

Consulta 1

SI - Banco de Dados

Novembro, 2020

Nome	Nr. USP
Edgard van Tol Taver	10686575
Guilherme Vaz de Sousa Ribeiro	10300320
Pedro Raul Taborga da Costa	4537076

Enunciado

O time de Marketing do Evento quer analisar as melhores palavras para colocar nos anúncios. Para isso, eles gostariam de saber o custo total de anúncios que contém uma determina palavra ou expressão no título.

SQL

Para, por exemplo, pegar anúncios que contenham a palavra **cultivate**:

```
WITH anuncio_preparado as (
  SELECT idanuncioonline, CONCAT(' ', titulo, ' ') as titulo, descricao, custoporclique * acessos as custo_total
  FROM public.anuncioonline
)
SELECT idanuncioonline, SUBSTR(titulo, 2, LENGTH(titulo) - 2) as titulo, descricao, custo_total FROM anuncio_preparado
WHERE titulo LIKE '%cultivate %'
```

É possível mudar o trecho **%cultivate %** para qualquer outra palavra diferente de **vision**. Exemplos:

```
WITH anuncio_preparado as (
  SELECT idanuncioonline, CONCAT(' ', titulo, ' ') as titulo, descricao, custoporclique * acessos as custo_total
  FROM public.anuncioonline
)
SELECT idanuncioonline, SUBSTR(titulo, 2, LENGTH(titulo) - 2) as titulo, descricao, custo_total FROM anuncio_preparado
WHERE titulo LIKE '%grow %'
```

```
WITH anuncio_preparado as (
  SELECT idanuncioonline, CONCAT(' ', titulo, ' ') as titulo, descricao, custoporclique * acessos as custo_total
  FROM public.anuncioonline
)
SELECT idanuncioonline, SUBSTR(titulo, 2, LENGTH(titulo) - 2) as titulo, descricao, custo_total FROM anuncio_preparado
WHERE titulo LIKE '%matrix %'
```

Veja que fazemos uma transformação no campo **titulo**, para adicionar um espaço no começo e outro espaço no fim do valor do campo.

Isso foi feito pois usamos os espaços para delimitar palavras. Sem usar os espaços, correríamos o risco de selecionar anuncios com palavras similares às buscadas.

Um exemplo seria a palavra **vision**, que está contida em **visionary** e **envisioneer**. Se buscássemos por **LIKE '%vision%'**, sem usar espaços, então retornaríamos, além de anúncios que possuam de fato a palavra **vision**, mas também anúncios com as palavras **visionary** e **envisioneer**, o que não é desejado. Para evitar isso, buscamos com espamos: **LIKE '%vision%'**.

Mas como estamos adicionando esse espaço na busca, então teríamos problemas caso a palavra desejada esteja no começo ou no fim do título do anúncio. Para contornar isso, preparamos o campo **titulo** adicionando espaços nos extremos e fazemos a busca em cima disso.

E, para remover essa transformação, aplicamos o **SUBSTR**, de forma que as tuplas retornadas não apresentem os espaços adicionais nas extremidades.

Resultados esperados

A tabela **anuncioonline** está populada com as seguintes tuplas:

idanuncioonline	titulo	descricao	url	acessos	custoporclique	fk_event
1	brand seamless applications	Fully- configurable logistical system engine	https://storify.com/proin/eu/mi/nulla/ac/enim/in.xml	10251	0.84	2

id	anuncioonline	titulo	descricao	url	acessos	custoporclique	fk_event
2		cultivate proactive initiatives	Monitored content-based system engine	http://scribd.com/quam/fringilla/rhoncus/mauris/enim/leo.xml	11508	0.34	4
3		morph efficient systems	Mandatory non-volatile Graphical User Interface	https://yahoo.co.jp/id/pretium/iaculis/diam/erat/fermentum/justo.html	11309	3.8	1
4		monetize interactive networks	Open-source needs-based leverage	http://nytimes.com/blandit/nam/nulla/integer/pede/justo/lacinia.png	5747	2.66	2
5		facilitate turn-key infomediaries	Customer-focused tertiary focus group	https://cbc.ca/massa/donec.png	11311	1.74	3
6		optimize cross-platform content	Enterprise-wide 24 hour matrix	http://wunderground.com/integer/non.html	4452	4.63	1
7		grow visionary e-markets	Object-based interactive alliance	http://yahoo.com/in/hac/habitasse.png	586	4.41	1
8		envisioneer mission-critical supply-chains	Enhanced encompassing forecast	https://hp.com/mus.json	11254	6.19	5
9		transform world-class schemas	Reduced well-modulated product	https://blogspot.com/quam.xml	6488	1.19	4
10		deploy sticky experiences	Total asymmetric model	http://pen.io/ut.jsp	9767	4.01	3
11		recontextualize revolutionary e-services	Multi-channelled upward-trending contingency	http://symantec.com/habitasse/platea/dictumst/morbijsp	2281	4.3	2
12		matrix integrated models	Balanced zero administration challenge	http://economist.com/mauris/ullamcorper/purus/sit/amet/nulla.xml	4966	3.63	1
13		incentivize e-business web-readiness	Reverse-engineered client-driven challenge	http://lycos.com/in/leo/maecenas.json	5235	0.98	3
14		envisioneer sticky schemas	Quality-focused background array	https://privacy.gov.au/porta/volutpat/quam/pede/lobortis/ligula.xml	10821	2.98	5
15		cultivate strategic infomediaries	Streamlined context-sensitive local area network	https://scientificamerican.com/felis/fusce.jpg	6606	1.44	3

Destas tuplas, apenas 3 possuem a palavra **cultivate** no campo **titulo**. Estas são as tuplas que queremos que a consulta retorne:

id	anuncioonline	titulo	descricao	url	acessos	custoporclique	fk_evento_idevento
2		cultivate proactive initiatives	Monitored content-based system engine	http://scribd.com/quam/fringilla/rhoncus/mauris/enim/leo.xml	11508	0.34	4

idanuncioonline	titulo	descricao	url	acessos	custoporclique	fk_evento_idevento
15	cultivate strategic infomediaries	Streamlined context-sensitive local area network	https://scientificamerican.com/felis/fusce.jpg	6606	1.44	3

As tuplas restantes NÃO possuem a palavra **vision** no campo **titulo**. Não queremos estas tuplas no resultado da consulta:

idanuncioonline	titulo	descricao	url	acessos	custoporclique	fk_event
1	brand seamless applications	Fully-configurable logistical system engine	https://storify.com/proin/eu/mi/nulla/ac/enim/in.xml	10251	0.84	2
3	morph efficient systems	Mandatory non-volatile Graphical User Interface	https://yahoo.co.jp/id/pretium/iaculis/diam/erat/fermentum/justo.html	11309	3.8	1
4	monetize interactive networks	Open-source needs-based leverage	http://nytimes.com/blandit/nam/nulla/integer/pede/justo/lacinia.png	5747	2.66	2
5	facilitate turn-key infomediaries	Customer-focused tertiary focus group	https://cbc.ca/massa/donec.png	11311	1.74	3
6	optimize cross-platform content	Enterprise-wide 24 hour matrix	http://wunderground.com/integer/non.html	4452	4.63	1
7	grow visionary e-markets	Object-based interactive alliance	http://yahoo.com/in/hac/habitasse.png	586	4.41	1
8	envisioneer mission-critical supply-chains	Enhanced encompassing forecast	https://hp.com/mus.json	11254	6.19	5
9	transform world-class schemas	Reduced well-modulated product	https://blogspot.com/quam.xml	6488	1.19	4
10	deploy sticky experiences	Total asymmetric model	http://pen.io/ut.jsp	9767	4.01	3
11	recontextualize revolutionary e-services	Multi-channelled upward-trending contingency	http://symantec.com/habitasse/platea/dictumst/morbi.jsp	2281	4.3	2
12	matrix integrated models	Balanced zero administration challenge	http://economist.com/mauris/ullamcorper/purus/sit/amet/nulla.xml	4966	3.63	1
13	incentivize e-business web-readiness	Reverse-engineered client-driven challenge	http://lycos.com/in/leo/maecenas.json	5235	0.98	3
14	envisioneer sticky schemas	Quality-focused background array	https://privacy.gov.au/porta/volutpat/quam/pede/lobortis/ligula.xml	10821	2.98	5

Para obter o custo total, precisamos multiplicar **acessos** por **custoporclique**. Desta forma, a tupla final deve ter esta forma:

idanuncioonline	substr	descricao	custo_total
2	cultivate	proactive initiatives	Monitored content-based system engine
			3912.72

idanuncioonline	substr	descricao	custo_total
15	cultivate strategic infomediaries	Streamlined context-sensitive local area network	9512.64