

Závislost dvou kvalitativních proměnných

Martina Litschmannová

Motivační příklad



Pro diferencovaný přístup v personální politice potřebuje vedení podniku vědět, zda spokojenost v práci závisí na tom, jedná-li se o pražský závod či závody mimopražské. Šetření se účastnilo 100 pracovníků z Prahy a 200 pracovníků z venkova. Výsledky šetření jsou v následující tabulce.

místo/stupeň spokojenosti	velmi nespokojen	spíše nespokojen	spíše spokojen	velmi spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

Pomocí explorační analýzy analyzujte výsledky šetření.

V jakém formátu obvykle získáváme tento typ dat?



Místo	Stupeň spokojenosti
Praha	velmi spokojen
Praha	spíše spokojen
Venkov	spíše nespokojen
Praha	spíše spokojen
Venkov	velmi spokojen
⋮	⋮
Venkov	spíše spokojen



Tento převod lze provést pomocí většiny tabulkových procesorů i statistického software.

Standardní datový formát

místo/stupeň spokojenosti	velmi nespokojen	spíše nespokojen	spíše spokojen	velmi spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

Kontingenční tabulka

Co je to kontingenční tabulka?



$X \setminus Y$	$y_{[1]}$	$y_{[2]}$	\dots	$y_{[s]}$	Celkem
$x_{[1]}$	n_{11}	n_{12}	\dots	n_{1s}	$n_{1\cdot}$
$x_{[2]}$	n_{21}	n_{22}	\dots	n_{2s}	$n_{2\cdot}$
\vdots	\vdots	\vdots	\dots	\vdots	\vdots
$x_{[r]}$	n_{r1}	n_{r2}	\dots	n_{rs}	$n_{r\cdot}$
Celkem	$n_{\cdot 1}$	$n_{\cdot 2}$	\dots	$n_{\cdot s}$	n

Schéma rozšířené kontingenční tabulky

Dvourozměrná tabulka četností,

z jejichž hodnot můžeme usoudit na závislost či nezávislost mezi dvěma kategoriálními proměnnými.

Jak posoudit míru závislosti mezi dvěma kvalitativními proměnnými pomocí metod explorační analýzy?



- Grafická analýza
 - ✓ shlukový sloupcový graf,
 - ✓ skládaný sloupcový graf,
 - ✓ mozaikový graf,
 - ✓ 100% skládaný sloupcový (pruhový) graf

Motivační příklad



Pro diferencovaný přístup v personální politice potřebuje vedení podniku vědět, zda spokojenost v práci závisí na tom, jedná-li se o pražský závod či závody mimopražské. Šetření se účastnilo 100 pracovníků z Prahy a 200 pracovníků z venkova. Výsledky šetření jsou v následující tabulce.

místo/stupeň spokojenosti	velmi nespokojen	spíše nespokojen	spíše spokojen	velmi spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

Pomocí explorační analýzy analyzujte výsledky šetření.



	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.03333333	0.08333333	0.16666667	0.05000000
Venkov	0.06666667	0.03333333	0.43333333	0.13333333

sdužené relativní četnosti

- Sdužené relativní četnosti: $\frac{n_{ij}}{n}$
- Součet všech sdužených relativních četností je roven 1.

Interpretace:

Např. Mezi 300 respondenty je 10 (3 %) velmi nespokojených pracovníků z podniku alokovaného v Praze.



	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.1	0.25	0.50	0.15
venkov	0.1	0.05	0.65	0.20

řádkové relativní četnosti

- Řádkové relativní četnosti: $\frac{n_{ij}}{n_{i.}}$
- Součet řádkových relativních četností **odpovídajících jednomu řádku** kontingenční tabulky je roven 1.

Interpretace:

Např. Ze 100 respondentů, kteří pracují v podniku alokovaném v Praze, je 10 (10 %) velmi nespokojeno s prací.



	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
venkov	20	10	130	40

pozorované četnosti

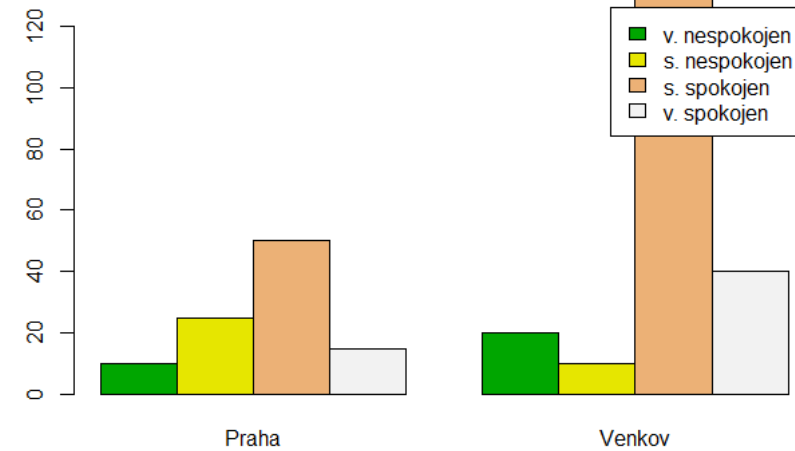
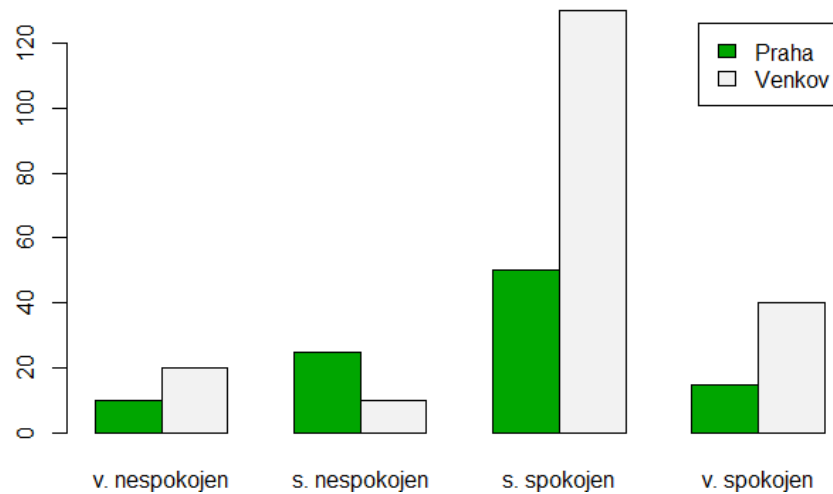
	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.3333333	0.7142857	0.2777778	0.2727273
venkov	0.6666667	0.2857143	0.7222222	0.7272727

sloupcové relativní četnosti

- Sloupcové relativní četnosti: $\frac{n_{ij}}{n_{.j}}$
- Součet sloupcových relativních četností **odpovídajících jednomu sloupci** kontingenční tabulky je roven 1.

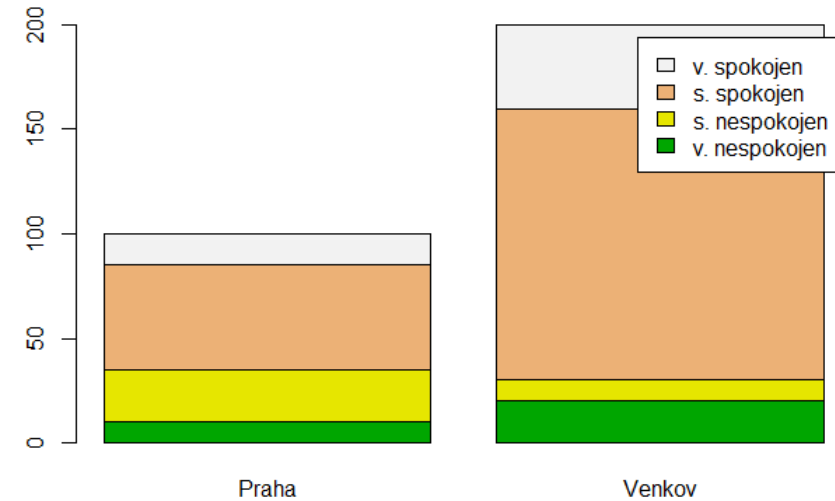
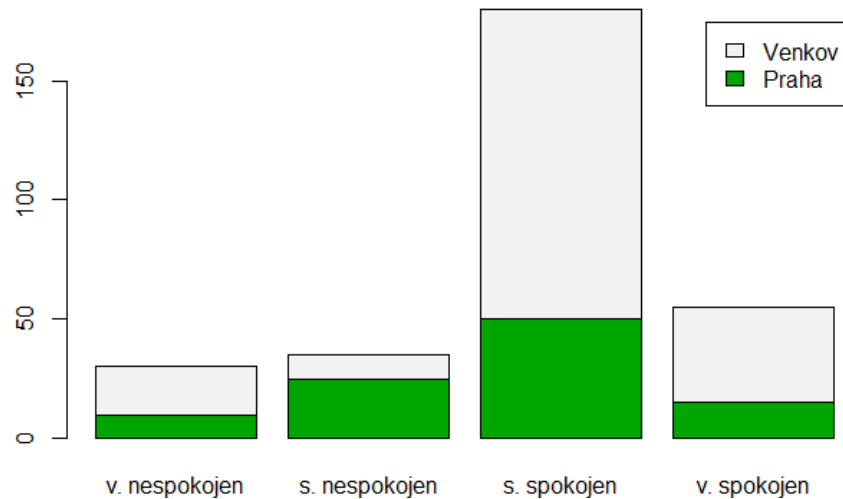
Interpretace:

Např. Ze 30 respondentů, kteří jsou velmi nespokojeni s prací, 10 (33 %) pracuje v podniku alokované v Praze.



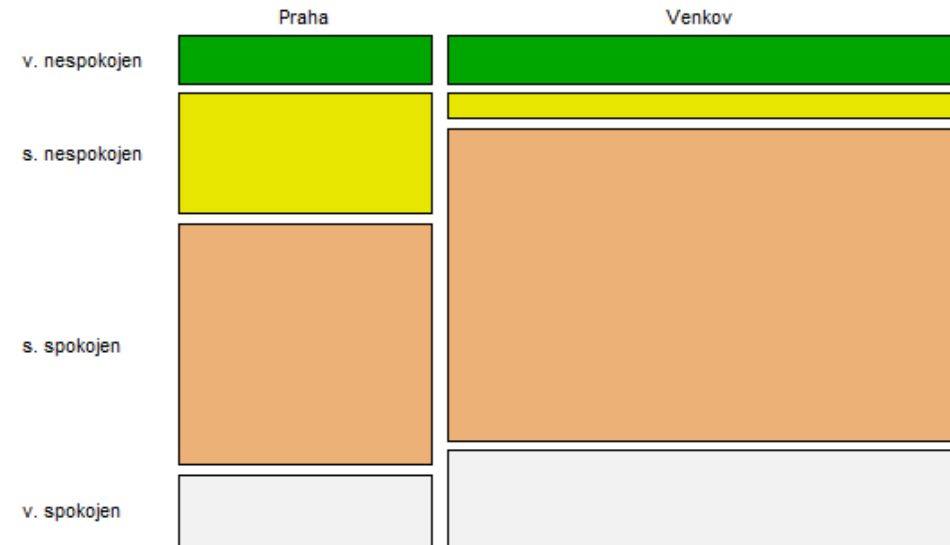
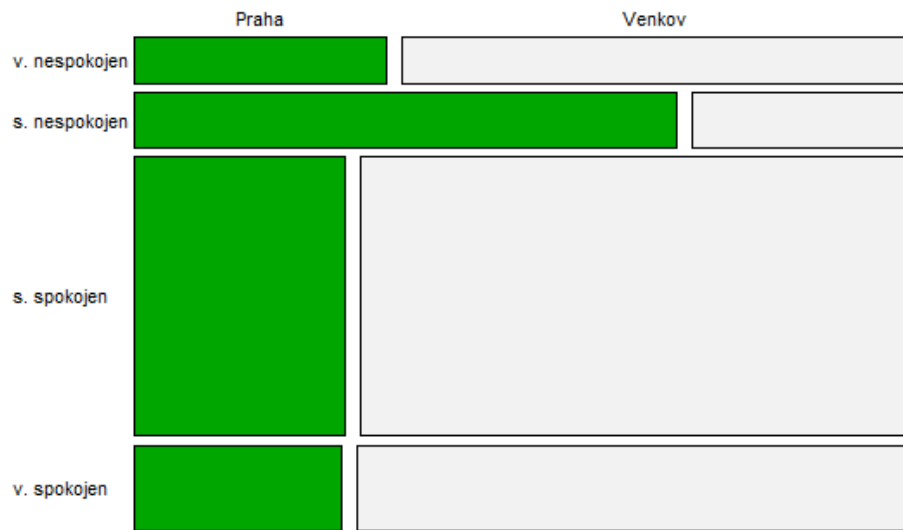
Shlukový sloupcový graf

Rozmyslete si, který z grafů byste pro vizualizaci použili.



Skládaný sloupcový graf

Rozmyslete si, který z grafů byste pro vizualizaci použili.



Mozaikový graf

Rozmyslete si, který z grafů byste pro vizualizaci použili.

Explorační analýza pomocí R – vizualizace

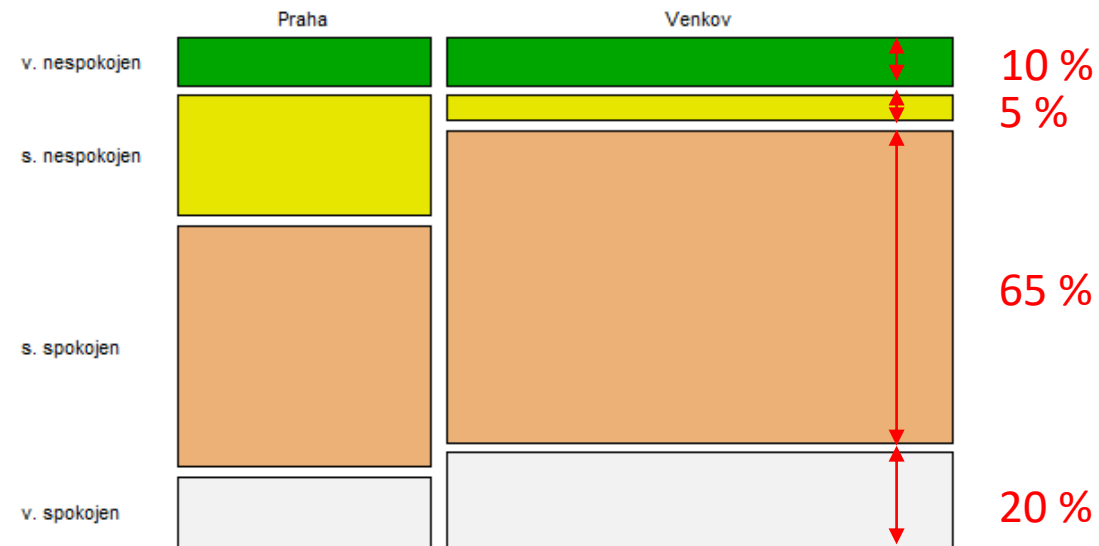


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.1	0.25	0.50	0.15
venkov	0.1	0.05	0.65	0.20

řádkové relativní četnosti



Mozaikový graf

Explorační analýza pomocí R – vizualizace

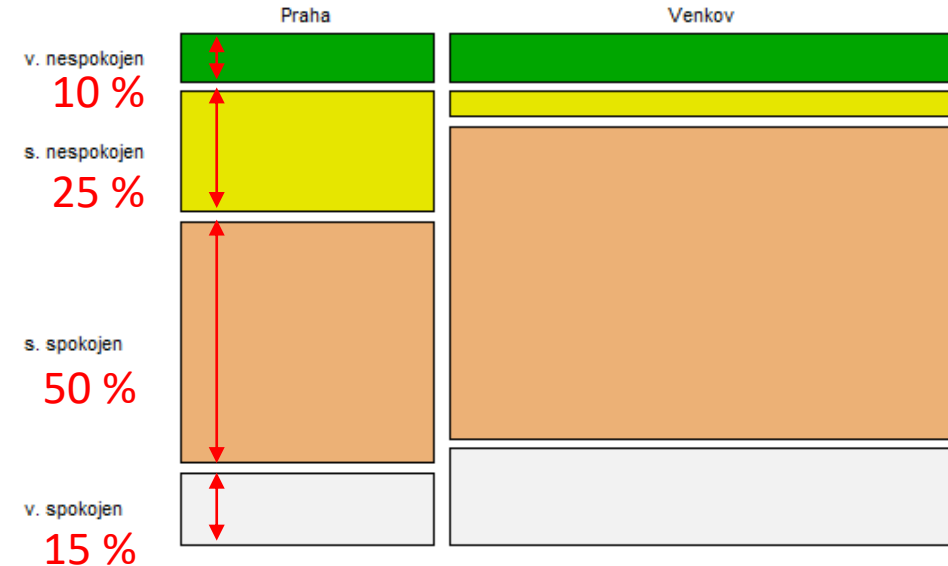


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.1	0.25	0.50	0.15
venkov	0.1	0.05	0.65	0.20

řádkové relativní četnosti



Mozaikový graf

Explorační analýza pomocí R – vizualizace

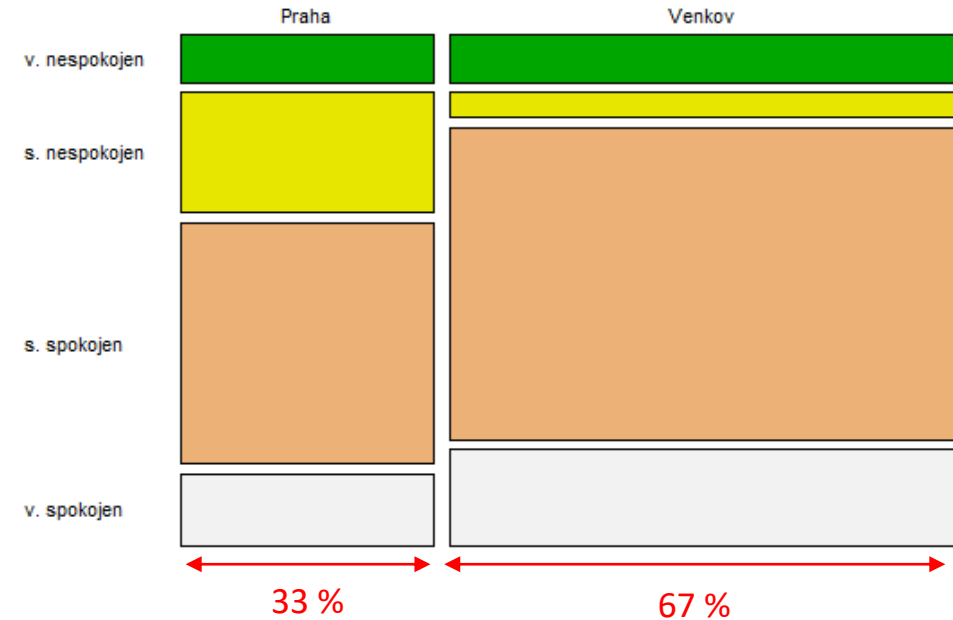


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen	Celkem
Praha	10	25	50	15	100 (33 %)
Venkov	20	10	130	40	200 (67 %)

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.1	0.25	0.50	0.15
Venkov	0.1	0.05	0.65	0.20

řádkové relativní četnosti



Mozaikový graf

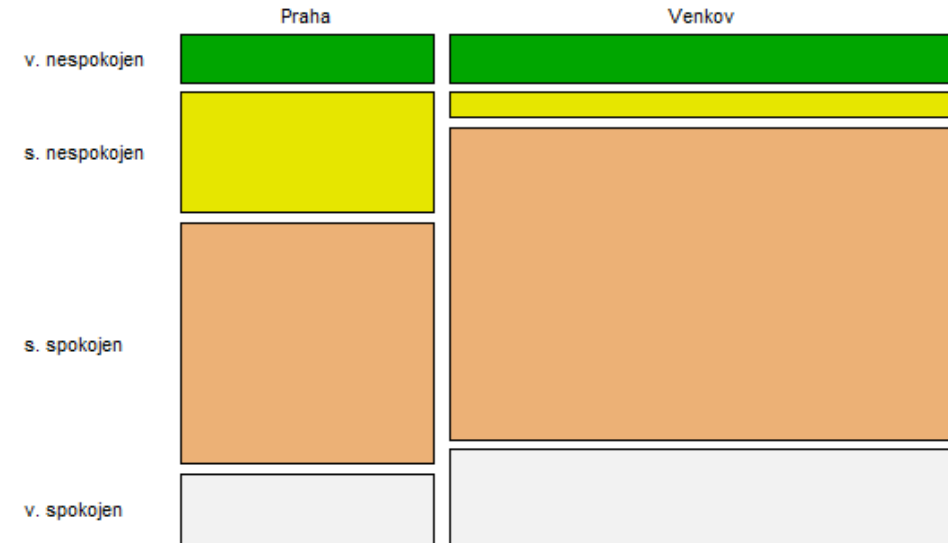


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.1	0.25	0.50	0.15
Venkov	0.1	0.05	0.65	0.20

řádkové relativní četnosti



Mozaikový graf

Jak by graf vypadal kdyby spokojenost v práci nesouvisela s tím, kde je alokván podnik, v němž respondent pracuje?

Na grafu bychom viděli „vodorovné barevné pásy“. Procentuální zastoupení respondentů dle úrovně spokojenosti v zaměstnání by bylo pro „Prahu“ i „Venkov“ stejné.

Čím členitější „mozaiku“ vidíme, tím silnější závislost mezi kvalitativními znaky vnímáme.

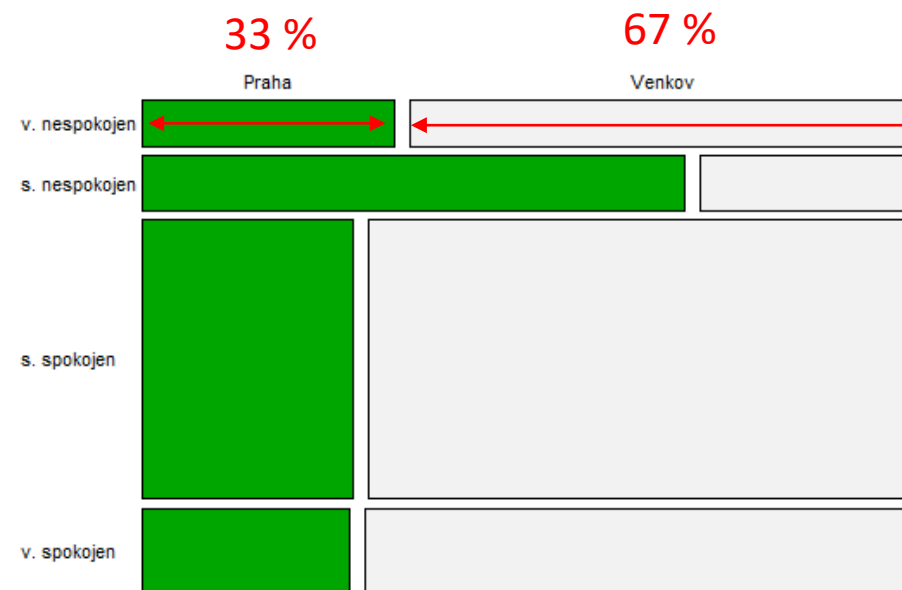


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.3333333	0.7142857	0.2777778	0.2727273
Venkov	0.6666667	0.2857143	0.7222222	0.7272727

sloupcové relativní četnosti



Mozaikový graf

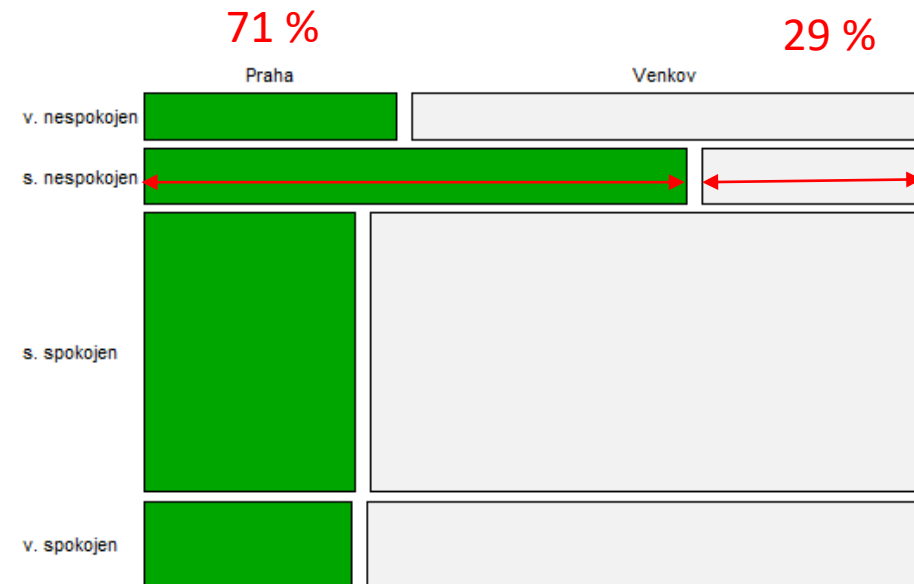


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.3333333	0.7142857	0.2777778	0.2727273
Venkov	0.6666667	0.2857143	0.7222222	0.7272727

sloupcové relativní četnosti



Mozaikový graf

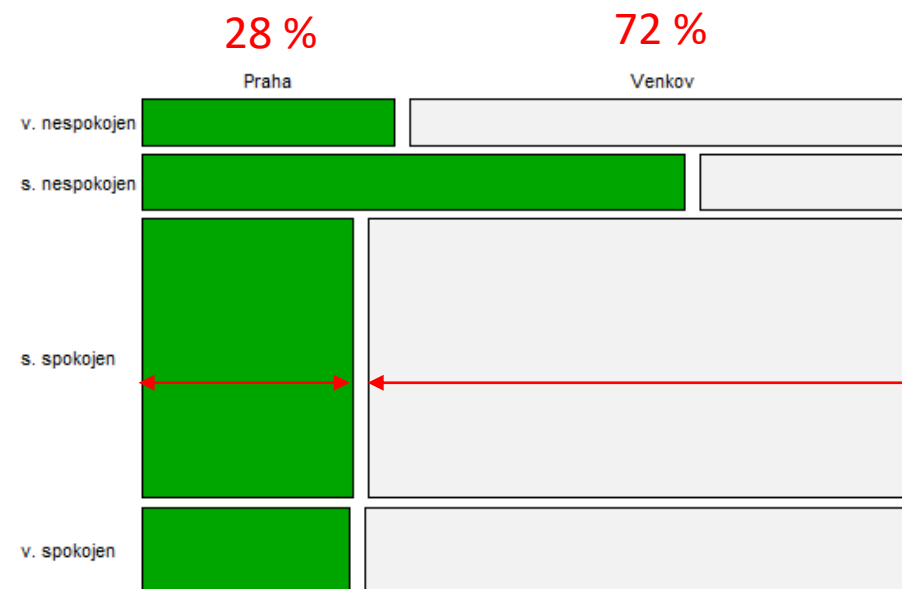


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.3333333	0.7142857	0.2777778	0.2727273
Venkov	0.6666667	0.2857143	0.7222222	0.7272727

sloupcové relativní četnosti



Mozaikový graf

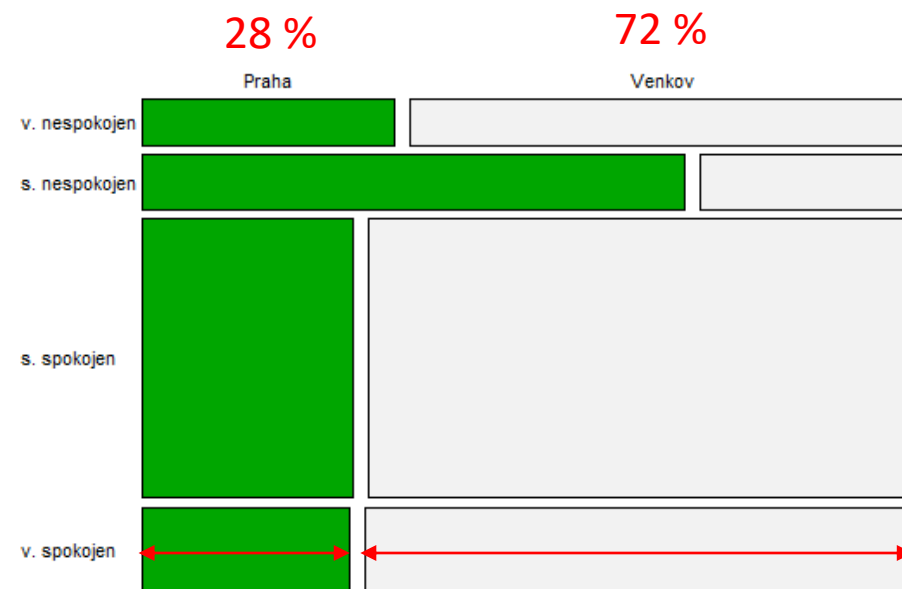


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.3333333	0.7142857	0.2777778	0.2727273
Venkov	0.6666667	0.2857143	0.7222222	0.7272727

sloupcové relativní četnosti



Mozaikový graf

Explorační analýza pomocí R – vizualizace

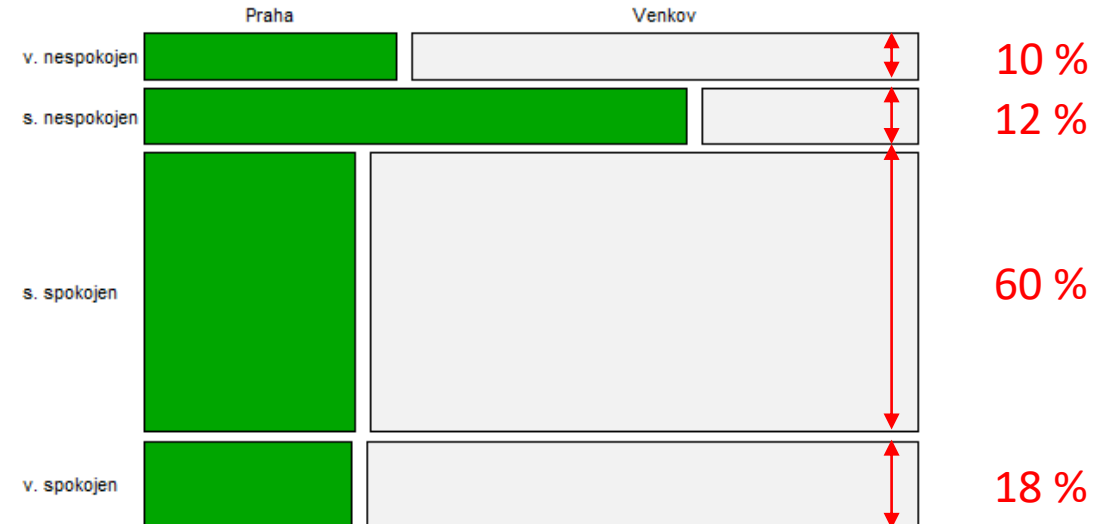


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40
Celkem	30 (10 %)	35 (12 %)	180 (60 %)	55 (18 %)

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.3333333	0.7142857	0.2777778	0.2727273
Venkov	0.6666667	0.2857143	0.7222222	0.7272727

sloupcové relativní četnosti



Mozaikový graf

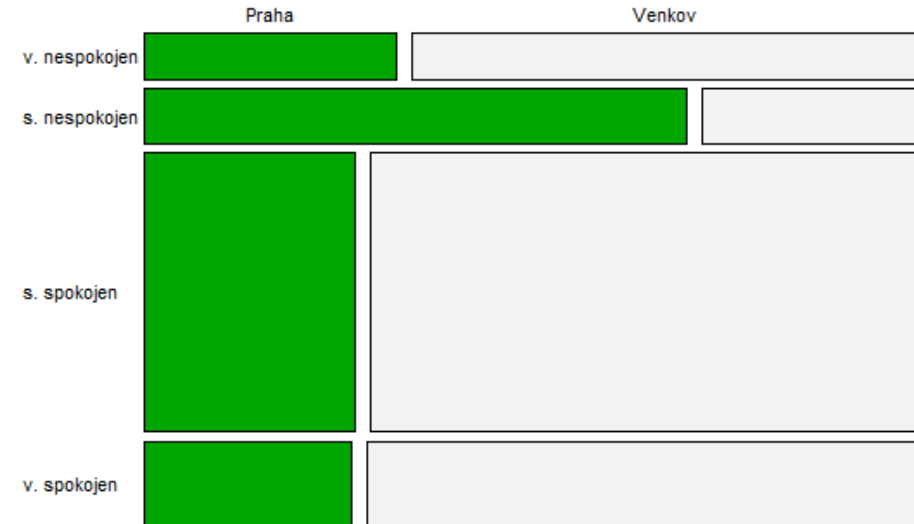


	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	10	25	50	15
Venkov	20	10	130	40

pozorované četnosti

	v. nespokojen	s. nespokojen	s. spokojen	v. spokojen
Praha	0.3333333	0.7142857	0.2777778	0.2727273
Venkov	0.6666667	0.2857143	0.7222222	0.7272727

sloupcové relativní četnosti

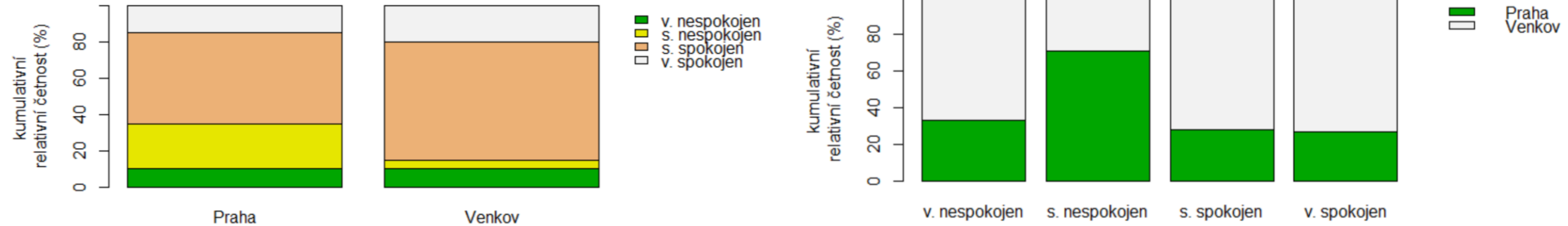


Mozaikový graf

Jak by graf vypadal kdyby alokace zaměstnavatele nesouvisela s tím, jak jsou respondenti spokojeni v práci?

Na grafu bychom viděli „svislé barevné pásy“. Procentuální zastoupení respondentů dle alokace zaměstnavatele by pro všechny úrovně spokojenosti v práci bylo stejné.

Čím členitější „mozaiku“ vidíme, tím silnější závislost mezi kvalitativními znaky vnímáme.



100% skládaný sloupcový graf

Rozmyslete si, který z grafů byste pro vizualizaci použili.



Děkuji za pozornost!

martina.litschmannova@vsb.cz



VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA

FAKULTA
ELEKTROTECHNIKY
A INFORMATIKY

KATEDRA
APLIKOVANÉ
MATEMATIKY