



Da variação a media móvel

Transcrição

Primeiramente, podemos observar é a grande variação do gráfico. Vemos um "sobe e desce" constante. Até aqui, vimos em exemplos anteriores curvas mais simples, mas o gráfico atual é mais próximo da realidade. Encontraremos no nosso cotidiano o que chamamos de **variâncias**, estas subidas e descidas no gráfico.

Então, podemos perceber que o gráfico de *views* do blog tem mais variância que os anteriores. Isto não é um problema. Analisando o gráfico podemos dizer que o número de visitas tem uma variância de 600 acessos. Isto significa se fizermos uma campanha e tiver um aumento de acesso de 400 pessoas, o comportamento não foi excepcional.

Mas caso exista o desejo de criar uma curva suave, temos uma forma de fazer isto, usando o `AVERAGE` do Google Spreadsheets, que irá nós fornecer uma média de um determinado período. No nosso caso, iremos começar tirando a média de três dias anteriores a 8 de Julho de 2008 e três dias posteriores.

Data	Views do Blog
21-Jun-2008	1740
22-Jun-2008	1585
23-Jun-2008	2308
24-Jun-2008	1990
25-Jun-2008	1813
26-Jun-2008	1705
27-Jun-2008	2247
28-Jun-2008	2068
29-Jun-2008	2038
30-Jun-2008	1953
1-Jul-2008	2074
2-Jul-2008	2109
3-Jul-2008	2248
4-Jul-2008	2338
5-Jul-2008	1941
6-Jul-2008	2045
7-Jul-2008	2251
8-Jul-2008	2210
9-Jul-2008	2225
10-Jul-2008	2336
11-Jul-2008	2341
12-Jul-2008	2300

2192,714286 x

=AVERAGE(B16:B22)

A média da semana será 2.192,714286 . Podemos fazer o mesmo processo com as demais células, por isso chamamos de **médias móveis**, porque conseguimos gerar uma média que "escorrega" junto com os valores.

4-Jul-2008	2338	
5-Jul-2008	1941	
6-Jul-2008	2045	
7-Jul-2008	2251	
8-Jul-2008	2210	2192.714286
9-Jul-2008	2225	2244
10-Jul-2008	2336	2290.285714
11-Jul-2008	2341	2320.714286
12-Jul-2008	2300	2332.142857
13-Jul-2008	2369	2322.571429
14-Jul-2008	2464	2332.571429
15-Jul-2008	2290	2340.142857
16-Jul-2008	2158	2343.571429
17-Jul-2008	2406	2348.285714
18-Jul-2008	2394	2331.857143
19-Jul-2008	2324	2343.285714
20-Jul-2008	2402	2386.714286
21-Jul-2008	2349	2439.142857
22-Jul-2008	2370	2442.285714
23-Jul-2008	2462	2461.428571
24-Jul-2008	2773	2481.142857
25-Jul-2008	2416	2487.571429
26-Jul-2008	2458	2530.714286
27-Jul-2008	2540	2512.857143

Vamos completar a coluna da Média Móvel de 7 dias.

Data	Views do Blog	Media Móvel - 7
21-Jun-2008	1740	
22-Jun-2008	1585	
23-Jun-2008	2308	
24-Jun-2008	1990	1912.571429
25-Jun-2008	1813	1959.428571
26-Jun-2008	1705	2024.142857
27-Jun-2008	2247	1973.428571
28-Jun-2008	2068	1985.428571
29-Jun-2008	2038	2027.714286
30-Jun-2008	1953	2105.285714
1-Jul-2008	2074	2118.285714
2-Jul-2008	2109	2100.142857
3-Jul-2008	2248	2101.142857
4-Jul-2008	2338	2143.714286
5-Jul-2008	1941	2163.142857
6-Jul-2008	2045	2179.714286
7-Jul-2008	2251	2192.285714
8-Jul-2008	2210	2192.714286
9-Jul-2008	2225	2244
10-Jul-2008	2336	2290.285714
11-Jul-2008	2341	2320.714286
12-Jul-2008	2300	2332.142857
13-Jul-2008	2369	2322.571429
14-Jul-2008	2464	2332.571429
15-Jul-2008	2290	2340.142857
16-Jul-2008	2158	2343.571429
17-Jul-2008	2408	2348.285714
18-Jul-2008	2394	2331.857143

Nós estamos interessados no formato do nosso gráfico, por isso é irrelevante que os dados dos três primeiros dias não estejam preenchidos. Também criaremos o gráfico das médias móveis.

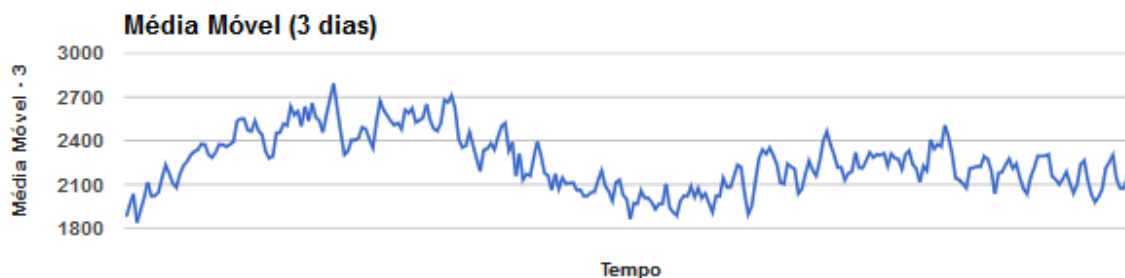


Vemos que este gráfico tem uma curva mais suave que o anterior.

Mais adiante, testaremos como seria o nosso gráfico, se fizéssemos a média móvel de 3 dias, em que consideraremos a data, o dia anterior e o posterior.

Data	Views do Blog	Media Móvel - 7	Média Móvel - 3
21-Jun-2008	1740		
22-Jun-2008	1585		1877.666667
23-Jun-2008	2308		1961
24-Jun-2008	1990	1912.571429	2037
25-Jun-2008	1813	1959.428571	1836
26-Jun-2008	1705	2024.142857	1921.666667
27-Jun-2008	2247	1973.428571	2006.666667
28-Jun-2008	2068	1985.428571	2117.666667
29-Jun-2008	2038	2027.714286	2019.666667
30-Jun-2008	1953	2105.285714	2021.666667
1-Jul-2008	2074	2118.285714	2045.333333
2-Jul-2008	2109	2100.142857	2143.666667
3-Jul-2008	2248	2101.142857	2231.666667
4-Jul-2008	2338	2143.714286	2175.666667
5-Jul-2008	1941	2163.142857	2108
6-Jul-2008	2045	2179.714286	2079
7-Jul-2008	2251	2192.285714	2168.666667
8-Jul-2008	2210	2192.714286	2228.666667
9-Jul-2008	2225	2244	2257
10-Jul-2008	2336	2290.285714	2300.666667
11-Jul-2008	2341	2320.714286	2325.666667
12-Jul-2008	2300	2332.142857	2336.666667
13-Jul-2008	2369	2322.571429	2377.666667
14-Jul-2008	2464	2332.571429	2374.333333
15-Jul-2008	2290	2340.142857	2304
16-Jul-2008	2158	2343.571429	2284.666667
17-Jul-2008	2406	2348.285714	2319.333333
18-Jul-2008	2394	2331.857143	2374.666667

Veremos que agora, a curva no gráfico ficará menos suave.



No gráfico de 7 dias, diminuimos a variância. O mesmo aconteceria se fizéssemos um gráfico de 9 dias. Porém, existe um problema em fazer uma média móvel de um período muito longo. Vamos testar a média de 21 dias.

Data	Views do Blog	Media Móvel - 7	Média Móvel - 3	Média Móvel - 21
21-Jun-2008	1740			
22-Jun-2008	1585		1877.666667	
23-Jun-2008	2308		1961	
24-Jun-2008	1990	1912.571429	2037	
25-Jun-2008	1813	1959.428571	1836	
26-Jun-2008	1705	2024.142857	1921.666667	
27-Jun-2008	2247	1973.428571	2008.666667	
28-Jun-2008	2068	1985.428571	2117.666667	
29-Jun-2008	2038	2027.714286	2019.666667	
30-Jun-2008	1953	2105.285714	2021.666667	
1-Jul-2008	2074	2118.285714	2045.333333	2074.52381
2-Jul-2008	2109	2100.142857	2143.666667	2101.190476
3-Jul-2008	2248	2101.142857	2231.666667	2138.52381
4-Jul-2008	2338	2143.714286	2175.666667	2145.952381
5-Jul-2008	1941	2163.142857	2108	2160.238095
6-Jul-2008	2045	2179.714286	2079	2178.666667
7-Jul-2008	2251	2192.285714	2188.666667	2210.047619
8-Jul-2008	2210	2192.714286	2228.666667	2217.047619
9-Jul-2008	2225	2244	2257	2229.238095
10-Jul-2008	2336	2290.285714	2300.666667	2246.571429
11-Jul-2008	2341	2320.714286	2325.666667	2265.428571
12-Jul-2008	2300	2332.142857	2336.666667	2279.52381
13-Jul-2008	2369	2322.571429	2377.666667	2296.333333
14-Jul-2008	2464	2332.571429	2374.333333	2321.333333
15-Jul-2008	2290	2340.142857	2304	2325.047619
16-Jul-2008	2158	2343.571429	2284.666667	2349.666667
17-Jul-2008	2406	2348.285714	2319.333333	2373.238095
18-Jul-2008	2394	2331.857143	2374.666667	2380.047619

Vemos que o desenho do gráfico é o mais suave de todos, principalmente quando comparado com o primeiro. No entanto, a desvantagem é que perdemos a chance de analisar detalhes e não conseguimos analisar comportamentos diferenciados em determinados períodos.

Desta forma, variações pontuais, como picos em um dia, irá desaparecer. Por outro lado, enxergamos melhor as variações de um mês. Então, a forma correta de escolher o recorte de tempo é se perguntar: "o que eu quero responder sobre um determinado período?"

Análise de Dados é um grande quebra-cabeça. Cada gráfico irá trazer uma nova informação. E com cada um deles, nós preenchemos um peça que faltava, e compreendemos melhor o todo da nossa empresa. Com isto, conseguimos tomar melhores decisões para o negócio, que no fim é o objetivo da análise de dados: auxiliar e fundamentar decisões de negócios empresariais

