

Požadavky zadavatele na Návrh řešení

Dodavatel je ve svém Návrhu řešení (součást nabídky) povinen respektovat požadavky zadavatele uvedené v této příloze.

Zadavatel stanoví, že Návrh řešení bude dodavatelem členěn do tří samostatných částí (kapitol), a to:

- 1) Specifikace počtu aktivních hlasových SIM karet, které bude dodavatel schopen pro plnění předmětu veřejné zakázky využít (analyzovat)
- 2) Popis navrhovaného řešení
- 3) Postup získávání dat

Zadavatel stanoví, že maximální rozsah Návrhu řešení veřejné zakázky nesmí přesáhnout 30 normostran A4.

1) Požadavky zadavatele na specifikaci počtu aktivních SIM karet

Zadavatel stanoví, že dodavatel je v rámci Návrhu řešení povinen uvést počet aktivních hlasových SIM karet¹, které bude mít pro plnění veřejné zakázky k dispozici (minimálně 3 miliony). Mít k dispozici aktivní hlasové SIM karty se rozumí možnost dodavatele zajistit požadovaná geolokační data vztahující se ke konkrétní SIM kartě za období pěti (5) měsíců předcházejících a/nebo šesti (6) měsíců následujících dni určenému zadavatelem za podmínek vymezených v Příloze č. 4 této ZD – Závazném návrhu smlouvy na plnění veřejné zakázky.

2) Požadavky zadavatele na Popis navrhovaného řešení

Zadavatel stanoví, že dodavatel je povinen uvést v této části Návrhu řešení podrobný popis postupu dodavatele při plnění této veřejné zakázky včetně obsahu databáze, který se zavazuje dodat.

Zadavatel stanoví, že dodavatelem navrhované řešení musí být navrženo v souladu s požadavky zadavatele na specifikaci požadovaných dat uvedenými dále v této příloze.

Zadavatel stanoví, že dodavatel v rámci Popisu navrhovaného řešení musí reflektovat všechny dále uvedené požadavky, které představují minimální požadavky na obsah databáze.

¹ Aktivními hlasovými SIM kartami se rozumí soubor všech aktivních SIM karet využívajících hlasové služby, tj. bez započtení tzv. M2M SIM karet (SIM karty určené pro senzory, vzdálené ovládání apod.) a SIM karet využívaných výlučně pro poskytování datových služeb (připojení k internetu bez současného využívání hlasových služeb).

Zadavatel stanoví, že dodavatel je povinen v části „Popis navrhovaného řešení“ podrobně popsat řešení, které dodavatel nabízí, tj. postup zpracování, popis obsahu výsledných databází a všech požadovaných parametrů.

Zadavatelem stanovené požadavky představují minimální požadavky na obsah databáze. Soulad s požadavky zadavatele bude předmětem hodnocení nabídky.

A. Obecné požadavky

Dodavatel při tvorbě výsledných databází statistických údajů musí postupovat dle níže zadaných požadavků.

Rozdělení požadavků na fixní a variabilní je učiněno pro potřeby testování dat v rámci PoC (blíže specifikováno v příloze 1 ZD, která je současně přílohou č. 1 smlouvy na plnění veřejné zakázky).

Není-li výslovně uvedeno jinak, jsou veškeré parametry popisované v tomto dokumentu závaznými (fixními) požadavky zadavatele.

Objem dodávaných dat – Dodavatel musí mít po celou dobu plnění veřejné zakázky (po celou dobu trvání smlouvy na plnění veřejné zakázky) zajištěnu možnost analyzovat základní soubor o velikosti minimálně **3 milionů aktivních hlasových SIM karet² reprezentativně rozmístěných po celém území České republiky a odrážejících demografickou strukturu populace**. Celkový soubor SIM karet musí umožnit analýzu na úrovni obcí. SIM karet tedy musí být dostatečné množství a musí být prostorově natolik rozptýlené, aby popsaly všechny požadované ukazatele (dále v textu) za každou obec v České republice.

Redukce základního souboru

Ze základního souboru analyzovaných SIM karet je v první fázi analýzy nutné vyřadit zahraniční SIM karty v síti, duální SIM karty³, SIM karty, u nichž je evidentní jejich využití jiné než v mobilních telefonech (ačkoli se nejedná o M2M SIM karty, které základní soubor již neobsahuje), dále pak SIM karty, jež jsou po většinu měřeného úseku neaktivní, nebo z jiného důvodu nejsou data z nich získaná validní pro účely této zakázky. V těchto případech je nutné využít specifických nástrojů, aby data o těchto SIM kartách neovlivňovala (nezkreslovala) výslednou databázi.

² Aktivními hlasovými SIM kartami se rozumí soubor všech aktivních SIM karet využívajících hlasové služby, tj. bez započtení tzv. M2M SIM karet (SIM karty určené pro senzory, vzdálené ovládání apod.) a SIM karet využívaných výlučně pro poskytování datových služeb (připojení k internetu bez současného využívání hlasových služeb).

³ Duálními SIM kartami je myšlen stav, kdy jedinec má více aktivních SIM karet, jejichž prostorové vzorce chování jsou totožné. Technickými SIM kartami jsou myšleny karty, které se nenacházejí v mobilních telefonech, ale například v osobních automobilech, technických zařízeních fungujících na bázi sítě GSM (chytrá domácnost, ovládání, dveří, bran aj.)

Územní dimenze – zadavatel požaduje dodání požadovaných statistických údajů za celé území České republiky. Všechny údaje jsou požadovány za každou z 6 254, popř. 6 258 obcí⁴. Praha a ostatní územně členěná města České republiky jsou ve výsledné databázi uvažovány jako jedna obec. **Granularita neboli územní detail je tedy požadována do úrovně obcí.**

Časová dimenze – požadované statistické údaje (databáze) jsou výsledkem analýz pohybu a pobytu SIM karet v signalizační síti. **Pohyb a pobyt SIM karet je sledován celkem ve čtyřech měřeních, každé měření sleduje období čtyř po sobě následujících týdnů.** První dvě měření jsou vybírána tak, aby nebyla ovlivněna sezónními vlivy (prázdniny, svátky apod.). Jedno měření je požadováno v jarním období⁵, druhé v podzimním⁶. Další dvě měření by pak měla odrážet specifickou pohybu během léta a zimy⁷. Nicméně vzhledem k současné epidemické situaci v České republice není v době vypisování veřejné zakázky jasné, jaká případná omezení pohybu osob budou v daných obdobích platit. Vymezení konkrétních termínů uskutečnění měření bude stanoveno zadavatelem postupem vymezeným v Příloze č. 4 ZD – Závazném návrhu smlouvy na plnění veřejné zakázky, kdy konkrétní termín plnění bude projednán s dodavatelem s ohledem na aktuální epidemickou situaci a snahu o nalezení vhodného řešení minimalizujícího ovlivnění výpovědní hodnoty dat v důsledku nestandardních podmínek.

Anonymizace dat – Požadované databáze je nutné anonymizovat tak, aby nebylo možné ztotožnit konkrétního jedince s konkrétním údajem. Anonymizace lokalizačních dat je primárně na straně provozovatele sítě, popřípadě primárního zpracovatele (poskytovatele) těchto dat. Ti jsou povinni údaje anonymizovat ještě před jejich prvotním zpracováním.

Pro potřeby anonymizace se obecně využívá omezení výpovědní hodnoty dat v případě velmi detailních údajů⁸. Z toho důvodu nejsou poskytovány údaje pod určitou kritickou hodnotou pro zveřejnění⁹. Zadavatel veřejné zakázky v tomto směru požaduje, aby v případě údajů pod touto kritickou hodnotou bylo odlišeno, zda hodnota daného ukazatele je 0, nebo se jedná o nízkou nezveřejnitelnou hodnotu. Svě řešení dodavatel vysvětlí v návrhu řešení.

⁴ V Česku je v současnosti 6 254 samostatných obcí a 4 vojenské újezdy (bez stálého obyvatelstva). Sám dodavatel zvolí soubor buď 6 254, nebo 6 258 jednotek, dle nastavení svých nástrojů pro získávání dat.

⁵ Ideálně od května přibližně do poloviny června.

⁶ V ideálním případě období září a října (nutné vhodně nastavit vzhledem ke svátkům 28. září a 28. října).

⁷ V těchto dvou měřeních tedy nebude nutné se vyhýbat obdobím specifických dojížděkových anomálií (př. letní prázdniny), jelikož účelem měření je postihnout právě tyto extrémní případy.

⁸ Například popisující chování malého počtu osob na malém území, kdy je teoreticky možné jedince s údajem ztotožnit.

⁹ Hodnota daného ukazatele je tak nízká, že nemůže být zveřejněna, aby nebylo údaj možné ztotožnit s konkrétní osobou (jedná se o údaje v řádech jednotek).

B. Specifické požadavky na postup plnění a výslednou formu dodávaných dat

Sledování pohybu – Dodavatel specifikuje vlastní řešení, jak zaznamenává pohyb SIM karet/osob v síti. Požadovaný časový detail je 30 minut, tzn. zvolená technika sledování tak musí umožňovat identifikaci pobytu v těchto intervalech.

Dodavatel je povinen v analýzách mobility jednotlivých SIM karet uvažovat jen taková místa výskytu, kde osoba/SIM karta strávila alespoň 30 minut. Výskyty na kratší dobu jsou považovány za nahodilé (například tranzit přes obec aj.) a nebude s nimi v níže popsaných analýzách počítáno.

Požadavky na Popis navrhovaného řešení

Dodavatel je povinen v části „Popis navrhovaného řešení“ zahrnout následující požadavky zadavatele. **Pokud není uvedeno jinak, jedná se o fixní požadavky.**

Dodavatel analyzuje všechna místa výskytu každé SIM karty v síti (v období 4 po sobě následujících týdnů; výskyty na kratší dobu než 30 minut nezahrnuté; buňky BTS agregovány do obcí, tj. obce výskytu).

Na základě znalosti míst výskytu každé osoby/SIM karty během sledovaného období jsou každé SIM kartě/osobě přiřazovány atributy. Atribut vyjadřuje vztah jedince k určitému místu/obci. Každý jedinec může mít více atributů vůči různým obcím, avšak k jedné obci může mít jen jeden atribut.

Atributy:

Rezident – jedná se o základní atribut určující obec, kterou lze považovat za místo bydliště. **Každá SIM karta/osoba ve sledovaném souboru musí mít přiřazeno právě jedno místo, kde je rezidentem.** Je-li obec pro jedince rezidenční, nemůže již vůči této obci mít přidělen jiný atribut (dojíždějící či druhé bydlení).

Rezidenční obec je ta obec, kde jedinec **tráví ve sledovaném 4týdenním měření nejvíce času v nočních hodinách určených intervalem 0:00–4:00¹⁰.** *[variabilní parametr]*

Dojíždějící – jsou definovány 3 typy dojíždějící rozlišené dle jejich intenzity (počet návštěv a strávený čas). *[variabilní parametr]*

¹⁰ Z celkových 28 sledovaných dnů v době od 0:00–4:00 každá osoba strávila nejvíce času v určité lokalitě, kterou dále považujeme za domov.

Typ 1 = dojíždka za prací a do škol – nejintenzivnější typ dojíždky. Každý jedinec/SIM karta může mít přiřazenu pouze jednu či žádnou obec dojíždky za prací/do škol.

Atribut dojíždějího za prací/školou do dané obce je přidělen těm osobám, které ve sledovaném období **danou obec navštíví alespoň 16krát a v dané lokalitě celkem stráví alespoň 80 hodin**¹¹. *[variabilní parametr]* V případě, že pro jedince neexistuje žádná obec (vyjma rezidenční obce), ve které by se vyskytoval s takto stanovenou intenzitou, není mu tento atribut udělen vůči žádné obci. V případě, že stanovené podmínky splňuje jedinec vůči většímu množství obcí, atribut dojíždějího za prací je přiřazen k té obci, kde jedinec strávil nejvíce času během celého sledovaného období. Ostatní místa jsou pak zahrnuta do dojíždky typu 2.

Typ 2 = intenzivní dojíždka za službami – typ dojíždky, který lze připodobnit k pravidelné dojíždce za službami, a to zejména nákupy, dále pak kulturní, sociální a sportovní služby, popřípadě návštěvy příbuzných a přátel aj. Každá osoba může mít přiřazen atribut vůči většímu množství obcí.

Atribut intenzivně dojíždějího za službami je přiřazen těm jedincům, kteří již nejsou zahrnuti v dojíždce typu 1 a zároveň ve sledovaném období **danou obec navštíví alespoň 4krát a jedinec v dané lokalitě celkem stráví alespoň 8 hodin**¹². *[variabilní parametr]*

Typ 3 = občasná dojíždka za službami – nejslabší dojíždková vazba, u které lze identifikovat pravidelnost v chování a lze ji proto považovat za cílenou a periodickou dojíždku. Lze ji připodobnit k nedenní dojíždce za specifickými službami, jako jsou kulturní, sociální, sportovní a jiné služby, popřípadě lékařská péče či návštěva úřadů či návštěvy příbuzných a přátel. Obdobně jako v případě typu 2 může mít osoba více takových obcí. Jedna osoba tak může mít přidělen tento atribut vůči několika obcím zároveň.

Atribut občasně dojíždějího za službami je přiřazen těm jedincům, kterým nebyl přidělen atribut 1 či 2 vůči dané obci a zároveň ve sledovaném období **danou obec navštívili alespoň 2krát a jedinec v dané lokalitě celkem strávil alespoň 4 hodiny**¹³. *[variabilní parametr]*

Typy dojíždky 1, 2 a 3 – pro všechny typy jsou **započítávány jen ty pobyty v dané obci, které jsou delší než 30 minut**. Kratší pobyty nejsou do analýzy vůbec zahrnuty.

¹¹ Odpovídá intenzitě dojíždky 4krát týdně po dobu 5 hodin v dané obci, přičemž srovnávacím obdobím je celé sledované období 4 týdnů.

¹² Odpovídá intenzitě dojíždky jednou týdně po dobu 2 hodin v dané obci, přičemž srovnávacím obdobím je celé sledované období 4 týdnů.

¹³ Představuje nejnižší intenzitu dojíždky, u které je prokazatelná periodicitu a nejedná se tak o nepravidelné, nahodilé cesty či pouze tranzitující osoby.

Pro všechny typy dojíždky (1–3) jsou stejná pravidla ohledně měření frekvence (četnosti) návštěv dané obce. **Jednotlivé návštěvy cílové obce se sčítají po dnech**¹⁴.

Druhé bydlení – specifický typ dojíždky, kdy rezident jedné obce pravidelně nocuje v jiné obci. Charakter druhého bydlení může mít různou intenzitu i periodicitu (většinou týdenní rytmus) a liší se v průběhu roku. Typově se jedná zejména o víkendové pobyty na chalupách či chatách, dojíždka studentů či pracovníků vícedenního charakteru apod. Každý jedinec může mít stanovenou jen jednu či žádnou obec druhého bydlení. V případě, že níže stanoveným parametrům vyhovuje více obcí, jedinci je přiřazen atribut druhého bydlení vůči té obci, kde přenocoval častěji, v případě shody, kde strávil více času. Atribut druhého bydlení je nezávislým na ostatních attributech. Jedinci může být přiřazen atribut druhého bydlení vůči stejné obci, kde již má přiřazen atribut dojíždky typu 1–3.

Atribut druhého bydlení je přiřazen těm osobám, u kterých je intenzita druhého nejčastějšího místa přenocování alespoň 6 přenocování během sledovaného období. Za přenocování se považuje, že jedinec v daném místě stráví většinu intervalu 0:00–4:00. Zmíněných 6 přenocování v obci musí být uskutečněno v pravidelné periodicitě, tj. všechna přenocování musí být uskutečněna ve víkendovém období (PÁ–SO, SO–NE, NE–PO) nebo ve všední dny (PO–ÚT, ÚT–ST, ST–ČT, ČT–PÁ)¹⁵. *[variabilní parametr]*

Nocující návštěvník – osoba, která v dané obci není rezidentem ani pro ni není místem pravidelné dojíždky či druhým bydlením, avšak přesto zde přenocuje. Jedné osobě/SIM kartě může být přiřazeno více míst, kde je nocujícím návštěvníkem.

Nocujícím návštěvníkem je ta osoba, která v obci ve sledovaném čtyřtýdenním období alespoň 1x přenocovala (rozhodný interval 0:00–4:00), avšak není rezidentem dané obce ani daná obec není jeho místem pravidelné dojíždky či druhého bydlení. *[variabilní parametr]*

Návštěvník – osoba jednorázově navštěvující danou lokalitu a setrvávající zde při jedné návštěvě alespoň 3 hodiny. Zároveň se však nesmí jednat o žádnou z předešlých tří kategorií pravidelně dojíždějících. Jedné osobě/SIM kartě může být přiřazeno více míst, kde je návštěvníkem.

Návštěvníkem je osoba, která obec ve sledovaném čtyřtýdenním období přesně 1x navštívila a alespoň 3 hodiny zde setrvala a zároveň zde není rezidentem, nejedná se o místo jeho druhého bydlení ani není dojíždějícím prvního, druhého či třetího typu. *[variabilní parametr]*

¹⁴ To znamená, že mohl jedinec v daný den několikrát pendlovat mezi rezidenční obcí a například obcí, kde pracuje, V takovém případě budou záznamy z daného dne započteny jako jedna cesta a jejich celková délka bude odpovídat součtu času tráveného v dané obci během dílčích cest (započítávají jen ty cesty, které trvaly déle než 30 minut).

¹⁵ Tím dojde k částečnému odstínění rekreatantů (velký počet přenocování v kuse), a odstínění nočních směnných provozů (směny v různých částech týdne).

Projekce na populaci:

Zadavatel požaduje, aby výsledné databáze vypovídaly o prostorových vzorcích chování obyvatel České republiky. Výsledky analyzovaného vzorku je proto nutné vztáhnout k celkové populaci. K zobecnění/projekci na populaci je nutné využít vhodných specifických nástrojů¹⁶.

Veškeré operace provedené s daty musí být popsány a vysvětleny v doprovodném metodickém dokumentu („Metodický postup řešení“) k dodávaným Finálním databázím.

Ze strany zadavatele je požadována dodávka 3 kompletních databází sledujících 3 skupiny jevů za každé měření.

I. Data o počtu obvykle bydlicího obyvatelstva

Jedná se o tabulku obsahující statistické údaje o předpokládaném počtu rezidentů každé obce Česka. Obsah databáze je navrhnout v tabulce 1.

Sloupec C – počet osob s atributem rezident v této obci. Tj. počet osob, pro které je obec definovaná ve sloupci A rezidenční obcí.

Sloupec D – nevyjíždějící rezidenti. Jedná se o počet rezidentů obce, kteří nemají vůči jiné obci přidělen žádný atribut dojížděky 1–3.

Sloupec E – občasně vyjíždějící za službami. Jedná se o počet rezidentů obce, kteří mají vůči jiné obci atribut dojížděky 3. typu, ale nemají vůči žádné obci atributy 2. či 1. typu.

Sloupec F – intenzivně vyjíždějící za službami. Jedná se o počet rezidentů obce, kteří mají vůči jiné obci atribut dojížděky 2. typu, ale nemají vůči žádné obci atribut 1. typu.

Sloupec G – intenzivně vyjíždějící za službami. Jedná se o počet rezidentů obce, kteří mají vůči jiné obci atribut dojížděky 1. typu (příčemž může, ale nemusí mít vůči dalším obcím atributy 2. či 3. typu).

Sloupec H – druhé bydlení. Jedná se o počet rezidentů jiných obcí, pro které je obec definovaná sloupcem A obcí druhého bydlení.

Sloupec I – počet nocujících návštěvníků. Jedná se o počet rezidentů jiných obcí, kteří vůči obci definované sloupcem A naplňují parametry pro označení za nocujícího návštěvníka.

¹⁶ Ty zohledňují parametry jako např. odhady podílu vzorku na populaci odvislé od podílu konkrétního operátora na místním trhu, lokální demografické struktury, evidenční počet obyvatel dle konvenčních zdrojů aj. Tj. součet hodnot všech cest typu 1 pro jednu zdrojovou obec.

Sloupec J - počet návštěvníků. Jedná se o počet rezidentů jiných obcí, kteří vůči obci definované sloupcem A naplňují parametry pro označení návštěvníka.

Sloupec C je roven součtu hodnot ve sloupcích D–G.

Tabulka 1: Návrh výsledné databáze údajů o počtu rezidentů

Struktura dat požadovaných pro analýzu dojížděkových vztahů v Česku									
A	B	C	D	E	F	G	H	I	J
Data o obvykle bydlícím obyvatelstvu – počty rezidentů									
Obec	ZUI obec	Počet rezidentů (počet obyvatel)	z toho				Druhé bydlení	Nocující návštěvníci	Návštěvníci
			nevyjíždějící za službami ani za prací/školou	občasně vyjíždějící za službami	intenzivně vyjíždějící za službami	vyjíždějící za prací/školou			
Říčany	123456	15 650	10 000	4 000	1 000	650	750	550	250
Praha	654321	1 253 641	1 200 000	40 000	10 000	3 641	57 000	25 000	110 000
Kolín	987654	31 154	25 000	4 000	1 000	1 154	1 500	750	1 200

U této databáze dodavatel následně provede porovnání výstupů (počtu rezidentů) s konvenčními zdroji popisujícími obdobné jevy (evidence obyvatel MV, počet obyvatel ČSÚ popř. jiné). K porovnání bude využito statistických metod a případné neúměrné rozdíly oproti konvenčním zdrojům budou popsány a vysvětleny. Neopodstatněné rozdíly pak povedou k úpravám výpočtu.

II. Data o meziobecní dojízdce

Jedná se o **OD matici popisující směrově určenou pravidelnou meziobecní vyjížděku/dojížděku**. V databázi je identifikována **zdrojová obec, cílová obec a typ uskutečňované cesty**. Databáze tak reprezentuje počet osob, které daný typ cesty (daný atribut) pravidelně uskutečňují. Databáze zahrnuje všechny meziobecní relace (hlavní i vedlejší dojížděkové/vyjížděkové směry). Obsah databáze je navrhnut v tabulce 2.

Sloupec E – stanovuje typ cesty dle definovaných 3 typů dojížděky a dalších atributů (druhé bydlení, nocující návštěvník, návštěvník).

Sloupec F – typ cesty 1, 2, 3 určuje počet rezidentů obce (sloupec A), pro které je cílová obec (sloupec C) cílem dojížděky 1. typu, 2. typu a 3. typu. Jedná se tedy o **počet rezidentů dané obce, kteří mají přiřazen dojížděkový atribut (typ cesty 1, 2, 3) k cílové obci**.

Typ cesty 4, 5, 6 určuje počet rezidentů obce (sloupec A), pro které je cílová obec (sloupec C) místem druhého bydlení (typ cesty 4), místem kde je nocujícím návštěvníkem (typ cesty 5) nebo návštěvníkem (typ cesty 6).

Sloupec G – počty osob uskutečňujících cesty během obvyklé pracovní doby. Zobrazuje stejný ukazatel jako sloupec F, s tím rozdílem, že v tomto případě jsou započítávány jen ty cesty, které jsou alespoň z části uskutečňovány ve **všední dny v době od 8 do 17 hodin**, přičemž pobyt v cílové obci může být zahájen před tímto intervalem a ukončen po intervalu. Jinými slovy do této statistiky jsou počítány jen ty osoby, které daná kritéria pro přidělení atributů splňují, přesto že jsou započítávány pouze ty cesty/pobyty, které byly uskutečňovány ve stanoveném intervalu.

Tabulka 2: Návrh výsledné databáze údajů o meziobecní dojížděcí

Struktura dat požadovaných pro analýzu dojížděčkových vztahů v Česku						
A	B	C	D	E	F	G
Data o meziobecní dojížděcí						
zdrojová obec	ZUJ ZO	cílová obec	ZUJ CO	typ cesty	počet osob	počet osob (PO-PÁ, 8-17)
Říčany	123456	Praha	654321	1	1000	750
Říčany	123456	Praha	654321	2	2500	2100
Říčany	123456	Praha	654321	3	3500	2900
Říčany	123456	Praha	654321	4	150	50
Říčany	123456	Praha	654321	5	300	100
Říčany	123456	Praha	654321	6	3000	200
Říčany	123456	Kolín	987654	1	500	350
Říčany	123456	Kolín	987654	2	1500	1300
Říčany	123456	Kolín	987654	3	2500	2100
Říčany	123456	Kolín	987654	4	120	50
Kolín	987654	Říčany	123458	1	200	150
Kolín	987655	Říčany	123458	2	300	200
Kolín	987656	Říčany	123458	3	500	350

Součet počtu rezidentů vyjíždějících z dané obce za prací/školou (typ 1) do všech směrů¹⁷, musí být roven počtu vyjíždějících za prací/školou definovaných v tabulce 1 sloupec G.

III. Denní chod počtu přítomného obyvatelstva

Jedná se o rozsáhlou databázi, která **v každé obci Česka sleduje počet obvykle přítomného obyvatelstva během jednotlivých dnů v týdnu**. Za každou z 6 254 či 6 258 obcí je v databázi k dispozici denně 24 hodinových intervalů, 7 dní v týdnu, tj. celkem 168 záznamů ke každé obci a každému sledovanému údaji. Databáze obsahuje statistické údaje na základě sledování během celého měření, tedy 4 po sobě následujících týdnů (celkově jsou v rámci zakázky plánována čtyři taková měření). Každý jedinec je za každý interval přiřazen maximálně k jedné obci, a to k té, kde stráví dominantní část intervalu. Počet obvykle přítomných osob v obci je dále členěn dle jednotlivých atributů přiřazených konkrétním jedincům. V této databázi budou uváděny

¹⁷ Tj. součet hodnot všech cest typu 1 pro jednu zdrojovou obec.

jen ty osoby, jejichž pobyt v obci trval minimálně 30 minut (viz část věnovaná sledování pohybu/pobytu uvedená výše v textu).

Obsah databáze je navrhnut v tabulce 3.

Sloupec E – Celkový počet (obvykle) přítomných osob. Jedná se o **počet osob přítomných v obci v daném intervalu**, přičemž jejich pobyt je v délce minimálně 30 minut (viz část věnovaná sledování pohybu/pobytu uvedená výše v textu).

Sloupec F – Jedná se o počet přítomných osob v obci v daném časovém intervalu, které **mají atribut rezidenta dané obce**.

Sloupce G-I – Jedná se o počet přítomných osob v obci v daném časovém intervalu, které **mají atribut dojíždějího (typ 1, 2 a 3)**

Sloupec J – Jedná se o počet přítomných osob v obci v daném časovém intervalu, **kterí mají vůči obci status nocujícího návštěvníka**.

Sloupec K – Jedná se o počet přítomných osob v obci v daném časovém intervalu, **kterí mají vůči obci status návštěvníka**.

Sloupec L – Jedná se o počet přítomných osob v obci v daném časovém intervalu, **pro které je obec místem druhého bydlení**

Sloupec M – Jedná se o počet přítomných osob v obci v daném časovém intervalu, které **nemají přiřazen žádný atribut vůči této obci**.

Tabulka 3: Návrh výsledné databáze údajů o denním chodu počtu přítomného obyvatelstva

Struktura dat požadovaných pro analýzu dojížděkových vztahů v Česku													
A	B	C	D	E	F	G	H	I	J	K	L	M	
Denní chod počtu přítomných osob/obyvatelstva													
Obec	ZUJ obec	den v týdnu	časový interval	Celkový počet přítomných osob	z toho							druhé bydlení	tranzitující a jiné
					rezidenti	dojíždějící za prací/školou	intenzivní dojížděka za službami	občasná dojížděka za službami	nocující návštěvníci	návštěvníci			
Říčany	123456	pondělí	00:00 - 01:00	15 231	13 514	610	326	438	50	30	112	151	
Říčany	123456	pondělí	01:00 - 02:00	15 652	13 850	690	312	480	50	25	98	147	
Říčany	123456	pondělí	02:00 - 03:00	15 134	13 451	584	351	385	51	35	105	172	
Říčany	123456											
Říčany	123456	pondělí	23:00 - 24:00	15 205	13 389	681	321	435	60	40	121	158	
Říčany	123456	úterý	00:00 - 01:00	15 286	13 405	685	326	451	55	45	103	216	
Říčany	123456											
Praha	654321	pondělí	00:00 - 01:00	1 195 456	1 152 351	23 512	13 462	2 485	750	600	1 268	1 028	
Praha	654321											
Kolín	987654	pondělí	00:00 - 01:00	31 154	27 201	1 523	658	845	60	110	364	393	

Součet hodnot ve sloupcích F, G, H, I, J, K, L a M musí být roven hodnotě sloupce E.

3) Požadavky zadavatele na Postup získávání dat

Dodavatel je povinen předložit v rámci Návrhu řešení část „Postup získávání dat“, kde podrobně popíše způsob získávání dat a jednotlivé operace na nich učiněné.

Postup získávání dat musí obsahovat zcela konkrétní odpovědi dodavatele na níže popsané 4 tematické okruhy, které budou předmětem hodnocení.

V rámci části „Postup získávání dat“ musí dodavatel poskytnout vysvětlení k následujícím tematickým okruhům:

- (i) Tvorba sítě servisních polygonů BTS
- (ii) Technika zpracování záznamu
- (iii) Řešení problémů vycházející ze samotného fungování sítě
- (iv) Projekce na populaci

A. Požadavky zadavatele na popis tvorby sítě servisních polygonů BTS

Dodavatel je povinen uvést část označenou jako „Tvorba sítě servisních polygonů BTS“, ve které popíše následující oblasti řešení veřejné zakázky:

a) Vymezení sítě servisních polygonů (buněk) BTS

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak dochází k vymezení sítě servisních polygonů (buněk) BTS?
- jaké aspekty síť zohledňuje (např. orientace vysílače, terén, půdní pokryv)?
- jaké konkrétní nástroje k tomu využívá?¹⁸

b) Řešení problematiky překryvu polygonů (buněk) BTS (jedna lokalita v dosahu více BTS)

Území obsluhovaná jednotlivými BTS se v realitě překrývají. Jednotlivé buňky modelu servisní sítě BTS však musí být bez překryvu.

Dodavatel zcela konkrétně zodpoví následující otázky:

- na základě jakých parametrů je území překryvu přidělováno k jednotlivým buňkám modelu servisní sítě BTS?
- jak je zajištěno, že SIM karty reálně se vyskytující v této oblasti (oblasti překryvu) se budou převážně (či s větší pravděpodobností) připojovat k BTS, ke které je oblast přiřazena?

¹⁸ Podstatu těchto nástrojů a metod podrobně popíše. Není dostačujícím pouze tvrzení, že vymezení zpřesňuje například místním šetřením, pomocí metod GPS lokátorů či jiných nástrojů.

c) Nesoulad buněk modelu servisní sítě BTS s administrativními hranicemi obcí

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak dodavatel řeší problém nesouladu sítě buněk BTS se správními obvody obcí?
- jak dochází k přerozdělení údajů z jednotlivých buněk sítě na území jednotlivých obcí, jakých metod a nástrojů je k tomu využíváno, v čem spočívá jejich funkčnost a které aspekty zohledňují?
- jak ovlivňuje zvolená metoda/technika validitu dat a zda se validita dat v důsledku této metody prostorově liší (např. nižší přesnost měření při hranicích státu)?

d) Řešení konkrétního případu nesouladu modelu servisních buněk a administrativních hranic obcí

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak bude postupováno v případě, kdy jedna servisní buňka obsahuje kompletní intravilán dvou samostatných obcí?
- jak je v takovém případě řešeno rozdělení údajů mezi tyto obce?

B. Požadavky zadavatele na popis Techniky zpracování záznamu

Zadavatel stanoví, že dodavatel je v rámci Popisu získávání dat povinen uvést část označenou jako „Technika zpracování záznamu“, ve které popíše následující oblasti řešení veřejné zakázky:

a) Popis techniky měření v požadovaných intervalech

Dodavatel zcela přesně popíše techniku, jak dochází ke zpracování primárních záznamů do časových intervalů (max. 60min. intervaly).

Zadavatel stanoví, že využívaná varianta řešení musí být dodavatelem jednoznačně popsána a vysvětlena.

b) Technika měření na konkrétním příkladu

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak jsou řešeny tzv. migrační intervaly¹⁹, tj. čas mezi záznamy během kterého došlo k pohybu SIM karty?

¹⁹ Jedná se o období mezi dvěma po sobě následujícími připojeními ke dvěma různými BTS.

- jak se konkrétně projeví v databázi příklad popsany obrázkem níže²⁰



- jak zvolená technika měření řeší naznačený migrační interval mezi 7:06–7:50, ke které lokalitě je tento interval přiřazen?
- k jaké lokalitě bude sledovaný jedinec (SIM karta) přiřazen v případě 60minutových intervalů (snímků), tj. čas mezi 7:00 a 8:00 bude přiřazen k jaké lokalitě a proč?

C. Požadavky zadavatele na popis Řešení problémů vycházejících ze samotného fungování sítě

Zadavatel stanoví, že dodavatel je v rámci Postupu získávání dat povinen uvést část označenou jako „Řešení problémů vycházejících ze samotného fungování sítě“, ve které popíše, jak jsou řešeny konkrétní problematické jevy, jež se v rámci samotného fungování sítě projevují a mohou negativně ovlivnit validitu dat, a to prostřednictvím popisu následujících oblastí řešení veřejné zakázky:

a) Řešení tzv. překmitávání

Tzv. překmitáváním se rozumí situace, kdy může SIM karta změnit vysílač BTS, ke kterému se hlásí, aniž by však došlo k jejímu fyzickému pohybu²¹.

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak je problém tzv. překmitávání řešen a jak je zajištěno, aby překmitávání nenarušilo výpovědní hodnotu dat?

b) Pohyby SIM karet mimo dosah servisní sítě

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak je řešena problematika lokalit na okraji servisní sítě (příhraničních oblastí, periferní oblasti aj.), ve kterých mohou jedinci často během dne zmizet z dosahu sítě a nejsou k jejich pohybu/pobytu jakékoli záznamy²²?
- jak jsou takové případy na základě zvolené techniky/metody přiřazovány ke konkrétním lokalitám v jednotlivých časových intervalech?
- jak je zajištěno, aby tím nedošlo ke zkrácení dat v těchto oblastech?

²⁰ Příklad je teoretický a nemusí tak být zcela v souladu s jinými výše definovanými požadavky jako například délka základního intervalu, minimální doba pobytu v lokalitě apod.

²¹ Překmitávání má mnoho možných důvodů např. pobyt na okraji servisní buňky, vyhodnocení signálu sousední BTS jako silnějšího, momentální přetížení některých BTS aj.

²² To se může týkat například osob pracujících v zahraničí, které na významnou část dne zcela zmizí ze sítě a opětovně se přihlásí až večer.

Obdobně se projevují i případy, kdy jedinec na určitou část dne vypíná své mobilní zařízení. Řešení těchto situací musí být rovněž v popisu zahrnuto.

D. Požadavky zadavatele na popis Projekce na populaci

Zadavatel stanoví, že dodavatel je v rámci Postupu získávání dat povinen uvést část označenou jako „Projekce na populaci“, ve které popíše následující oblasti řešení veřejné zakázky:

a) Technika přepočtu sledovaného vzorku na populaci

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak dochází k přepočtu ze sledovaného vzorku (SIM karty) na statistické údaje popisující chování celé populace?
- jak konkrétně je přepočet uskutečňován, jaké parametry do něj vstupují a jaké aspekty jsou zohledněny?

b) Konvenční datové sady, ke kterým je zkoumaný vzorek vztahován či kalibrován

Dodavatel zcela konkrétně zodpoví následující otázky:

- k jakým ukazatelům jsou data vztahována, popřípadě kalibrována?
- jak se výsledné údaje liší vůči konvenčním zdrojům obdobných dat (ČSÚ, evidence obyvatel aj.)?

c) Reprezentativnost výsledných dat

Mobilní zařízení jsou využívána zejména částí populace s vysokou osobní mobilitou, naopak osoby nevlastnící mobilní zařízení jsou osoby s nízkou každodenní mobilitou (typově děti přibližně do 10 let, senioři aj.).

Dodavatel zcela konkrétně zodpoví následující otázky:

- jak zohledňuje navržené řešení přepočtu (projekce na populaci) problematiku reprezentativnosti vzorku?
- jak je v tomto smyslu zajištěna zobecnitelnost výsledků na za celou populaci?

