

# Wykorzystanie SQL i pakietów MS Business Intelligence do budowy aplikacji bazodanowej

Adrian Mateusz Nowosielski 319080

## 1. Wstępny zarys projektu

Celem projektu jest zbudowanie aplikacji bazodanowej, która na wejściu otrzyma dwa zbiory danych w formie tekstowej/Excel, przetworzy dane, następnie waliduje i na końcu wytworzy raport na podstawie wczytanych danych. Mój projekt realizuje problem rozliczania się z klientem w firmach informatycznych na koniec miesiąca – dana firma ma x projektów, na których pracują różne osoby / niektóre na pełny etat / niektóre na pół etatu i na różnych projektach. Projekt ma na celu zagregować wszystkie dane oraz wygenerować ogólny koszt projektów dla klienta oraz w ramach możliwości wygenerować koszty dla poszczególnych osób (osoba procedura SQL).

## 2. Opis działania aplikacji bazodanowej

Działanie aplikacji bazodanowej opiera się na dostarczeniu na początku dwóch plików ze zbiorem danych:

- na wejściu do programu otrzymujemy dwa pliki w Excel / tekstowe / format csv, w których jeden z nich zawiera informacje o projektach klientów realizowanych w firmie w danym miesiącu. W nim otrzymujemy podstawowe informacje o kliencie – nazwa projektu, miasto, nip, nazwę projektu itp.
- w drugim pliku otrzymujemy dane odnośnie do przepracowanych przez pracowników godzin na danych projektach w wybranym miesiącu. Dane muszą zawierać imię i nazwisko pracownika, liczbę przepracowanych godzin danego dnia nad projektami. Liczba godzin przepracowanych danego dnia nie może być większa niż 12.

Następnie przechodzimy do fazy importowania zbioru danych do odpowiednich tabel tymczasowych w SQL oraz odpowiedniej walidacji poszczególnych elementów:

- akceptujemy oczywiście tylko takie projekty, które firma realizowała w danym miesiącu oraz przyjmujemy, że anomalią jest, jak pracownik pracował więcej niż 160 godzin w miesiącu.
- przyjmujemy, że pracowników mamy wczytanych już w bazie danych oraz ich stawki godzinowe, stanowisko oraz projekty, do których należą w danym miesiącu.
- przyjmujemy, że dane są tylko dla jednego miesiąca, więc wszystkie pozycje powinny pochodzić z jednego miesiąca
- przyjmujemy, że klienci mogą mieć więcej niż jeden projekt w jednej firmie

Po sukcesie walidacji danych dane są eksportowane do odpowiednich tabel oraz konwertowane na odpowiednie typy danych. Następnie uruchamiana jest procedura SQL, która tworzy raport projektowy dla klienta za ostatni miesiąc zawierający liczbę godzin przepracowanych przez poszczególne osoby na projekcie, kwoty itp.

### 3. Architektura techniczna aplikacji bazodanowej

W celu utworzenia wcześniej opisanej aplikacji bazodanowej mam zamiar użyć następujących technologii:

- baza danych MS-SQL – baza danych będzie wykorzystywana do przeprowadzania procedur walidacyjnych, do generowania raportów XML, oraz do przechowywania danych z walidowanych zbiorów danych
- Python 3.11 – język programowania python będzie wykorzystany do przeprowadzania wstępnej obróbki danych. Na początku Python wywoła procedurę usunięcia wszystkich tabel tymczasowych – następnie wczyta wszystkie dane tymczasowe z plików do odpowiednich tabel i wywoła procedurę walidacji tych danych oraz następnie usunięcia odpowiednich plików wejściowych w przypadku sukcesu.
- pliki źródłowe – pliki z danymi wejściowymi opisanymi w ostatnim punkcie. Pliki mogą być w formacie text Tab Delimited Unicode lub csv. Do wczytywania danych do tabel tymczasowych będzie wykorzystana biblioteka pandas.
- procesowanie i eksport raportu do pliku XML – procedury wbudowane MSSQL

### 4. Lista wymagań powstała podczas tworzenia projektu

Wymagania funkcjonalne	Opis pojedynczego wymagania
W01	Generacja pliku XML, posiadającego informacje o kosztach na poszczególne projekty informatyczne (podział na osoby, projekty, sumę)
W02	Wymaganie pliku wraz z projektami realizowanymi w poszczególnym miesiącu w firmie informatycznej (informacje o kliencie)
W03	Wymaganie pliku wraz z odpowiednimi danymi o przepracowanych godzinach na projektach przez pracowników w danym miesiącu
W04	Słownik klientów oraz projektów, aby nie powtarzać danych (osobna tabela w bazie danych)
W05	Słownik pracowników w firmie, aby nie powtarzać danych (osobna tabela w bazie danych)
W06	Procedura generująca plik XML dla dowolnego miesiąca i dowolnego projektu realizowanego w firmie w określonym miesiącu
W07	Wydatki na projekty są generowane wyłącznie w złotych(PLN)
W08	Wszystkie błędy walidacyjne mają być przechowywane w odpowiedniej tabeli LOG
W09	Wszystkie błędy powinny wypisywać się na konsolę w przypadku złej walidacji wraz z powodem błędnej walidacji
W10	Raporty za projekty przeprowadzane na przestrzeni jednego miesiąca
W11	Walidacja danych wejściowych za pomocą metryk opisanych na początku dokumentu
W12	Możliwość wygenerowania raportu miesięcznego XML na podstawie nazwy projektu i określonego miesiąca
W13	Sprawdzenie danych pracowników, które wczytujemy, tak żeby się nie powtarzały w przypadku wstawiania
W14	Mamy tabelę klientów, która jest powiązana z tabelą projektów realizowanych przez firmę

W15	Do każdego projektu przypięci są pracownicy oraz pracownicy mają powiązanie z tabelą log_praca
W16	Pracownicy mogą być przypięci do więcej niż jednego projektu oraz klienci mogą mieć również wiele projektów