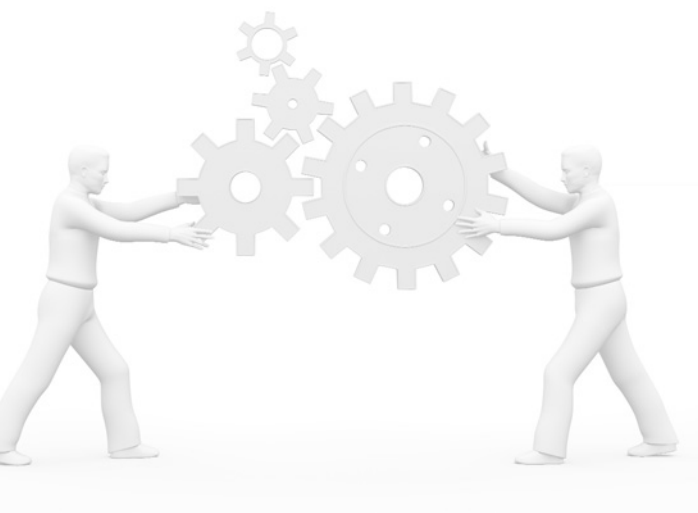


Funkcje bazy danych

Mariusz Jankowski analityk, programista rozwiązań wsparcia biznesowego;

Pytania: czytelnicy.controlling@infor.pl

W niniejszym opracowaniu omawiamy sposób działania i przeznaczenie funkcji bazy danych. Artykuł zawiera liczne przykłady dotyczące najważniejszych funkcji z tej grupy jak: BD.SUMA, BD.ŚREDNIA, BD.MAX, BD.MIN, BD.ILE.REKORDÓW.



Funkcje bazodanowe są specyficzną kategorią funkcji arkusza kalkulacyjnego. Jest to grupa funkcji jednolitych, zaprojektowanych w konkretnym celu, będąca także po części alternatywą dla funkcji statystycznych i matematycznych. Pierwszą rzeczą, która rzuca się w oczy, jest pewna konwencja dotycząca nazw funkcji. Każda z dwunastu ma ten sam przedrostek *BD* zakończony kropką (od słowa *Baza Danych*) – w wersji angielskiej jest to przedrostek *D* (*Database*). Drugą rzeczą świadczącą o pewnej jednolitości tej grupy jest fakt, że wszystkie funkcje posiadają dokładnie taką samą budowę składającą się z trzech jednakowych argumentów.

Należy zdecydowanie podkreślić, że żadna kategoria funkcji arkuszowych nie jest zorganizowana w taki lub podobny sposób. Funkcje bazodanowe operują na wybranych wierszach tabeli, takich, które spełniają określone warunki; działają także na wierszach ukrytych, zatem ich widoczność nie ma znaczenia.

Korzystanie z funkcji baz danych jest bardzo wydajną, skuteczną i przyjemną metodą dokonywania obliczeń na tabelach. Funkcje te działają bardzo szybko (w przeciwieństwie do formuł tablicowych) i są bardzo czytelne – nawet w przypadku konstruowania skomplikowanych kryteriów wyszukiwania.

Budowa funkcji bazy danych

Jak wspomniałem wyżej, każda z dwunastu funkcji bazodanowych posiada taką samą konstrukcję składającą się z trzech argumentów.

- 1) *baza* – zakres tabeli (listy) danych, łącznie z wierszem nagłówkowym,

- 2) *pole* – kolumna, na której wierszach przeprowadzane są obliczenia; argument ten może zostać wprowadzony do formuły na trzy różne sposoby:
- odwołanie do komórki zawierającej nagłówek właściwej kolumny (zalecane przez wzgląd na łatwą parametryzację),
 - tekst z nazwą kolumny zapisany w cudzysłowie,
 - numer kolumny w tabeli danych,
- 3) *kryteria* – zakres komórek, zawierający warunki określone przez użytkownika.

Kryteria – podstawowe zasady tworzenia

Widok 1 pokazuje listę zarobków pracowników zatrudnionych w pewnej firmie. W celach analitycznych potrzebne jest uzyskanie odpowiedzi na kilka prostych pytań:

- Jaka jest średnia wieku dla kobiet i mężczyzn?
- Ile jest osób w firmie z konkretnym wykształceniem?
- Jaki jest minimalny, średni i maksymalny staż pracy w firmie?
- Ile jest w firmie osób zarabiających powyżej 3000 zł?

Tabela zawiera dane z zakresu A4:F18 – wiersz pierwszy jest wierszem powielonym, pod którym określone będą kryteria. Umieszczenie kryteriów funkcji powyżej tabeli danych jest bardzo dobrą praktyką głównie ze względu na czytelność projektu. Zakres kryteriów nie powinien być umieszczony pod listą danych, aby nie ograniczać możliwości dopisywania nowych rekordów. Ze względu na trudny dostęp nie powinien być także określony obok wielokolumnowej tabeli źródłowej (**widok 1**).

Widok 1. Tabela danych z powielonym wierszem nagłówkowym dla kryteriów

	A	B	C	D	E	F	
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie	
2							
3							
4	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie	
5	Banaś	Kobieta	37	pomaturalne	14	2 450 zł	
6	Brzezicka	Kobieta	28	średnie zawodowe	3	2 280 zł	
7	Chodała	Męczyzna	34	średnie	5	1 379 zł	
8	Ditrych	Męczyzna	41	wyższe	12	3 570 zł	
9	Drzał	Kobieta	23	zawodowe	2	1 970 zł	
10	Dudek	Kobieta	35	średnie zawodowe	7	2 325 zł	
11	Galicki	Męczyzna	37	wyższe	9	4 759 zł	
12	Grzebyk	Męczyzna	35	pomaturalne	3	1 276 zł	
13	Haba	Męczyzna	66	średnie	35	2 560 zł	
14	Hunczak	Kobieta	28	wyższe	1	3 430 zł	
15	Król	Kobieta	21	średnie	2	1 058 zł	
16	Król	Męczyzna	47	średnie	17	1 740 zł	
17	Królik	Męczyzna	32	średnie zawodowe	10	2 105 zł	
18	Machocki	Męczyzna	27	wyższe	1	2 820 zł	
19							

Zobaczymy, w jaki sposób definiować kryteria i jakie formuły zastosować, aby otrzymać odpowiedź na pytania postawione wcześniej.

1. Jaka jest średnia wieku kobiet pracujących w firmie?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie		Wynik
2		=Kobieta						28,7

- Kryterium wpisane do komórki B2 to formuła: `"=Kobieta"`.
- Formuła wynikowa w komórce H2 to:

$$=BD.ŚREDNIA(A4:F18;C1;A1:F2)$$

W pierwszym argumencie funkcji pobierany jest zakres komórek dla całej tabeli. Drugi parametr zawiera odwołanie do komórki C1 ("*Wiek*"), trzeci odwołuje się do zakresu kryteriów A1:F2.

- Chcąc obliczyć średnią wieku dla mężczyzn, należy w komórce B2 wpisać formułę: "*=Mężczyzna*". Można także wpisać bezpośrednio do komórki słowo *Mężczyzna* bez cudzysłowów. W takim przypadku formuła również zwróci prawidłowy wynik, jednak bezpieczniej jest podawać kryterium tekstowe w postaci formuły (w cudzysłowie), o czym przekonamy się, analizując następny przykład.

2. Ile jest osób w firmie z wykształceniem średnim?

	A	B	C	D	E	F	G	H	
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie		Wynik	
2				średnie					7

	A	B	C	D	E	F	G	H	
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie		Wynik	
2				=średnie					4

- Tym razem celem było zliczenie osób mających średnie wykształcenie. Jak widać, wyniki różnią się od siebie, pomimo że w komórce H2 znajduje się taka sama formuła:

$$=BD.ILE.REKORDÓW(A4:F18;;A1:F2)$$

- W pierwszym przypadku do komórki D2 wpisane zostało bezpośrednio słowo *średnie* bez żadnych cudzysłowów. Takie kryterium Excel interpretuje jako ciąg tekstu zaczynający się od słowa *średnie*, a nie tekst dokładnie równy słowu *średnie*. W tym przypadku oznacza to tyle, że Excel zliczył osoby posiadające wykształcenie *średnie* i osoby posiadające wykształcenie *średnie zawodowe*. Stąd w wyniku otrzymaliśmy nieprawidłową wartość 7 (trzy osoby posiadają wykształcenie *średnie zawodowe*).
- W drugim przypadku otrzymaliśmy prawidłowy wynik 4. Excel zliczył tylko osoby, które mają wpisane wykształcenie *średnie*. Taka też była nasza intencja. Kryterium wpisane w komórce D2 ma postać formuły "*=średnie*".

3. Jaki jest minimalny, średni i maksymalny staż pracy w firmie?

	A	B	C	D	E	F	G	H	I	J	
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie		MIN	MAX	AVG	
2								1	35	8,64	

W trzecim przypadku nie wprowadzamy żadnych kryteriów, ponieważ nie interesuje nas wynik w zobiciu na konkretną zmienną, tylko dane dotyczące wszystkich pracowników. Zerknięcie na tabelę źródłową wystarczy, aby dostrzec, że osoba z najmniejszym stażem pracuje w firmie niecały rok, zaś osoba z najdłuższym stażem – 35 lat. Po zaznaczeniu wszystkich danych w kolumnie E i wybraniu na pasku stanu opcji *Średnia*, widzimy, że średni staż w firmie to 8,64 lat, i właśnie taką wartość pokazuje ostatnia formuła.

- Wszystkie trzy formuły wyliczające posiadają identyczne wartości argumentów (A4:F18;E1;A1:F2). Wykorzystano odpowiednio formuły: BD.MIN, BD.MAX i BD.ŚREDNIA.

4. Ile jest w firmie osób zarabiających powyżej 3000 zł?

	A	B	C	D	E	F	G	H	I	J	
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie		MIN	MAX	AVG	
2								1	35	8,64	

- Kryterium wpisane do komórki F2 to formuła:

$$=">"&3000$$

- Formuła wynikowa w komórce H2 to:

$$=BD.ILE.REKORDÓW(A4:F18;F1;A1:F2)$$

Koniunkcja i alternatywa warunków

Koniunkcja (iloczyn) oznacza konieczność spełnienia dwóch lub większej liczby warunków. Warunek koniunkcji zazwyczaj możemy zapisać symbolicznie za pomocą funkcji ORAZ. Alternatywa (suma) oznacza natomiast, że przynajmniej jeden z dwóch lub kilku warunków musi być spełniony. Warunek alternatywy możemy zapisać symbolicznie za pomocą funkcji LUB.

Nawiązując do przykładu z widoku 1, założmy, że firma chce odpowiedzieć sobie na bardziej szczegółowe pytania.

5. Czy spośród pracowników mających więcej niż 30 lat to mężczyźni zarabiają więcej od kobiet?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie	Wynik	
2		=Kobieta	>30					2387,5

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie	Wynik	
2		=Mężczyzna	>30					2484,14

- Koniunkcja warunków oznacza, że wszystkie warunki muszą być spełnione. Jeżeli któryś z nich zwraca wartość FAŁSZ, to wynikiem działania funkcji (testu logicznego) jest również wartość FAŁSZ. W przypadku koniunkcji (iloczynu) wszystkie warunki muszą być zapisane w tym samym wierszu, tak jak w przykładach powyżej.
- W przypadku kobiet mających więcej niż 30 lat, tylko dwie osoby spełniają te kryteria (wiersz 5 i 10), w przypadku mężczyzn jest to aż siedem osób (wiersze: 7, 8, 11, 12, 13, 16, 17).
- Obydwie formuły mają taką samą konstrukcję:

$$=BD.ŚREDNIA (A4:F18;F1;A1:F2)$$

6. Ile jest w firmie osób mających wykształcenie pomaturalne lub wyższe?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie	Wynik	
2				=pomaturalne				6
3				=wyższe				

- Alternatywa warunków oznacza, że przynajmniej jeden z warunków musi być spełniony. Jeżeli któryś z nich zwraca wartość PRAWDA, to wynikiem działania funkcji (testu logicznego) jest również wartość PRAWDA. W przypadku alternatywy (sumy) wszystkie warunki muszą być zapisane w osobnych wierszach, tak jak w przykładzie powyżej.

- Formuła w komórce H2 to:

$$=BD.ILE.REKORDÓW(A4:F18;;A1:F3)$$

Należy zwrócić uwagę na dwie rzeczy. Po pierwsze – zakres z kryteriami został powiększony o jeden wiersz, po drugie – drugi argument funkcji został pominięty, ponieważ nie dokonujemy tutaj operacji arytmetycznej, a jedynie zliczamy niepuste komórki w tabeli, które spełniają kryteria określone w trzecim argumentcie funkcji.

7. Jaki jest łączny staż pracy (w latach) kobiet mających mniej niż 30 lat i mężczyzn mających ponad 60 lat?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie	Wynik	
2		=kobieta	<30					43
3		=mężczyzna	>60					

- W tym przypadku mamy do czynienia z warunkami mieszanymi, tzn. z jednoczesnym wystąpieniem koniunkcji i alternatywy warunków. Pierwszy warunek dotyczy kobiet mających mniej niż 30 lat. Excel zlokalizował cztery takie kobiety (wiersze: 6, 9, 14, 15). Ich łączny staż pracy w firmie wynosi 8 lat. Drugi warunek odnosi się do mężczyzn mających ponad 60 lat. W firmie pracuje jeden taki mężczyzna – jego staż pracy wynosi 35 lat. Zsumowanie tych dwóch wyników daje liczbę 43.
- Formuła w komórce H2 to:

$$=BD.SUMA(A4:F18;E1;A1:F3)$$

8. Jaka jest średnia wieku pracowników z wykształceniem zawodowym lub takich, którzy zarabiają powyżej 3000 zł?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie		Wynik
2				=zawodowe				32,25
3						>3000		

- W firmie pracuje jedna osoba z wykształceniem zawodowym (wiersz 9), która ma 23 lata. W firmie pracują także trzy osoby zarabiające powyżej 3000 zł (wiersze: 8, 11, 14). Ich wiek to odpowiednio 41, 37, 28. Średnia z tych czterech liczb to 32,25 i taką też wartość podaje nam w wyniku formuła:

$$=BD.ŚREDNIA(A4:F18;C1;A1:F3)$$

Kryteria formułowe

Kryteria formułowe (wyliczeniowe) – jak wskazuje nazwa, odwołują się do formuł utworzonych za pomocą funkcji arkuszowych. Niezrozumienie zasad tworzenia takich kryteriów sprawia, że są one raczej rzadko stosowane w praktyce. W niektórych przypadkach, przy konstruowaniu bardziej szczegółowych zapytań, są jednak niezbędne, poza tym dają bardzo duże możliwości.

9. Ile jest osób w firmie zarabiających powyżej średniej?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Wykształcenie	Staż pracy	Kryterium!		Wynik
2						TRUE		6

- Na pierwszy rzut oka widać, że zmienił się nagłówek w komórce F1. Zamiast słowa *Wynagrodzenie* pojawiło się tajemnicze *Kryterium!*. Chodzi o to, że jeżeli warunek jest zapisany w postaci formuły, to nie należy używać nagłówka kolumny. Należy go zastąpić innym dowolnym tekstem (niebędącym nagłówkiem żadnej innej kolumny z listy danych). Ewentualnie
- Formuła w komórce F2 to:

$$=F5>ŚREDNIA(F5:F18)$$

Formuła zwraca wartość PRAWDA, ponieważ wartość w komórce F5 (2450 zł) jest większa od średniej zarobków w firmie (2409 zł).

- Formułę dla kryterium wpisujemy, używając adresu z pierwszego wiersza danych (F5). Adres ten musi być wpisany w formule w postaci względnej. Tylko taka forma umożliwia porównanie wartości z następnymi wierszami ze średnią wynagrodzeń. W naszym przypadku formuła:

$$=F5>ŚREDNIA(F5:F18)$$

oznacza, że każda pojedyncza wartość z zakresu *\$F\$5:\$F\$18* będzie porównywana ze średnią arytmetyczną z tego zakresu (2409 zł). Jeżeli wynagrodzenie pracownika będzie wyższe od średniej – warunek zostanie spełniony i formuła zwróci wartość PRAWDA. W przeciwnym wypadku będzie to FAŁSZ.

- Formuła wynikowa w komórce H2 to:

$$=BD.ILE.REKORDÓW(A4:F18;;A1:F2)$$

Formuła zlicza pracowników, którzy zarabiają powyżej średniej (PRAWDA), i zwraca prawidłową wartość 6.

10. Czy są w firmie pracownicy posiadający takie samo nazwisko?

	A	B	C	D	E	F	G	H
1	Kryterium!	Płeć	Wiek	Wykształcenie	Staż pracy	Wynagrodzenie		Wynik
2	FALSE							2

- W tym przypadku kryterium formułowe wprowadzamy do komórki A2. Zgodnie z omówionymi wcześniej regułami, zmieniamy nazwę nagłówka dla kolumny, w której określiliśmy warunek.
- Formuła w komórce A2 to:

$$=LICZ.JEŻELI(\$A\$5:\$A\$18;A5)>1$$

Jej celem jest sprawdzenie każdego nazwiska na liście. Jeżeli nazwisko znajduje się na liście przynajmniej dwa razy, wówczas warunek zostaje spełniony, a to oznacza, że w wyniku działania funkcji zostanie zwrócona wartość większa od zera. I tak jest w istocie. Formuła w komórce H2:

$$=BD.ILE.REKORDÓW(A4:F18;;A1:F2)$$

zwraca wartość 2, a to oznacza, że w firmie pracują osoby posiadające takie samo nazwisko.

11. Ile przeciętnie zarabiają mężczyźni powyżej 30 lat, z wykształceniem zawodowym lub średnim zawodowym?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Kryterium!	Staż pracy	Wynagrodzenie		Wynik
2		=Mężczyzna	>30	FALSE				2105

- Tym razem mamy do czynienia z koniunkcją trzech warunków, dodatkowo ostatni z nich ma postać alternatywy. Pierwszy warunek odnosi się do płci pracownika, drugi do jego wieku, trzeci do wykształcenia. O ile w pierwszych dwóch przypadkach kryteria są proste i zrozumiałe, o tyle w trzecim przypadku zastosowano kryterium formułowe.
- Formuła zawarta w komórce D2 to:

$$=LUB(D5="zawodowe";D5="średnie zawodowe")$$

- Formuła w komórce H2 to:

$$=BD.ŚREDNIA(\$A\$4:\$F\$18;F1;A1:F2)$$

Wynikiem jej działania jest wartość 2105. Istnieje bowiem tylko jeden pracownik, który spełnia jednocześnie wszystkie trzy kryteria (wiersz 17).

- Skonstruowanie odpowiedniej formuły tablicowej lub formuły opartej na funkcjach logicznych byłoby w tym przypadku zadaniem trudniejszym.

12. Ile jest w firmie osób, których staż pracy jest krótszy niż 4 lata, a wynagrodzenie zawiera się w przedziale 2000–3000 zł?

	A	B	C	D	E	F	G	H
1	Nazwisko	Płeć	Wiek	Kryterium!	Staż pracy	Kryterium!		Wynik
2					<4	TRUE		2

- W tym przypadku mamy do czynienia z koniunkcją dwóch warunków. Po pierwsze – staż pracy pracownika musi być mniejszy niż 4 lata. Po drugie – wynagrodzenie pracownika musi się zawierać w przedziale 2000–3000 zł (i jest to kryterium formułowe zapisane w postaci koniunkcji).
 - Formuła określająca warunek w komórce F2 to:
- $$=ORAZ(F5>=2000;F5<=3000)$$
- Formuła wynikowa w komórce H2 to:
- $$=BD.ILE.REKORDÓW(A4:F18;;A1:F2)$$
- Formuła zwraca w wyniku wartość 2, ponieważ dwie osoby spełniają kryteria określone w trzecim argumentzie funkcji (wiersze 6 i 18).

Wnioski

Funkcje bazodanowe stanowią grupę mocno powiązanych ze sobą funkcji. Świadczy o tym nie tylko pewna konwencja przyjęta w nazewnictwie, ale także jednakowa konstrukcja i ogólne przeznaczenie. Każda funkcja bazy danych wymaga określenia osobnego zakresu kryteriów, który jest podawany jako ostatni argument. Parametr ten jest

W takim przypadku należy pamiętać o kilku sprawach:

- 1) kryteria formułowe zwracają w wyniku zawsze wartości logiczne PRAWDA/FALSZ,
- 2) w odwołaniach do kolumn należy zawsze stosować adres pierwszej komórki zawierającej dane w interesującym nas polu (postać względna),

Korzystanie z funkcji baz danych jest bardzo wydajną, skuteczną i przyjemną metodą dokonywania obliczeń na tabelach.

kluczowy, ponieważ na jego podstawie określa się, które rekordy należy brać pod uwagę przy obliczeniach. Kryteria te są wykorzystywane także przez filtr zaawansowany do pobrania wierszy spełniających konkretne warunki. Excel pozwala na stosowanie kryteriów formułowych nie tylko na podstawie funkcji logicznych, ale całego zestawu funkcji arkuszowych.

- 3) konieczna jest zmiana nazwy nagłówka lub pozostawienie komórki pustej,
- 4) można stosować koniunkcję i alternatywę warunków, także mieszać warunki ze sobą,
- 5) w wielu przypadkach wygodniejsze może być dodanie do tabeli kolumny z wyliczonymi wartościami niż zastosowanie kryterium formułowego.

dwutygodnik BIULETYN RACHUNKOWOŚCI

BIULETYN RACHUNKOWOŚCI

– dwutygodnik przygotowywany z myślą o księgowych, samodzielnych księgowych, specjalistach ds. rachunkowości, głównych księgowych, osobach prowadzących księgi rachunkowe.



Zamów prenumeratę
„BIULETYNU RACHUNKOWOŚCI”
na 2011 rok!



Informacje i zamówienia: Biuro Obsługi Klienta, ul. Okólna 40, 05-270 Marki,
801 626 666, 22 761 30 30, 22 212 30 07, e-mail: bok@infor.pl, GG 5117020

