

Pytania i odpowiedzi do wyzwania Govtech – Gmina Miejska Świdnik.

1. *Jak należy rozumieć sytuację w której mamy powielenie rekordu? W bazie 05_SP5 mam 4KkuhYErVz - 17 numer budynku i powielenie rekordu. Tak jak by podwójnie w różnych odstępach czasu powielono wpis? Np. w odstępie pół roku? Nie ma daty, która to potwierdza. Co robimy w takiej sytuacji?*

W bazach oświatowych może wystąpić wiele wierszy dotyczące tego samego adresu. Rozwiązanie powinno zaakceptować i przetworzyć wprowadzanie uczniów pojedynczo (każdy wiersz to 1 uczeń z danego adresu), lub sumarycznie (każdy wiersz to suma uczniów z danego adresu). Może się zdarzyć np. tak, że jak w gospodarstwie domowym jest 4 dzieci, to 3 może chodzić do jednej szkoły i 1 do innej szkoły. W takim wypadku liczba opiekunów zawsze jest taka sama, liczba dzieci to ich suma (niezależnie od tego, czy wprowadzający sumował uczniów, czy wpisywał każdego w osobnym wierszu).

2. *Mamy sytuacje, gdzie deklaracja śmieciowa zawiera 2, a żadna z baz pod danych adresem nie ma wpisu. Rozumiem, że taką sytuację traktujemy jako prawidłową. Nie mamy jak zweryfikować adresu bo nie występuje w innej bazie – chyba że mamy większe zużycie wody.*

TAK – to prawidłowa/możliwa sytuacja

3. *Co oznacza pewność wnioskowania wyrażona w procentach? Jeżeli jakakolwiek baza wskazuje większą liczbę niż w deklaracji to mamy 100% pewności że dochodzi do zaniżania osób w deklaracji śmieciowej. Czy dla Państwa pewność wnioskowania to jak największe różnice osobowe między bazami? Czyli deklaracja np. 2 osoby – bazy inne 15 osób wskazują. Czyli wielokrotne przekroczenie więc duże prawdopodobieństwo nieprawidłowości.*

NIE – pewność wnioskowania oznacza jak największe rozbieżności pomiędzy bazą deklaracji o pozostałymi bazami (im większe i im więcej rozbieżności, tym większa szansa, że deklaracja została złożona nieprawidłowo)

4. *Co rozumiecie Państwo pod pojęciem „szablon danych”? Jak rozumiem ma być to osobny plik Excel do wprowadzania danych przed jednostkę samorządu lub podmiot zewnętrzny?*

TAK, powinien być to osobny plik, który będzie zaimportowany do głównego arkusza przetwarzającego dane

5. *Czy szablony danych mają przewidywać hurtowe wprowadzenie danych z rejestrów własnych jednostek np. poprzez CTRL+C/CTRL+V czy też umożliwiać wprowadzanie danych rekord po rekordzie?*

Szablon powinien posiadać obie funkcjonalności.

6. *W jakiej wersji Excel będą testowane rozwiązania?*

Microsoft Office 2016

7. *Rozumiem, że wymiennosc pojęć arkusz, system, aplikacja dopuszcza możliwość stosowania makr VBA w arkuszu Excel?*

TAK

8. *Rozumiem, że każda jednostka będzie wypełniać swój plik z danymi i dostarczać go do referatu/wydziału zajmującego się sprawdzeniami deklaracji śmieciowych a nie, że będzie krążył jeden plik pomiędzy wszystkimi? Czyli taki arkusz jak mamy teraz ma być wynikiem scalenia wszystkich otrzymanych szablonów?*

TAK, każda z jednostek otrzyma szablon, który uzupełni danymi i przekaże do urzędu

9. *W 01_DEKLARACJE występują dane od zarządców nieruchomości tj. wspólnot, spółdzielni etc. ale także od pojedynczych właścicieli. Rozumiem, że dane od pojedynczych właścicieli będą trafiały do tej bazy z Państwa własnych rejestrów na podstawie złożonych deklaracji. Czy własne dane będziecie traktować jako „zarządcę nieruchomości” i wprowadzać je tym samym szablonem czy też chcecie je uzupełniać w jakiś inny sposób?*

TAK – własne dane będziemy traktować jako „zarządcę nieruchomości”

10. *W danych szkolnych nawet Uwzględniając pkt 11 z dokumentu „Pytanie odpowiedź” jest pewien błąd założeń. Błędem tym jest kluczowanie danych po samym adresie. Załóżmy hipotetyczną sytuację:*

W domu jednorodzinnym mieszka dwóch braci z żonami (4 osoby). Jedno małżeństwo ma 3kę dzieci, drugie 2kę. Daje to łącznie 9 osób pod jednym adresem.

Niezależnie od tego czy w danych szkolnych będzie każde dziecko podane jako osobny wiersz czy też będą jedynie dwa rekordy:

Adres | Liczba dzieci | Liczba opiekunów

Adres X | 3 | 2

Adres X | 2 | 2

Przy przyjęciu, że kluczem jest Adres suma wyjdzie: 3 dzieci + 2ka dzieci + 2 opiekunów co da łącznie 7 osób a nie 9.

TAK, to możliwa sytuacja i nie da się jej uniknąć bez identyfikacji/porównywania osób (a takie było założenie). Zakładamy jednak, że opisany przypadek występuje w bazie sporadycznie i nie będzie miał kluczowego znaczenia dla wyniku działania rozwiązania.

11. Nasunęło nam się jeszcze jedno pytanie. W punkcie 3.5 regulaminu konkursu istnieje zapis: "analiza danych z deklaracji oraz wytypowanie ewentualne nieprawidłowości (adres i liczba osób) na podstawie baz danych zamawiającego oraz innych, proponowanych."

Co oznaczają inne proponowane bazy? Czy mogą być to bazy zewnętrzne? Czy po dokonaniu oceny, że któreś z danych byłyby adekwatne i pomocne w wykrywaniu nieprawidłowości możemy prosić o ich udostępnienie przez Państwa?

Rozwiązanie powinno umożliwiać w przyszłości powiększenie liczby źródeł danych, ale na obecnym etapie zamówienia nie jest to możliwe.

12. Odnośnie dostarczonej bazy danych - jakie dane zawiera zakładka 02_MOPS? Czy są to dane z pkt. 3.10.b "dane o liczbie osób zamieszkałych pod danym adresem korzystających z 500+ oraz 300+". Jeśli tak, to co oznacza na tej zakładce kolumna "liczba osób" - liczbę świadczeniobiorców ("rodzice"), czy liczbę świadczeń ("dzieci"). Czy świadczenia 500+ i 300+ są jakoś rozdzielone czy zsumowane?

Liczba osób stanowi sumę zarówno opiekunów jak i dzieci, świadczenia są sumowane.

13. Czy opisaną w mejlu od p. Tomasza Komsty strukturę oczekiwanego rozwiązania należy rozumieć tak, że plik ma zawierać arkusze (szablony), do których będzie się wprowadzać dane, a wynik "algorytmu" ma być umieszczony na oddzielnym arkuszu, czy też wynik ma być oznaczeniem "podejrzanych" deklaracji na tym arkuszu, na którym są deklaracje (dodatkowa kolumna)?

W celu uzyskania lepszej przejrzystości oczekiwanym rozwiązaniem jest oddzielny arkusz lub osobna zakładka.