

**Aaaby używać INDEXów
nawet gdy klauzula WHERE
zawiera warunek typu OR**

Bartosz Jabłoński

Citibank Europe PLC / MiNI PW

✉ yabwon@gmail.com

Wtorek, 8. października 2019

PolSUG



<https://www.linkedin.com/groups/8675451>

communities.sas.com/polsug

Czelendź

Jak poradzić sobie w sytuacji, w której warunek filtrujący WHERE zawiera łącznik logiczny OR między dwoma poINDEXowanymi zmiennymi, tak aby móc użyć INDEXu do szybkiego wyfiltrowania danych?

Czelendź

Jak poradzić sobie w sytuacji, w której warunek filtrujący WHERE zawiera łącznik logiczny OR między dwoma poINDEXowanymi zmiennymi, tak aby móc użyć INDEXu do szybkiego wyfiltrowania danych?

the answer

SAS Proceedings' paper 3722-2019



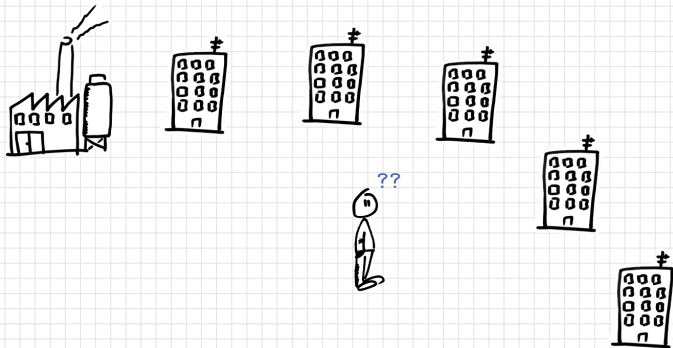


```
WHERE prod_from_Factory = "ABC";
```



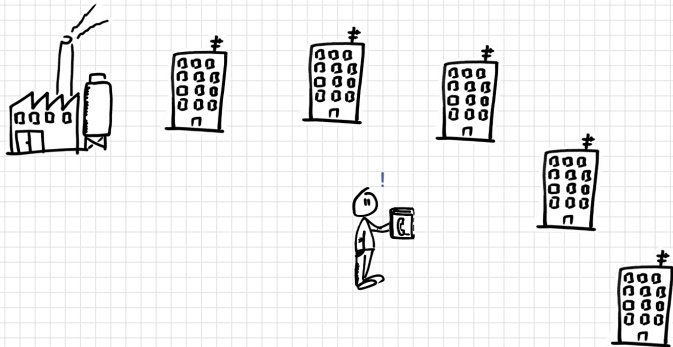
```
WHERE prod_from_Factory = "ABC";
```

małe dane? - przeczytaj sekwencyjnie



```
WHERE prod_from_Factory = "ABC";
```

duże dane?



```
WHERE prod_from_Factory = "ABC";
```

duże dane? - użyj INDEXu



```
WHERE prod_from_FactOne = "ABC" OR prod_from_FactTwo = "XYZ";  
/* niestety, INDEX tu nie pomoże... */
```



małe dane?
czytaj sekwencyjnie

```
WHERE prod_from_FactOne = "ABC" OR prod_from_FactTwo = "XYZ";
```

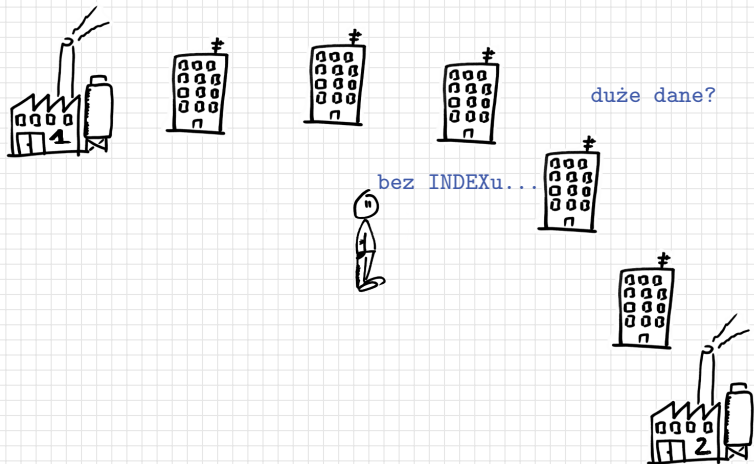


małe dane?
czytaj sekwencyjnie

index na małym zbiorze
jest jak autoalarm
we Fiacie Multipla...

```
WHERE prod_from_FactOne = "ABC" OR prod_from_FactTwo = "XYZ";
```





```
WHERE prod_from_FactOne = "ABC" OR prod_from_FactTwo = "XYZ";
```



duže dane?



```
set ... ;  
WHERE prod_from_FactOne = "ABC";
```

```
output;
```

```
set ... ;  
WHERE prod_from_FactTwo = "XYZ";
```

```
output;
```



duże dane?



```
Declare HASH List();  
List.DefineKey("CurObs");  
List.DefineDone();
```



```
set ... curobs = CurObs;  
WHERE prod_from_FactOne = "ABC";  
List.ADD();  
output;
```

```
set ... curobs = CurObs;  
WHERE prod_from_FactTwo = "XYZ";  
if List.FIND() then output;
```

Obrazek Poglądowy

Zerknijmy na mały przykład, który wyjaśni wszystko.

Zbiór z dwiema poindeksowanymi zmiennymi

Obs.	VarL	VarN
1	A	10
2	A	10
3	A	20
4	B	20
5	B	20
6	B	30
7	C	30
8	C	30
9	C	40

WHERE VarL="B" or VarN=20

WHERE VarL="B"

WHERE VarN=20

Zbiór w obserwacjami wczytanymi przez warunki
WHERE VarL="B" oraz WHERE VarN=20

Current Obs.	VarL	VarN	
4	B	20	✓
5	B	20	✓
6	B	30	✓
3	A	20	✓
4	B	20	✗
5	B	20	✗

Obserwacje przeczytane warunkiem WHERE VarL="B"

Current Obs.	VarL	VarN	
4	B	20	✓
5	B	20	✓
6	B	30	✓

Lista numerów obserwacji

numer obserwacji	"4"	"5"	"6"
------------------	-----	-----	-----

Obserwacje dodane warunkiem WHERE VarN=20

Current Obs.	VarL	VarN	
4	B	20	✓
5	B	20	✓
6	B	30	✓
3	A	20	✓
4	B	20	✗
5	B	20	✗

Live Demo?

Zajrzyj pod adres www poniżej,
przeczytaj i odpal kody, pobaw się nimi!

www: `http://www.mini.pw.edu.pl/~bjablons/SASpublic/`

kody zaczynają się od:

`OR-condition-in-WHERE-clause-with-INDEX-*.sas`
(and `Countries.sas`)

dziękuję!



- Bartosz Jabłoński, "Use the Advantage of INDEXes Even If a WHERE Clause Contains an OR Condition", SAS GF 2019 Proceedings, <https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3722-2019.pdf>
- Art Carpenter, "Carpenter's Guide to Innovative SAS Techniques", SAS Press
- Paul M. Dorfman, Koen Vyverman, "Data Step Hash Objects as Programming Tools", SUGI 30 Proceedings, www2.sas.com/proceedings/sugi30/236-30.pdf
- Chris Schacherer, "Introduction to SAS Hash Objects", SAS GF 2015 Proceedings, <https://support.sas.com/resources/papers/proceedings15/3024-2015.pdf>
- Paul M. Dorfman, Don Henderson, "Data Management Solutions Using SAS Hash Table Operations: A Business Intelligence Case Study", SAS Press
- Paul M. Dorfman, "Fundamentals of the The SAS Hash Object", SESUG 2016 Proceedings, http://analytics.ncsu.edu/sesug/2016/H0W-195_Final.PDF.pdf
- Billy Clifford, "Frequently Asked Questions about SAS Indexes", SUGI 30 Proceedings, <http://www2.sas.com/proceedings/sugi30/008-30.pdf>
- Michael A. Raithel, "The Complete Guide to SAS Indexes", SAS Press
- Mark Keintz, "A Faster Index for Sorted SAS Datasets", SAS GF 2009 Proceedings, <https://support.sas.com/resources/papers/proceedings09/024-2009.pdf>