

# **Sprawozdanie z projektu „Analiza wybranych witamin, składników mineralnych w składzie żywności wybranych krajów Europejskich”**

## **Baza danych**

W bazie danych istnieje 236995 wierszy, oraz 12 kolumn zawierających różne dane. Pierwsza, druga, oraz czwarta kolumna zawierają jedynie kody produktów oraz ich ID. Kolumna trzecia określa, dla którego państwa Unii Europejskiej dany rekord się odnosi. Jest ich siedem w naszym projekcie. Kolumna piąta określa szczegółową nazwę produktu (jest ich 2503), a kolejne 3 coraz mniejsze podgrup, do których on należy (np. produkty zbożowe -> Zboża -> Ziarna zbóż). Następnie zawarte są składniki odżywcze (16) w danych produktach, jednostka, w której podajemy wartość i ilość danego składnika w danym produkcie.

## **Etapy realizacji projektu, używane funkcje w naszym skrypcie**

Na początku zainstalowano odpowiednie pakiety umożliwiające na pracę na zmiennych tekstowych, oraz wczytanie posiadanej tabeli w formacie xlsx (tabeli programu Excel). Po wczytaniu z odpowiedniego katalogu roboczego w celu przejrzystości i czytelności informacji, zmieniono wybrane kolumny nagłówkowe. Po wyświetleniu informacji o tabeli wyświetlono odpowiednie listy państw, składników, oraz innych grup danych dzięki usunięciu ich duplikatów. Następnie w programie wszystkie działania zawarte są w kolejnych funkcjach zależnych od wybranych parametrów z następującymi wynikami poniżej. Dla każdej z nich pokazano na kilku przykładach ich działanie w programie, najprościej i najszybciej wybierając numer składnika / państwa z utworzonej specjalnie w tym celu listy, przeciętny użytkownik nie powinien więc mieć problemu z ich korzystaniem.

### **Funkcja nr 1 (Nutrient\_Content\_In\_Country)**

Zawartość wybranego składnika w każdym z produktów wybranego państwa w kolejności malejącej względem zawartości danego składnika odżywczego w produktach. Parametry danej funkcji początkowo to: nazwa składnika, oraz kraju.

### **Funkcja nr 2 (Nutrient\_Content\_In\_Country\_MinMax)**

Produkt, dla którego zawartość składnika odżywczego w wybranym kraju jest największa, bądź najmniejsza. Po wywołaniu funkcji otrzymujemy odpowiedź na pytanie, oraz odpowiedni wiersz z bazy danych. Parametry danej funkcji początkowo to: nazwa składnika, oraz kraju.

### **Funkcja nr 3 (Total\_Nutrient\_In\_Products)**

Zsumowane zawartości składników odżywczych dla produktów każdego z badanych państw w celu ich porównania. Oprócz użycia poprzedniej funkcji w celu optymalizacji kodu dodany został także wiersz zawierający sumę wartości dla wszystkich państw. Wynikowe wartości zostały posortowane

w porządku malejącym, a dzięki niewielkiej liczbie państw można w wyniku odczytać także wartość maksymalną i minimalną. Parametrem danej funkcji początkowo była nazwa składnika.

#### **Funkcja nr 4 (Mean\_Of\_Nutrient\_In\_Products)**

Dzięki użyciu poprzedniej funkcji nr 3 pokazano następnie średnią zawartość danego składnika odżywczego w produktach dla każdego z badanych państw (w kolejności malejącej względem zawartości), oraz zaokrągloną do 4 miejsc po przecinku, a dzięki niewielkiej liczbie państw można w wyniku odczytać także wartość maksymalną i minimalną. Dodany został wiersz zawierający średnią ogólną dla wszystkich badanych państw. Parametrem danej funkcji początkowo była nazwa składnika.

#### **Funkcja nr 5 (PercentageDifference)**

Po obliczeniu średniej porównano inne państwa do panującej normy (średniej ogólnej) pod względem zawartości danego składnika w produktach, w formie procentowej odchyłki (posortowane malejąco). Wykorzystano wewnątrz danej funkcji funkcję poprzednią w celu zmniejszenia ilości kodu. Parametrem danej funkcji początkowo była nazwa składnika.

#### **Funkcja nr 6 (Healthier\_Countries\_Relative\_To\_Nutrient)**

Kraje, gdzie zawartość wybranego składnika odżywczego jest powyżej panującej normy (od najbardziej znaczących, „zdrowych”). W celu zmniejszenia ilości kodu wykorzystano poprzednią funkcję. Parametrem danej funkcji początkowo była nazwa składnika.

#### **Funkcja nr 7 (Less\_healthy\_Countries\_Relative\_To\_Nutrient)**

Kraje, gdzie zawartość wybranego składnika odżywczego jest poniżej panującej normy (od najbardziej znaczących, „niezdrowych”). W celu zmniejszenia ilości kodu wykorzystano funkcję nr 5, analogicznie do poprzedniej. Parametrem danej funkcji początkowo była nazwa składnika.

#### **Funkcja nr 8 (Nutrients\_Content\_In\_Product)**

Nazwy oraz ilość składników odżywczych dla danego produktu w wybranym państwie, poprzez wybranie odpowiednich wierszy z bazy danych. Parametrem danej funkcji początkowo była nazwa produktu.

#### **Funkcja nr 9 (Nutrients\_Content\_In\_Product\_In\_Country)**

Składniki odżywcze w wybranym produkcie, oraz państwie (posortowane malejąco). Użyto tutaj poprzedniej funkcji w celu uproszczenia składni, oraz optymalizacji kodu. Parametrem danej funkcji początkowo była nazwa produktu, oraz kraju.

#### **Funkcja nr 10 (Average\_Nutrients\_In\_Product)**

Lista składników odżywczych, zawartych w danego rodzaju produkcie w danym państwie, wraz z ich ilością. Wykorzystano tutaj wewnątrz funkcję nr 8 w celu optymalizacji kodu, uproszczenia składni. Parametrem danej funkcji początkowo była nazwa produktu.

#### **Funkcja nr 11 (Nutrient\_In\_Products)**

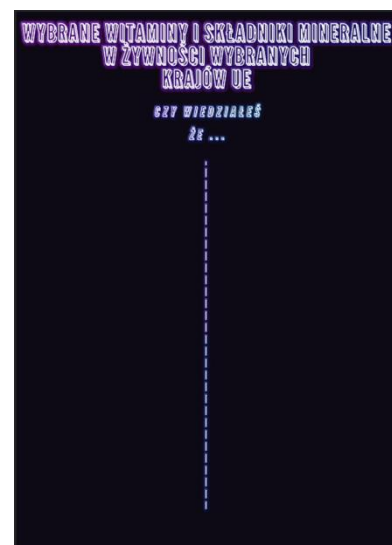
Sprawdzenie, w którym produkcie zawartość danego składnika jest największa, a także wyszukanie kraju, gdzie ta wartość osiąga maksymalną wartość. Parametrem danej funkcji początkowo była nazwa składnika, oraz kraju.

Po utworzeniu wszystkich funkcji następnie nastąpiło ich udoskonalenie – przy wywoływaniu odpowiednich funkcji zmieniono nazwę kolumny z wartościami liczbowymi wynikowymi, aby zawierała w nazwie operację, w wyniku której powstały (suma, itp.). Zauważyliśmy, że niektóre rekordy z nazwami składników posiadają napisy „Total”, które pogorszały przejrzystość dla takich przypadków. Postanowiono więc je wyszukać w kolumnie bazy, oraz dokładnie, bez pozostawienia zbędnych np. przecinków usunąć. Z powodu długich nazw niektórych składników użyto także poleceń, które podzieliły nazwy oddzielone średnikami, następnie pozostawiły jedynie ich pierwszą część z podstawową nazwą. Z powodu ilości zmienianych danych operacja ta trwała bardzo długo. W celu optymalizacji postanowiono posłużyć się typem czynnikowym, gdzie w poziomach danej kolumny zamieniono dosłownie kilka nazw, zamiast ponad kilku, czy nawet kilkudziesięciu tysięcy. Podczas wybierania poszczególnych rekordów z bazy często posługiwano się zamiennie typem ramki, oraz macierzy, która skracała potrzebny czas i znacznie upraszczała zadanie. Podczas pracy spotkano się z drobnym problemem wczytywania wartości liczbowych z ramki danych, przez co sortowanie było nieefektywne. Skorzystano wtedy z zamiany na typ wektorowy, gdy wartości liczbowe były traktowane, jako wyrazy. Po posortowaniu odpowiednio wartości w wybranych funkcjach, do porównujących kraje dodano także rekord uwzględniający wszystkie z nich, w celu porównania państw do panujących norm, bądź całkowitych sum. Po wykonaniu tych czynności następnie stworzono możliwość zapisu odpowiednich wyników działań funkcji, a także pokazania ich na wykresie wewnątrz programu, bądź, jako pliki wyjściowe w formacie jpg. Postanowiono, aby użytkownik sam miał możliwość wyboru, w jaki sposób chce uzyskać potrzebne mu informacje – dzięki odpowiednim parametrom logicznym, które dodano do większości z posiadanych skryptów. W wykresach poprawiono czytelność, przypisano odpowiednie nazwy nagłówkowe, oraz poszczególnych wartości, oraz dodano komunikaty informujące o pomyślności utworzenia wybranego wyniku pracy.

Ostatnim elementem realizacji projektu było utworzenie plakatu.

## Przygotowanie plakatu projektowego

- Szukanie inspiracji, wybranie najciekawszych wniosków do przedstawienia
- Zebranie referencji, plakatów i infografik, którymi można się inspirować
- Wybór stylu graficznego plakatu. Wybór padł na styl neonowy.
- Wybór kolorystyki – fiolet z niebieskim
- Wybór czcionek do plakatu
  1. Warsaw gothic – nagłówek i liczby na wykresach
  2. Krona One do informacji
- Wykonanie „szkieletu” plakatu



- Pobranie darmowych ikon do niektórych informacji ze strony [flaticon.com](https://flaticon.com)
- Odrysowanie dwóch wykresów

- Dodanie neonów do obrazków
- Nałożenie poprawek, polishing plakatu
- Wyeksportowanie pliku o wymiarach A3, czyli 210 x 297 mm o rozdzielczości 300 dpi zgodnie z wytycznymi na projekcie. Plakat jest profesjonalnie przygotowany do druku.

## Ostateczna wersja naszego plakatu projektowego

