roku stoupá. Přibývá školení Bezpečnosti a ochrana zdraví při práci, dále jen BOZP, zvyšuje se bezpečnost na pracovištích, ať různými pracovními pomůckami, stroji, ale i zmiňovaným proškolením pracovníků. Nikdo si nepřeje, aby právě na jeho pracovišti vznikl jakýkoli pracovní úraz nebo úraz smrtelný.

I přes všechny snahy snížit smrtné pracovní úrazy ve světě je trend smrtelnosti takový, že se počet smrtelných pracovních úrazů převážně zvyšuje. Je patrné, že v některých státech klesá, ale v převážné většině, někde nepatrně a někde skokem, stoupá. Podle oficiálních údajů Mezinárodní organizace práce zemře ročně na následky pracovních úrazů a nemocí z povolání přes dva miliony lidí. Statistiky ukazují, že každý rok dojde na světě ke 270 milionům pracovních úrazů.

Následující obrázek (obrázek 1) znázorňuje porovnání smrtelných pracovních úrazů v letech 2000 – 2001 ve vybraných evropských zemích. Udávané počty smrtelných pracovních úrazů jsou u mužů a na 100 tisíc zaměstnaných osob. Z obrázku je patrné, ve kterých zemích došlo k růstu a naopak ve kterých došlo k poklesu pracovních úrazů. Uvedené hodnoty smrtelných úrazů jsou absolutní čísla.



Obrázek 1 - Meziroční nárůsty smrtelných pracovních úrazů v Evropě za roky 2000 a 2001, upraveno s využitím [6]

Následující obrázek (obrázek 2) vyjadřuje procentuální nárůsty či poklesy smrtelných pracovních úrazů v evropských státech v letech 2000 a 2001. Hodnoty jsou srovnány podle velikosti od nejvyššího poklesu po nejvyšší nárůst.

Pokles smrtelných úrazů o 75 % nastal v Lucembursku, z původních 149 na 37 případů. Dalším příkladným státem, kde smrtelná úrazovost meziročně poklesla, je Nizozemí a Maďarsko. V Nizozemí klesla smrtelná úrazovost o 25 % a v Maďarsku také o 25 %.

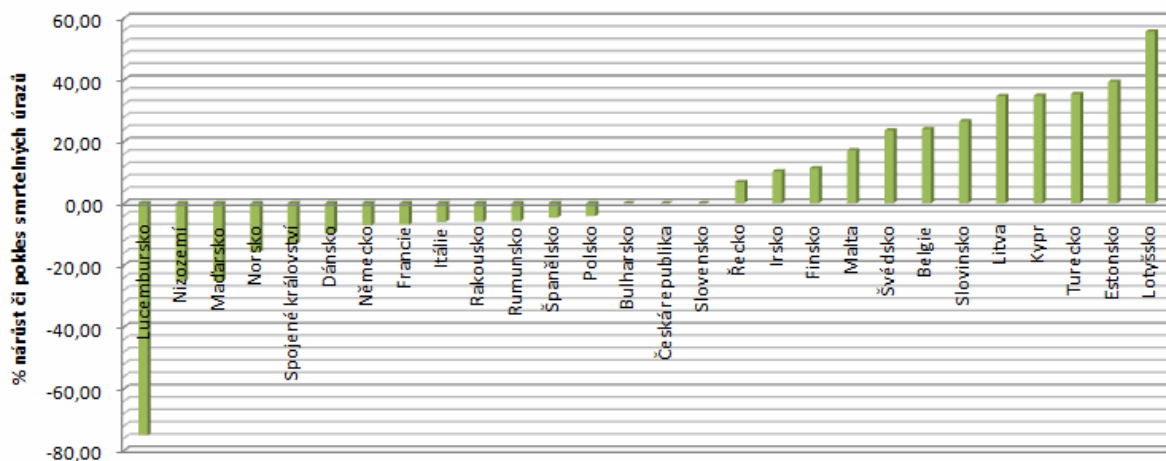
Další státy, ve kterých poklesl počet smrtelných pracovních úrazů, jsou [6]:

- Norsko (16 %),
- Spojené království (13 %),
- Dánsko (10 %),
- Německo (7 %),
- Francie (7 %),
- Itálie (6 %),
- Rakousko (6 %),
- Rumunsko (6 %),
- Španělsko (5 %),
- Polsko (4 %).

V opačném případě, kdy došlo k nárůstu smrtelných úrazů, je na špici žebříčku Lotyšsko, kde se počet smrtelných pracovních úrazů zvýšil o 50, což je 56 %.

Dalšími státy, kde meziročně velmi stoupla smrtelná pracovní úrazovost, jsou [6]:

- Estonsko (39 %),
- Turecko (35 %),
- Kypr (35 %),
- Litva (35 %),
- Slovinsko (26 %),
- Belgie (24 %),
- Švédsko (24 %),
- Malta (17 %),
- Finsko (11 %),
- Irsko (11 %),
- Řecko (7 %).



Obrázek 2 – Procentuální nárůst či pokles smrtelných pracovních úrazů v roce 2001 oproti roku 2000 ve vybraných evropských státech, upraveno s využitím [6]

V Bulharsku, České republice a na Slovensku byla v letech 2000 a 2001 stejná smrtelná úrazovost. V Bulharsku se jednalo o 100 případů, v České republice o 96 případů a na Slovensku o 71 případů smrtelných pracovních úrazů.

4.3 Shrnutí kapitoly

Byla nastíněna situace smrtelných pracovních úrazů v České republice za rok 2008, kdy došlo k celkem 174 případům. Největší pracovní úmrtnost byla v Ústeckém kraji, kde zahynulo 27 pracovníků. Nejvíce pracujících zemřelo, dle OKEČ, ve zpracovatelském průmyslu a stavebnictví. Co se týče velikosti organizace, tak nejvíce rizikové byly v roce 2008 podniky s počtem pracovníků 0 – 49.

Statistika smrtelných pracovních úrazů ve světě byla provedena za roky 2000 – 2001, kde meziročně nejvyšší nárůst případů mělo Lotyšsko a meziročně nejlepší pokles Lucembursko.

5 Data mining a metodologie CRISP-DM

Dnešní společnost lze charakterizovat jako informační společnost. Každý má doma počítač a bez něj si nedokáže představit den. Mnoho lidí využívá i ke své profesi výpočetní techniky a zpracovává na ní určitá data, vyhledává a analyzuje. Výpočetní technika nás neoddelitelně obklopuje a je jisté, že bez ní už to v žádném případě nepůjde. Je možné říci, že nám ulehčuje a usnadňuje práci.

Většina lidí, uživatelů výpočetní techniky, zpracovává a analyzuje data, aniž by o tom vlastně věděla. Představme si pouhou domácnost, která si zpracovává tabulku příjmů a

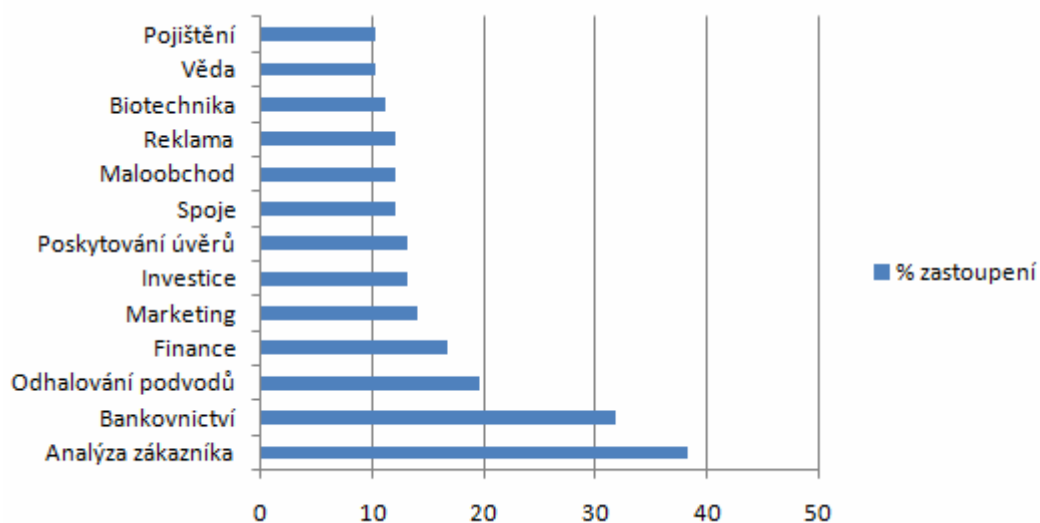
výdajů. Z těchto nasbíraných dat, která se průběžně do tabulky zapisují, potom domácnost hodnotí, např. kolik procent příjmů utratí, kde by mohla snížit náklady, vypracovat si graf výdajů a příjmů za jednotlivé měsíce, získat informace proč daný měsíc byly vyšší výdaje než měsíc předchozí. Možností je spousta. Právě v tento okamžik začínají „dolovat“, i když na nízké uživatelské úrovni, informace z nashromážděných dat. I v tomto případě se dá hovořit o data miningu.

5.1 Data mining

Data mining (DM) můžeme charakterizovat jako netriviální získávání implicitních, dříve neznámých a potenciaálně užitečných informací z dat [17]. Představuje proces, při kterém se shromažďují a vyhodnocují data. Tento pojem je z oblasti Business Intelligence a cílem je průběžné shromažďování co možná největšího počtu dat a jejich následné analýzy, která může pomoci najít nové trendy, na základě nichž lze dělat důležitá marketingová a obchodní rozhodnutí. Napomáhá také objevovat v datech strategické informace.

Dnes se můžeme s DM setkat v mnoha oborech a odvětvích. Do těchto oborů, které pracují s DM, patří telekomunikace, vědecký výzkum, pojišťovnictví, bankovníctví, cestovní ruch, farmaceutický průmysl, energetika, řízení lidských zdrojů, veřejná správa a mnoho dalších. Největší využití technik DM je patrné v oblasti bankovníctví, analýze zákazníků a odhalování podvodů [12].

Následující obrázek (obrázek 3) vypovídá o tom, v kterých odvětvích se nejvíce DM technik využívá.



Obrázek 3 - Oblasti s nejvyšším využitím technik DM, upraveno s využitím [12]

DM projekt je proces, který vyžaduje značné množství zdrojů. Od lidských zdrojů, přes hmotné a datové zdroje až po softwarové. Avšak společným jmenovatelem jim jsou peněžní prostředky.

Mezi základní matematické a statistické techniky, na kterých je DM založen, jsou například [17]:

- rozhodovací stromy,
- neuronové sítě,
- genetické algoritmy,
- logistická regrese,
- seskupovací analýza a klasifikace,
- korelační analýza,
- faktorová analýza.

Seskupovací analýza a klasifikace

Seskupovací analýza je technika, která slouží k rozdělení dat do skupin s obdobnými charakteristikami. Klasifikace definuje podstatné atributy skupin v podobě klasifikačních kritérií. Umožňují identifikovat a charakterizovat různé segmenty v datech [17].

Klasifikaci rozumíme rozklad množiny na třídy. Přitom každou z těchto tříd můžeme považovat za výchozí množinu objektů a v ní rovněž provádět klasifikaci [20].

Protože seskupovací analýza má sloužit jako prostředek k získávání klasifikace, nabízí se možnost rozlišovat shlukovací metody nikoli dle matematických prostředků, ale podle cílů, k nimž směřují. Budeme tedy rozlišovat hierarchické a nehierarchické metody seskupovací analýzy [20].

Hierarchický způsob shlukování se dále dělí na [20]:

- aglomerativní: vychází z jednotlivých objektů a jejich postupným seskupováním docházím až ke konečnému stavu – spojení všech objektů do jedné množiny,
- divizní: shlukování prováděno tak, že se vychází z množiny objektů určených ke klasifikaci jako celku a jejím postupným rozdělováním se získává hierarchický systém podmnožin.

Nehierarchický způsob shlukování se dále dělí na [20]:

- optimalizační: hledá takový rozklad množiny objektů určených ke klasifikaci, který je optimální podle vhodně zvoleného kritéria optimality rozkladu,
- analýza modelů: využívá pravděpodobnostní přístup.

Shluková analýza slouží především jako prostředek generování hypotéz o klasifikaci objektů nebo znaků.

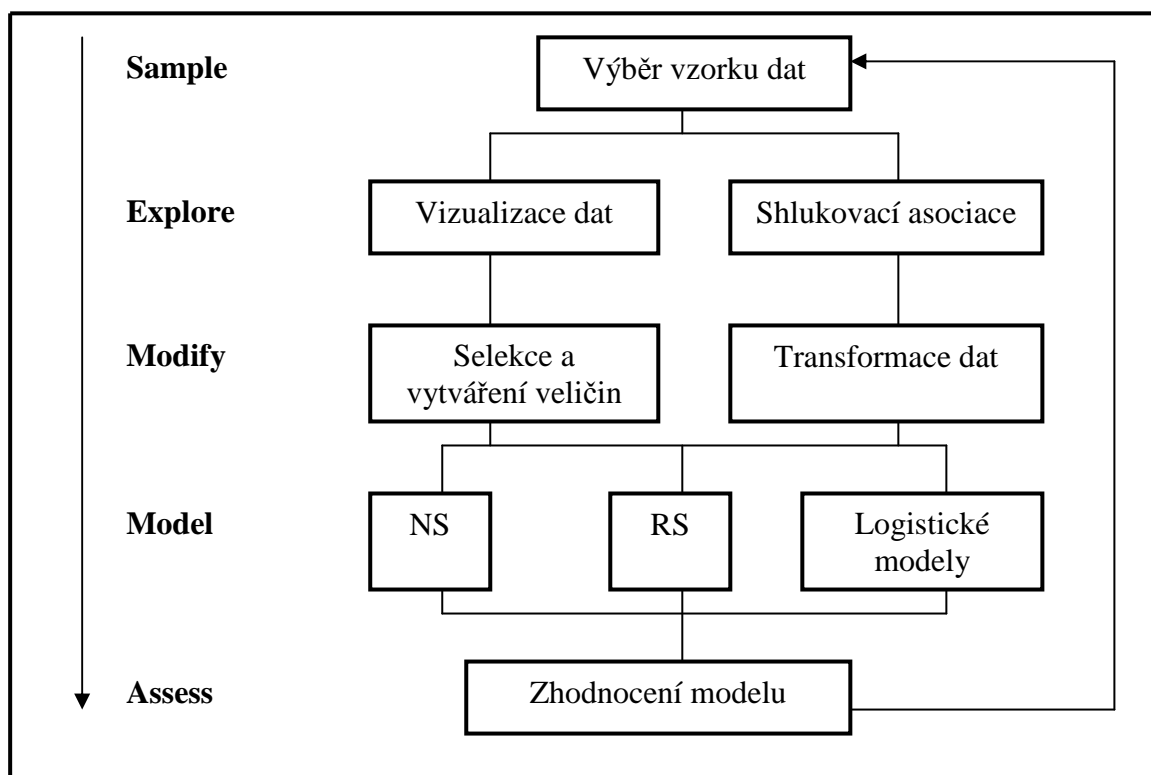
5.2 CRISP-DM

Jelikož DM zahrnuje velkou šíři způsobů práce a metod, je velmi obtížné ukázat a nastínit jednoznačný návod k postupu. Ale i přes všechnu velkou škálu metod a způsobů práce během 90. let vykrytalizovaly dvě obecné metodologie, které alespoň v hrubých rysech popisují jednotlivé kroky. Jsou to metodologie SEMMA a CRISP-DM.

Metodologie SEMMA se skládá z pěti, po sobě jdoucích, kroků, jimiž jsou [1]:

- sample – vybírání vhodných objektů,
- explore – vizuální explorace a vizualizace dat,
- modify – seskupování objektů a hodnot atributů, datové transformace,
- model – analýza dat,
- assess – porovnání modelů a interpretace.

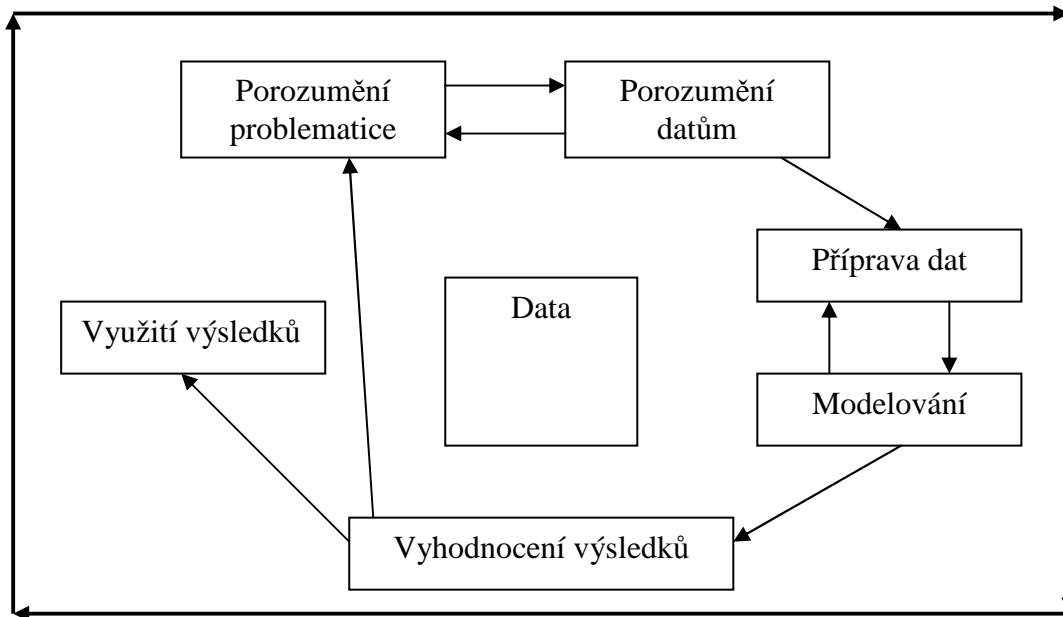
Následující obrázek (obrázek 4) vypovídá o pěti krocích metodologie SEMMA.



Obrázek 4 – SEMMA, upraveno s využitím [1]

Zkratka CRISP-DM znamená Cross-Industry Standard Proces for Data Mining. CRISP je souhrnná DM metodologie. Její model nabízí návody krok po kroku, úkoly a cíle pro každou část celého procesu. CRISP-DM umožňuje provádět rozsáhlé DM projekty rychleji, efektivněji a méně nákladně prostřednictvím osvědčených postupů. Tento model pomáhá vyhnout se běžným chybám [14].

Další obrázek (obrázek 5) charakterizuje metodologii CRISP-DM.



Obrázek 5 - CRISP-DM, upraveno s využitím [1]

Metodologie CRISP-DM byla využita pro celou práci. Jednotlivé etapy CRISP-DM nejlépe odpovídaly práci s daty a jejich následné modelování. Software Clementine pracuje s využitím této metodologie, což bylo také dalším důvodem výběru CRISP-DM.

Metodologie CRISP-DM rozděluje celý proces DM projektu do šesti po sobě jdoucích základních etap, v rámci nichž dále rozlišuje další kroky. Základními etapami jsou [9]:

1. porozumění problému (Business understanding),
2. porozumění datům (Data understanding),
3. příprava dat (Data preparation),
4. modelování (Modeling),
5. vyhodnocení výsledků (Evaluation),
6. implementace vytvořeného modelu, využití v praxi (Deployment).

1. Porozumění problému

První (vstupní) fáze je zaměřena na definování a pochopení cílů projektu a požadavků z manažerského a obchodního hlediska. Poté se musí manažerské požadavky převést do zadání úlohy pro dobývání znalostí z databází a předběžně navrhnout plán, jak cílů dosáhnout [17].

Při definování cíle je též zapotřebí definovat kritéria, z podnikatelského hlediska, pro hodnocení užitečnosti či úspěšnosti výstupu projektu. Tato kritéria mohou mít dvě odlišné formy, mohou být objektivně měřitelná nebo subjektivně vnímatelná [17].

Obchodní cíl

Pro pracovníky Výzkumného úřadu pro BOZP znamenají kontroly BOZP velkou zátěž. Nejedná se pouze o zátěž časovou, ale také finanční zátěž. Proto je třeba s těmito lidskými a finančními zdroji na poli dodržování BOZP zacházet efektivně. Efektivně v tomto případě znamená správně vybrat podnik, kam pracovníky na tyto kontroly vysílat. Je velmi důležité zaměřit se s kontrolami na nejrizikovější oblasti a skupiny.

Výchozí situace

V oblasti BOZP lze pozorovat několik situací. Těmito situacemi jsou nedostatek lidských zdrojů pro všechny naplánované kontroly a nedostatek finančních zdrojů pro provedení všech kontrol. Proto je nutné vybrat opravdu ty nejrizikovější ekonomické subjekty pro dané kontroly.

V zájmu Výzkumného úřadu BOZP je snížení nákladů, ale i kontrol, které jsou ve výsledku zbytečné.

Výzkumný ústav poskytl data o smrtelné pracovní úrazovosti v letech 2002 – 2007 v souboru s formátem .xls.

Náklady a výnosy projektu

V případě, že kontrolní pracovník provádí analýzu a kontrolu v odvětví, které je evidentně bezúrazové, je tato kontrola zbytečná, tudíž i náklady na mzdu tohoto pracovníka zbytečně vynaložené.

Výnosy projektu jsou tedy náklady, které se ušetří na zbytečných a neúčelných kontrolách BOZP.

Cíl projektu DM

Z hlediska DM je cílem projektu nalézt určitou charakteristiku ekonomického subjektu (firmy). Pomocí této charakteristiky by bylo možné včas rozpoznat firmu, kde by k smrtelnému případu s největší pravděpodobností došlo. Do této charakteristiky by mohla patřit velikost podniku, kategorie OKEČ, odpracované roky zaměstnance, aj.

Plán projektu

Plán projektu je rozdělen do několika následujících fází [17]:

a) Popis dat:

- vytvoření srozumitelných číselníků,
- popis datového souboru pomocí analýzy v prostředí MS Excel,
- popsání kvality dat.

b) Příprava dat:

- výběr relevantních a použitelných dat,
- odvozování nových proměnných v prostředí MS Excel,
- určení formátu dat v prostředí MS Excel,
- vytvoření souboru s upravenými daty.

c) Modelování

Pro modelování dat byla zvolena technika seskupovací analýzy. Fáze modelování má tyto kroky [17]:

- volba modelovacích technik,
- export dat do software Clementine,
- vytvoření modelů danými modelovacími technikami,
- posouzení vytvořených modelů.

d) Hodnocení

Obecně je dáno několik požadavků na modely a jsou to jednoduchost, rychlost, snadná pochopitelnost, nenákladnost a přiměřená přesnost. Hodnocení modelu je založené na tom, zda splňuje dané požadavky s důrazem na požadavek přesnost.

V této části se ale také hodnotí dosažené výsledky z pohledu manažerů, zda byly splněny cíle zadané na počátku.

e) Uvedení do praxe

Takto vytvořené modely může pracovník Výzkumného ústavu pro BOZP využít v praxi. Na základě údajů o předchozí smrtelné pracovní úrazovosti a výsledků z ní, bude moci pracovník určit, na které odvětví se s kontrolami zaměřit.

Délka jednotlivých fází

Jednoznačně časově nejnáročnější fází je příprava dat, v procentuálním vyjádření se jedná přibližně o 55 % z celkového času. Pro fázi porozumění problému by mělo stačit 5 % času, další fáze projektu, porozumění datům, zabere přibližně 10 % času projektu. 15 % času zabere modelování a dalších 15 % případně na sepsání závěrečné zprávy.

Požadované zdroje

Mezi požadované zdroje budeme řadit:

- datové zdroje, které poskytne Výzkumný ústav pro BOZP,
- lidské zdroje (pracovníci Výzkumného ústavu, s nimiž budou prováděny konzultace ohledně problému, ale také ohledně dat),
- číselníky potřebné pro identifikaci.

Programové vybavení

Mezi programové vybavení patří nástroje, kterými budou data zpracována. Jedná se o softwarové vybavení tohoto typu:

- MS Excel 2007,
- SPSS Clementine 11.1.

2. Porozumění datům

Druhá fáze, porozumění datům, začíná prvotním převzetím dat. Po převzetí dat, následují činnosti, které umožní získat základní představu o těchto datech a jejich pochopení.

V této části se také provádí zhruba průzkum dat. Tato analýza je zaměřena na DM otázky, které mohou být zodpovězeny použitím dotazů, vizualizací a reportů. Toto zahrnuje rozdělení klíčových atributů, vazby mezi páry, jednoduché statistické analýzy, četnosti hodnot různých atributů, průměrné hodnoty, minima a maxima apod. Tyto analýzy se mohou zaměřit přímo na cíl DM projektu a sloužit tak pro formulaci hypotéz, nebo pouze přispívat k popisu dat. Pokud je to vhodné, tak mohou být součástí diagramy a grafy, které poukazují na datové charakteristiky nebo které mohou být určitým vodítkem k zajímavým podskupinám v datech [17].

Převzetí prvotních surových dat

Data byla obdržena v elektronické podobě ve formátu .xls s obsahem dvou listů. Jedná se o data popisující průběh smrtelných pracovních úrazů v letech 2002 – 2007. Na prvním listu datového souboru byla surová data. Na druhém datovém listě byla legenda. Následující obrázek (obrázek 6) nastiňuje, v jaké podobě byla data převzata.

	A	B	C	D	E	F	G	H	I	J
1	8	1	1	2 A	-		1	2	1	CZ
2	4	1	0	1 A	-		2	0	2	CZ
3	4	1	2	1 D	H		25	2	1	CZ
4	6	1	1	2 D	J		28	1	1	CZ
5	7	1	0	1 D	L		31	2	0	CZ
6	5	1	2	1 D	M		34	3	0	CZ
7	7	1	1	2 F	-		45	2	1	CZ
8	6	1	0	1 G	-		50	2	0	CZ
9	6	1	0	1 G	-		50	2	0	CZ
10	10	1	1	2 G	-		51	6	2	CZ
11	4	1	1	2 G	-		52	1	2	CZ
12	4	1	1	3 H	-		55	3	0	CZ
13	10	1	1	2 I	-		60	2	4	CZ
14	8	1	0	1 I	-		60	2	4	CZ
15	40	1	0	1 I	-		60	2	3	CZ
16	6	1	2	1 J	-		65	1	2	CZ
17	6	2	3	1 K	-		70	3	1	CZ
18	5	1	2	1 K	-		74	8	4	CZ
19	7	1	1	2 K	-		74	3	0	CZ
20	7	8	0	1 L	-		75	1	1	CZ
21	7	8	0	1 L	-		75	1	1	CZ
22	9	8	0	1 L	-		75	1	1	CZ
23	40	8	0	1 L	-		75	1	1	CZ
24	9	1	2	1 N	-		85	1	2	CZ
25	3	1	1	2 D	J		28	6	3	CZ
26	3	1	1	2 D	J		28	6	3	CZ
27	3	1	1	2 D	J		28	6	3	CZ
28	10	1	0	3 A	-		1	3	0	CZ

Obrázek 6 - Prvotní data neupravená [zdroj: vlastní]

Data vyjadřují smrtelnou pracovní úrazovost v České republice v letech 2002 – 2007 na všech pracovních pozicích dle OKEČ.

Ve zdrojovém souboru dat se nachází celkem 1092 případů o 103 proměnných, které identifikují určitý smrtelný pracovní úraz. Ke zdrojovým datům byl vytvořen datový slovník, který tvoří přílohu č. 1.

Legendu, získanou spolu s daty, bylo nutno přidat do zdrojových dat, aby bylo možné provést analýzu dat a bylo pochopitelné, co který sloupec vyjadřuje.

Obrázek (obrázek 7) vyjadřuje prvotní neupravená data doplněná o legendu. Takto upravená data již mají malou vypovídací schopnost. Tato malá vypovídací schopnost je dána tím, že pro úplnou vypovídací schopnost je nutné využít daný číselník.

	A	B	C	D	E	F	G	H	I	J
1	OrganDozoru	PravniForma_1	PravniForma_2	PravniForma_3	katOKEC1	katOKEC2	oddilOKEC3	OKEC_4	OKEC_5	UzemniJednotka_1
2	8	1	1	2	A	-	1	2	1	CZ
3	4	1	0	1	A	-	2	0	2	CZ
4	4	1	2	1	D	H	25	2	1	CZ
5	6	1	1	2	D	J	28	1	1	CZ
6	7	1	0	1	D	L	31	2	0	CZ
7	5	1	2	1	D	M	34	3	0	CZ
8	7	1	1	2	F	-	45	2	1	CZ
9	6	1	0	1	G	-	50	2	0	CZ
10	6	1	0	1	G	-	50	2	0	CZ
11	10	1	1	2	G	-	51	6	2	CZ
12	4	1	1	2	G	-	52	1	2	CZ
13	4	1	1	3	H	-	55	3	0	CZ
14	10	1	1	2	I	-	60	2	4	CZ
15	8	1	0	1	I	-	60	2	4	CZ
16	40	1	0	1	I	-	60	2	3	CZ
17	6	1	2	1	J	-	65	1	2	CZ
18	6	2	3	1	K	-	70	3	1	CZ
19	5	1	2	1	K	-	74	8	4	CZ
20	7	1	1	2	K	-	74	3	0	CZ
21	7	8	0	1	L	-	75	1	1	CZ
22	7	8	0	1	L	-	75	1	1	CZ
23	9	8	0	1	L	-	75	1	1	CZ
24	40	8	0	1	L	-	75	1	1	CZ
25	9	1	2	1	N	-	85	1	2	CZ
26	3	1	1	2	D	J	28	6	3	CZ
27	3	1	1	2	D	J	28	6	3	CZ
28	2	1	1	2	D	J	28	6	3	CZ

Obrázek 7 - Prvotní data doplněná o názvy sloupců (atributů) [zdroj: vlastní]



Popis a analýza dat

Popis a analýza dat byla provedena jak v prostředí MS Excel, tak v Clementine s využitím výstupního modulu Data audit. Pouze u některých atributů bylo možno zpracovat prvotní analýzu z důvodu absence daných číselníků.

Atribut Sex

Z dat bylo možné vyjádřit počet smrtelných pracovních úrazů u žen a mužů. Počet smrtelných úrazů u žen bylo v letech 2002 – 2007 celkem 56 a smrtelných pracovních úrazů u mužů bylo v těchto letech celkem 1036. Hodnota 0 představuje pohlaví muž a hodnota 1 pohlaví žena.

Situaci ohledně počtu mužů a žen, kteří zahynuli v důsledku smrtelného pracovního úrazu, naznačuje následující obrázek (obrázek 8). Procentuální vyjádření v následujícím obrázku se vztahuje k celkovému počtu smrtelných pracovních úrazů v letech 2002 – 2007.

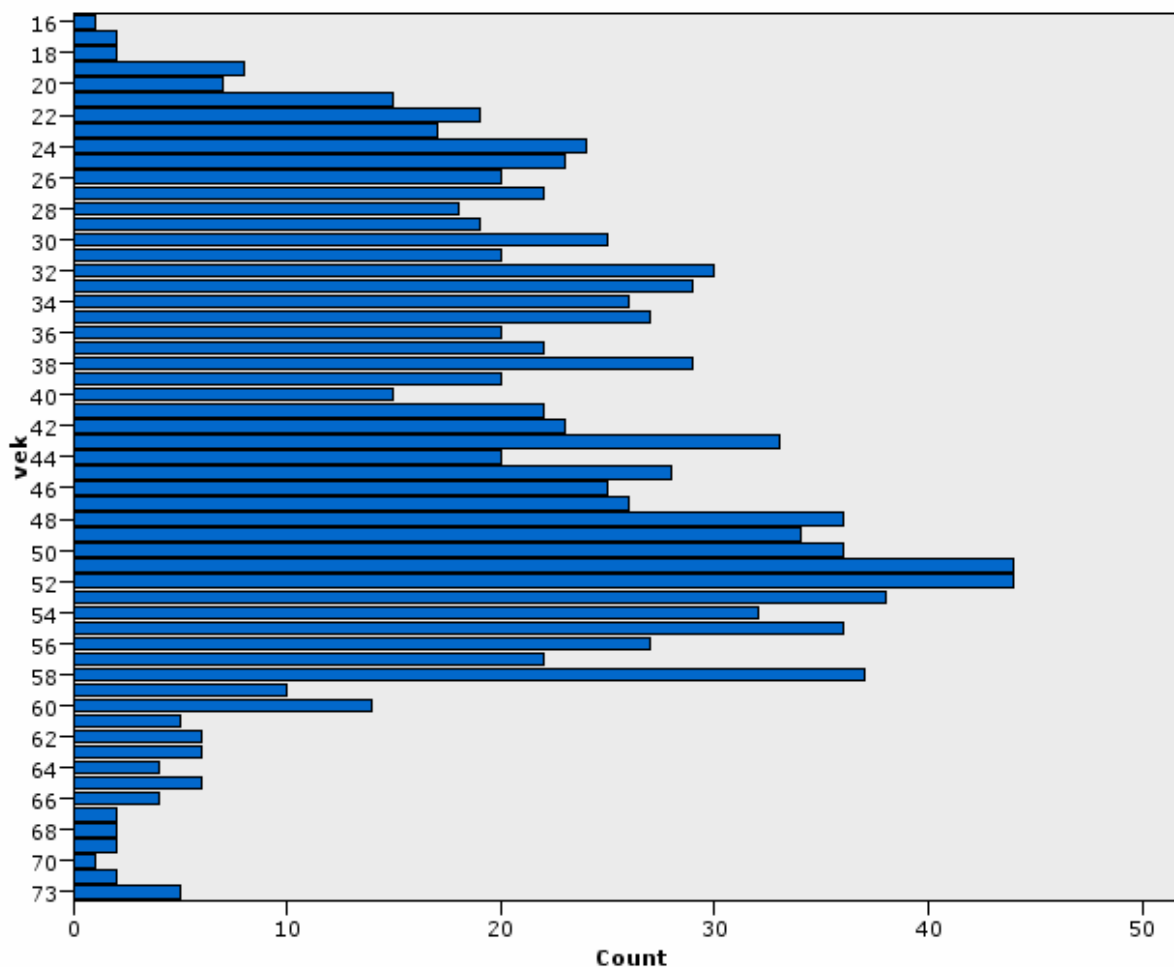
Value ▲	Proportion	%	Count
0		94,87	1036
1		5,13	56

Obrázek 8 - Počty a procentuální vyjádření smrtelných pracovních úrazů v letech 2002 - 2007, [zdroj: vlastní]

V letech 2002 – 2007 bylo průměrně zaměstnáno v každém z let celkem 2 737 800 mužů a 2 080 600 žen. Průměrný počet smrtelných pracovních úrazů v letech 2002 – 2007 byl 182 případů. Průměrné počty smrtelných pracovních úrazů za dané roky lze procentuálně vyjádřit na celkovou zaměstnanost. V letech 2002 – 2007 utrpělo smrtelný pracovní úraz 0,038 % zaměstnaných mužů a 0,0027 % žen.

Atribut Věk

Dále bylo možno zjistit průměrný věk, ve kterém docházelo k smrtelným úrazům. Tento věk byl 43 let. Z věku bylo také možno zjistit minimální a maximální hodnotu. Tyto hodnoty věku byly min. 16 let věku a max. 73 let věku. Tuto situaci popisuje následující obrázek (obrázek 9).



Obrázek 9 - Celkový počet smrtelných pracovních úrazů u atributu Věk za roky 2002 - 2007, [zdroj:vlastní]

Je patrné, že nejvíce smrtelných pracovních úrazů se stalo zaměstnancům s věkem 51 a 52, v každé věkové skupině po 44 případech. Další věkovou skupinou s nejvyšším počtem úmrtí je věk 53 s 38 případy a věk 58 s 37 případy. Mezi 48 a 55 rokem věku je počet smrtelných pracovních úrazů nad 30 případů. Ostatní roky života, s výjimkou věku 58, mají počty smrtelných úrazů pod 30 případů. Nejméně smrtelných úrazů nastává po 59 roce věku. V žádném případě nepřesáhl hranici deseti případů.

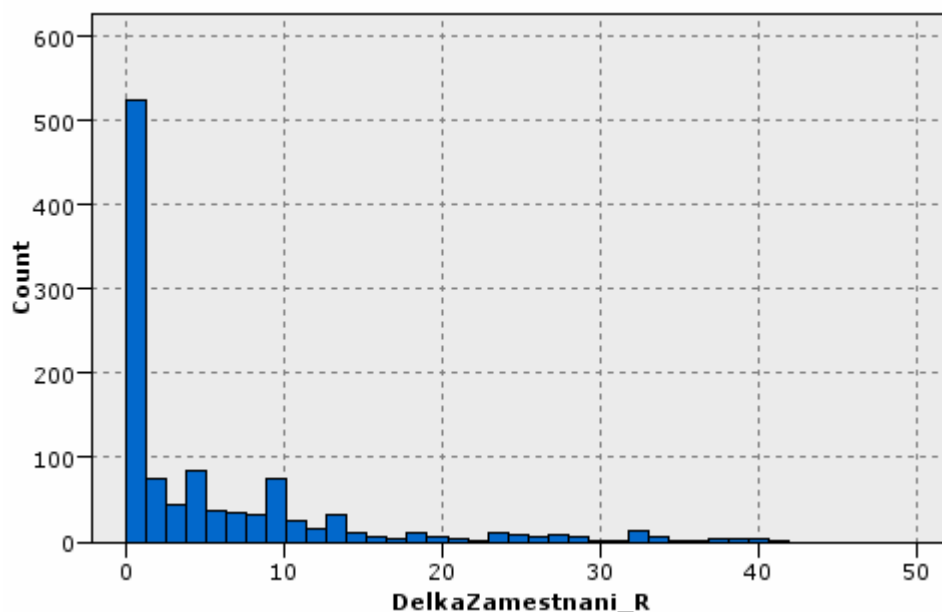
Bylo by dobré se s kontrolami zaměřit na ekonomické subjekty se zaměstnanci okolo 50 let věku.

Atribut DelkaZamestnani_R

Atribut DelkaZamestnani_R vyjadřuje, jak dlouho byl pracovník zaměstnán, když nastal smrtelný pracovní úraz.

Bylo možno určit délku pracovního vztahu v letech u dané organizace. Nejkratší délka pracovního vztahu byl necelý rok. Nejdelší doba pracovního vztahu byla 42 let.

Z následujícího obrázku (obrázek 10), lze vyčíst délku zaměstnání v letech a počet smrtelných pracovních úrazů.



Obrázek 10 - Délka zaměstnání v letech a jejich četnost pro roky 2002 – 2007, [zdroj: vlastní]

Při pracovním vztahu trvajícím kratším než jeden rok se stalo celkem 407 smrtelných pracovních úrazů, což je z celkového počtu 37,27 %. Další velkou skupinou je pracovní vztah kratší než 2 roky. Zde se stalo 117 případů, to představuje 10,71 %.

Můžeme tedy říci, že nejvíce rizikovou skupinou je délka zaměstnání do 1 roku. Může to být způsobeno několika faktory, fluktuací, nedostatečnou kvalifikací, nedostatečným školením BOZP, nízkou zkušeností zaměstnanců. Tyto nedostatky je potřeba ve firmách eliminovat např. častějšími školeními, lepšími pracovními pomůckami či motivování zaměstnanců, aby se zamezilo fluktuaci.

Další rizikovou skupinou je délka pracovního vztahu od dvou do deseti let. V této skupině se stává nejvíce smrtelných pracovních úrazů. Může to být dáno malou zkušeností zaměstnanců, kteří se v této skupině nachází.

Je patrné, že čím větší počet let pracovního vztahu, tím je nižší počet smrtelných pracovních úrazů. To potvrzuje tvrzení, že při prvních deseti letech pracovního vztahu nemají zaměstnanci ještě velmi dostatečné znalosti a zkušenosti.

Tímto je tedy dán fakt, že kontroly BOZP by měly směřovat do ekonomických subjektů s vyšší fluktuací.

Atribut StatniPrislusnost

Atribut StatniPrislusnost udává, jaké státní příslušnosti byl člověk, který zemřel. Převážně se jedná o občany České republiky, ale jsou zde také případy občanů Slovenské republiky v počtu 49 případů, dále 17 případů polských občanů zaměstnaných na území ČR, ve 23 případech se jednalo o osoby ukrajinské státní příslušnosti, ve 4 případech o bulharské pracovníky a po jednom případě o Bělorusa a Rusa.

Tento fakt nastiňuje následující tabulka (obrázek 11), kde jsou počty přepočítané na procenta. Jednoznačně nejvíce jsou zastoupeni čeští občané.

Value ▲	Proportion	%	Count
BG		0,37	4
BY		0,09	1
CZ		91,3	997
PL		1,56	17
RU		0,09	1
SK		4,49	49
UA		2,11	23

Obrázek 11 - Počty úmrtí dle státní příslušnosti v letech 2002 – 2007, [zdroj: vlastní]

V České republice bylo v letech 2002 – 2007 průměrně zaměstnáno celkem 4 818 400 lidí, v průměru se stalo 166 smrtelných pracovních úrazů, což představuje 0,0035 %.

V letech 2002 – 2007 u nás bylo v průměru zaměstnáno [upraveno s využitím [3]]:

- 60 624 slovenských pracovníků,
- 11 185 polských pracovníků,
- 48 112 ukrajinských pracovníků,
- 2 884 bulharských pracovníků,
- 1 387 běloruských pracovníků,
- 3 023 ruských pracovníků.

Za rok v České republice zahyne při pracovním úrazu průměrně 16 cizinců.

Při analýze smrtelných pracovních úrazů pouze u cizinců v ČR je procentuální zastoupení národností takovýto [3]:

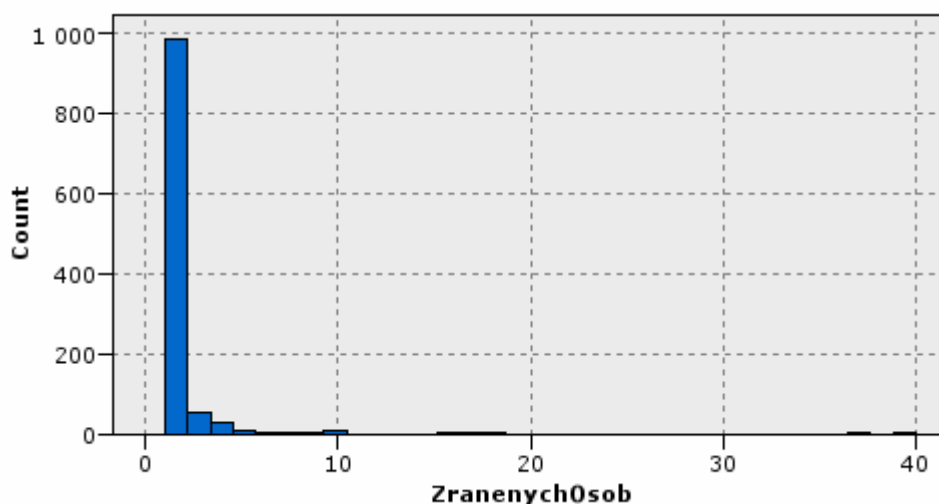
- 0,52 % slovenských pracovníků,
- 0,18 % polských pracovníků,
- 0,24 % ukrajinských pracovníků,
- 0,044 % bulharských pracovníků,
- 0,013 % běloruských pracovníků,
- 0,013 % ruských pracovníků.

Při tomto zjištění by bylo dobré kontroly bezpečnosti práce cílit na podniky zaměstnávající slovenské a ukrajinské pracovníky.

Atribut ZranenychOsob

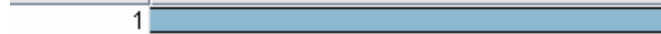

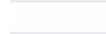

Atribut ZranenychOsob vypovídá o celkovém počtu zraněných osob u jednotlivých smrtelných pracovních úrazů.

Jedná se o počty od 1 do 40 zraněných osob, přičemž alespoň u jedné z osob nastal smrtelný pracovní úraz. Následující obrázek (obrázek 12) naznačuje počty zraněných osob a jejich četnost.



Obrázek 12 - Počet zraněných osob a jejich četnost v letech 2002 - 2007, [zdroj: vlastní]

Stejně tak tuto skutečnost, ale také s procentuálním vyjádřením, ukazuje další obrázek (obrázek 13).

Value ▲	Proportion	%	Count
1		79,21	865
2		10,99	120
3		4,67	51
4		2,38	26
5]		0,73	8
6]		0,46	5
7]		0,09	1
8]		0,27	3
9]		0,09	1
10]		0,64	7
16]		0,09	1
17]		0,09	1
18]		0,09	1
37]		0,09	1
40]		0,09	1

Obrázek 13 - Počet zraněných osob u smrtelného pracovního úrazu v letech 2002 - 2007, [zdroj: vlastní]

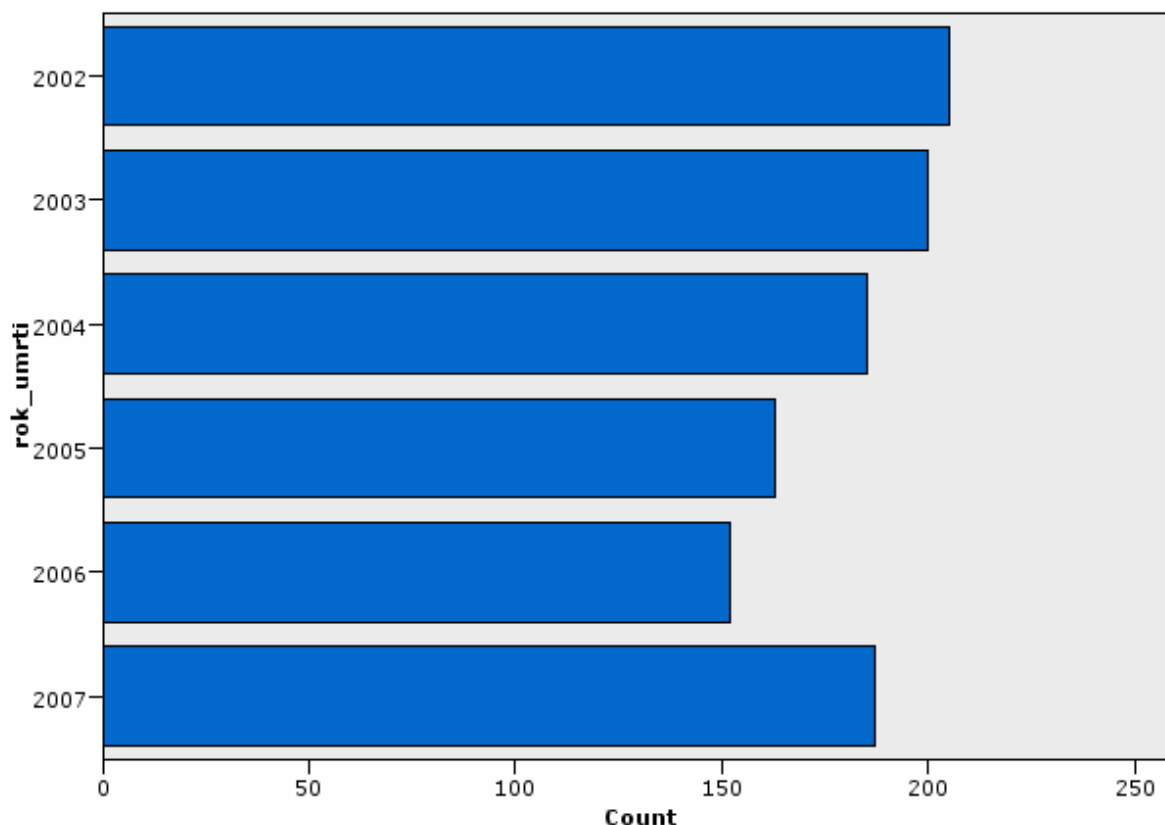
Lze říci, že v 80 % případů smrtelných pracovních úrazů se jednalo pouze o jednu zraněnou osobu, která však těmto zraněním podlehla. Ovšem jsou i případy, kdy zraněných osob je více, ale smrtelný pracovní úraz nastane pouze u některých z nich. Pro případ uveďme požár na pracovišti, kde zraněných osob je mnoho, ale ne všichni zemřou.

V jednom případě, kdy zraněných osob bylo celkem 40, se jednalo o dopravní nehodu autobusu při dálkové přepravě osob. Řidič autobusu chybně odhadl vzdálenost překážky a do této narazil.

Atribut RokUmrti

Co se týče roku a počtu případů, tak v roce 2002 se stalo 205 případů. V roce 2003 se jednalo o 200 případů. V dalším roce, v roce 2004, nastalo 185 smrtelných pracovních úrazů. V roce 2005 zahynulo celkem 163 pracovníků, dále potom v roce 2006 152 úmrtí a v roce 2007 se jednalo o 187 smrtelných pracovních úrazů. Můžeme tedy říci, že od roku 2002 počet smrtelných pracovních úrazů klesal až do roku 2006, avšak v roce 2007 opět vzrost a to o 35 případů.

V následujícím obrázku (obrázek 14) je zachycen rok úmrtí a četnost smrtelných úrazů vztahujících se k rokům 2002 - 2007.



Obrázek 14 - Četnost smrtelných úrazů v daném roce v letech 2002 - 2007, [zdroj: vlastní]

Následující obrázek (obrázek 15) zachycuje počet úmrtí v daných letech 2002 až 2007 i s procentuálním vyjádřením.

Value	Proportion	%	Count
2002		18,77	205
2003		18,32	200
2004		16,94	185
2005		14,93	163
2006		13,92	152
2007		17,12	187

Obrázek 15 - Četnost smrtelných úrazů v letech 2008 - 2007, [zdroj: vlastní]

Tento atribut obsahuje i dva záznamy, které vyjadřují, že k úrazu došlo ještě v roce 2001, ale úmrtí nastalo až v roce 2002.

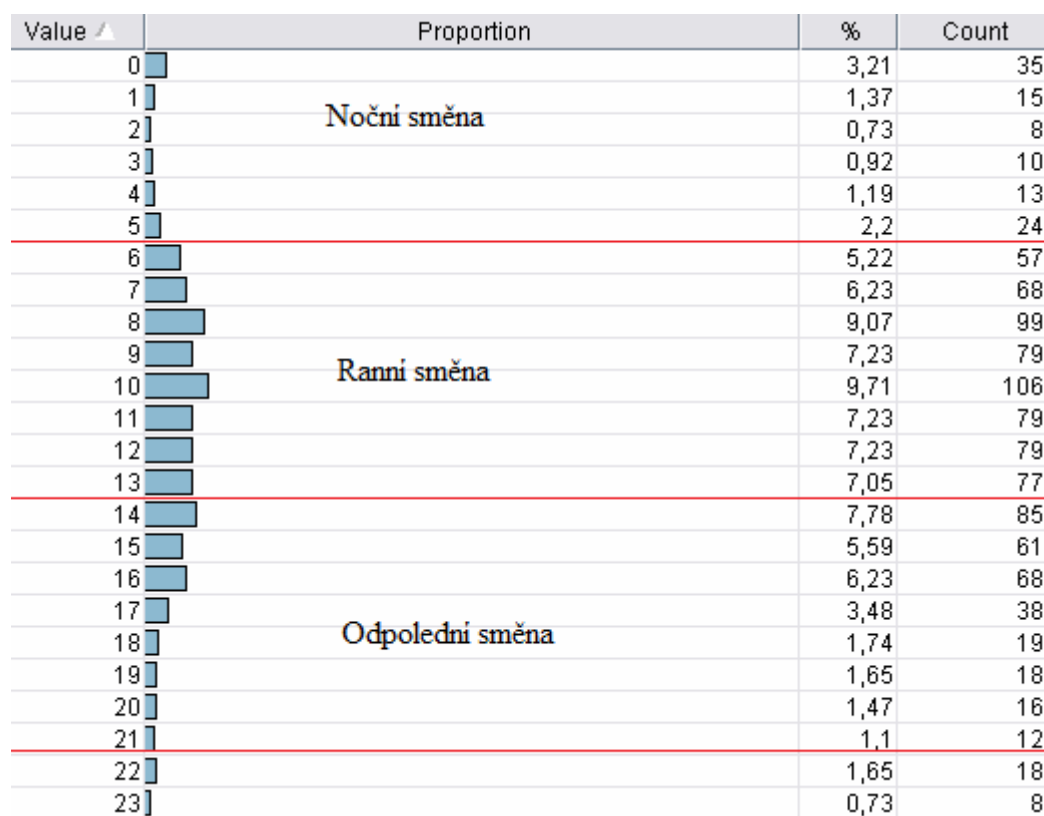
Atribut CasUrazu_H

Dalším atributem, který má vypovídací schopnost, je čas, udáván v hodinách, kdy ke smrtelnému úrazu došlo.

V případě, že bychom si soubor hypoteticky rozdělili na tři části dle směn, na ranní směnu od 6:00 do 14:00, odpolední směnu od 14:00 do 22:00 a noční směnu od 22:00

do 6:00, je možno tvrdit, že k nejvíce smrtelným pracovním úrazům dochází při ranní směně, méně úrazů při odpolední směně a nejméně při noční směně. Nejsou k dispozici údaje o organizacích, zda pracují v jednosměnném či vícesměnném provozu, jde pouze o teoretické tvrzení a usuzování z dat.

Následující obrázek (obrázek 16) vyjadřuje čas úrazu v hodinách a četnost smrtelných pracovních úrazů v časové době. Je k dispozici také procentuální vyjádření smrtelných úrazů vztahující se k času úrazu. Je patrné, že nejvíce smrtelných pracovních úrazů se stalo mezi 6:00 – 16:00. Nejvíce však okolo 10:00, téměř 10 % případů.



Obrázek 16 - Čas úrazu v třísměnném provozu v letech 2002 - 2007, [zdroj: vlastní]

Při dalším uvažování můžeme zavést další hypotézu, že podnik pracuje pouze ve dvousměnném provozu s dvanácti hodinovou pracovní dobou. Ranní směna od 6:00 do 18:00 a noční směna od 18:00 do 6:00. V tomto případě lze říci, podle následujícího obrázku (obrázek 17), že více smrtelných pracovních úrazů se stává při denní směně. Ovšem znovu je nutné říci, že není k dispozici údaj o ekonomickém subjektu a jeho jednosměnném či vícesměnném provozu.

Value ▲	Proportion	%	Count
0	Noční směna	3,21	35
1		1,37	15
2		0,73	8
3		0,92	10
4		1,19	13
5		2,2	24
6	Denní směna	5,22	57
7		6,23	68
8		9,07	99
9		7,23	79
10		9,71	106
11		7,23	79
12		7,23	79
13		7,05	77
14		7,78	85
15		5,59	61
16		6,23	68
17		3,48	38
18	Noční směna	1,74	19
19		1,65	18
20		1,47	16
21		1,1	12
22		1,65	18
23		0,73	8

Obrázek 17 – Čas úrazu v dvousměnném provozu v letech 2002 – 2007, [zdroj: vlastní]

Je tedy možné říci, z výše uvedeného, že převážná část smrtelných pracovních úrazů se stává při denních (ranních) směnách. Můžeme také tvrdit, že poskytnutá data o smrtelných pracovních úrazech v letech 2002 – 2007, jsou převážně v případech jednosměnných provozů.

Bylo by dobré, kdyby se do záznamu o hlášení pracovního úrazu zapisovala směnnost podniku. Pomohlo by to k další charakteristické vlastnosti rizikového ekonomického subjektu a tím i k lepšímu plánování preventivních kontrol.

Atribut katOKEC1

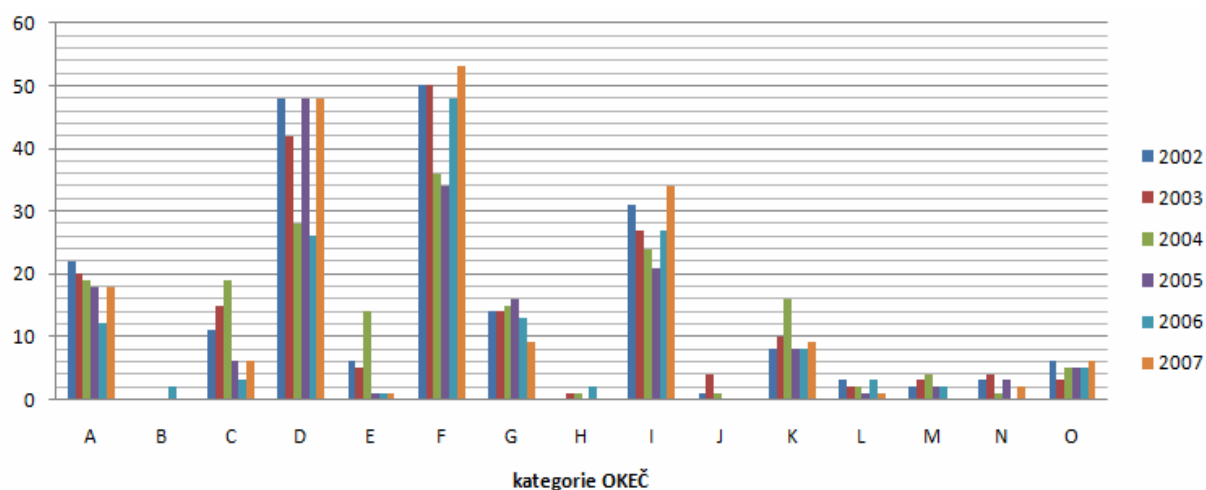
Tento atribut vypovídá o tom, ve kterém odvětví OKEČ se daný smrtelný úraz stal a to v letech 2002 – 2007. Následující obrázek (obrázek 18) představuje jednotlivé kategorie, dle OKEČ od A do O, s četnostmi i s procentuálním vyjádřením. Z tabulky je patrné, že nejvíce smrtelných úrazů se stalo v daných letech v odvětví F, což je stavebnictví, dále v odvětví D, které představuje zpracovatelský průmysl. Třetím nejvíce rizikovějším odvětvím bylo v daných letech odvětví I – doprava, skladování a spoje.

Nejméně smrtelných pracovních úrazů bylo v odvětví B – rybolov, chov ryb, kde se jednalo o dva případy, dále se čtyřmi případy v odvětví H – ubytování, stravování.

Value ▲	Proportion	%	Count
A		9,98	109
B		0,18	2
C		5,49	60
D		21,98	240
E		2,56	28
F		24,82	271
G		7,42	81
H		0,37	4
I		15,02	164
J		0,55	6
K		5,4	59
L		1,1	12
M		1,19	13
N		1,19	13
O		2,75	30

Obrázek 18 - Smrtelné pracovní úrazy dle kategorie OKEČ v letech 2002 – 2007, [zdroj: vlastní]

Následující obrázek (obrázek 19) vyjadřuje počty jednotlivých smrtelných pracovních úrazů pro roky 2002 – 2007 samostatně a v jednotlivých kategoriích OKEČ.



Obrázek 19 - Smrtelné úrazy dle OKEČ za jednotlivé roky 2002 - 2007, [zdroj: vlastní]

Nejvyšší smrtelná úrazovost byla v letech 2002 – 2007 v České republice v odvětvích D (zpracovatelský průmysl), F (stavebnictví) a I (doprava, skladování a spoje).

Kategorie A (zemědělství, myslivost a lesnictví) měla v letech 2002 – 2006 klesající tendenci smrtelné úrazovosti, ovšem v roce 2007 byl zaznamenán vzestup oproti rokům minulým.

Kategorie C (těžba nerostných surovin) od roku 2002 do roku 2004 zaznamenávala nárůst smrtelných pracovních úrazů. Od roku 2005 je trend kolísavý.

Kategorie D (zpracovatelský průmysl) měl od roku 2002 do roku 2004 klesající tendenci, ovšem v roce 2005 se počet smrtelných úrazů velmi zvýšil. V dalším roce, 2006, prudce počet úrazů klesnul a v následujícím roce opět rapidně stoupl.

Kategorie E (výroba a rozvod elektřiny, plynu a vody) měla téměř ve všech letech 2002 – 2007 stejnou četnost smrtelných úrazů, kromě roku 2004, kde je zaznamenán velký nárůst.

Kategorie F (stavebnictví) má klesající tendenci od roku 2002 – 2005, bohužel v dalších dvou letech byl zaznamenán nárůst smrtelných úrazů.

Kategorie G (obchod, opravy motorových vozidel a výrobků pro osobní potřebu a převážně pro domácnost) zaznamenává klesající tendenci smrtelných pracovních úrazů od roku 2006.

Kategorie I (doprava, skladování a spoje) v letech 2002 – 2005 zaznamenávala každoroční poklesy smrtelných úrazů. Od roku 2006 má vzrůstající tendenci.

Kategorie K (činnosti v oblasti nemovitostí a pronájmu, podnikatelské činnosti) má nejvyšší počet smrtelných pracovních úrazů v roce 2004. V dalších letech je možné hovořit o průměru.

Ostatní kategorie OKEČ v letech 2002 – 2007 v České republice zaznamenávají pouze mírné výkyvy mezi jednotlivými sledovanými roky.

Na základě analýzy atributu katOKEC1 by se cílené kontroly dodržování bezpečnosti práce měly zaměřit na kategorie F (stavebnictví), D (zpracovatelský průmysl) a I (doprava, skladování a spoje).

Z dalších atributů nelze vyčíst žádné určité informace, jelikož se jedná o číselné hodnoty, které by daly smysl až v okamžiku, kdy bude k dispozici pomoc v podobě číselníků.

3. Příprava dat

Tato fáze je obvykle jednou z nejnáročnějších fází, protože data bývají často v různých formátech, v různých tabulkách, obsahují chybějící hodnoty, některé atributy potřebné pro analýzu chybějí úplně.

Musíme se umět rozhodnout, která data budou využita pro analýzu. Kritéria pro výběr dat jsou [1]:

- souvislost s cíli DM,
- kvalita a technické podmínky (např. omezení objemu dat nebo typů).

Nezbytným krokem ve fázi přípravy dat je jejich naformátování. Formátování se týká v první řadě syntaktických změn prováděných na datech, které nezmění jejich význam. Některé z nástrojů mají své požadavky na pořadí atributů, např. první pole obsahuje jedinečný identifikátor záznamu. Může být nezbytná i změna pořadí záznamů [17].

Příprava dat zahrnuje činnosti, které vedou k vytvoření datového souboru, jenž bude zpracováván jednotlivými analytickými metodami. Data by měla [17]:

- obsahovat údaje vyznačené pro danou úlohu,
- mít podobu, která je vyžadována vlastními analytickými algoritmy.

Tato fáze tedy zahrnuje několik kroků [1]:

- selekci dat,
- čištění dat,
- transformaci dat,
- vytváření dat,
- integrování dat,
- formátování dat.

Získaná data, jak už bylo řečeno, byla ve formátu .xls, avšak ve velmi nesrozumitelné podobě. Bylo tedy nutné data zpracovat tak, aby bylo možné s nimi dále pracovat a získat z nich potřebné informace.

Do zdrojového souboru s daty byl vložen řádek s legendou. Ovšem ani tato operace nebyla postačující. Za každou skupinu, spolu jednoznačně souvisejících údajů, např. DruhZraneni_1, DruhZraneni_2, DruhZraneni_3, byl vložen sloupec. Tento sloupec sloučil hodnoty spolu souvisejících předcházejících sloupců. V konkrétním případě pro nový atribut Druh_zraneni bylo nutno sloučit sloupce DruhZraneni_1, DruhZraneni_2, DruhZraneni_3.

Každý nově vytvořený atribut je tedy možné hierarchicky rozložit. Pro konkrétní atribut DruhZraneni_1 máme (podle vhodného číselníku) hodnoty 0 – 9, kde:

- 0: druh zranění není znám nebo blíže specifikován,
- 1: zlomeniny, vymknutí, podvrtnutí, natažení kloubů a vazů,
- 2: poranění mozku, míchy a vnitřních orgánů,
- 3: poranění povrchová, zhmoždění, rány otevřené, amputace, rozdrčení
- 4: popálení, poleptání, opaření, omrznutí,
- 5: otravy a toxické účinky látek, dušení a účinky tlaku, tonutí,
- 6: účinky elektrického proudu, blesku,
- 7: následky námahy, šok, cizí těleso,
- 8: poranění mnohočetná,
- 9: ostatní dříve neuvedená.

Další atribut (DruhZraneni_2) upřesňuje každou kategorii atributu DruhZraneni_1. Kategorie 1 (zlomeniny, vymknutí, podvrtnutí, natažení kloubů a vazů) atributu DruhZraneni_1 má v atributu DruhZraneni_2 další podkategorie. Jsou to hodnoty 0 – 2. Sloučením atributů DruhZraneni_1 a DruhZraneni_2 získáme, pro kategorii 1, hodnoty 10 = zlomeniny, vymknutí, natažení kloubů a vazů, 11 = zlomenina, 12 = vymknutí, podvrtnutí, natažení kloubů a vazů.

Posledním atributem této hierarchie je DruhZraneni_3. Tento doplňuje o další informace předchozí dva atributy. Pro atribut DruhZraneni_2 má ve skupině 11 hodnoty 0, 1, 2, 5. Hodnota celého atributu 110 = zlomenina, 111 = zlomenina zavřená (vč. roztržštěné, našťípnuté, nalomené, spirální apod.), 112 = zlomenina otevřená (vč. komplikované, stělné a způsoben bodnutím), 115 = zlomenina (obratle) s poškozením míchy.

Následující obrázek (obrázek 20) hovoří o upravení dat pro jejich lepší srozumitelnost.

DruhZraneni_1	DruhZraneni_2	DruhZraneni_3	Druh_zraneni
2	1	3	213
0	2	0	020
8	1	2	812
0	1	0	010
6	1	4	614
8	1	1	811
2	3	1	231
2	3	1	231
8	1	2	812
8	1	2	812
8	1	3	813
8	1	3	813
8	1	3	813
8	1	1	811
8	1	3	813
8	1	3	813
8	1	3	813
8	1	3	813
8	1	3	813

Obrázek 20 - Upravená data pro práci [zdroj: vlastní]

Konečná podoba dat, tak jak se budou vkládat do Clementine, je sestavena z 26 atributů, které vycházejí z původních dat získaných od Výzkumného ústavu pro BOZP. Z původních 103 atributů jich bude potřeba pouze 26. Budou využity pouze nově vytvořené atributy, které vznikly sloučením. Jedná se o 20 atributů. Dále se bude pracovat s atributy, které nebylo potřeba upravovat (slučovat), což je věk, pohlaví a rok úmrtí. Některé atributy nebudou pro práci potřeba vůbec, nemají na práci vliv.

Popis atributů hotových dat

Tyto nové atributy vycházejí odvozením z atributů prvotních, tak jak byly obdrženy prvotně. Příprava a jejich odvozování byly prováděny v MS Excel. Tyto budou použity jako vstupy do prostředí Clementine. Původní atributy, z kterých se vytvořily nové, jsou popsány v příloze 1, v datovém slovníku.

1. Pravni_forma

Atribut Pravni_forma byl vytvořen ze třech spolu souvisejících atributů a to Pravni_forma1, Pravni_forma2 a Pravni_forma3.

2. katOKEC

Hodnoty atributu katOKEC byly složeny ze dvou souvisejících atributů (katOKEC1, katOKEC2), kde byly písmenné hodnoty nahrazeny číselnými.

3. oddilOKEC

Nový atribut oddilOKEC byl odvozen ze tří původních atributů (oddilOKEC1, oddilOKEC2, oddilOKEC3). Hodnoty jednotlivých atributů nebylo potřeba měnit, byly ve vyhovujícím číselném formátu.

4. Uzemni_jednotka

Hodnoty nového atributu Uzemni_jednotka bylo nutno vytvořit z pěti souvisejících hodnot atributů (Uzemni_jednotka1, Uzemni_jednotka2, Uzemni_jednotka3, Uzemni_jednotka4, Uzemni_jednotka5) sloučením.

5. Velikost_podniku

Hodnoty atributu Velikost_podniku vycházejí ze sloučení dvou původních atributů (VelikostPodniku1, VelikostPodniku2).

6. Fu_Uzemni_jednotka

Atribut Fu_Uzemni_jednotka měl být složen z pěti atributů, ale pro lepší práci byl jeden atribut (Fu_Uzemni_jednotka1), kde byly hodnoty písmenné „CZ“, vynechán.

7. Sex

Hodnoty atributu sex byly převedeny na hodnoty 0 (muž) a 1 (žena) z původních 1 (muž) a 2 (žena).

8. Vek

Tento atribut nevyžadoval žádné úpravy.

9. StatniPrislusnost

Hodnoty tohoto atributu byly v písmenné formě. Bylo nutno je přeformátovat do číselné podoby. Státní příslušnost CZ má hodnotu 1, BG 2, BY 3, PL 4, RU 5, SK 6 a UA 7.

10. KZAM

Atribut KZAM byl složen z pěti původních atributů (Zamestnani_1, Zamestnani_2, Zamestnani_3, Zamestnani_4, Zamestnani_5).

11. Postaveni

Hodnoty atributu Postaveni bylo nutno složit ze dvou původních atributů (Postaveni_1, Postaveni_2)

12. Delka_zam_mesice

Tento atribut byl sloučen ze dvou atributů. První atribut byl délka zaměstnání v letech (DelkaZamestnani_R) a druhý délka zaměstnání v měsících (DelkaZamestnani_M). Roky byly převedeny na měsíce a sečteny s počtem měsíců.

13. ZranenychOsob

Hodnoty tohoto atributu nebylo potřeba slučovat či formátovat.

14. DatumUrazu

Hodnoty atributu DatumUrazu byly přeformátovány na šestimístné číslo oddělené po dvou místech tečkou.

15. Rok_umrti

Původní hodnoty byly pouze jednomístné, číslo vyjadřovalo rok úmrtí (např. 3=2003). Hodnoty byly převedeny na čtyřmístné číslo.

16. OdpracovanyCas_H

Hodnoty nového atributu byly sloučeny ze dvou atributů, odpracovaný čas v hodinách a odpracovaný čas v minutách a tyto převedeny na hodiny. Výsledkem byla hodnota s dvěma desetinnými místy.

17. Druh_zraneni

Hodnoty nově vytvořeného atributu Druh_zraneni byly vytvořeny ze tří původních atributů (DruhZraneni_1, DruhZraneni_2, DruhZraneni_3).

18. ZranenaCast

Atribut ZranenaCast byl vytvořen ze čtyř původních atributů (ZranenaCast_1, ZranenaCast_2, ZranenaCast_3, ZranenaCast_4).

19. Pracoviste

Hodnoty atributu nebylo potřeba formátovat.

20. Misto

Atribut Misto vytvořen ze čtyř původních atributů (Misto_1, Misto_2, Misto_3, Misto_4).

21. Cinnost

Hodnoty nového atributu Cinnost byly složeny ze čtyř původních hodnot atributů (Cinnost_1, Cinnost_2, Cinnost_3, Cinnost_4).

22. Ukon

Atribut Ukon byl složen ze šesti původních atributů (Ukon_1, Ukon_2, Ukon_3, Ukon_4, Ukon_5, Ukon_6).

23. Cinitel_zdroj

Hodnoty atributu Cinitel_zdroj byly zpracovány z původních pěti atributů (Cinitel_Zdroj_1, Cinitel_Zdroj_2, Cinitel_Zdroj_3, Cinitel_Zdroj_4, Cinitel_Zdroj_5).

24. Nehoda

Nově vytvořený atribut Nehoda byl zpracován sloučením pěti původních atributů (Nehoda_1, Nehoda_2, Nehoda_3, Nehoda_4, Nehoda_5) vztahujících se k novému atributu Nehoda.

25. Udalost

Hodnoty atributu Udalost jsou vytvořeny ze tří původních atributů (Udalost_1, Udalost_2, Udalost_3) sloučením.

26. Pricina

Tento nově vytvořený atribut byl vytvořen z původních pěti atributů (P1_Pricina_1, P1_Pricina_2, P1_Pricina_3, P1_Pricina_4, P1_Pricina_5).

Avšak ani tyto údaje nejsou úplně srozumitelné. Bylo nutné obstarat odpovídající číselníky, díky kterým jsou údaje v nově vytvořených atributech jasně pochopitelné a slovně ohodnocené. Ke každému nově vytvořenému atributu bylo nutné mít jeden daný číselník. Jednotlivé číselníky byly propojeny s novými atributy.

Následující obrázek (obrázek 21) představuje ukázkou číselníku pro atribut Udalost_1.

1	Udalost_1	Popis
2	00	Neuvedena
3	11	Havárie letadla
4	12	Havárie plavidla
5	13	Havárie silničního a zvláštního vozidla
6	14	Havárie zdvihacího zařízení
7	15	Havárie kolejového vozidla
8	16	Exploze pracovního, hnacího stroje, vloženého zařízení, předmětu
9	17	Exploze ostatní
10	20	Ztráta kontroly ns.
11	21	Zachycení, sražení, přejetí, přiražení osoby pohybujícím se strojem
12	22	Udeření, navinutí, říznutí, stisknutí, přiražení, vtažení osoby částí stroje
13	23	Udeření, říznutí, stisknutí, přiražení osoby nářadím
14	24	Zasažení osoby předměty, částicemi odletujícími od stroje
15	25	Zasažení osoby předměty, částicemi odletujícími od nářadí
16	26	Zasažení osoby úlomky z rozbitého předmětu, jeho povrchu
17	27	Zanesení drobných částic do očí a jiných dutin
18	28	Zasažení osoby elektrickým proudem, popálení elektrickým obloukem
19	29	Udeření, přiražení osoby dveřmi, vraty, oknem, poklopem, roletou, zásuvkou
20	30	Únik, výtok, vysplíchnutí, rozlití kapaliny ns.
21	31	Popálení osoby roztavenou látkou ze stroje, zpracovávaného materiálu
22	32	Popálení osoby roztavenou látkou z pánve, formy, nádoby
23	33	Popálení osoby roztavenou látkou při sáhnutí, šlápnutí do ní
24	34	Popálení osoby plamenem, opaření párou, tekutinou ze stroje
25	35	Popálení osoby plamenem, opaření, párou, tekutinou z nádoby, předmětu
26	36	Popálení osoby o horký povrch stroje, nářadí, předmětu
27	37	Ozáření osoby zářením ze stroje (i el. oblouku) a z látek
28	38	Poleptání osoby látkami ze stroje, nádoby, předmětu, hromady

Obrázek 21 - Číselník pro atribut Udalost_1 [zdroj: vlastní]

Pro číselníky s takto zadanými atributy jako např. Udalost_1, Udalost_2 a Udalost_3 bylo nutno vytvořit společný číselník. Tento nový číselník vysvětluje hodnoty ve vytvořeném atributu Udalost v předzpracovaných datech. Následující obrázek (obrázek 22) představuje společný číselník pro Udalost_1, Udalost_2 a Udalost_3.

1	Udalost	Popis
2	0000	neuvedena
3	1101	nežádoucí událost ns.
4	1111	porucha řídicích orgánů letadla (vč. řídicích ploch)
5	1112	porucha nosné soustavy letadla
6	1113	porucha trupu
7	1114	porucha pohonné jednotky (motoru, vrtule, trysky)
8	1115	porucha přistávacího zařízení (podvozku)
9	1116	porucha palubních přístrojů
10	1120	závady letištní plochy (vč. návěstidel) a radionavigačního systému ns.
11	1130	nepříznivé meteorologické podmínky ns.
12	1131	vítr - náraz, poryv větru
13	1132	zhoršená viditelnost pro mlhu
14	1133	námraza
15	1134	zhoršená viditelnost pro déšť, průtrž mračen
16	1140	chybná pilotáž ns.
17	1141	vzlet, přistávání bez povolení
18	1142	chybné vyhodnocení údajů, chybná navigace
19	1143	chyba personálu pro řízení letového provozu
20	1144	přehlédnutí překážky při nízkém letu, letecké práci
21	1145	chybný přistávací manévr
22	1146	akrobacie při nedostatečné výšce letu

Obrázek 22 - Číselník pro vytvořený atribut Udalost [zdroj: vlastní]

S takto upravenými číselníky pro jednotlivé atributy se již dá bezproblémově pracovat a identifikovat jednotlivé hodnoty společně se slovním popisem. Nyní je již absolutní vypovídací schopnost všech atributů, s kterými se bude pracovat.

4. Modelování

V úvodním kroku čtvrté fáze – modelování – se doporučuje vybrat konkrétní algoritmus, který se bude používat pro analýzu.

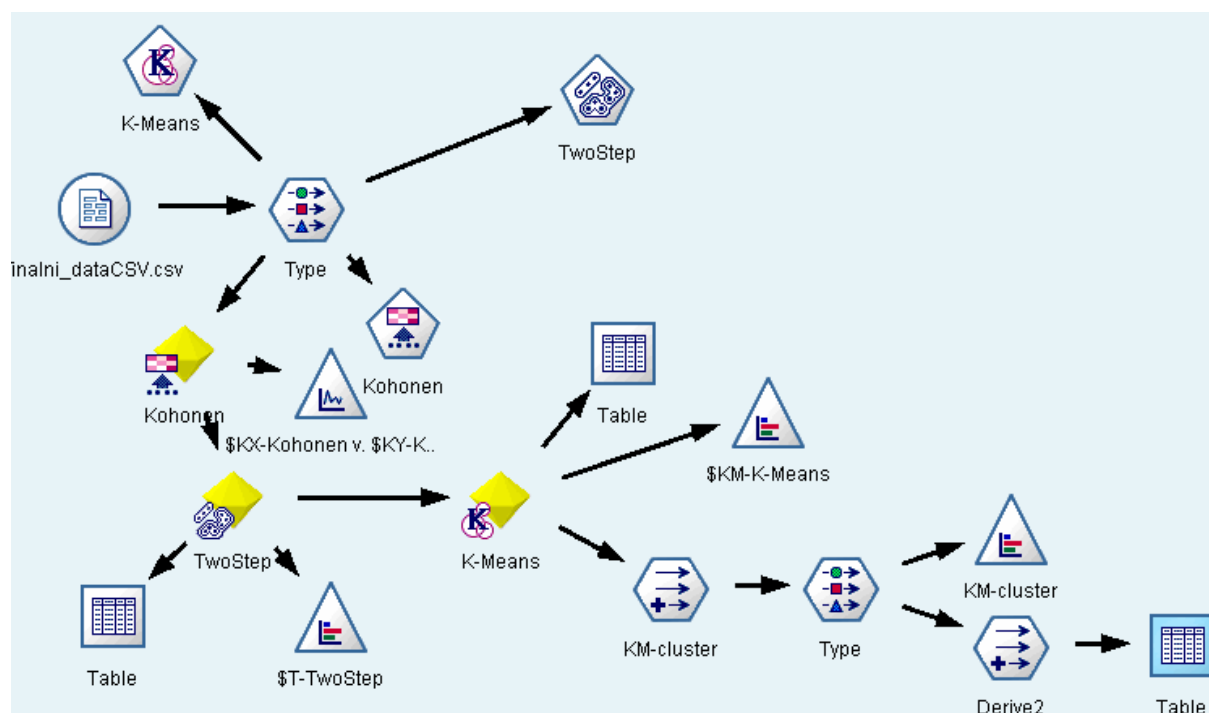
Před vlastním sestavením modelu je potřeba vytvořit mechanismus nebo postup, který by testoval kvalitu a sílu (správnost) modelu [17].

V průběhu samotného modelování je vytvořen model či více modelů. Důležitou a nezbytnou součástí této etapy je ocenění modelů. Analytik, informatik, ohodnotí modely podle hledisek, kterými jsou kritéria pro přesnost definovaná v první fázi. Jestliže je to možné, berou se také v úvahu obchodní cíle a kritéria hodnotící obchodní úspěšnost. Protože analytikův pohled je spíše techničtější, spojuje se s obchodním analytikem a expertem v dané oblasti, aby interpretovali výsledky modelování v obchodních souvislostech [17].

Pro modelování dat byl využit software Clementine. Data byla do software exportována ve formátu CSV oddělený středníkem. Bylo nutno nastavit pro jednotlivé atributy datový typ, a to množina, uspořádaná množina či příznak.

Při samotném modelování byly využity seskupovací algoritmy, kterými jsou, v prostředí Clementine, Kohonenova mapa, K – Means a Two Step.

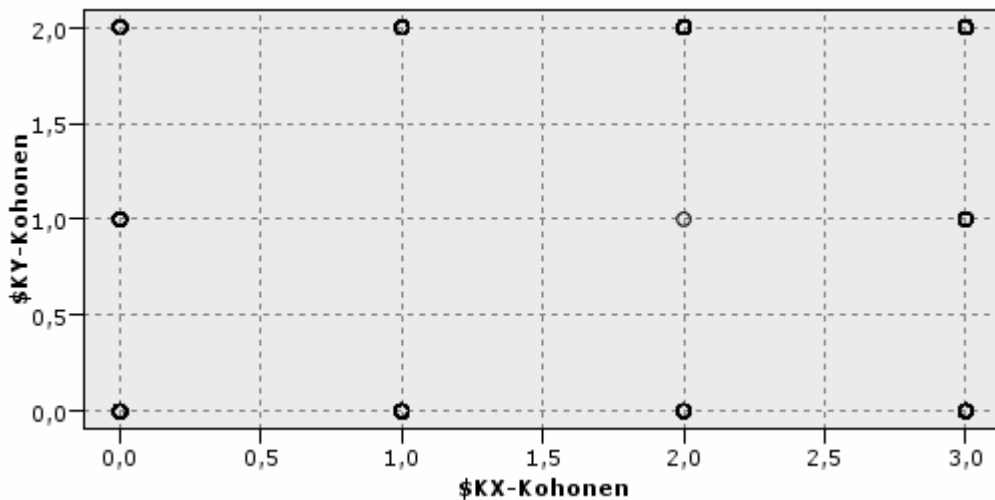
Následující obrázek (obrázek 23) představuje stream pro tvorbu modelu pomocí seskupovací analýzy (segmentace) s využitím algoritmů Kohonenovy mapy, K – Means a Two Step. Tyto tři algoritmy byly propojeny pro výběr nejlepší metody.



Obrázek 23 - Stream pro tvorbu seskupovací analýzy, [zdroj: vlastní]

Kohonenova mapa je seskupovací neuronová síť určená k odhalení vztahů, které mohou být skryté v mnohorozměrných datech.

U tohoto algoritmu není možnost výběru počtu shluků. Algoritmus vybere sám správnou variantu, do kolik shluků údaje rozdělí. Na následujícím obrázku (obrázek 24) je vidět, že údaje byly rozděleny do 11 shluků. Shluky jsou na mapě prezentovány body. Každý shluk má na mapě své souřadnice [x, y]. Jednotlivé shluky obsahují případy smrtelných pracovních úrazů v letech 2002 – 2007.



Obrázek 24 - Výstup z algoritmu "Kohonenova mapa" znázorňující počet shluků, [zdroj: vlastní]

Z dalšího obrázku (obrázek 25) je patrné, že nejvíce zastoupený je shluk 9, který má souřadnice na Kohonenově mapě [3; 0].



Obrázek 25 - Četnost zastoupení shluků v algoritmu Kohonenova mapa, [zdroj: vlastní]

Shluky byly seskupeny na základě hodnot atributů sex, postaveni, pracoviste, rok_umrti, statniPrislusnost, velikost_podniku a zranenychOsob, což je patrné z dalšího obrázku (obrázek 26).



Obrázek 26 - Atributy v nejčetnějším shluku algoritmu Kohonenova mapa, [zdroj: vlastní]

Atributy shluku se slovním ohodnocením za využití příslušných číselníků u metody Kohonenova mapa:

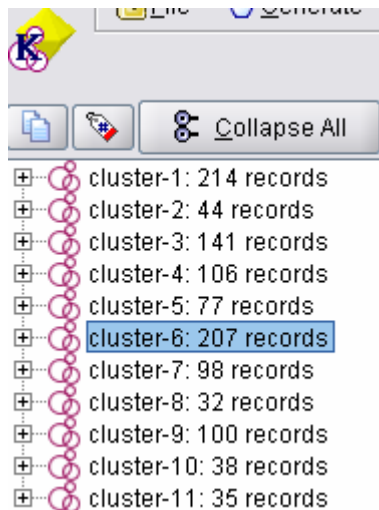
- sex: 0 = muž,
- postaveni: 11 = zaměstnanec v pracovním poměru,
- pracoviste: 2 = příležitostné nebo pohyblivé (mobilní) pracoviště nebo úraz na pracovní cestě,
- rok_umrti: 2002,
- statniPrislusnost: 1 = CZ,
- velikost_podniku: 120 = do pěti zaměstnanců,
- zranenychOsob: 1 osoba.

K – Means se snaží rozdělit objekty do zadaného počtu shluků tak, aby byla minimální variabilita uvnitř shluků a maximální mezi shluky. Výstupem z této procedury je skupinová příslušnost, tedy každý případ je zařazen do jednoho ze shluků.

V případě algoritmu K – Means bylo nutno definovat počet shluků. Pro lepší porovnání algoritmů bylo nastaveno 11 možných shluků.

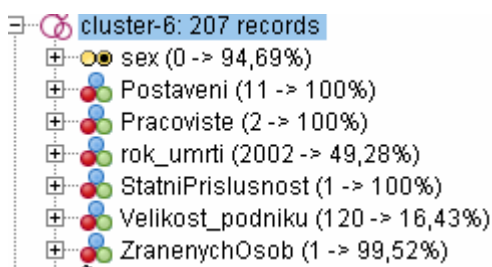
Následující obrázek (obrázek 27) představuje počet shluků a četnosti údajů ve shlucích. Nejvíce zastoupený je shluk 6.

Shluky byly seskupeny na základě hodnot atributů sex, postaveni, pracoviste, rok_umrti, statniPrislusnost, velikost_podniku a zranenychOsob.



Obrázek 27 - Počet jednotlivých shluků u metody K - Means, [zdroj: vlastní]

Dalším obrázkem (obrázek 28) jsou nejvíce zastoupené hodnoty atributů shluku 6, který měl největší četnost případů.



Obrázek 28 - Atributy v nejčetnějším shluku algoritmu K - Means, [zdroj: vlastní]

Atributy shluku se slovním ohodnocením za využití příslušných číselníků u metody K – Means:

- sex: 0 = muž,
- postaveni: 11 = zaměstnanec v pracovním poměru,
- pracoviste: 2 = příležitostné nebo pohyblivé (mobilní) pracoviště nebo úraz na pracovní cestě,
- rok_umrti: 2002,
- statniPrislusnost: 1 = CZ,
- velikost_podniku: 120 = do pěti zaměstnanců,
- zranenychOsob: 1 osoba.

Two step je založen na hierarchických metodách používaných k získávání nejlepšího počtu shluků pro daný problém.

Při definování algoritmu Two Step je nutné vybrat ze dvou možností nastavení počtu shluků. Jednou možností je automatické nastavení, kde lze zadat maximální a minimální počet shluků. Druhou možností je možnost vlastního počtu shluků.

Počet shluků byl tedy nastaven na minimální hodnotu 11, z důvodu lepšího porovnávání algoritmů. U algoritmu Kohonenova mapa, kde není možnost nastavení počtu shluků, jich vzniklo 11, tudíž u této metody využijeme výsledný počet shluků z Kohonenovy mapy.

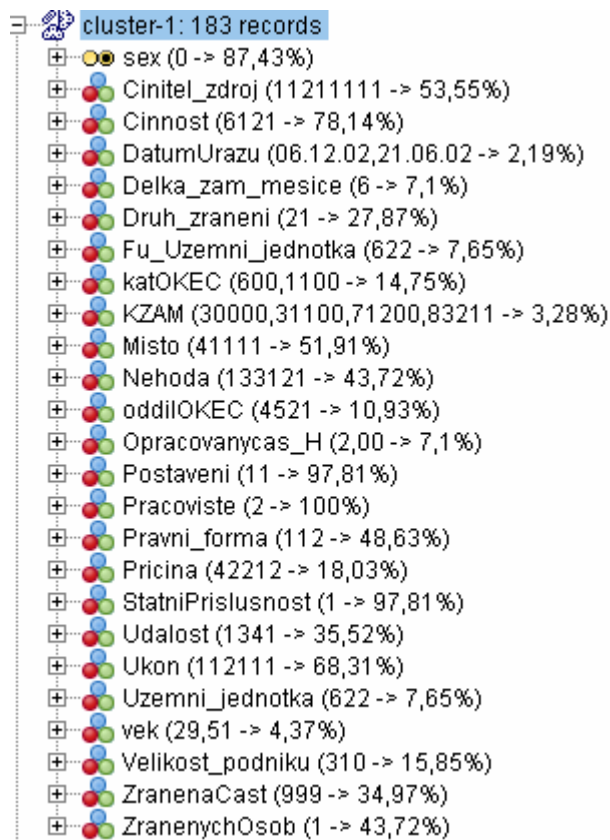
Následující obrázek (obrázek 29) ukazuje počet vytvořených shluků s počtem zastoupených údajů.



Obrázek 29 - Počet shluků algoritmu Two Step s četnostní údajů, [zdroj: vlastní]

Jednotlivé shluky byly seskupeny na základě hodnot všech atributů, kromě hodnot atributu rok_umrti.

Další obrázek (obrázek 30) představuje nejvíce zastoupené hodnoty atributů shluku 1. Tento shluk má největší četnost případů smrtelné úrazovosti.



Obrázek 30 - Atributy v nejčetnějším shluku algoritmu Two Step, [zdroj: vlastní]

Z atributů ve shluku bude pro potřeby charakteristiky rizikového podniku potřeba pouze `delka_zam_mesice`, `katOKEC`, `KZAM`, `oddilOKEC`, `postaveni`, `pracoviste`, `pravni_forma`, `vek`, `velikost_podniku`. Vhodné atributy shluku se slovním ohodnocením za využití příslušných číselníků u metody Two Step:

- `delka_zam_mesice`: 6 měsíců,
- `katOKEC`: 600 = stavebnictví, 1100 = činnosti v oblasti nemovitostí a pronájmu; podnikatelské činnosti,
- `KZAM`: 30000 = technik, 31100 = technik v geologii, 71200 = zedník, 83211 = řidič,
- `oddilOKEC`: 4521 = výstavba pozemních a inženýrských staveb,
- `postaveni`: 11 = zaměstnanec v pracovním poměru,
- `pracoviste`: 2 = příležitostné nebo pohyblivé (mobilní) pracoviště nebo úraz na pracovní cestě,
- `pravni_forma`: 112 = společnost s ručením omezením,
- `vek`: 29, 51 let,
- `velikost_podniku`: 310 = 100 – 199 zaměstnanců.

5. Vyhodnocení výsledků

V této fázi se hodnotí úroveň, s jakou model dosahuje obchodních cílů a snaží se určit, zda je nějaký důvod, taktéž obchodní, aby tento model byl nedostatečný.

Vytvořený model je možné vyhodnotit takovým způsobem, že ho užitíme na reální situaci a poté sledujeme jeho kvalitu. Nutností je však posoudit a zvážit rozpočtové a časové podmínky, jestli vůbec takovéto hodnocení umožňují.

V případě, že se výsledný model prokáže jako schopný uspokojit potřeby, následuje velmi podrobná a důkladná kontrola celé DM úlohy a určuje se, jestli nebyl zanedbán či přehlédnut nějaký důležitý faktor či úkol. Tato kontrola zahrnuje ujistění se o kvalitě (tj. o správném sestavení modelu; o použití atributů, které budou dostupné i pro další následné analýzy).

Rozhodne se, s ohledem na výsledky hodnocení a kontroly procesu, jak se bude dále pokračovat. Je nutné rozhodnout, zda ukončit tento projekt a přesunout ho do další fáze, kterou je implementace, či zahájit další opakování některých fází nebo dokonce začít úplně nový DM projekt. Zahájení nového DM projektu zahrnuje analýzu zbývajících zdrojů a rozpočtu, které ovlivní rozhodnutí.

DM projekt, charakterizovat ekonomický subjekt pro kontrolu BOZP, byl shledán jako vhodným pro zavedení do praxe.

Výsledky, jakož celý postup, byly znovu důsledně překontrolovány. Nebyly vykazovány žádné nedostatky či opomenuté kroky při sestavování modelu, ani při přípravě vhodných dat pro modelování. Modelovací techniky seskupovací analýzy byly dostatečné, nebylo nutné doplnit model o další algoritmy. Výsledky se opírají též o výstupy z popisné statistiky.

Byly využity tři algoritmy seskupovací analýzy, Kohonenova mapa, K – Means a Two Step. Kohonenova mapa a K – Means předkládaly stejné výsledky, shluky byly provedeny na základě stejných atributů s těmito hodnotami:

- sex: 0 = muž,
- postaveni: 11 = zaměstnanec v pracovním poměru,
- pracoviste: 2 = příležitostné nebo pohyblivé (mobilní) pracoviště nebo úraz na pracovní cestě,
- rok_umrti: 2002,
- statniPrislusnost: 1 = CZ,
- velikost_podniku: 120 = do pěti zaměstnanců,
- zranenychOsob: 1 osoba.

Na základě stejných výsledků tedy budeme hovořit o dvou vhodných metodách, ovšem musíme také zohlednit, která metoda je jednodušší. V tomto případě je jednodušší metoda Kohonenova mapa. Není nutné nastavovat počet shluků jako u dalších metod seskupovací analýzy. Pro další využití modelu se tedy bude využívat algoritmus Kohonenova mapa.

Na základě výsledků z algoritmu Kohonenova mapa můžeme tedy charakterizovat rizikový ekonomický subjekt takto:

- velikost podniku: do pěti zaměstnanců,
- postavení zaměstnance: zaměstnanec v pracovním poměru,
- pracoviště zaměstnance: příležitostné nebo pohyblivé (mobilní) pracoviště nebo úraz na pracovní cestě,
- pohlaví: muž.

Je však nutné přihlédnout k výsledkům popisné statistiky. Zde byly nalezeny další charakteristické rysy:

- věk: 50 let,
- délka zaměstnání v letech: do 1 roku,
- státní příslušnost (krom českých pracovníků): slovenští, polští a ukrajinští pracovníci,
- kategorie OKEČ: stavebnictví, zpracovatelský průmysl, doprava, skladování a spoje.

Takto charakterizované ekonomické subjekty je nutné cíleně kontrolovat z hlediska BOZP. Tyto ekonomické subjekty jsou velmi rizikové z hlediska pracovních úrazů, následně smrtelných pracovních úrazů a vyžadují odborné kontroly.

6. Implementace

Pro zavedení DM modelů do obchodních činností využívá tento úkol výsledky hodnocení a vyvozuje z nich strategii pro implementaci. V případě, že byl identifikován obecný postup pro vytvoření platného modelu či více modelů, je zde tento postup dokumentován pro další jeho použití [17].

Pro další využití tohoto modelu je nutné do zdrojových dat neustále vkládat nové záznamy o smrtelných pracovních úrazech. Poté bude možné opět s přesností určit ekonomický subjekt, který je rizikový. V případě, že se datový soubor nebude aktualizovat, hrozí neaktuálnost výsledků. V praxi to bude znamenat opět špatně cílené ekonomické subjekty a špatné zacházení s lidskými i finančními zdroji.

V procesu zavádění modelů by se nemělo zapomenout na vytvoření plánů pro kontrolu a údržbu. Význam kontroly a údržby roste, pokud se výsledky DM analýz mají stát součástí běžných a každodenních aktivit společnosti, organizace. Precizní příprava strategie údržby pomáhá vyhnout se přehnaně dlouhým obdobím, po která jsou DM výsledky špatně užívány [17].

Na závěr projektu by měla být sepsána závěrečná zpráva. Tato zpráva může být jako stručné shrnutí nebo se může jednat o závěrečné a vyčerpávající prezentaci všech výsledků, kterých bylo dosaženo během celého procesu [17].

Finálním vypracovávaným dokumentem je revize projektu. Tato zpráva se nepředává zákazníkovi, ale slouží pro vnitřní potřeby DM firmy, tudíž se může zdát jako zbytečná. Tento dokument by však chybět neměl, protože jde o shromažďování podnikových znalostí. Pracovníci by měli zhodnotit, které fáze šly dobře, které špatně, co bylo uděláno dobře a co by bylo potřeba do budoucna vylepšit. Sepisují se zde velmi důležité zkušenosti získané během celého projektu. Ukazují na nebezpečná místa v analýze. Zaznamenané individuální postřehy umožňují pracovníkům sdílet své zkušenosti a tím pracovat v dalších analýzách efektivněji a vyvarují se zbytečných chyb a nedostatků [22].

Výsledky analýzy v posledním kroku využívají právě zákazníci, nikoli analytici. Je tedy velmi důležité, aby zákazník pochopil, co je nezbytné a důležité učinit pro to, aby mohly být získané výsledky efektivně využívány [22].

5.3 Shrnutí kapitoly

V této kapitole byl definován DM jako netriviální získávání implicitních, dříve neznámých a potencionálně užitečných informací z dat. Byly popsány také DM techniky. Celý projekt je založen na DM metodologii CRISP-DM. Tato metodologie má šest fází. V první fázi, porozumění problému, byl definován především obchodní a DM cíl. Další fáze, porozumění problému, byla věnována popisu a analýze dat spolu s představením surových dat. Třetí fáze, příprava dat, zahrnovala práci se surovými daty a číselníky. Data byla formátována pro další fázi, modelování. Zde byla využita seskupovací analýza s algoritmy Kohonenova mapa, K – Means a Two Step pro vytvoření modelu. V dalším kroku (vyhodnocení výsledků) byla určena charakteristika rizikového ekonomického subjektu. Poslední fází je implementace vytvořeného modelu a jeho využití v praxi.

6 Závěr

Kontroly BOZP jsou stále více diskutovaným tématem z důvodu četnosti pracovních, ale také smrtelných úrazů. V České republice má klesající tendenci, i když nepatrnou. Stále se však ročně stane v průměru okolo 180 smrtelných pracovních úrazů, což je alarmující číslo. Z tohoto důvodu byla cílem práce potřeba sestavit přesnou charakteristiku ekonomického subjektu, který je k pracovním úrazům náchylný, tudíž rizikový. Tato charakteristika bude využívána pracovníky kontrolního oddělení BOZP, aby se minimalizovalo plýtvání lidských a finančních zdrojů na zbytečných kontrolách.

V práci byla zmíněna legislativa upravující BOZP a rozčleněno pojištění podle závaznosti (zákonné a smluvní), druhu (životní a neživotní) a míry rizikovosti (rizikové a rezervotvorné). Pro případ pracovního úrazu přichází v úvahu úrazové pojištění a pojištění nemoci. Tato pojištění se řadí do pojištění komerčních. V případě zákonného pojištění hovoříme o sociálním pojištění, kde se uplatňuje princip solidarity, nikoli ekvivalence jako u komerčního pojištění.

Další část práce se zabírala statistikou smrtelných pracovních úrazů v České republice za rok 2008 a v Evropě v letech 2000 – 2001. V ČR v roce 2008 nastalo celkem 174 smrtelných pracovních úrazů, z toho v 8 případech se jednalo o ženy. Největší smrtelná pracovní úrazovost byla v roce 2008 v Ústeckém kraji, kde došlo k 27 případům. Nejrizikovější kategorií OKEČ je stavebnictví a zpracovatelský průmysl. Největší meziroční pokles smrtelných pracovních úrazů v Evropě v letech 2000 – 2001 zaznamenalo Lucembursko. Zde počet smrtelných pracovních úrazů poklesl o 75 %.

V práci byl dále definován DM jako netriviální získávání implicitních, dříve neznámých a potenciaálně užitečných informací z dat. DM se nejvíce využívá v analýze zákazníka, bankovníctví, odhalování podvodů a v neposlední řadě také v pojišťovnictví. Byly probrány DM techniky, jako rozhodovací strom, neuronová síť, seskupovací analýza.

Jako jednou z metodologií DM byla charakterizována metodologie CRISP-DM a její fáze.

Celý projekt, jehož cílem bylo vytvořit model pro plánování preventivních kontrol dodržování BOZP a následně najít charakteristiku rizikového ekonomického subjektu, je postaven právě na CRISP-DM. Fáze této metodiky jsou porozumění problému, porozumění datům, příprava dat, modelování, vyhodnocení výsledků a implementace modelu a jeho využití v praxi.

Bylo využito třech algoritmů pro vytvoření modelu v prostředí Clementine. Jednalo se o Kohonenovu mapu, K – Means a Two Step. Nejvhodnějším algoritmem byl určen algoritmus Kohonenova mapa, jelikož podává stejné výsledky jako K – Means, ale je jednodušší. Dále byla určena přesná charakteristika rizikového ekonomického subjektu. Mezi tyto charakteristiky patří firmy s počtem zaměstnanců nepřesahující 5, pracovním vztahem je zaměstnanec zařazen jako zaměstnanec v pracovním poměru. Důležitým charakteristickým rysem je též délka pracovního vztahu kratší než 1 rok. Kromě českých pracovníků by se kontroly měly zaměřit na firmy zaměstnávající slovenské, polské a ukrajinské pracovníky. Charakteristickým rysem je také pohlaví muž a věk okolo 50 let. Další charakteristikou je také místo výkonu práce, kde velkým rizikem je pracoviště příležitostné či pohyblivé nebo úraz na pracovní cestě. Rizikovým odvětvím OKEČ je stavebnictví, zpracovatelský průmysl, doprava, skladování a spoje.

7 Použitá literatura

- [1] BERKA, Petr. *Dobývání znalostí z databází*. 1. vyd. Praha : [s.n.], 2003. 366 s. ISBN 80-200-1062-9.
- [2] BERRY, Michael J. A., LINOFF, Gordon. *Data mining techniques : For marketing, sales, and customer support*. 1st edition. [s.l.] : [s.n.], 1997. 454 s. ISBN 0-471-17980-9.
- [3] Český statistický úřad. *Cizinci v České republice* [online]. 2009 [cit. 2009-08-10]. Dostupný z WWW: <[http://www.czso.cz/csu/2002edicniplan.nsf/publ/1414-02-\(casova_rada\)>](http://www.czso.cz/csu/2002edicniplan.nsf/publ/1414-02-(casova_rada)>).
- [4] Český statistický úřad. *OKEČ - SYSTEMATICKÁ ČÁST PLATNÁ OD 1.1.2003* [online]. 2008 , 20.3.2008 [cit. 2009-03-19]. Dostupný z WWW: <[>](http://www.czso.cz/csu/klasifik.nsf/i/okec).
- [5] Český statistický úřad. *Pracovní neschopnost pro nemoc a úraz v ČR* [online]. 2009 , 19.5.2009 [cit. 2009-05-25]. Dostupný z WWW: <[>](http://www.czso.cz/csu/2008edicniplan.nsf/kapitola/3305-08-za_rok_2008-01).
- [6] Český statistický úřad. *Smrtelné pracovní úrazy* [online]. 2006 [cit. 2009-03-23]. Dostupný z WWW: <[>](http://www.czso.cz/csu/redakce.nsf/i/i_6_2_smrtelne_pracovni_urazy).
- [7] DUCHÁČKOVÁ, Eva. *Principy pojištění a pojišťovnictví*. 2. aktualiz. vyd. Praha : Ekopress, 2005. 178 s. ISBN 80-86119-92-0.
- [8] Euromise. *Metody dobývání znalostí* [online]. 2002 [cit. 2009-05-24]. Dostupný z WWW: <[>](http://euromise.vse.cz/kdd/index.php?page=metody#stromy).
- [9] European Commission under the ESPRIT program. *CRISP-DM* [online]. 2007 [cit. 2009-04-26]. Dostupný z WWW: <[>](http://www.crisp-dm.org/Process/index.htm).
- [10] Finance.cz. *Sociální pojištění OSVČ : sazby a vyměřovací základ* [online]. 2000-2009 [cit. 2009-06-07]. Dostupný z WWW: <<http://www.finance.cz/dane-a-mzda/informace/socialni-pojisteni-osvc/sazby/>>.
- [11] JAKUBKA, Jaroslav, et al. *Zákoník práce : Prováděcí nařízení vlády a další související předpisy s komentářem*. 1. vyd. Olomouc : ANAG, 2007. 1023 s. ISBN 978-80-7263-370-8.
- [12] KDnuggets. *Poll: Data Mining Applications in 2008* [online]. 2008 [cit. 2009-08-10]. Dostupný z WWW: <[>](http://www.kdnuggets.com/polls/2008/data-mining-applications.htm).

- [13] Ministerstvo vnitra. *309/2006 Sb. - Zákony - Vyhledávání - Portál veřejné správy České republiky* [online]. 2003-2009 [cit. 2009-03-15]. Dostupný z WWW: <http://portal.gov.cz/wps/portal/_s.155/696/_s.155/701?l=309/2006>.
- [14] *NEJPOUŽÍVANĚJŠÍ METODOLOGIE* [online]. 2008 [cit. 2009-03-31]. Dostupný z WWW: <<http://www1.osu.cz/studium/dozna/crispdm.htm>>.
- [15] OBITKO, Marek. *Učení neuronové sítě : Předpovídání pomocí neuronových sítí* [online]. 1998 [cit. 2006-05-24]. Dostupný z WWW: <<http://www.obitko.com/tutorials/predpovidani-neuronovou-siti/uceni-neuronove-site.html>>.
- [16] PARR RUD, Olivia. *Data mining : Praktický průvodce dolováním dat pro efektivní prodej, cílený marketing a podporu zákazníků (CRM)*. 1. vyd. [s.l.] : [s.n.], 2001. 329 s. ISBN 80-7226-577-6.
- [17] PETR, Pavel. *Data Mining Díl I.* 1. vyd. Pardubice : Univerzita Pardubice, 2006. 144 s. ISBN 80-7194-886-1.
- [18] Podnikatel.cz. *Narřízení vlády* [online]. 2007-2009 [cit. 2009-04-04]. Dostupný z WWW: <<http://www.podnikatel.cz/zakony/narizeni-vlady-c-494-2001-sb-kterym-se-stanovi-zpusob-evidence-hlaseni-a-zasilani-zaznamu-o-urazu-vzor/>>.
- [19] SPSS Inc.. *AnswerTree : klasifikační a rozhodovací stromy* [online]. 2008 [cit. 2009-04-26]. Dostupný z WWW: <http://www.spss.cz/sw_answertree.htm>.
- [20] ŠARMANOVÁ, Jana, LUKASOVÁ, Alena. *Metody shlukové analýzy*. [s.l.] : [s.n.], 1985. 210 s.
- [21] Topinfo. *Zákoník práce a související předpisy* [online]. 2001-2009 [cit. 2009-05-04]. Dostupný z WWW: <<http://www.tzb-info.cz/t.py?t=15&i=539>>.
- [22] VÍTEK. *Metodologie CRISP-DM* [online]. 2002 [cit. 2009-04-01]. Dostupný z WWW: <<http://datamining.xf.cz/view.php?cisloclanku=2002102807>>.
- [23] Výzkumný ústav bezpečnosti práce. *BOZP info - Zákon č. 251/2005 Sb. o inspekci práce* [online]. 2002-2008 [cit. 2008-03-15]. Dostupný z WWW: <http://www.bozpinfo.cz/knihovna-bozp/citarna/tematicke_prilohy/inspekce_prace/zoip251clanek.html>. ISSN 1801-0334.

Seznam tabulek

Tabulka 1 - Přijaté pojistné a vyplacené poj. plnění v rámci odpovědnostního pojištění, upraveno s využitím [7].....	12
Tabulka 2 - Úhrn smrtelných úrazů v krajích a ČR celkem v roce 2008, upraveno s využitím [5]	15
Tabulka 3 - Smrtelné úrazy dle velikosti organizace v roce 2008, upraveno s využitím [5]... ..	15
Tabulka 4 - Smrtelné pracovní úrazy dle OKEČ za celou ČR v roce 2008, upraveno s využitím [5]	17
Tabulka 5 - Smrtelné pracovní úrazy dle OKEČ za kraj Ústecký v roce 2008, upraveno s využitím [5]	17
Tabulka 6 - Největší úmrtnost v OKEČ za kraje v roce 2008, upraveno s využitím [5]	18

Seznam obrázků

Obrázek 1 - Meziroční nárůsty smrtelných pracovních úrazů v Evropě za roky 2000 a 2001, upraveno s využitím [6].....	19
Obrázek 2 – Procentuální nárůst či pokles smrtelných pracovních úrazů v roce 2001 oproti roku 2000 ve vybraných evropských státech, upraveno s využitím [6].....	21
Obrázek 3 - Oblasti s nejvyšším využitím technik DM, upraveno s využitím [12].....	22
Obrázek 4 – SEMMA, upraveno s využitím [1]	24
Obrázek 5 - CRISP-DM, upraveno s využitím [1].....	25
Obrázek 6 - Prvotní data neupravená [zdroj: vlastní]	29
Obrázek 7 - Prvotní data doplněná o názvy sloupců (atributů) [zdroj: vlastní]	30
Obrázek 8 - Počty a procentuální vyjádření smrtelných pracovních úrazů v letech 2002 - 2007, [zdroj: vlastní]	31
Obrázek 9 - Celkový počet smrtelných pracovních úrazů u atributu Vek za roky 2002 - 2007, [zdroj:vlastní]	32
Obrázek 10 - Délka zaměstnání v letech a jejich četnost pro roky 2002 – 2007, [zdroj: vlastní]	33
Obrázek 11 - Počty úmrtí dle státní příslušnosti v letech 2002 – 2007, [zdroj: vlastní]	34
Obrázek 12 - Počet zraněných osob a jejich četnost v letech 2002 - 2007, [zdroj: vlastní].....	35
Obrázek 13 - Počet zraněných osob u smrtelného pracovního úrazu v letech 2002 - 2007, [zdroj: vlastní]	36
Obrázek 14 - Četnost smrtelných úrazů v daném roce v letech 2002 - 2007, [zdroj: vlastní].	37
Obrázek 15 - Četnost smrtelných úrazů v letech 2008 - 2007, [zdroj: vlastní].....	37
Obrázek 16 - Čas úrazu v třisměnném provozu v letech 2002 - 2007, [zdroj: vlastní].....	38
Obrázek 17 – Čas úrazu v dvousměnném provozu v letech 2002 – 2007, [zdroj: vlastní].....	39
Obrázek 18 - Smrtelné pracovní úrazy dle kategorie OKEČ v letech 2002 – 2007, [zdroj: vlastní]	40
Obrázek 19 - Smrtelné úrazy dle OKEČ za jednotlivé roky 2002 - 2007, [zdroj: vlastní]	40
Obrázek 20 - Upravená data pro práci [zdroj: vlastní].....	44
Obrázek 21 - Číselník pro atribut Udalost_1 [zdroj: vlastní].....	48
Obrázek 22 - Číselník pro vytvořený atribut Udalost [zdroj: vlastní].....	49
Obrázek 23 - Stream pro tvorbu seskupovací analýzy, [zdroj: vlastní]	50
Obrázek 24 - Výstup z algoritmu "Kohonenova mapa" znázorňující počet shluků, [zdroj: vlastní]	51
Obrázek 25 - Četnost zastoupení shluků v algoritmu Kohonenova mapa, [zdroj: vlastní].....	51
Obrázek 26 - Atributy v nejčetnějším shluku algoritmu Kohonenova mapa, [zdroj: vlastní].	52
Obrázek 27 - Počet jednotlivých shluků u metody K - Means, [zdroj: vlastní].....	53
Obrázek 28 - Atributy v nejčetnějším shluku algoritmu K - Means, [zdroj: vlastní]	53
Obrázek 29 - Počet shluků algoritmu Two Step s četnostní údajů, [zdroj: vlastní]	54
Obrázek 30 - Atributy v nejčetnějším shluku algoritmu Two Step, [zdroj: vlastní]	55

Seznam příloh

Příloha č. 1 Datový slovník

Datový slovník

OrganDozoru

- popis atributu: orgán státní správy dohlížející na dodržování BOZP,
- možnosti: zjištění orgánu státní správy,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 3, 4, 5, 6, 7, 8, 9, 10, 20, 40.

PravniForma_1

- popis atributu: uspořádání ekonomického subjektu dle právní formy,
- možnosti: zjištění typu ekonomického subjektu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

PravniForma_2

- popis atributu: uspořádání ekonomického subjektu dle právní formy,
- možnosti: zjištění typu ekonomického subjektu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

PravniForma_3

- popis atributu: uspořádání ekonomického subjektu dle právní formy,
- možnosti: zjištění typu ekonomického subjektu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 7.

katOKEC1

- popis atributu: kategorie Odvětvová klasifikace ekonomické činnosti,
- možnosti: zjištění kategorie,
- klasifikace: kvalitativní,
- datový typ: textový řetězec,
- nabývá hodnot: A – O.

katOKEC2

- popis atributu: kategorie Odvětvová klasifikace ekonomické činnosti,
- možnosti: zjištění kategorie,
- klasifikace: kvalitativní,
- datový typ: textový řetězec,
- nabývá hodnot: A – N.

oddilOKEC3

- popis atributu: kategorie Odvětvová klasifikace ekonomické činnosti,
- možnosti: zjištění kategorie,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1, 2, 5, 10, 14, 15, 17 – 37, 40, 41, 45, 50 – 52, 55, 60, 62 – 66, 70 – 75, 80, 85, 90 – 93.

OKEC_4

- popis atributu: kategorie Odvětvová klasifikace ekonomické činnosti,
- možnosti: zjištění kategorie,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

OKEC_5

- popis atributu: kategorie Odvětvová klasifikace ekonomické činnosti,
- možnosti: zjištění kategorie,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

UzemniJednotka_1

- popis atributu: stát, kde ekonomický subjekt prováděl práci,
- klasifikace: kvalitativní,
- datový typ: textový řetězec,
- nabývá hodnot: CZ, SK.

UzemniJednotka_2

- popis atributu: pouze doplňuje soustavu pro upřesnění územní jednotky,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0.

UzemniJednotka_3

- popis atributu: kraj, kde ekonomický subjekt pracoval,
- možnosti: zjištění kraje,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

UzemniJednotka_4

- popis atributu: kraj, kde ekonomický subjekt pracoval,
- možnosti: upřesnění kraje,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 3.

UzemniJednotka_5

- popis atributu: okres, kde ekonomický subjekt pracoval,
- možnosti: zjištění okresu,
- klasifikace: kvantitativní i kvalitativní,
- datový typ: celé číslo, textový řetězec,
- nabývá hodnot: 0 – 9, A – F.

VelikostPodniku_1

- popis atributu: velikost podniku dle počtu zaměstnanců,
- možnosti: zjištění počtu zaměstnanců,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 5,
- kódování: 1 = 1 – 9 zaměstnanců, 2 = 10 – 99 zaměstnanců, 3 = 100 – 999 zaměstnanců, 4 = 1.000 – 9.999 zaměstnanců, 5 = 10.000 a více zaměstnanců.

VelikostPodniku_2

- popis atributu: velikost podniku dle počtu zaměstnanců,
- možnosti: zjištění počtu zaměstnanců,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 10, 20, 30, 40, 50, 60, 70.

Fu_UzemniJednotka_1

- popis atributu: země odkud ekonomický subjekt pochází,
- možnosti: zjištění státu,
- klasifikace: kvalitativní,
- datový typ: textový řetězec
- nabývá hodnot: BY, CZ, DE, ES, HR, HU, IT, JE, MN, PL, LO, RU, SK, SR, UA.

Fu_UzemniJednotka_2

- popis atributu: země odkud ekonomický subjekt pochází,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0.

Fu_UzemniJednotka_3

- popis atributu: kraj, kde ekonomický subjekt podniká v ČR,
- možnosti: zjištění kraje,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Fu_UzemniJednotka_4

- popis atributu: kraj, kde ekonomický subjekt podniká v ČR,
- možnosti: upřesnění kraje,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 3.

Fu_UzemniJednotka_5

- popis atributu: okres, kde ekonomický subjekt podniká v ČR,
- možnosti: zjištění okresu,
- klasifikace: kvantitativní, kvalitativní,
- datový typ: celé číslo, textový řetězec,
- nabývá hodnot: 0 – 9, A – F.

Sex

- popis atributu: pohlaví,
- možnosti: zjištění pohlaví,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 2,
- kódování: 1 = M, 2 – Ž.

Vek

- popis atributu: věk pracovníka,
- možnosti: zjištění věku pracovníka,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 16 – 73.

StatniPrislusnost

- popis atributu: národnost pracovníka,
- možnosti: zjištění národnosti pracovníka,
- klasifikace: kvalitativní,
- datový typ: textový řetězec,
- nabývá hodnot: BG, BY, CZ, PL, RU, SK, UA.

Zamestnani_1

- popis atributu: pracovní pozice,
- možnosti: zjištění pracovní pozice zaměstnance,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9,
- kódování: 1 = zákonodárci, vedoucí a řídicí pracovníci, 2 = vědečtí a odborní duševní pracovníci, 3 = techničtí, zdravotničtí, pedagogičtí pracovníci a pracovníci v příbuzných oborech, 4 = nižší administrativní pracovníci, 5 = provozní pracovníci ve službách a obchodě, 6 = kvalifikovaní dělníci v zemědělství, lesnictví a příbuzných oborech, 7 = řemeslníci a kvalifikovaní výrobci, zpracovatelé, opraváři, 8 = obsluha strojů a zařízení, 9 = pomocní a nekvalifikovaní pracovníci.

Zamestnani_2

- popis atributu: upřesnění pracovní pozice,
- možnosti: zjištění pracovní pozice zaměstnance,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 4.

Zamestnani_3

- popis atributu: upřesnění pracovní pozice,
- možnosti: zjištění pracovní pozice zaměstnance,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Zamestnani_4

- popis atributu: upřesnění pracovní pozice,
- možnosti: zjištění pracovní pozice zaměstnance,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Zamestnani_5

- popis atributu: upřesnění pracovní pozice,
- možnosti: zjištění pracovní pozice zaměstnance,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Postaveni_1

- popis atributu: pracovní vztah,
- možnosti: zjištění pracovního vztahu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 6,
- kódování: 1 = zaměstnanci, 2 = zaměstnavatelé, 3 = osoby pracující na vlastní účet, 4 = členové produkčních družstev, 5 = pomáhající rodinní příslušníci, 6 = osoby nekvalifikované podle postavení v zaměstnání.

Postaveni_2

- popis atributu: pracovní vztah (upřesnění),
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 4.

DelkaZamestnani_R

- popis atributu: délka pracovního vztahu v rocích,
- možnosti: zjištění počtu odpracovaných let u dané organizace,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 42.

DelkaZamestnani_M

- popis atributu: délka pracovního vztahu v měsících,
- možnosti: doplnění atributu DelkaZamestani_R o počet měsíců,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 11.

ZranenychOsob

- popis atributu: počet zraněných osob,
- možnosti: udává počet zraněných osob při smrtelném pracovním úrazu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 40.

DatumUrazu

- popis atributu: datum úrazu,
- možnosti: zjištění přesného data úrazu,
- klasifikace: kvantitativní,
- datový typ: datum,
- nabývá hodnot: 29.11.2001 – 28.12.2007.

RokUmrti

- popis atributu: rok úmrtí,
- možnosti: zjištění roku úmrtí,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 2 – 7,
- kódování: 2 = 2002, 3 = 2003, 4 = 2004, 5 = 2005, 6 = 2006, 7 = 2007.

CasUrahu_H

- popis atributu: čas úrazu v hodinách,
- možnosti: zjištění kdy se úraz stal,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 23.

CasUrahu_M

- popis atributu: čas úrazu v minutách,
- možnosti: zjištění kdy se úraz stal, doplňuje atribut CasUrahu_H o minuty,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 59.

OdpracovanyCas_H

- popis atributu: odpracovaný čas v hodinách,
- možnosti: zjištění jak dlouho zaměstnanec pracoval, než nastal úraz,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 23.

OdpracovanyCas_M

- popis atributu: odpracovaný čas v minutách,
- možnosti: doplnění atributu OdpracovanyCas_H o počet minut,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 59.

DruhZraneni_1

- popis atributu: druh zranění,
- možnosti: zjištění o jaký druh zranění se jedná,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9,
- kódování: 1 = zlomeniny, vymknutí, podvrtnutí, natažení kloubů a vazů, 2 = poranění mozku, míchy a vnitřních orgánů, 3 = zhmoždění, otevřené rány, rozdrčení, amputace, 4 = popálení, opaření, poleptání, omrznutí a účinky záření, horka/chladu, 5 = otravy a toxické účinky látek, dušení a účinky tlaku, tonutí, 6 = účinky elektrického proudu, blesku, 7 = následky námahy, šok, cizí těleso, 8 = poranění mnohočetná, 9 = jiná zranění.

DruhZraneni_2

- popis atributu: druh zranění,
- možnosti: upřesnění atributu DruhZraneni_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

DruhZraneni_3

- popis atributu: druh zranění,
- možnosti: upřesnění atributu DruhZraneni_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

ZranenaCast_1

- popis atributu: zraněná část těla,
- možnosti: zjištění, na které části těla byl pracovník zraněn,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 8,
- kódování: 1 = část těla (blíže neurčeno), 2 = krk, hrudník, břicho, pánev, 3 = horní končetina, 4 = dolní končetina, 5 = více částí těla, 6 = ostatní části těla dříve neuvedené.

ZranenaCast_2

- popis atributu: zraněná část těla,
- možnosti: upřesnění atributu ZranenaCast_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 4, 8, 9.

ZranenaCast_3

- popis atributu: zraněná část těla,
- možnosti: upřesnění atributu ZranenaCast_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 4, 8, 9.

ZranenaCast_4

- popis atributu: zraněná část těla,
- možnosti: upřesnění atributu ZranenaCast_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 5, 7 – 9.

Pracoviste

- popis atributu: pracoviště,
- možnosti: pracoviště, kde má pracovník výkon práce,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 2,
- kódování: 1 = obvyklé pracoviště nebo pracoviště v provozovně, kde obvykle pracuje, 2 = příležitostné nebo pohyblivé (mobilní) pracoviště nebo úraz na pracovní cestě.

Misto_1

- popis atributu: místo na pracovišti,
- možnosti: místo, kde se pracovník pohyboval, když nasral pracovní úraz,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 8,
- kódování: 1 = prostory pro výrobu, údržbu, skladování, 2 = staveniště, povrchové doly, lomy, pískovny, 3 = prostory a objekty pro zemědělství a lesnictví, pole a terén, 4 = komunikace, speciální prostory určené pro dopravu, 5 = prostory obchodů, služeb, kulturních, sportovních aj. zařízení a obytných domů,

6 = prostory pro inženýrské, administrativní práce, ve zdravotnických a školských zařízeních, 7 = místa ve výšce, pod zemí, pod vodou, 8 = prostory v dopravních prostředcích, strojích a zařízeních.

Misto_2

- popis atributu: místo na pracovišti,
- možnosti: upřesnění atributu Misto_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Misto_3

- popis atributu: místo na pracovišti,
- možnosti: upřesnění atributu Misto_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Misto_4

- popis atributu: místo na pracovišti,
- možnosti: upřesnění atributu Misto_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 10 – 18, 21 – 28, 30 – 36, 41 – 45, 51, 53, 54, 64, 72, 80, 81.

Činnost_1

- popis atributu: činnost, při které se úraz stal,
- možnosti: zjištění činnosti,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9,
- kódování: 1 = práce v zemědělství, lesnictví, chov ryb a jiných živočichů, 2 = činnost při těžbě surovin a energetice, 3 = výroba, zpracování, 4 = montáž, demontáž, opravy, údržba a revize, 5 = stavební práce, 6 = doprava, skladování, pomocné práce, ostraha, úklid, 7 = činnost spojená se službami, sportem a kulturou, 8 = pohyb, duševní činnost a pasivní účast, 9 = jiné.

Cinnost_2

- popis atributu: činnost, při které se úraz stal,
- možnosti: upřesnění atributu Cinnost_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Cinnost_3

- popis atributu: činnost, při které se úraz stal,
- možnosti: upřesnění atributu Cinnost_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Cinnost_4

- popis atributu: činnost, při které se úraz stal,
- možnosti: upřesnění atributu Cinnost_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Ukon_1

- popis atributu: úkon, který pracovník prováděl při vzniku pracovního úrazu,
- možnosti: zjištění pracovního úkonu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 8,
- kódování: 1 = obsluha strojů a zařízení, 2 = práce s náradím, náčiním, 3 = chůze, výstup a jiný pohyb, stravování, 4 = ruční manipulace s břemeny, práce bez pomoci náradí, 5 = práce s výbušninami, v zamořeném a nebezpečném prostoru, 6 = přítomnost v blízkosti zdroj, 7 = ošetřování lidí, kosmetika (i pasivní účast), 8 = ošetřování živočichů a jiná práce s nimi.

Ukon_2

- popis atributu: úkon, který pracovník prováděl při vzniku pracovního úrazu,
- možnosti: upřesnění atributu Ukon_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 5.

Ukon_3

- popis atributu: úkon, který pracovník prováděl při vzniku pracovního úrazu,
- možnosti: upřesnění atributu Ukon_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Ukon_4

- popis atributu: úkon, který pracovník prováděl při vzniku pracovního úrazu,
- možnosti: upřesnění atributu Ukon_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Ukon_5

- popis atributu: úkon, který pracovník prováděl při vzniku pracovního úrazu,
- možnosti: upřesnění atributu Ukon_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Ukon_6

- popis atributu: úkon, který pracovník prováděl při vzniku pracovního úrazu,
- možnosti: upřesnění atributu Ukon_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Cinitel_Zdroj_1

- popis atributu: pracovní pomůcka či zdroj pracovního úrazu,
- možnosti: zjištění zdroje (původu) pracovního úrazu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9,
- kódování: 1 = dopravní stroje, 2 = stroje hnací a vložená zařízení, 3 = stroje výrobní, měřicí, zkušební, zbraně, 4 = nářadí, náčiní a nástroje, mechanizované ruční nářadí a strojky, pracovní pomůcky, 5 = terén, přírodní předměty, budovy, stavební objekty, 6 = pevné materiály a předměty, 7 = chemikálie, výbušniny, tekuté a plynné látky, 8 = plamen, statická elektřina, záření, zvuk, 9 = lidé, živočichové, přírodní vlivy.

Cinitel_Zdroj_2

- popis atributu: pracovní pomůcka či zdroj pracovního úrazu,
- možnosti: upřesnění atributu Cinitel_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Cinitel_Zdroj_3

- popis atributu: pracovní pomůcka či zdroj pracovního úrazu,
- možnosti: upřesnění atributu Cinitel_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 11, 12, 14 – 16, 21 – 23, 25, 27, 31, 32, 35, 37, 41 – 43, 51.

Cinitel_Zdroj_4

- popis atributu: pracovní pomůcka či zdroj pracovního úrazu,
- možnosti: upřesnění atributu Cinitel_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 11 – 18, 21 – 27, 30 – 34, 36 – 38, 41 – 45, 51 – 55, 61, 71, 72, 74, 81, 82, 84, 91.

Cinitel_Zdroj_5

- popis atributu: pracovní pomůcka či zdroj pracovního úrazu,
- možnosti: upřesnění atributu Cinitel_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 10 – 16, 20 – 25, 28, 31 – 35, 37, 41, 42, 45, 51, 71.

Nehoda_1

- popis atributu: čím byl daný úraz způsoben,
- možnosti: zjištění příčiny úrazu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9,
- kódování: 1 = havárie, exploze, 2 = zachycení, zasažení, udeření osoby, 3 = popálení, poleptání, ozáření, 4 = pád předmětu, přiražení, 5 = pád osoby, 6 = náraz, namožení, vdechnutí/požítí toxické látky, infarkt, 7 = nehoda způsobená jinou osobou nebo živočichem, 8 = účinky počasí, živlů, prostředí, tonutí, 9 = jiný způsob.

Nehoda_2

- popis atributu: čím byl daný úraz způsoben,
- možnosti: upřesnění atributu Nehoda_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

Nehoda_3

- popis atributu: čím byl daný úraz způsoben,
- možnosti: upřesnění atributu Nehoda_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 11 – 18, 21 – 27, 31 – 33, 41, 43, 51, 52, 55.

Nehoda_4

- popis atributu: čím byl daný úraz způsoben,
- možnosti: upřesnění atributu Nehoda_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 6, 9.

Nehoda_5

- popis atributu: čím byl daný úraz způsoben,
- možnosti: upřesnění atributu Nehoda_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Udalost_1

- popis atributu: jak, jakým způsobem se nehoda udála,
- možnosti: zjištění průběhu úrazu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 83.

Udalost_2

- popis atributu: jak, jakým způsobem se nehoda udála,
- možnosti: upřesnění atributu Udalost_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 7.

Udalost_3

- popis atributu: jak, jakým způsobem se nehoda udála,
- možnosti: upřesnění atributu Udalost_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Nehoda_Zdroj_1

- popis atributu: nehoda a zdroj úrazu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9.

Nehoda_Zdroj_2

- popis atributu: nehoda a zdroj úrazu,
- možnosti: upřesnění atributu Nehoda_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 8.

Nehoda_Zdroj_3

- popis atributu: nehoda a zdroj úrazu,
- možnosti: upřesnění atributu Nehoda_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 11 – 16, 21 – 25, 27, 31, 32, 35, 37, 41 – 43, 51.

Nehoda_Zdroj_4

- popis atributu: nehoda a zdroj úrazu,
- možnosti: upřesnění atributu Nehoda_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 10 – 18, 21 – 23, 25 – 28, 30 – 38, 41 – 45, 51 – 52, 55, 61, 81 – 84, 91.

Nehoda_Zdroj_5

- popis atributu: nehoda a zdroj úrazu,
- možnosti: upřesnění atributu Nehoda_Zdroj_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 11 – 18, 20 – 23, 25, 26, 28, 31 – 33, 36, 37, 41, 42, 45, 51, 52, 71, 81.

P1_Pricina_1

- popis atributu: důvod vzniku úrazu,
- možnosti: zjištění důvodu vzniku pracovního úrazu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9,
- kódování: 1 = nebezpečné pracovní podmínky, 2 = závady v pracovním prostředí, 3 = chybějící osobní ochranné prostředky, špatné ustrojení, 4 = nesprávné a nebezpečné chování zraněného, 5 = nesprávné a nebezpečné jednání jiných osob, 6 = nedostatek předpokladů zraněného pro provádění daného úkonu, 7 = nedostatek předpokladů u jiných osob pro provádění úkonu, 8 = neovlivnitelné okolnosti (živočich, přírodní účinky, autonehoda), 9 = jiné.

P1_Pricina_2

- popis atributu: důvod vzniku úrazu,
- možnosti: upřesnění atributu P1_Pricina_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9.

P1_Pricina_3

- popis atributu: důvod vzniku úrazu,
- možnosti: upřesnění atributu P1_Pricina_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

P1_Pricina_4

- popis atributu: důvod vzniku úrazu,
- možnosti: upřesnění atributu P1_Pricina_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 10 – 18, 20 – 24, 99.

P2_Pricina_1

- popis atributu: další důvody vzniku úrazu než P1_Pricina,
- možnosti: upřesnění atributu P1_Pricina,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9.

P2_Pricina_2

- popis atributu: další důvody vzniku úrazu než P1_Pricina,
- možnosti: upřesnění atributu P1_Pricina,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9.

P2_Pricina_3

- popis atributu: další důvody vzniku úrazu než P1_Pricina,
- možnosti: upřesnění atributu P1_Pricina,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 9.

P2_Pricina_4

- popis atributu: další důvody vzniku úrazu než P1_Pricina,
- možnosti: upřesnění atributu P1_Pricina,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 10 – 18, 20 – 25, 31, 99.

DruhPredpisu_1

- popis atributu: druh předpisu (právního či jiného),
- možnosti: zjištění, který předpis byl porušen při pracovním úrazu,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 4,
- kódování: 0 = předpis neuveden, 1 = zákon, 2 = prováděcí předpis, 3 = interní předpis, 4 = technická norma.

DruhPredpisu_2

- popis atributu: druh předpisu (právního či jiného),
- možnosti: upřesnění atributu DruhPredpisu_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0 – 4.

DruhPredpisu_3

- popis atributu: druh předpisu (právního či jiného),
- možnosti: upřesnění atributu DruhPredpisu_1,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 0, 11 – 14, 21, 26.

Citace

- popis atributu: přesné znění druhu předpisu, číslo zákona,
- klasifikace: kvantitativní, kvalitativní,
- datový typ: celé číslo, textový řetězec,
- chybějící hodnoty: 56.

Ustanoveni_Paragraf

- popis atributu: číslo paragrafu,
- možnosti: upřesnění atributu Citace,
- klasifikace: kvantitativní,
- chybějící hodnoty: 56.

Ustanoveni_Odstavec

- popis atributu: číslo odstavce,
- možnosti: upřesnění atributu Citace,
- klasifikace: kvantitativní,
- chybějící hodnoty: 56.

Ustanoveni_Pismeno

- popis atributu: číslo písmene odstavce,
- možnosti: upřesnění atributu Citace,
- klasifikace: kvantitativní,
- chybějící hodnoty: 56.

Clanek_Bod

- popis atributu: článek,
- možnosti: upřesnění atributu Citace,
- klasifikace: kvantitativní,
- chybějící hodnoty: 56.

Clanek_Pismeno

- popis atributu: článek,
- možnosti: upřesnění atributu Citace,
- klasifikace: kvantitativní,
- chybějící hodnoty: 56.

Porusitel

- popis atributu: porušitel předpisu,
- možnosti: zjištění, která osoba porušila daný předpis, a tím byl úraz zapříčiněn,
- klasifikace: kvantitativní,
- datový typ: celé číslo,
- nabývá hodnot: 1 – 9,
- kódování: 1 = porušeno zraněným, 2 = por. jinou osobou, 3 = por. zaměstnavatelem, 4 = por. zraněným i zaměstnavatelem, 5 = por. cizí firmou (dodavatelskou), 6 = por. zraněným a jinou osobou, 7 = por. zaměstnavatelem a

cizí firmou, 8 = por. zraněným, zaměstnavatelem a cizí firmou, 9 = porušení
nebylo zjištěno.