

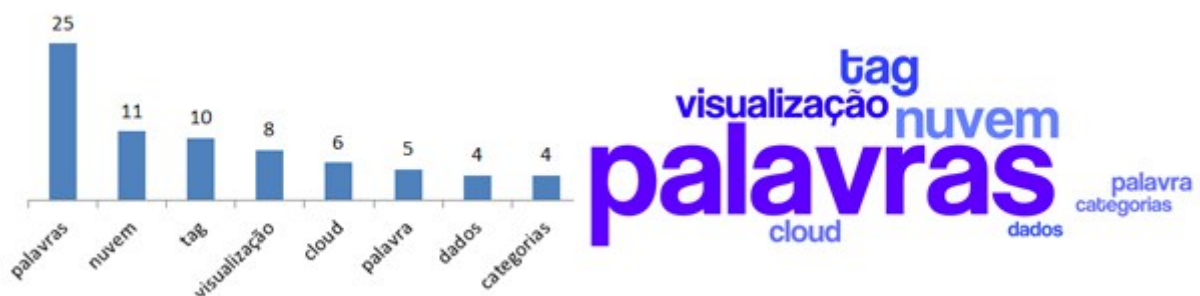


quantificada do número de ocorrência das palavras (ao invés das “categorias” ou “marcadores” no caso de tags).

### Nuvens de palavras representam um novo tipo de visualização

Em seu clássico artigo “What is Visualization”, Lev Manovich explica que visualização de dados, historicamente, sempre envolveu a redução. Ao invés de mostrarmos uma lista com cinco mil respostas a uma survey, transformamos essa lista em um gráfico como um histograma, mostrando a distribuição das respostas. Ou seja, uma redução. Esta redução envolve transformar um dado em outro formato visual: o volume de respostas a cada pergunta, por exemplo, se é traduzido/reduzido em tamanho de colunas.

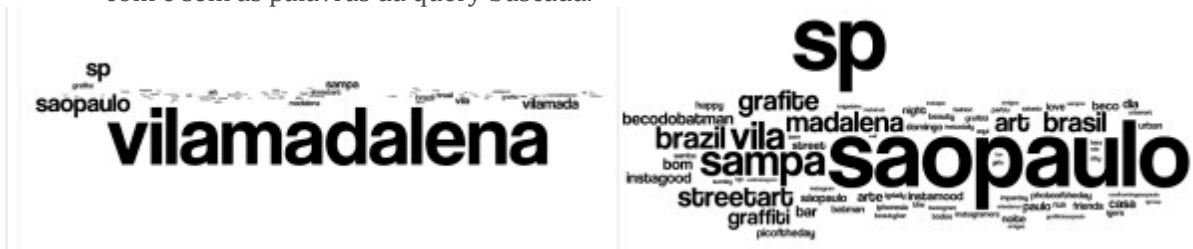
Mas quando falamos de texto e de nuvens de palavras, estamos falando de uma visualização direta. Utilizamos o mesmo objeto medido (palavras) para representar as relações entre o que estamos medindo (palavras!). Este tipo de visualização direta tem ocorrências muito mais elaboradas como o [Cinema Redux](#), mas as nuvens de palavras são de longe a mais conhecida. Para ilustrar, imagine que as diferenças entre as visualizações das frequências das palavras mais comuns no início deste post:



### Nuvens de palavras? Realmente simples?

Quando falamos de nuvens de palavras (e quando vejo alguns relatórios e infográficos por aí...), parece que a montagem de uma nuvem é algo tão simples quanto abrir a ferramenta, apertar ctrl+c e ctrl+v. Fazer uma nuvem de palavras eficaz é algo bem mais complexo do que isto e devemos ter alguns cuidados, como os listados abaixo.

- Nuvens de palavras baseadas em monitoramento de mídias sociais (ou baseados em *queries*): recomenda-se retirar ou relativizar o tamanho da palavra envolvida na query em questão. É óbvio que a palavra estará em todas as unidades de texto (menções), então seu tamanho será muito maior do que as outras palavras, prejudicando a clareza da visualização. Abaixo dois exemplos de nuvem brutas de tags, com e sem as palavras da query buscada.



Tagclouds com e sem query.

- Cores e Marcas de Clientes: assim como outros tipos de elementos visuais de um relatório, é mais eficiente utilizar as cores próprias da marca e branding da empresa ou produto do cliente ao se apresentar nuvens de palavras em torno da marca.
- Simbologia Afetiva das Cores: assim como ocorre com os gráficos de sentimento, pode ser útil mostrar nuvens de palavras segmentadas de acordo com o sentimento atribuído na análise. Neste caso, a simbologia das cores mais básica está relacionada à análise de sentimento: vermelho, verde e laranja para denotar negativo, positivo e neutro. Arbitrárias, mas são codificações consensuais (pelo menos na maioria dos países), então

não faça uma nuvem de palavras sobre “elogios de consumidores” usando tons de vermelho.

- Cores enquanto gradação: a intensidade da cor em cada palavra pode direcionar o olhar, mas ser uma dimensão que não traz dados. Cuidado com as configurações básicas das ferramentas de nuvem de palavras. Não deixe que suas limitações façam uma visualização confusa.
- Realmente dizem algo? Nem todas nuvens de palavras vão lhe dar informações relevantes. Em alguns casos, a nuvem de palavras vai reproduzir o óbvio ou, ainda, motivar interpretações errôneas. Imagine que você extraiu todas as *bios* dos seguidores de uma conta Twitter e usa uma nuvem de palavras pra medir e visualizar as mais comuns. Encontrará palavras como “apaixonado/a”, “mundo”, “deus” e outras bem comuns. Mas são termos comuns no Twitter brasileiro como um todo. Contexto e comparações são sempre bem vindos para refinar as interpretações.

### Onde fazer nuvens de palavras?

- A [Wordle](#) é o software online usado por 9 entre 10 analistas de monitoramento de mídias sociais. Basta copiar e colar um texto, editar as opções, ou ainda ir na seção avançada, e *voilà!*
- A [Wordcounter](#) conta as palavras de um texto, recurso muitas vezes necessário para se realizar nuvens de palavras customizadas e mais claras no Wordle, por exemplo.
- [Tagxedo](#) permite criar nuvens de palavras com formas abstratas e silhuetas representacionais, como o raio que caiu como uma luva para a letra de Aladdin Sane lá em cima.
- O projeto [Many Eyes](#) permite criar não só nuvens de palavras, mas também outras visualizações de texto como *phrase net* e *word tree*.
- Já o portal [Tapor](#) é mantido pela Universidade de Alberta e se dedica a manter, investigar e produzir conteúdo e ferramentas de estudo de textos.

### Nuvem de palavras formam apenas um cubinho de gelo do iceberg de text analytics e análise semântica

Análise semântica, análise de conteúdo e análise de discurso são metodologias amplamente consolidadas nas ciências sociais aplicadas e podem te ajudar a aprofundar bastante as interpretações realizadas. Dois excelentes livros, traduzidos para o português, tratam bastante do tema: *Análise de Conteúdo*, de Laurence Bardin; e *Pesquisa Qualitativa com Imagem, Texto e Som*, de Martin Bauer e George Gaskell.

E conheça a história e teoria por trás do popular Wordle no artigo [Participatory Visualization with Wordle](#), Fernanda B. Viégas, Martin Wattenberg e Jonathan Feinberg.

FONTE: <http://tarcizosilva.com.br/blog/o-que-se-esconde-por-tras-de-uma-nuvem-de-palavras/>