

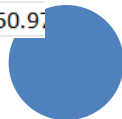


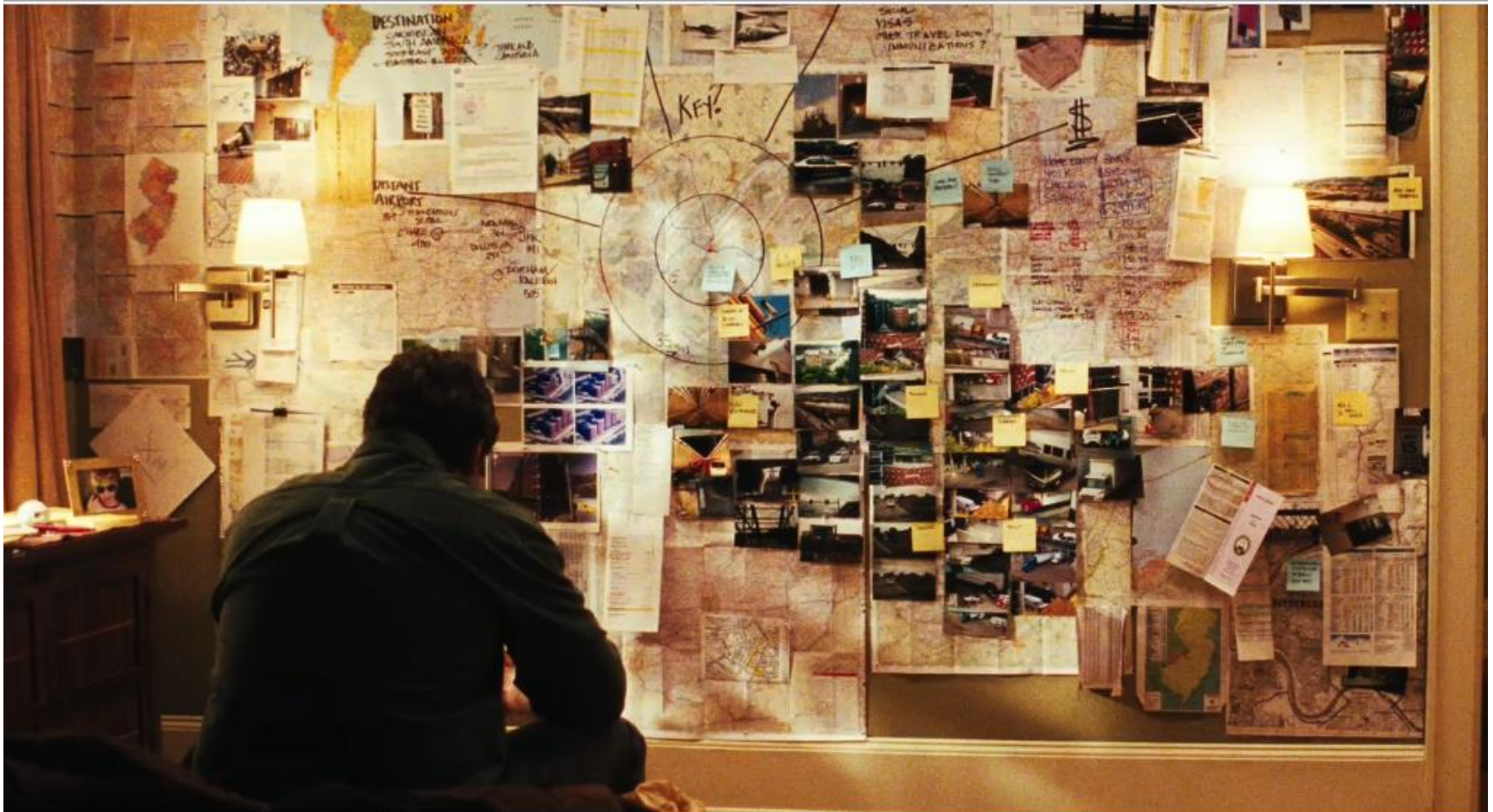
# **KVANTITATIVNÍ ANALÝZA**

**POL494 KONCEPTY A METODOLOGIE V POLITOLOGII**

**Mgr. Petr Voda, Ph.D.**

UZEMNI_JED	U111100101	U111100102	U111100103	U111100104	U111100301	U111100302	U111100303	U111100304	U111100401	U111100402	U111100403	U111100404	U111100501	U111100501
CR_19	10436560	100	5109766	5326794	10141553	97.173	4932640	5208913	244563	2.343	148293	96270	4912225	47.06
KRAJ_3018	1268796	100	613738	655058	1142817	90.071	542753	600064	103594	8.165	59075	44519	644039	50.7
KRAJ_3026	1289211	100	637252	651959	1250990	97.035	613717	637273	32247	2.501	19860	12387	511290	39.65
KRAJ_3034	628336	100	308296	320040	619609	98.611	302976	316633	7052	1.122	4386	2666	276985	44.08
KRAJ_3042	570401	100	282137	288264	551524	96.691	270331	281193	15660	2.745	9876	5784	256018	44.88
KRAJ_3051	295595	100	145483	150112	286184	96.816	139929	146255	7999	2.706	4741	3258	114832	38.84
KRAJ_3069	808961	100	397453	411508	791932	97.895	385787	406145	15273	1.888	10535	4738	345814	42.74
KRAJ_3077	432439	100	211537	220902	423205	97.865	206095	217110	7916	1.831	4643	3273	187480	43.35
KRAJ_3085	547916	100	268967	278949	539119	98.394	263342	275777	7450	1.36	4804	2646	249913	45.61
KRAJ_3093	511627	100	252310	259317	502389	98.194	246548	255841	7089	1.386	4435	2654	240899	47.08
KRAJ_3107	505565	100	250196	255369	500643	99.026	247109	253534	4139	0.819	2593	1546	257923	51.01
KRAJ_3115	1163508	100	567882	595626	1138598	97.859	552401	586197	19866	1.707	12562	7304	596245	51.24
KRAJ_3123	628427	100	305526	322901	623513	99.218	302607	320906	4047	0.644	2446	1601	302645	48.15
KRAJ_3131	579944	100	282500	297444	575811	99.287	279868	295943	3445	0.594	2228	1217	307372	51.24
KRAJ_3140	1205834	100	586489	619345	1195219	99.12	579177	616042	8786	0.729	6109	2677	620770	51.48
MC_50005	30561	100	15284	15277	23560	77.092	11453	12107	3192	10.445	1834	1358	12055	39.44
MC_50008	48773	100	23719	25054	40619	83.282	19249	21370	6394	13.11	3526	2868	22639	46.41
MC_50009	71769	100	34177	37592	63199	88.059	29504	33695	6880	9.586	3818	3062	34963	48.71
MC_50011	131793	100	62790	69003	121652	92.305	57234	64418	8510	6.457	4650	3860	68080	51.65
MC_50014	83968	100	40714	43254	74606	88.851	35507	39099	8131	9.683	4578	3553	42089	50.14
MC_50017	104185	100	49829	54356	93381	89.63	43956	49425	8306	7.972	4455	3851	52029	49.93
MC_50018	41429	100	20162	21267	36671	88.515	17400	19271	4062	9.805	2396	1666	20398	49.23
MC_50020	104918	100	50263	54655	95792	91.302	44982	50810	7633	7.275	4447	3186	53730	51.21
MC_50021	55602	100	27108	28494	47676	85.745	22466	25210	7135	12.832	4270	2865	25288	45.44
MC_50022	113200	100	53227	59973	103503	91.434	47927	55576	8076	7.134	4505	3571	58668	51.82
MC_50129	20500	100	9662	10838	20192	98.498	9483	10709	282	1.376	170	112	10316	50.31
MC_50208	23273	100	11262	12011	22665	97.388	10845	11820	567	2.436	394	173	11740	50.44
MC_50231	14041	100	6934	7107	13549	96.496	6530	7019	464	3.305	382	82	7158	50.91





o [http://3.bp.blogspot.com/\\_ozHIgiPMIOE/TPVGnEkP3OI/AAAAAAAAAEA/sybAd0IEABA/s1600/Picture%2B6.png](http://3.bp.blogspot.com/_ozHIgiPMIOE/TPVGnEkP3OI/AAAAAAAAAEA/sybAd0IEABA/s1600/Picture%2B6.png)





název	rok	šířka	délka	výška	věže	architekt	kód	cena	...
.									
.									
.									
Neusch...	1869	47°33'	10°44'	1008	7	Hoffmann	2	20	...
.									
.									
.									











# VZOREK

- Náhodný výběr
  - Reprezentativnost
  - Náročný na zdroje
- Nenáhodný výběr
  - Týká se všech způsobů sběru skrze internet
  - Záměrný výběr
  - Snaha maximalizovat varianci
- Populace (základní soubor) –  $N$
- Vzorek (výběrový soubor) –  $n$
- Jedinec –  $X$
- Příklad –  $x$





# PROMĚNNÁ

- Vlastnost (charakteristika) nějaké entity
- Reprezentace reality
  - Kvantitativní studií lze jen těžko zachytit všechny atributy jevu





# OPERACIONALIZACE

- Honosnost
- Skvělost, přepychovost, blyštivost, pompéznost, okázalost
- Investované prostředky? Množství odvedené práce? Obytná plocha? Soupis inventáře? Otázka návštěvníkům?





# OPERACIONALIZACE A TABULKA (DATOVÁ MATICE)

- Operacionalizace vytváří šablonu
- Do šablony jsou třízeny informace

Honosnost zámku			Koncept 2			Koncept 3
Náklady na stavbu	Obytná plocha	Počet věží	indikátor	indikátor	indikátor	indikátor
.	.	.	.	.	.	.
.	.	.	.	.	.	.

# TYPY PROMĚNNÝCH

- Nominální (architektonický styl)
- Ordinální (epocha architektonického stylu)
- Intervalové a poměrové (věk architekta)
  
- Kategorické (diskrétní) – nominální a krátké ordinální
  - Dichotomické – pouze dvě hodnoty
- Kardinální (spojité) – dlouhé ordinální a intervalové



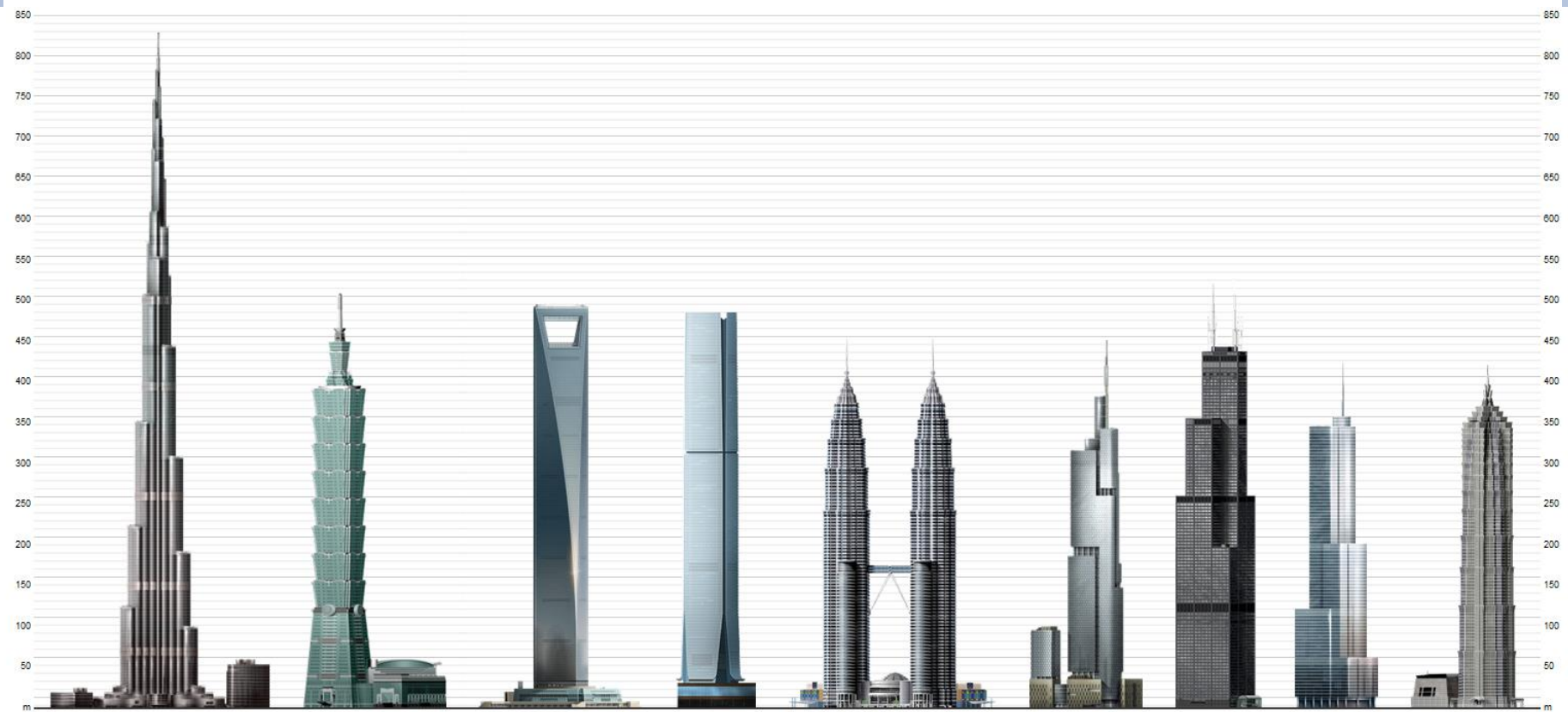
# TRANSFORMACE PROMĚNNÝCH

- Ordinální + ordinální + ... + ordinální = kardinální
- Kardinální = dichotomická + dichotomická + ...
- Nominální = dichotomická + dichotomická + ...









Name	Burj Khalifa	Taipei 101	Shanghai World Financial Center	International Commerce Center	Petronas Towers	Nanjing Greenland Financial Complex	Willis Tower	Trump International Hotel & Tower	Jin Mao Tower
City	Dubai	Taipei	Shanghai SH	Hong Kong HK	Kuala Lumpur	Nanjing JS	Chicago IL	Chicago IL	Shanghai SH
Country	United Arab Emirates	Taiwan	China	China	Malaysia	China	United States	United States	China
Illustrator	J. Eduardo Segundo Hernandez	Someformofhuman	Ka777	Mike K.	Cliff Tan	J. Eduardo Segundo Hernandez	Someformofhuman	J. Eduardo Segundo Hernandez	Cliff Tan
Status	Built	Built	Built	Built	Built	Built	Built	Built	Built
Built	2010	2004	2008	2010	1998	2010	1974	2009	1998
Floors	163	101	101	108	88	89	108	88	93
Use	Mixed use	Office	Mixed use	Mixed use	Office	Mixed use	Office	Mixed use	Mixed use
Antenna									
Spire	828 m	508 m			451.9 m	450 m	527.3 m	423.4 m	420.5 m
Roof		448 m	492 m	484 m		381 m	442.1 m	356.9 m	







# PLÁN VÝZKUMU

- Samotný výzkum musí být dopředu naplánován
  - Případná pochybení se jen obtížně napravují
- Otázka – hypotézy (teorie) – konceptualizace a operacionalizace – sběr dat (vzorek) – analýza
- otázka:
  - Popisná
  - Popisně explanatorní
  - Explanatorní



# HYPOTÉZY

- kvantitativní výzkum je deduktivní: postup teorie – hypotézy – „pozorování“ – závěry
- formulovány na základě teorie
- **Hypotézy používáme, jen když testujeme teorie**
- výzkumné
  - Formulace očekávaného vztahu mezi proměnnými
- Statistické
  - Hypotézy testované statistickými výpočty
  - Nulová hypotéza



# HYPOTÉZY - PŘÍKLAD

## ○ Výzkumná:

- Honosnost zámku se odvíjí od bohatství a prestiže vlastníků

## ○ Pracovní:

- Čím starší rodokmen rodu tím více peněz je pro stavbu použito.
- Čím vyšší je celkový přiznaný majetek rodu, tím obsáhlejší je soupis inventáře.

....

## ○ Nulová:

- Mezi délkou rodokmenu a investicemi do výstavby není vztah

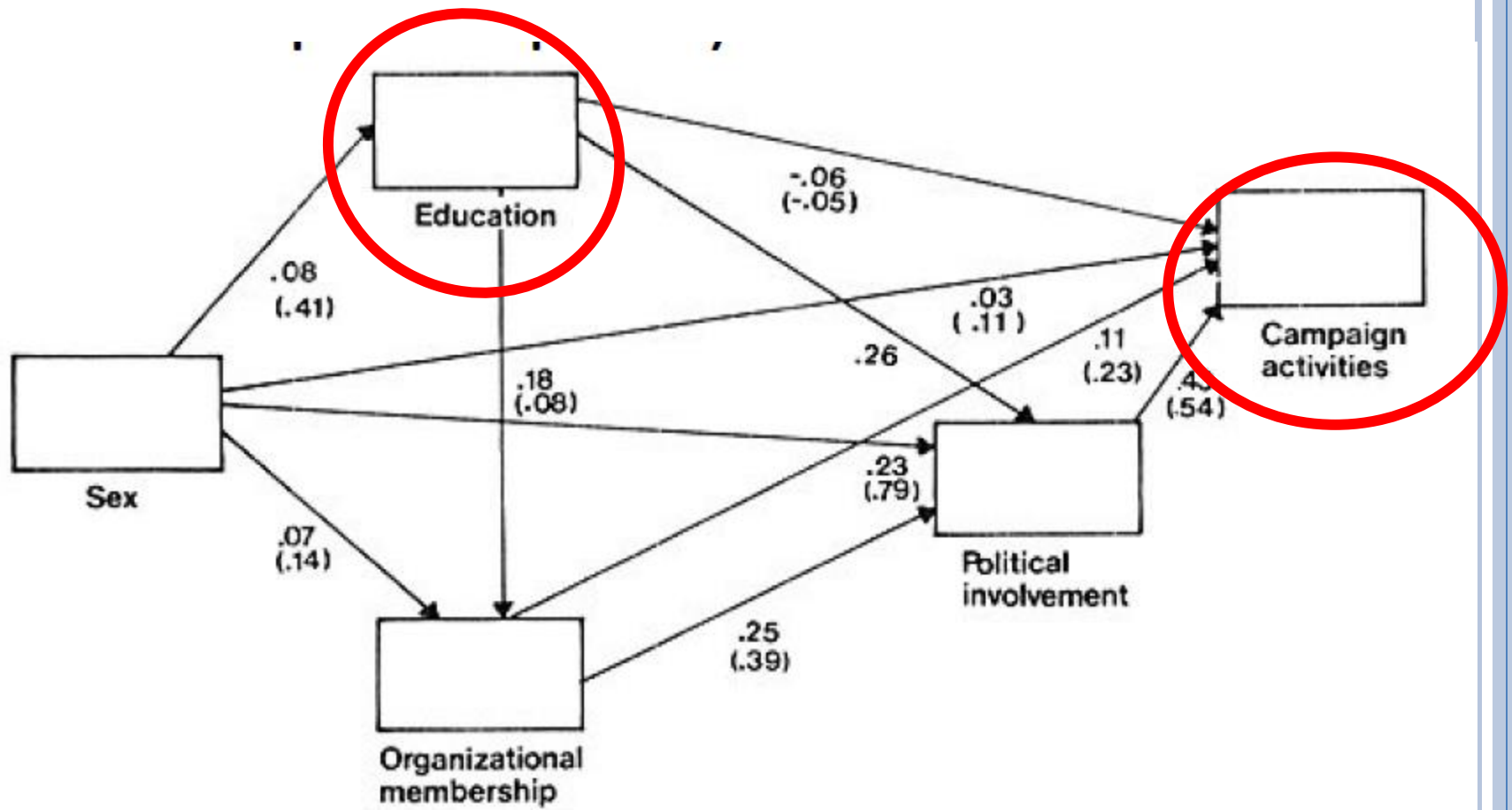


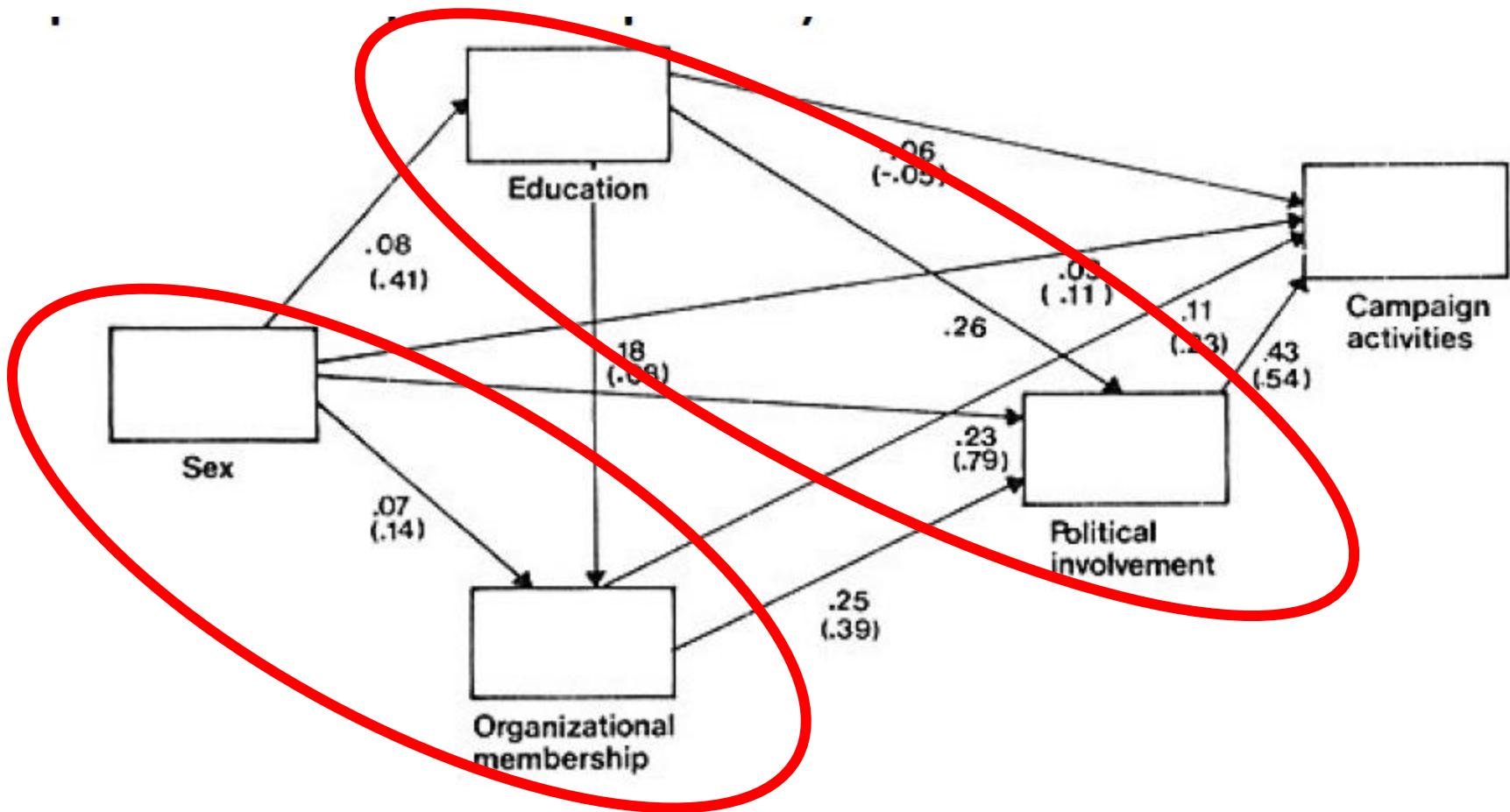


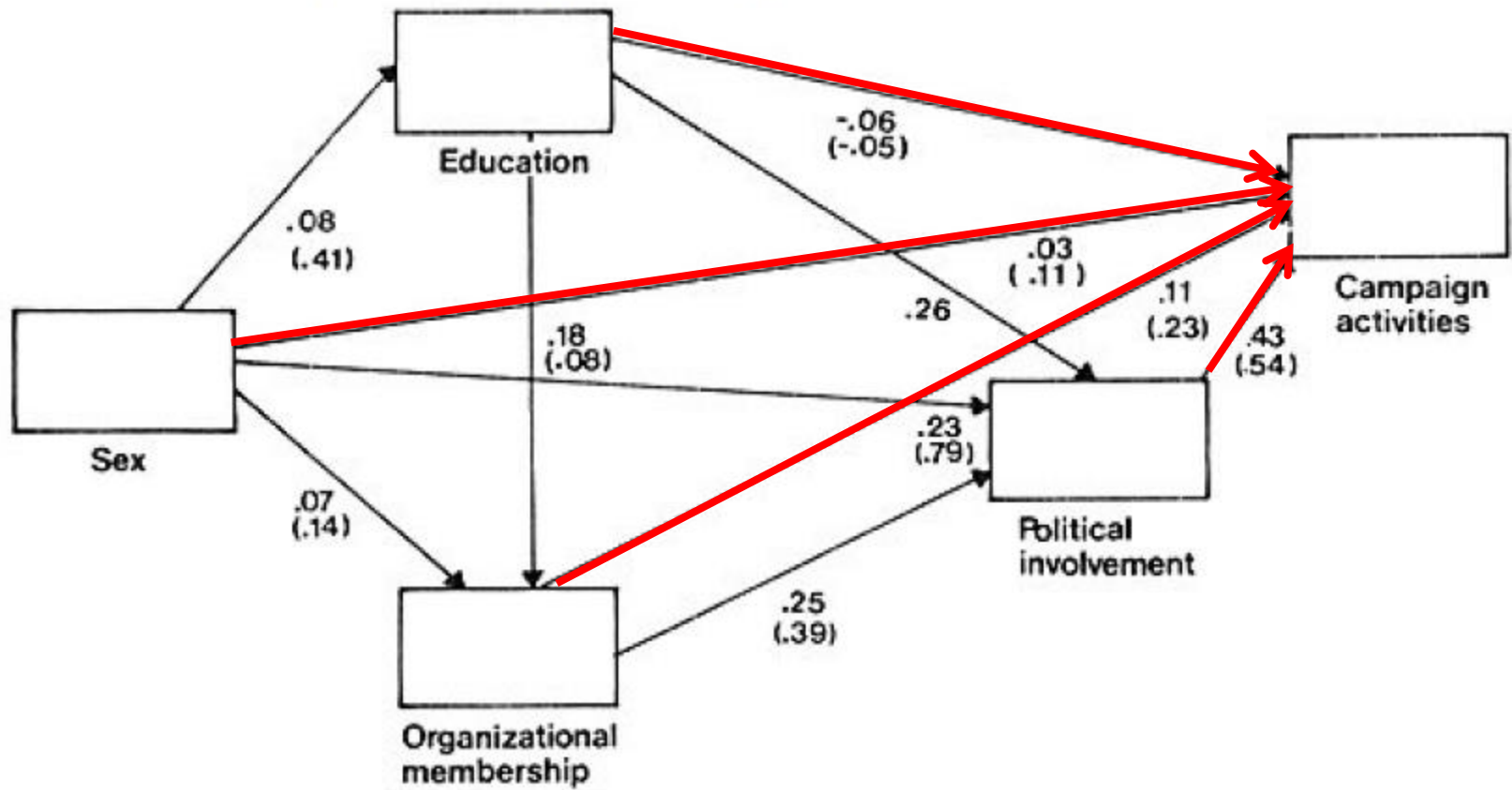
# TESTOVÁNÍ HYPOTÉZ

- Kontingenční tabulka
- Korelace
- Regrese
  
- A spousta rozšíření těchto principů

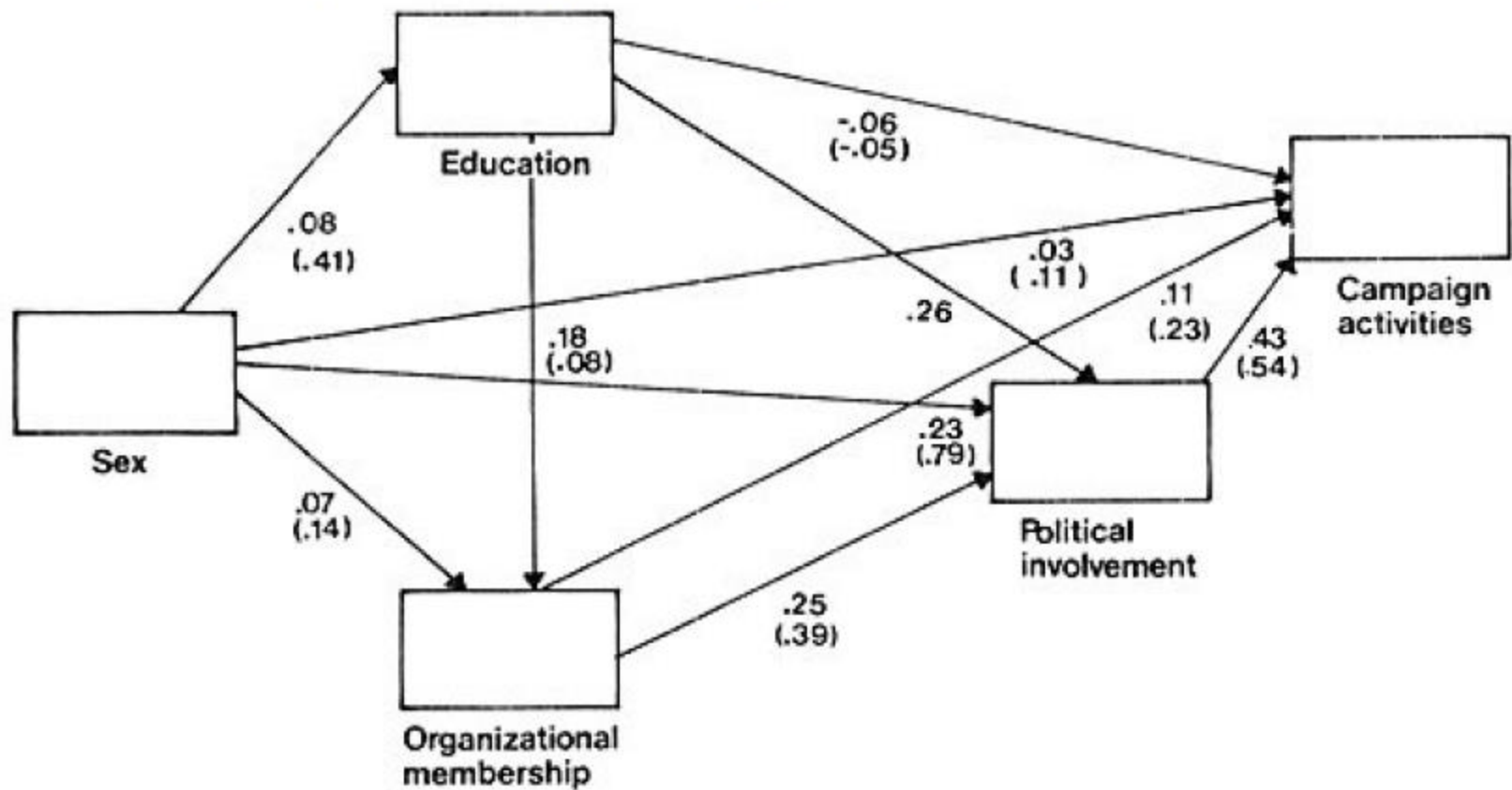












# PRINCIP KONTINGENČNÍ TABULKY

- Snažíme se přijít na to, proč případy patří do nějaké kategorie
- Dichotomická (nebo kategorická) závisle proměnná
- Zchátralý/udržovaný zámek
- Klasicistní/novorenesanční/empírový/secesní zámek
  
- Pokud rod vlastní o 1000 zlatých více než jiný rod, je šance že jejich zámek je zachovalý dvojnásobná oproti chudšímu rodu



# KONTINGENČNÍ TABULKA

	Zchátralý zámek	Udržovaný zámek	celkem
Bohatý rod	100	400	500
Chudý rod	400	100	500
celkem	500	500	1000

- Souvislost prokázána, pokud se pozorované četnosti liší od očekávaných
- za platnosti hypotézy nezávislosti
- Očekávaná hodnota zchátralých zámků ve vlastnictví bohatých rodů =  
 $(500/1000) \times (500/1000) = \frac{1}{4} (*1000)$



# KONTROLNÍ PROMĚNNÁ V KONTINGENČNÍ TABULCE

		Zchátralý zámek	Udržovaný zámek	celkem
Krátký rodokmen	Bohatý rod	50	200	250
	Chudý rod	200	50	250
Dlouhý rodokmen	Bohatý rod	50	200	250
	Chudý rod	200	50	250
	celkem	500	500	1000



# KONTROLNÍ PROMĚNNÁ V KONTINGENČNÍ TABULCE

		Zchátralý zámek	Udržovaný zámek	celkem
Krátký rodokmen	Bohatý rod	125	125	250
	Chudý rod	125	125	250
Dlouhý rodokmen	Bohatý rod	50	200	250
	Chudý rod	200	50	250
	celkem	500	500	1000

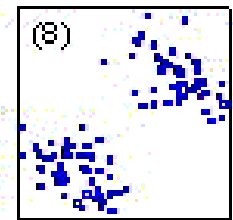
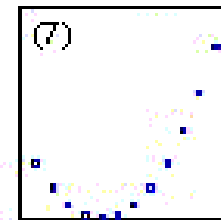
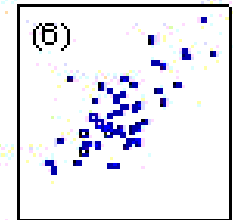
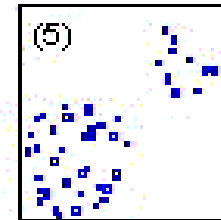
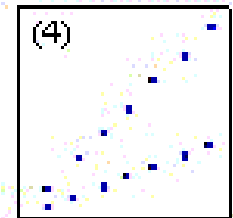
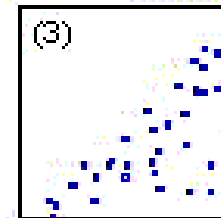
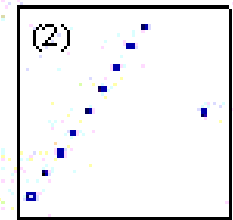
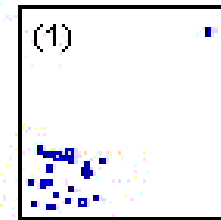
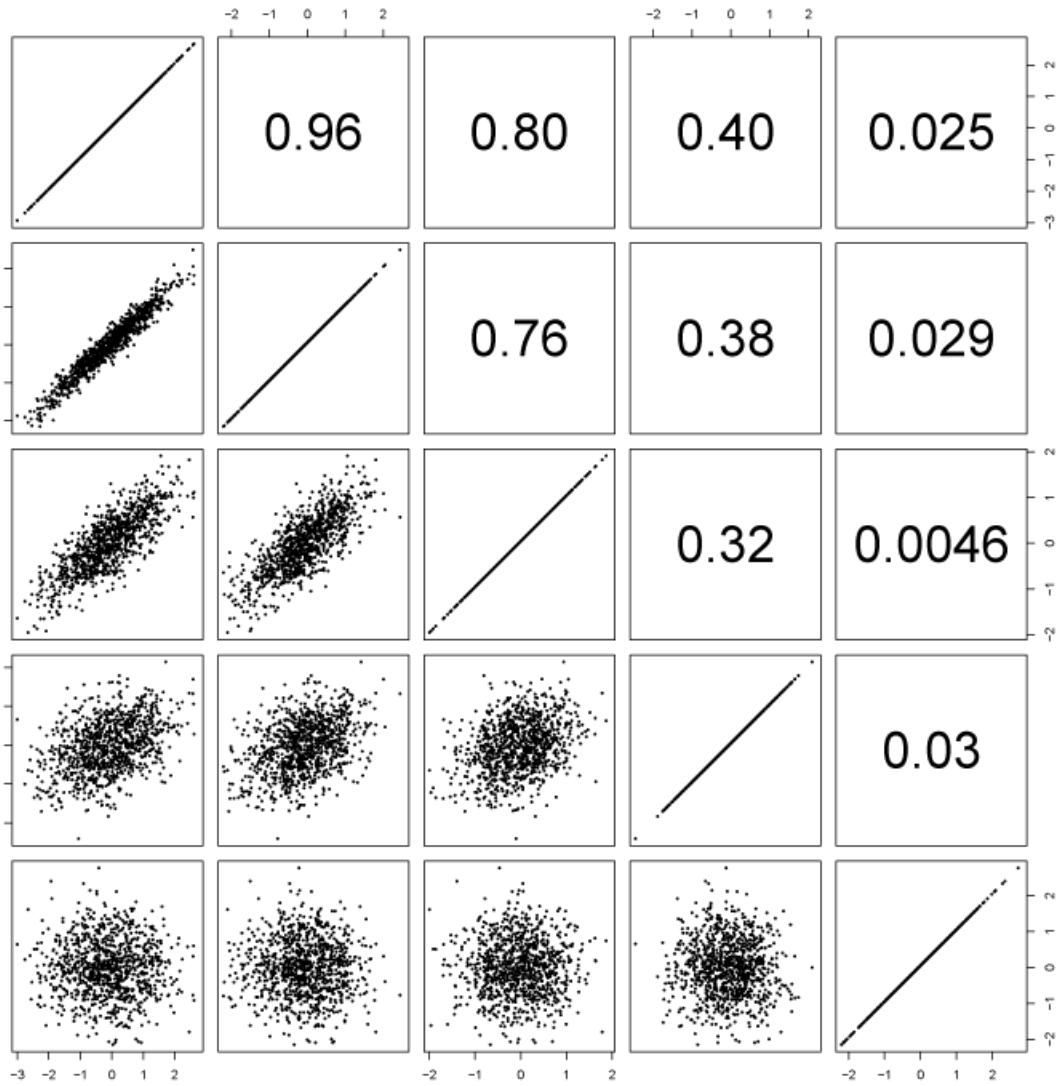




# PRINCIP KORELACE

- Čím víc, tím víc (míň)
- Jde jen o souvislost, protože není jasné, jestli vztah není dán jinou proměnnou
- Tradice šlechtického rodu a honosnost zámků spolu souvisí. Čím delší je rodokmen rodu, tím větší je obytná plocha zámků.





# PRINCIP STATISTICKÉ KONTROLY V REGRESI

- Kauzální vliv nekontrolujeme designem výzkumu, ale zahrnutím proměnných
- Matematicky modelováno, jaký je efekt nezávisle proměnné na závisle proměnnou, pokud ostatní charakteristiky zůstávají shodné
- Honosnost sídla, vyjádřená velikostí obytné plochy, roste s růstem bohatstvím šlechtického rodu. Pokud je šlechtický rod o 1 000 zlatých bohatší než jiný rod, a přitom mají stejně dlouhý rodokmen a stejně časté kontakty s královskou rodinnou, tak je obytná plocha jejich sídla o 200 čtverečních metrů větší



# PROČ MOHOU BÝT ZÁVĚRY KVANTITATIVNÍCH STUDIÍ CHYBNÉ?

	průměrný rok	průměr počtu věží
Francie	1759.8	6.7
Německo	1780	6.3
Anglie	1790	6
Česká republika	1803	5.8
Rakousko	1812	5.4
Polsko	1830	5
Litva	1859.8	4.7



Název	Rok	Počet Věží
Heidelberg	1730	1
Wurzburg	1742	3
Augustsburg	1746	3
Dresden	1751	4
Hohentübingen	1757	5
Falkenlust	1760	5
Langenzell	1763	6
Neuschwanstein	1769	7
Marienburg	1787	6
Hackhausen	1793	7





Název	Rok	Počet Věží
Slavkov	1760	1
Hluboká	1772	3
Litomyšl	1786	3
Jaroměř	1781	4
Náměšť	1797	4
Slatiny	1810	5
Lány	1813	7
Trója	1819	6
Jezeří	1827	7
Opočno	1824	8
	1803	5,8



# DATA

## ○ Individuální

- mikro-úroveň
- Data za jednotlivce, organizace
- Pracuje se s výběrovými soubory
- Potřeba statistické inference (pouze p-stní vzorky)
- Simpsonův paradox

## ○ Agregovaná

- Makro-úroveň
- Data za územní celky, skupiny obyvatel
- Možnost práce s celou populací
- Problém tzv. ekologické chyby



## DALŠÍ PROBLÉMY

- Nepravá korelace
- Nereprezentativita
- Chyby signifikance
- Různá zkreslení (hlavně v případě dotazníkových šetření)



# SHRNUTÍ

- Popis – popisné statistiky
- Souvislost – korelace, kontingenční tabulky
- Vysvětlení – různé formy regrese

