



FABIO CHALUB  
Universidade  
Nova de Lisboa  
chalub@fct.unl.pt

## QUAL NOME PARA O BEBÉ?

Se o leitor teve um filho em 2014, o mais provável é que tenha escolhido algum dos seguintes nomes: João, Rodrigo, Francisco, Martim ou Santiago. Para as meninas, as escolhas recaíram em Maria, Matilde, Beatriz, Leonor e Mariana. Se é evidente que uma escolha nunca se faz independentemente do meio em que se vive, o que será que o conjunto de escolhas pode mostrar sobre o nosso ambiente? Uma nova investigação mostra como a cultura americana evoluiu desde o princípio do século passado até ao momento atual, estudando apenas a evolução dos nomes dados aos recém-nascidos.

Por vezes, referimo-nos a duas diferentes culturas como "irmãs". Talvez porque partilhem traços – a língua ou a religião, por exemplo –, ou tenham um passado comum recente, ou ainda por terem uma grande interdependência económica. Mas seria possível quantificar, ou seja, determinar o quanto duas culturas são próximas ou distantes entre si?

Neste nível de generalidade, o problema é muito difícil. Mas seguindo os passos tradicionais do raciocínio analítico, devemos procurar algumas características que sejam fáceis de estudar – quantificáveis, com bases de dados completas e facilmente acessíveis. Podemos assim estudar a correlação de um determinado traço cultural tanto no espaço como no tempo.

Recentemente, uma equipa de físicos italianos resolveu usar as técnicas da estatística para saber como variam – no tempo e no espaço – os nomes que os pais dão aos seus filhos. Este é um traço cultural simples – não depende de interpretação do investigador, os registos são confiáveis e permitem uma clara quantificação. No entanto, a equipa resolveu debruçar-se sobre as bases de dados norte-americanas, provavelmente devido ao seu fácil acesso, mostrando a evolução temporal da cor-

relação entre os muitos nomes dados aos recém-nascidos nos seus diversos Estados.

Há regras explícitas e implícitas nos nomes que nós escolhemos para os filhos. Nalguns países, estes devem cingir-se a uma lista; noutros, os progenitores dispõem de ampla liberdade. De qualquer forma, sempre há certas "modas", nomes que, por uma razão ou outra, se tornam comuns e depois desaparecem. As "modas" são tão importantes que são facilmente perceptíveis quando estudamos quantitativamente a evolução dos nomes dados em longos períodos. Veja a figura 1. Mas não é o processo de escolha o objeto de estudo do artigo, mas sim como é que o conjunto de todas as escolhas feitas por uma população revela características do ambiente maior em que este grupo está inserido – é a isto que chamaremos um traço cultural.

Considerando duas diferentes regiões, a correlação entre as escolhas dos nomes indica o quanto estes dois Estados são culturalmente correlacionados. Mais especificamente, entre 1915 e 2014, 19.492 nomes femininos apareceram nos registos da Segurança Social, o que significa que em, pelo menos, um ano foram registadas, pelo menos, cinco crianças em algum dos 50 Estados ou no

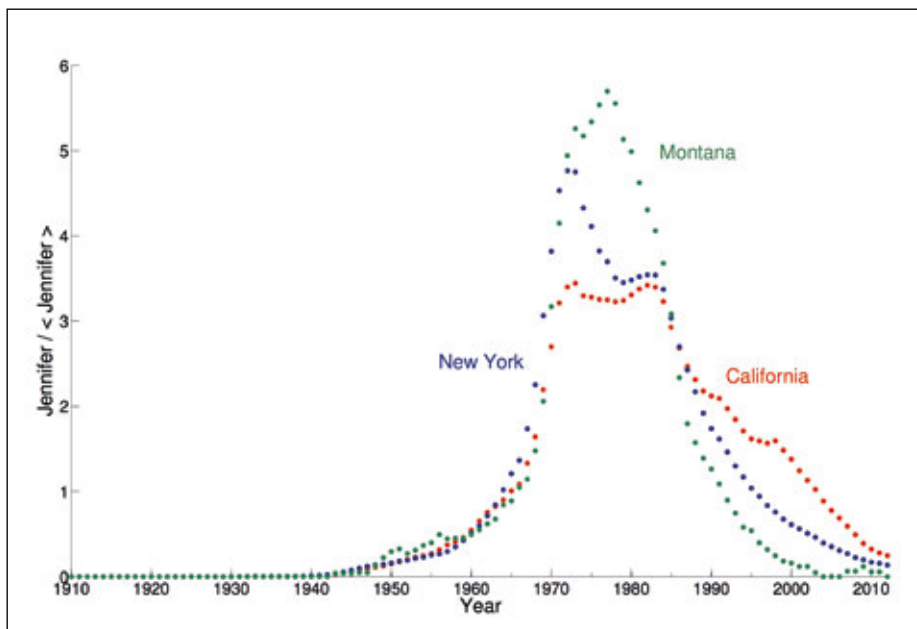


Figura 1. Nomes também têm "modas". Ascensão e queda do uso do nome *Jennifer* em três distintos Estados norte-americanos. O eixo y indica a fração de uso do referido nome em cada Estado de referência. Figura (assim como a figura 2) gentilmente cedida por Giorgio Parisi (Universidade Sapienza, Itália).

Distrito de Columbia. Assim, define-se  $f_i(n, t)$ , a fração de bebês com nome  $n$  no Estado  $i$  no ano  $t$ , e assim pode definir-se a matriz  $C$ , onde  $C_{ij}$  mede a correlação entre os diversos nomes nos estados  $i$  e  $j$  no ano  $t$ . Explicitamente,

$$C_{ij} = \frac{\sum_{n=1}^{19492} f_i(n, t) f_j(n, t)}{\sqrt{\sum_{n=1}^{19492} f_i(n, t)^2} \sqrt{\sum_{n=2}^{19492} f_j(n, t)^2}} .$$

É esta matriz que permite calcular a correlação entre dois Estados: obtém-se o vetor próprio associado ao maior valor próprio da matriz, o chamado "componente principal". Grosso modo, este indica uma forma unidimensional de extrair o máximo de informações de uma matriz bidimensional – a aproximação gerada é grosseira, sendo tanto melhor quanto maior for a diferença entre os dois primeiros valores próprios. É parte de um instrumento clássico de análise de dados, mas não entraremos em detalhes.

Usando as diversas entradas do referido vetor próprio, dito *dominante*, atribuímos aos diversos Estados norte-americanos um único número. Podemos ver quais são os mais próximos e os mais distantes a partir das diferenças entre os valores calculados. Para facilitar a interpretação, os autores do trabalho [1] usaram uma escala de cores que permite ver não apenas a similiaridade cultural entre os diferentes Estados, mas como estes se agrupam naturalmente em macrorregiões. Ver figura 2.

Sobressai nesta análise uma grande mudança cultural dentro dos Estados Unidos no século XX, que já

havia sido notada noutras instâncias. A tradicional divisão norte-sul, que marcou tanto a ex-colônia britânica, desaparece em meados do século passado, notando-se uma oposição mais acentuada entre Estados costeiros e centrais. Atualmente é mais fácil encontrar semelhanças entre os distantes Estados de Nova Iorque e Califórnia, cada um numa costa, do que entre estes e Montana, apesar de este último ser mais próximo de ambos do que os dois entre si. Já o Texas deixa paulatinamente de ser um Estado tipicamente sulista para se juntar aos da costa oeste. Veja a tabela 1 para os dados mais recentes. A formulação completa dos mesmos pode facilmente ser encontrada no *site* da Segurança Social norte-americana (<http://www.ssa.gov/oact/babynames/top5names.html>).

Muitos outros dados podem ser usados para corroborar esta mudança de clivagem: de norte-sul para dentro-fora. Veja, por exemplo, na figura 3, a comparação dos resultados eleitorais de 1916 com os de 2012 – não é tão marcante, mas o movimento registado é o mesmo.

São mudanças profundas, notáveis e detetadas com precisão quando usamos métodos quantitativos. Certamente outros estudos virão corroborar a mudança de *topologia*: de norte-sul para dentro-fora.

## REFERÊNCIA

[1] Paolo Barucca, Jacopo Rocchi, Enzo Marinari, Giorgio Parisi and Federico Ricci-Tersenghi. "Cross-correlations of American baby names." *Proc. Natl. Acad. Sci. USA*.

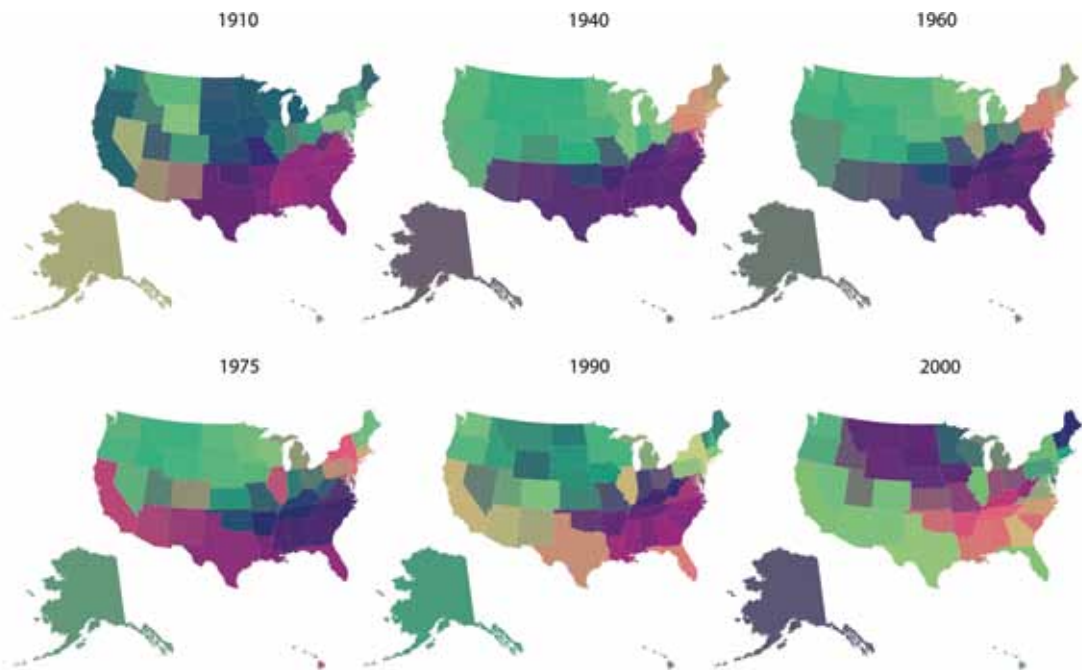


Figura 2. Estudo da correlação da escolha de nomes em cada Estado. Cores semelhantes indicam escolhas semelhantes para os recém-nascidos do sexo feminino. O estudo com nomes de meninos produziu resultados semelhantes, mas, para evitar repetições, não foi publicado.

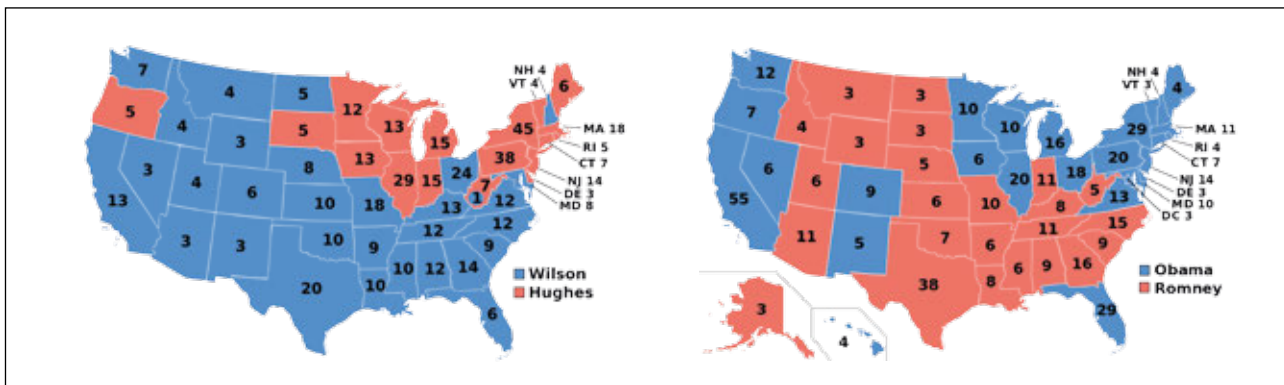


Figura 3. Divisão de votos nos Estados norte-americanos nas eleições de 1916 (esquerda), disputada entre os candidatos Woodrow Wilson (democrata) e Charles E. Hughes (republicano), e em 2012 (direita), quando Barack Obama (D) e Mitt Romney (R) disputaram os votos dos eleitores. Os números indicam a quantidade de votos dos colégios eleitorais e são irrelevantes para a presente discussão. Fonte: Wikimedia Commons.

Ordem	NY♂	MT♂	CA♂	NY♀	MT♀	CA♀
1	Jacob	William	Noah	Sophia	Emma	Sophia
2	Liam	Benjamin & Mason	Jacob	Olivia	Harper	Isabella
3	Ethan	...	Ethan	Emma	Sophia	Emma
4	Michael	Liam	Daniel	Isabella	Ava	Mia
5	Noah	Noah & Wyatt	Alexander	Mia	Emily & Isabella	Olivia

Tabela 1. Lista ordenada dos cinco nomes mais comuns dados a bebês nascidos em 2014 para meninos (♂) e meninas (♀) nos Estados norte-americanos de Nova Iorque (NY), Montana (MT) e Califórnia (CA). Compare com os nomes mais comuns de 1915 em todos os EUA: John, William, James, Robert e Joseph para os meninos e Mary, Helen, Dorothy, Margaret e Ruth para as meninas. Os empates estão marcados com '&'. Fonte: Wikipedia e sítio da Segurança Social norte-americana.