



Úřad vlády
České republiky

Druhý život dat: co díky datům státu víme a co tomu (ne)pomáhá

Petr Bouchal

Vládní analytický útvar



- 01** Analýzy pro rozhodování: VAÚ a RIA
- 02** V praxi: projekty a data v nich
- 03** Co funguje, co nefunguje



Analýzy pro rozhodování

Co VAÚ dělá

- Analytická podpora reforem
- Podklady pro rozhodování vlády
- Podpora procesu RIA



Analýzy pro rozhodování

Jak to děláme

- Analytická podpora reforem
 - Analytická kapacita: tým ekonomů, datových analytiků aj.
 - *Proces* RIA: definice problému → nástroje → varianty řešení → dopady
- Podklady pro rozhodování vlády
 - Rychlé přehledy literatury či dat
 - Důraz na rychlost a přímou využitelnost
- Podpora procesu RIA
 - Podpora tvůrců legislativy a ZZ RIA: metodika, školení, konzultace
 - Sekretariát komise RIA: podklady pro zpravodaje, organizace



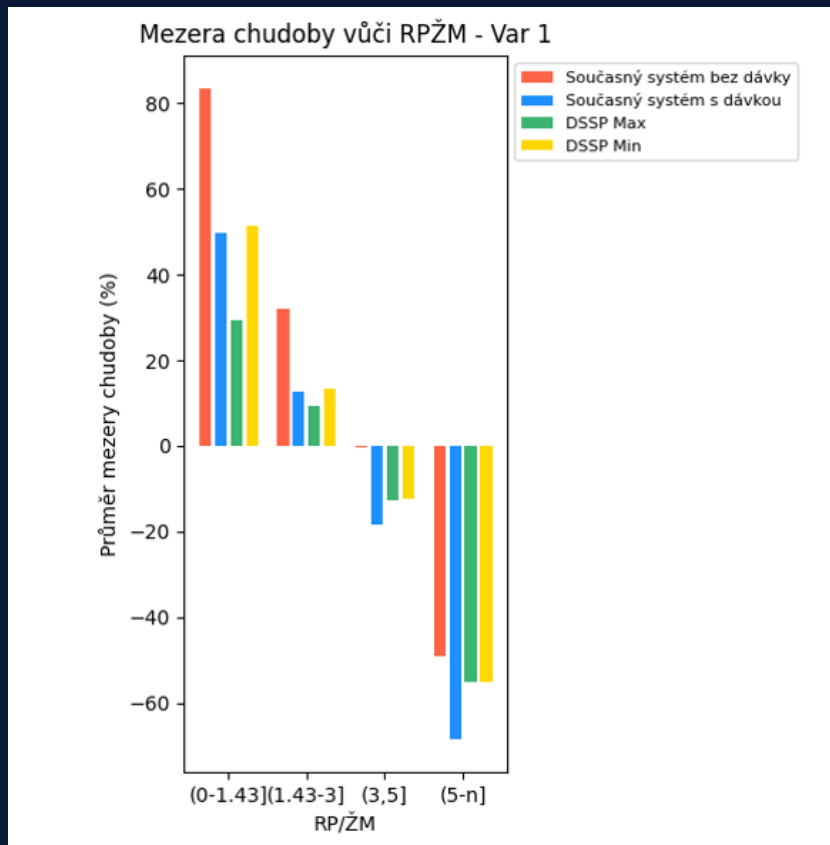
Data v analýzách

Jakou roli hrají

- **Proč administrativní data**
 - Už se sbírají kvůli službě/procesu
 - Zachycují kompletní populaci
- **Proč mikrodata**
 - Díky nim rozumíme nejen celku, ale i distribucím, částem celků a vztahům mezi nimi
 - Umožňují věrohodnou simulaci, srovnání jablek s jablky
 - Otevírají jinak nedostupné možnosti výzkumných designů a kauzálního usuzování
- **K čemu propojování dat**
 - Často jsou třeba data z více systémů, organizací, resortů
 - Nadresortní problémy

Data v analýzách

Příklad: simulace superdávky



- Cíl: odhadnout dopady reformy sociálních dávek
- Data: „case“ záznamy ČSSZ – spojené mezi dávkami
- Způsob přístupu k datům:
 - předání pseudonymizovaných dat ČSSZ
- Analýza: mikrosimulace dopadů různých variant změny na specifické skupiny cílové skupiny



Data v analýzách

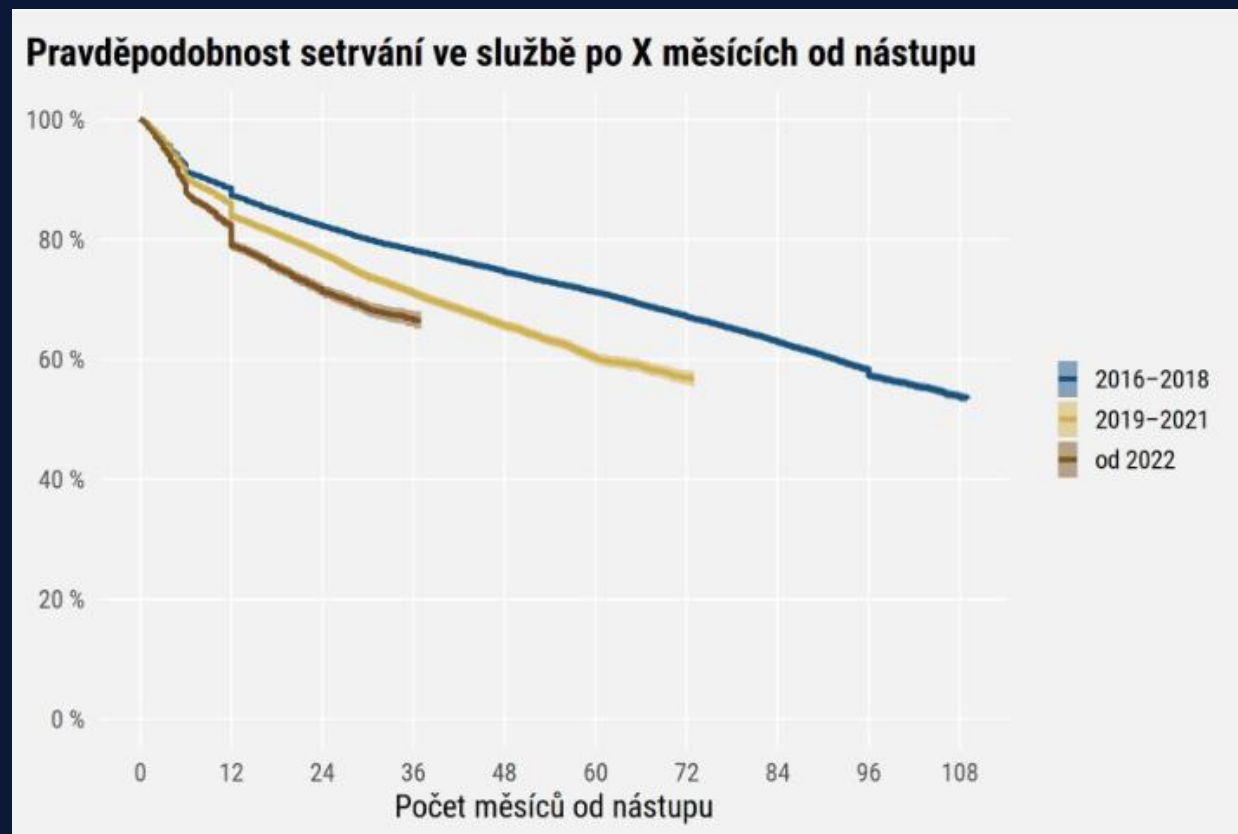
Příklad: podpora startupů

- Cíl: identifikovat záběr intervencí a odhadnout dopady zvažovaných opatření
- Data: údaje z daňových přiznání firem, OSVČ
- Způsob přístupu k datům:
 - Agregace → syntetická data → skripty
 - GFŘ použít skripty nad databází → agregace
- Analýza: simulace dopadů na dotčené firmy
fiskálních dopadů
celospolečenských čistých přínosů



Data v analýzách

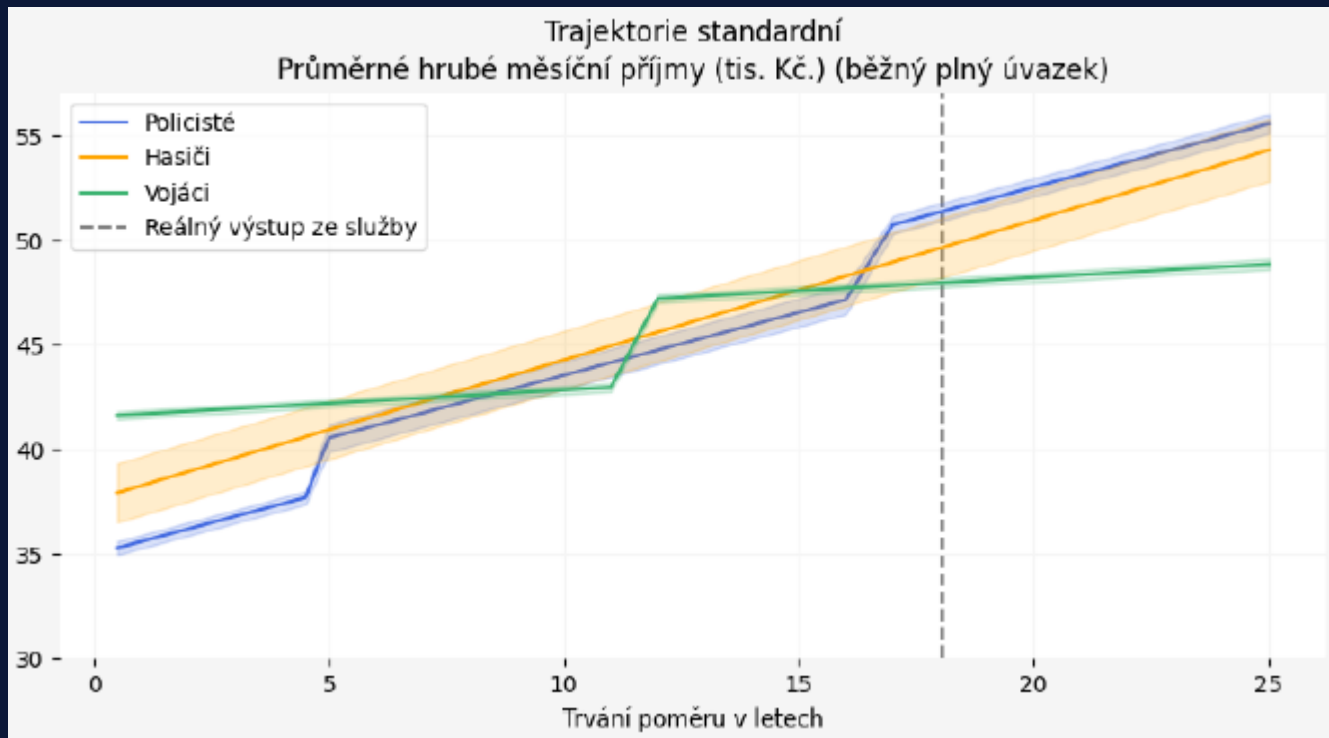
Příklad: dynamika pracovní síly státní služby



- Cíl: identifikovat stav a trendy v pracovní síle státní služby
- Data: platová (ISP – MF) a personální (MV/ÚV – ISoSS)
- Způsob přístupu k datům:
 - Předání pseudonymizovaných dat
- Analýza: kombinace ekonometrických modelů

Data v analýzách

Příklad: platy vojáků a příslušníků bezpečnostních sborů



- Cíl: popsat vztah platových podmínek svou skupin osob
- Data: platová (ISP – MF) a personální (MV, MO)
- Způsob přístupu k datům:
 - Předání pseudonymizovaných dat ze 4 systémů
- Analýza: modely kariérní dráhy, mikrosimulace dopadů navrhovaných změn



Data v analýzách

Cesty k bezpečnému analytickému využití dat

- Předání pseudonymizovaných dat ✓
- Syntetická data a zprostředkované výpočty ✓
- Agregace a charakteristiky na vyžádání ✓
- Spojení a zpřístupnění dat u jednoho ze subjektů ~
- (Poskytnutí syntetických dat ex ante) ✓
- (Sofistikovaná anonymizace: differential privacy aj.) ?
- (Přístup ke statistickým mikrodátům) ?



Co funguje a nefunguje

Co funguje

- Znalost dat na straně správců
- Memoranda o spolupráci (ale...)
- VAÚ jako subjekt mezi resorty
- Řízení přístupu a zabezpečení v analytickém prostředí
- Nacházení způsobů využití dat bez přístupu do živých systémů / náhledu na surová data

Co chybí

- Metadata, dokumentace a kvalita dat nejsou nikde zachycené
- Jasný právní titul pro využití dat
- Katalog: nevíme, zda chybí...
- Možnost spojení např. přes ISSS
- Sdílená infrastruktura pro datovou analýzu (code.gov.cz je dobrý začátek, ostatní chybí a každý resort tvoří nově a jinak)



Úřad vlády
České republiky

**Děkuji
za pozornost**

Petr Bouchal

petr.bouchal@vlada.gov.cz