Some observations on how the 20 most popular first names
combine with the 20 most popular surnames in the United States

by Lee Hartman

John Smith, according to [Wikipedia](), is "often regarded as the archetype of a common personal name" in the English-speaking world. Without a doubt, the most popular *surname* in the United States is indeed **Smith**, and—according to the [U.S. Census Bureau]() and WhitePages.com—the most popular *first* name is **John**. For this reason it may come as a surprise that **John Smith** is not even in the top *ten* popular first-name/last-name combinations in the U.S. According to the data at [WhitePages.com](), the 13 most common full names in the U.S. and their numbers of owners, in descending order, are the following:

1. James Smith 38,313
2. Michael Smith 34,810
3. Robert Smith 34,269
4. Maria Garcia 32,092
5. David Smith 31,294
6. Maria Rodriguez 30,507
7. Mary Smith 28,692
8. Maria Hernandez 27,836
9. Maria Martinez 26,956
10. James Johnson 26,850
11. William Smith 26,074
12. Robert Johnson 25,874
13. John Smith 25,255

The conclusion I draw is that **Smith** families name their sons **John** much less frequently than the population at large. In fact, if we look at the 20 most common first names and the 20 most common last names, we find that the frequencies with which they combine into full names *in practice* differ in some interesting ways from the "ideal" frequencies that would have resulted if the combinations had been made uniformly according to the frequency of each individual name in the population at large. Some pairs of first and last name attract each other; others, like **John** and **Smith**, repel.

At WhitePages.com—which is the main source of data for this study, in spite of some reservations described in [the sidebar]()—the 20 most popular first names, in decreasing order of popularity, are **John**, **Michael**, **James**, **Robert**, **David**, **Mary**, **William**, **Richard**, **Thomas**, **Jennifer**, **Patricia**, **Joseph**, **Linda**, **Maria**, **Charles**, **Barbara**, **Mark**, **Daniel**, **Susan**, and **Elizabeth**. The 20 most frequent surnames, in decreasing order, are **Smith**, **Johnson**, **Williams**, **Brown**, **Jones**, **Miller**, **Davis**, **Wilson**, **Anderson**, **Garcia**, **Rodriguez**, **Taylor**, **Thomas**, **Moore**, **Martin**, **Martinez**, **Jackson**, **Thompson**, **White**, and **Lee**. For each of the top 20 surnames, we have the number of individuals with that surname in the U.S. population. For example, there are 2,713,582 **Smiths**, 2,102,041 **Johnsons**, etc. When we add these 20 numbers we get the total of top-20-surname holders in the U.S. (22,343,136). When we divide the number for each surname by this total, we get a percentage for each surname, relative to the top-20-surname population (*not* to the entire population of the country). Using *P* for "percentage" and *s* for "surname", let's call this percentage *P(s)*. The *P(s)* for **Smith** is 2,713,582 divided by 22,343,136, or 12.15%. So almost an eighth of the people with top-20 surnames are named **Smith**.

Likewise for each of the top 20 *first* names, we have the number of individuals with that name in the U.S. population. For **John**, the number is 4,092,015. When we add these 20 numbers we get the total of top-20-first-name holders, which is 47,474,331. And similarly when we divide the number for each first name by this total, we get a percentage for each first name, relative to the top-20-first-name population. Again using *P* for "percentage" and now *f* for "first name", let's call this percentage *P(f)*. The P*(f)* for **John** is 4,092,015 divided by 47,474,331, or 8.62%.

Since we're looking only at the individuals with one of the 400 combinations of both a top-20 first name and a top-20 last name (the "20/20" population, if you will), let's use *A* to stand for "all" of that group. This number is 3,153,792.

Let's summarize:

- *P(s)* = a *surname*'s percentage of the top-20-*surname* population.

- *P(f)* = a *first name*'s percentage of the top-20-*first-name* population.

- *A* = 20/20 population, the number of holders of both a top-20 first name and a top-20 surname.

Given the above three variables, we can calculate for each full name an "ideal" number that tells how many individuals would have that combination if the first names were distributed in the same proportions among all the surnames. Let's call this ideal number *I*. For each full name, *I* will be the product of the first-name percentage, the surname percentage, and the 20/20 population:

- $I = P(f) \quad P(s) \quad A$

For **John Smith**, then, *I* = 0.0862 0.1215 3,153,792, or 33,031 (rounded to an integer).

For each full name we also have the *real* number of name-holders, as given at the WhitePages.com website. Let's call this number *R*. For **John Smith**, as we stated above, this number is 25,255.

For each full name—seen from the viewpoint of its *ideal* number of name-holders—the real number presents a discrepancy: a percentage by which the real number exceeds or falls short of the ideal number. Let's call this proportional discrepancy *D*, and calculate it as the difference between the real and ideal numbers divided by the ideal number:

- $D = \dfrac{(R - I)}{I}$

The discrepancy for **John Smith**, for example, is (25,255 - 33,031) / 33,031, or a negative 23.5%.

We are now almost ready to calculate the real-to-ideal discrepancy for each full name and display it in Table 2—in just a moment. But first, let me explain the organization of that table. Each row corresponds to a first name, and the rows are in the order of the popularity of the names, from most to least popular. Each column corresponds to a surname, but the columns are *not* in the order of popularity. Instead, they are in the order of "whiteness", according to the ethnic identity claimed by the name-holders in the U.S. census of 2000. The four rows of Table 1—labeled Asian, Black, Hispanic, and White (in alphabetical order)—show the percentages of surname-holders who described themselves in the census as follows:

- Non-Hispanic Asian and Pacific Islander Only ("Only" = not claiming two or more races)

- Non-Hispanic Black Only

- Hispanic Origin (of any race)

- Non-Hispanic White Only

Table 1: Ethnic Identity Claimed by Surname-Holders

	Miller	Anderson	Martin	Smith	Thompson	Wilson	Moore	White	Taylor	Davis	Johnson	Brown	Jones	Thomas	Williams	Jackson	Lee	Garcia	Martinez	Rodriguez
Asian	0.4	0.5	0.7	0.4	0.4	0.5	0.4	0.4	0.4	0.4	0.4	0.4	0.4	1.6	0.4	0.3	37.8	1.4	0.6	0.6
Black	10.4	18.1	15.3	22.2	22.5	25.3	26.9	27.4	27.7	30.8	33.8	34.5	37.7	38.2	46.7	53.0	17.4	0.5	0.6	0.5
Hispanic	1.4	1.6	4.0	1.6	1.6	1.7	1.5	1.6	1.6	1.6	1.5	1.6	1.4	1.7	1.6	1.5	1.3	90.8	91.7	92.7
White	85.8	77.6	77.5	73.4	72.5	69.7	68.9	67.9	67.8	64.7	61.6	60.7	57.7	55.5	48.5	41.9	40.1	6.2	6.0	5.5

So 85.8% of **Millers** classified themselves as white, 53.0% of **Jacksons** self-identified as black, etc. If you are wondering how even 1.6% of **Smiths** could identify themselves as Hispanic, or how a **Martinez** could self-identify as anything "non-Hispanic", remember (1) that the sample includes married women who adopted their husband's surname, and (2) that some individuals—for example former New Mexico Governor Bill Richardson—may have inherited an Anglo surname from their father and a Hispanic ethnic identity from their mother (actually, according to Wikipedia, Richardson has Mexican ancestry on both sides).

Now look at Table 2, showing the discrepancies between the "ideal" and the real numbers of persons holding each of the 400 combinations of top-20 first and last names. For ease in grasping the trends, we color-code each cell with progressively deeper shades of blue for greater positive discrepancies, and likewise with red for negative discrepancies (by analogy with the use of red ink in financial accounting).

Table 2: Percentage by Which Each Full Name Exceeds or Falls Short of Its Ideal Number
in the 20/20 Population

Key to Color-Coding

Male Name	< -30%	-30% to -20%	-20% to -10%	-10% to 10%	10% to 20%	20% to 30%	> 30%
Female Name							

	Miller	Anderson	Martin	Smith	Thompson	Wilson	Moore	White	Taylor	Davis	Johnson	Brown	Jones	Thomas	Williams	Jackson	Lee	Garcia	Martinez	Rodriguez
John	8.4	6.9	13.4	-23.5	3.6	2.5	3.6	1.7	0.4	-3.8	-35.2	-34.3	-26.5	5.1	-11.3	-27.7	-8.3	-66.4	-63.6	-69.5
Michael	18.1	2.4	13.0	5.6	6.7	2.6	15.7	-1.7	0.6	2.5	-4.8	4.2	-2.4	2.9	-8.1	-16.8	-9.8	-54.7	-54.4	-63.2
James	20.3	24.0	35.9	21.1	33.6	40.9	43.0	41.0	37.4	32.8	9.6	25.4	8.9	24.8	20.7	19.8	12.2	-84.4	-83.9	-87.4
Robert	29.6	27.0	16.8	8.7	18.0	22.0	17.2	18.1	39.2	7.2	5.9	14.2	12.6	4.8	-1.3	-1.8	12.1	-57.4	-55.1	-61.0
David	27.9	28.1*	15.8	10.8	11.0	12.6	8.8	16.2	-1.5	-44.7	3.5	10.3	5.0	0.9	-7.3	-15.4	10.1	-30.5	-28.5	-31.7
Mary	21.5	18.2	26.2	14.7	18.5	18.9	22.8	22.5	16.0	14.8	11.7	13.1	12.2	18.9	13.1	8.7	-16.4	-42.4	-41.3	-54.0
William	22.5	7.0	33.3	16.6	29.3	21.5	31.9	38.2	29.2	23.1	4.3	22.3	15.4	19.8	-59.1	9.6	-11.4	-82.3	-80.2	-74.2
Richard	24.1	22.9	12.6	6.2	8.4	1.7	7.3	12.2	5.0	5.2	-2.0	3.9	-4.6	-2.3	-17.0	-21.6	-6.6	-43.2	-43.0	-47.7
Thomas	12.9	3.0	22.1	-2.6	-3.5	9.7	32.2	13.2	4.3	-5.9	-15.0	1.8	-3.5	-81.1	-11.5	-11.0	-8.9	-77.8	-75.6	-84.2
Jennifer	24.9	17.3	22.5	16.7	15.2	14.8	11.5	14.8	14.9	8.7	10.3	6.5	24.0	-1.6	-7.4	-6.3	22.5	-32.9	-32.2	-33.2
Patricia	20.9	10.7	25.3	17.1	16.7	6.7	19.0	17.2	19.4	11.8	7.1	15.2	11.2	15.9	10.6	8.8	-22.9	-16.0	-16.9	-23.3
Joseph	-8.9	-34.8	0.1	-27.6	-24.8	-26.1	-20.4	-10.6	-23.9	-26.7	-33.9	-23.9	-39.0	-7.5	-26.4	-26.0	-26.4	-45.8	-40.3	-52.8
Linda	34.0	25.8	29.6	27.5	24.8	24.4	29.3	26.1	28.6	24.4	22.9	19.3	21.1	17.4	15.3	16.1	6.5	-49.2	-50.0	-56.5
Maria	-77.6	-77.3	-52.1	-78.8	-76.4	-77.4	-79.3	-77.7	-80.3	-79.7	-78.3	-78.5	-79.4	-75.6	-79.4	-80.9	-77.5	663.9	614.1	639.8
Charles	36.3	25.7	29.3	38.4	38.1	45.7	49.0	43.4	41.2	43.6	24.8	39.9	30.5	37.0	33.3	29.0	4.6	-83.7	-84.5	-88.5
Barbara	35.5	25.3	25.1	24.0	20.6	25.8	24.6	25.2	24.7	24.2	20.3	26.7	16.2	16.6	14.8	18.5	-18.7	-66.9	-67.5	-67.9
Mark	42.1	45.9	-38.5	4.1	32.4	22.0	-23.1	-16.8	11.0	7.3	14.7	-14.0	-16.7	2.5	-7.3	-23.4	-33.6	-67.0	-63.0	-71.3
Daniel	19.9	-19.1	10.3	-5.2	-6.4	-13.9	-8.5	-5.0	21.3	-22.2	-17.8	-15.5	-25.1	-25.3	-28.7	-32.3	-1.3	35.1	33.2	24.0
Susan	32.1	28.4	15.8	3.9	5.4	3.5	3.5	-0.9	3.0	-5.8	-6.5	-4.6	-17.1	-1.5	-24.2	-31.9	3.7	-68.2	-68.0	-71.7
Elizabeth	13.3	7.4	13.5	1.7	9.1	7.9	8.4	4.0	1.7	-2.4	-3.7	-3.4	-5.5	-3.3	-13.9	-18.2	-20.2	16.3	23.4	22.5

*For the name **David Anderson** I have used the figure from the 2010 pilot study, because the new figure for that name is suspect. Between the data collected in October 2010 and those collected in August 2013 (said to be "as of February 2011"), the number of **David Andersons** fell from 9,332 to 1,737, a loss of 81%. Such a decrease is not plausible.

Spanish Names

The most salient feature of Table 2 is the block of red cells—indicating negative correlation—on the righthand side, corresponding to the Spanish surnames **Garcia**, **Martinez**, and **Rodriguez**. It should not be surprising that Spanish surnames combine with English first names at far below "ideal" frequencies.

Spanish first names are available to Spanish-surnamed families in the U.S., and for every **John Garcia** there are about 3.8 individuals named **Juan Garcia** (12,983 3,407); for every **Mary Rodriguez** there are about 8.8 individuals named **Maria Rodriguez** (30,507 3,471).

The only unambiguously Spanish given name on the list is the female name **Maria**. This name in its combinations with the Spanish surnames shows by far the strongest positive correlations on the chart, at more than 600% in each case. Conversely, **Maria** shows strongly negative correlations with the non-Spanish surnames, ranging from minus 76% to minus 80%—with one apparent exception: **Martin**, at minus 52%. This less-negative figure is probably due to the fact that **Martin** can also be a Spanish surname, although only 4% of the **Martins** in the 2000 census claimed Hispanic ethnicity.

The list of popular *male* names from the 1990 U.S. census doesn't produce an unambiguously Spanish name until you reach its 28th name (**Jose**), and the next candidate is **Juan**, ranking at 52nd. I say "unambiguously Spanish" because the names **David**, **Patricia**, **Linda**, and **Daniel** are spelled the same in English and Spanish. Note that the real-to-ideal discrepancy for **David** in the three Spanish-surnamed columns is less negative than that for the unambiguously English male names like **Robert** and **Michael**. And **Daniel** actually registers positive percentages in all three Spanish-surnamed columns. Likewise, the traditional English name **Elizabeth**—in spite of its un-Spanish-like spelling—registers a positive showing with the Spanish surnames. I suspected this might be due to a fluke in the data until I learned that **Elizabeth** has been the seventh-most-popular name given to girl babies in Mexico from 1930 to 2008, according to the Registro Nacional de Población de México (cited at BabyCenter.com). On the same list, **Patricia** ranks at #13 in Mexico, making that name's negative numbers with Spanish surnames in Table 2 somewhat puzzling.

English Names

Note that the surnames that are not Spanish are all traditional English surnames, and that the given names (other than **Maria**) are also mostly English names. As shown in Table 1, the English surnames vary in the self-claimed ethnic identity of their holders, so that, for example, the black/white ratio of **Jackson** is 53/41, while that of **Miller** is only 10/86. As noted by Wikipedia, the names **Miller** and **Anderson** in the United States include some assimilated instances of German **Müller** and Swedish **Andersson**, and this explains their greater popularity in the U.S. than in Britain; it also partially explains their high ratings of "whiteness" on the ethnic scale in Table 1. Note also that the surname **Lee** includes not only those with this traditional English surname, but also a large "Asian and Pacific Islander" group, perhaps mostly of Korean and Chinese origin. In fact, the **Lee** surname is almost equally divided between "white" and Asian persons, at 40% and 38% respectively.

Clashes of Sound

Some of the remaining strongly negative percentages in Table 2 may be explained simply as due to the avoidance of repeated similar sounds, as in the cases of **John Johnson**, **John Jones** (perhaps), **David Davis**, **William Williams**, **Mark Martin**, and **Thomas Thomas**—the latter with only 13 individuals in the 2010 pilot study, and 808 in the newer data (by the way, **Thomas** is the only name that is both a top-20 first name and a top-20 last name). But, exceptionally, **Thomas Thompson** shows no particular sign of avoidance. In fact, Wikipedia gives a list of some two dozen notable individuals, mostly sports figures, named **Thomas Thompson**. About half of these, including Wisconsin's former governor **Tommy Thompson**, are listed as **Tommy**. Mere alliteration is not necessarily a deterrent factor in combining first and last names, as shown by the popularity of **Jennifer Jones**, with the second-highest positive percentage in both the **Jennifer** row and the **Jones** column. We may speculate as to whether the American actress **Jennifer Jones** (fl. 1943-1974) had a significant influence on naming trends. The U.S. Social Security Administration website charts the meteoric rise of **Jennifer** through the ranks of female baby names, from #98 in the 1950s to #1 in the 1970s and back down to #39 in the 2000s.

Where Have All the **Josephs** Gone?

Perhaps the most puzzling red-reddish-pink streak on Table 2 is the row for **Joseph**, the only row or column with *no* blue or bluish cells. The name is relatively unpopular with the common English and Spanish surnames, and yet must be sufficiently popular with some other surnames to figure among the top 20 names in the population at large. In order to explore this mystery further, I charted the same top-20 given names in combination with popular surnames of some selected ethnicities other than English represented in the U.S. population, namely German, Irish, Italian, Jewish, Polish, Scandinavian, and Scottish. The results are shown in Table 3.

For Table 3, I've combined figures for names with alternate spellings, shown here with their more frequent spellings. So **Novak** represents both **Novak** and **Nowak**. And likewise for the "-sen"/"-son" variants of the Scandinavian names.

I've included **Morgan** among the "Irish" surnames in memory of my old classmate, **Michael Morgan**, who used to revel in his Irishness. This, and its ending with "-gan"—like **Berrigan**, **Madigan**, etc.—led me to misclassify it as Irish. But, as you can see, it patterns differently from the other Irish surnames. Seeing the relative neutrality of **Morgan** between the bright-colored columns of **Murphy** and **Kelly**, I researched it further and found that, although some Irish people are indeed named **Morgan**, the name, according to Wikipedia, is actually of Welsh origin. And Roman Catholics make up less than 5% of the population of Wales—a fact whose significance will become clear in a moment.

In Table 3 we can see that the incidence of **Joseph** is somewhat popular with Irish surnames (except **Morgan**), and very popular with Polish and Italian ones. In fact, the positive correlations of **Joseph** with these latter two nationalities are the highest on the chart. And conversely, **Joseph** shows moderately-to-strongly negative percentages with German surnames, and large negative percentages with Scottish, Scandinavian, and Jewish ones. The common factor that suggests itself among Ireland, Italy, and Poland is their predominant Catholicism, in contrast to the predominant Protestantism of Germany, Scandinavia, and Scotland—as well as England, whose surnames in Table 2 are also marked by strongly negative numbers for **Joseph**. This apparent break along religious lines is the basis for the organization of the columns in Table 3.

In observing "statistical" trends like this (full disclosure: I have no training, formal or informal, in statistics), of course it is important to recognize that none of these generalizing statements is absolute. I refer to "Catholic surnames" only as shorthand for names that originate with nationalities that are predominantly Catholic—and similarly with "Protestant", "black", etc. From here on, I will dispense with the apologetic quotation marks on the terms I use for the name groupings and trust that the reader understands the loose way in which they are applied. Given this disclaimer, a number of additional gross general observations can be made. But first, I would set aside the Spanish surnames from the following comments, because of how weakly they combine with *any* of the English given names. And I am arbitrarily classifying as "neutral" all values of less than 10% above or below zero (color-coded white).

- **John, our "most common" first name, has no positive correlation with any of the common English surnames (except Martin). Otherwise, John is like Joseph in having a strong positive incidence with Catholic surnames, and a neutral-to-negative one with Protestant and Jewish ones.**

- **Michael is strongly positive with the Catholic surnames and mildly positive with the Jewish ones. Among the English surnames (Table 2), Michael shows a mild affinity for Miller, Martin, and Moore, suggesting that alliteration plays a positive role.**

- **James is positive with the surnames from the British Isles (English, Irish, and Scottish), and negative with the "continental" surnames (Italian, Polish, German, Scandinavian, and Jewish).**

- **Robert** pairs at neutral-to-positive frequencies with all surnames, including the Jewish ones.

- **David** is neutral to mildly positive with most of the surnames in both Tables 2 and 3, but negative with Celtic surnames and strongly negative with Italian ones.

Table 3: Relative Popularities of Full Names
That Combine the 20 Most Popular Given Names

with Popular Surnames of Some Non-English Ethnicities

| | Predominantly Catholic | | | | | | | | | | Predominantly Protestant* | | | | | | | | | Jewish | | |
| | Irish | | | | Italian | | | Polish | | | German | | | Scandinavian | | | Scottish | | | | | |
	Murphy	Morgan	Kelly	Sullivan	Russo	Rossi	Esposito	Novak	Kowalski	Kaminski	Wagner	Meyer	Schmidt	Hansen	Olson	Larson	Scott	Campbell	Stewart	Cohen	Goldstein	Goldberg
John	38.8	-4.4	43.5	80.8	58.5	53.7	56.9	40.9	52.8	50.1	7.3	8.1	3.0	-15.1	-18.4	-19.2	-20.1	-4.7	-10.9	-85.2	-89.3	-84.7
Michael	77.1	-3.8	57.8	57.5	73.1	65.6	104.0	36.1	48.4	52.1	5.0	-4.0	6.7	-18.0	-21.5	-25.6	-14.0	-18.5	-15.6	14.2	11.4	23.3
James	40.3	14.0	42.0	38.3	-13.5	-27.1	-16.0	-2.9	-23.0	-23.2	-5.1	-15.1	-21.1	-15.6	-24.0	-21.6	1.3	12.2	36.6	-74.2	-77.8	-82.5
Robert	3.4	-9.2	0.2	-4.2	4.5	13.3	5.8	26.6	22.8	19.7	45.0	18.8	24.4	1.1	6.3	0.6	-3.3	2.2	0.5	-9.6	2.2	-1.6
David	-25.4	-3.3	-24.5	-25.7	-43.8	-0.2	-49.0	2.5	16.0	17.9	11.0	7.6	6.2	7.1	19.0	14.9	-9.3	-11.2	-11.3	54.9	79.5	76.3
Mary	39.0	-2.4	28.5	39.4	20.5	21.0	9.9	23.8	23.2	25.9	9.2	5.7	4.3	-10.2	-8.8	-9.8	-5.1	-1.9	0.2	-64.4	-66.3	-62.8
William	26.1	8.1	20.4	18.8	-54.4	-43.0	-63.7	-16.7	-26.5	-39.8	19.2	7.9	17.5	-27.7	-37.4	-37.1	1.5	12.3	13.1	-60.7	-54.8	-57.4
Richard	-13.9	-13.4	-22.5	-13.7	18.7	36.7	18.1	33.2	64.1	36.7	59.6	26.2	23.6	21.2	29.4	19.3	-17.0	-15.7	-17.8	11.6	14.4	12.1
Thomas	71.2	-7.5	82.6	65.7	55.7	23.7	17.1	59.8	66.9	77.4	10.0	10.6	1.1	-11.2	-7.9	-20.4	-12.4	-5.6	-15.6	-84.3	-81.6	-84.1
Jennifer	-0.9	-3.3	-2.6	-0.2	-0.5	16.4	-7.0	10.6	-14.0	9.7	21.7	19.6	14.1	2.7	5.8	11.4	-13.5	-6.3	9.6	-17.0	-22.4	-15.5
Patricia	42.2	-3.0	41.0	42.2	25.3	22.2	22.1	23.3	41.4	18.0	3.2	2.5	-3.1	-6.0	-8.2	-4.7	-9.5	1.9	-8.5	-51.5	-59.5	-53.2
Joseph	15.5	-31.0	25.6	23.5	340.2	192.3	339.3	103.6	113.0	109.5	-5.8	-17.4	-7.5	-56.5	-64.8	-61.0	-39.6	-30.7	-39.9	-21.5	-31.8	-30.5
Linda	-1.7	8.7	-4.5	0.1	2.8	3.9	23.6	8.7	-19.2	-10.5	10.8	6.1	3.0	2.7	5.5	9.9	0.4	-3.2	1.9	-16.6	-7.0	-19.9
Maria	-78.0	-80.8	-78.7	-78.7	17.5	5.1	6.2	-60.4	-57.5	-61.4	-80.6	-75.2	-77.1	-82.3	-79.5	-81.0	-80.5	-82.7	-82.3	-78.2	-73.2	-84.1
Charles	-7.0	18.6	-7.0	-13.8	-13.5	-31.1	-21.6	-18.3	-47.9	-51.6	0.4	-1.1	-8.1	-28.0	-30.8	-34.0	3.6	9.9	24.3	-52.5	-42.2	-42.8
Barbara	9.2	4.3	11.6	19.6	11.0	24.2	17.7	34.0	50.4	12.0	18.9	12.1	10.3	5.8	2.1	5.1	3.0	3.7	6.6	24.2	44.8	25.4
Mark	-8.0	-21.6	-27.2	28.3	16.4	18.9	16.3	49.1	36.1	63.3	36.1	13.7	26.3	49.5	38.8	29.1	-20.9	-11.3	-10.8	19.3	87.1	61.0
Daniel	59.3	-12.5	20.1	110.8	3.4	-0.4	-2.4	41.2	45.6	58.2	14.1	16.5	21.4	-11.6	3.1	-1.5	-37.2	-30.2	-27.8	44.0	42.4	70.4
Susan	18.2	-8.1	4.6	25.9	22.1	21.9	26.0	21.7	40.6	52.0	15.9	21.0	23.2	19.9	24.5	18.3	-18.5	-4.0	-8.9	59.6	78.5	84.5
Elizabeth	10.7	-6.2	13.9	14.4	-7.1	23.8	9.1	19.7	9.0	35.5	5.0	-1.9	1.1	-5.3	-11.7	-10.8	-16.1	-10.2	-8.1	-22.2	-34.4	-35.0

*Catholic and Protestant church memberships in Germany today are approximately equal.

- **David is neutral to mildly positive with most of the surnames in both Tables 2 and 3, but negative with Celtic surnames and strongly negative with Italian ones.**

- **Mary scores positive with all English surnames except Lee. We don't know how the name Mary Lee is distributed among people with roots in England, Africa, and Asia. Among non-English surnames (Table 3) Mary ranges from strongly positive (Irish) to moderately positive (other Catholic surnames) to strongly negative (with Jewish surnames).**

- **William with English surnames has a profile similar to that of Mary: positive for all English names except Lee (and the rhyming Williams). Among the non-English surnames, William shows a strong negative correlation with the Italian, Polish, Scandinavian, and Jewish ones.**

- **Richard , generally neutral among the English surnames, shows its strongest affinities with the Polish and Germanic surnames, and negative figures among the Celtic ones (Scottish and Irish).**

- **Thomas has its strongest positive correlations with the Polish and Irish surnames, and its strongest negatives with the Jewish ones.**

- **Jennifer has negative correlations with the Jewish and Spanish surnames, but scores neutral to positive with most other surnames.**

- **Patricia has no negative correlations except with Lee and the Spanish and Jewish surnames. The negative showing of Patricia with Spanish surnames is puzzling, given the popularity of the name in Mexico, as noted above.**

- **Joseph, as noted above, is the most polarized of these first names along religious lines, with correlations that are positive with Catholic surnames and negative with Protestant and Jewish ones.**

- **Linda and Barbara have mild to strong positive correlations with all the English surnames except Lee.**

- **Maria is mainly a Spanish given name in the U.S. Its positive correlations with Spanish surnames are the largest absolute figures on the charts (over 600%). Maria also shows positive numbers with the Italian surnames (but no more so than Mary, Patricia, or Susan). With all other surnames Maria has negative percentages—some of the largest negative figures on the charts.**

- **Charles, with the English surnames, has the strongest positive interaction of all the first names, followed by James. But Charles is generally neutral to negative with the non-English ones.**

- **Mark has a strong affinity for some of the English surnames toward the whiter end of the scale, as well as for some of the Polish and Jewish names.**

- **Daniel has strong attraction for the Irish, Polish and Jewish surnames, but mostly a negative relationship with the Scottish ones.**

- **Susan and Elizabeth are both mostly neutral with regard to the English surnames, and mildly positive with most of the non-English names. Where they differ is with the Jewish surnames: Susan strongly positive, Elizabeth strongly negative.**

- Correlations of these 20 common first names with the Jewish surnames tend to be strong, either positively or negatively. Perhaps not surprising are the high figures for the Old Testament names **David** and **Daniel**. The other strong correlations are with **Mark** and the two female names **Susan** and **Barbara**.

The situation described here is temporary. The 20 first names that I have investigated are destined to be swept away in the next few decades. Surnames generally remain stable from generation to generation, but first names vary with fashion, and this is especially true of female names in the 21st century. According to the list maintained by the **U.S. Social Security Administration**, the 20 most popular male baby names in 2012 included only six from our list—**William**, **Michael**, **Daniel**, **James**, **David**, and **Joseph**. The 2012 list is headed by **Jacob**, **Mason**, **Ethan**, and **Noah**. The list of female names has changed even more: The top female baby names in 2012 were **Sophia**, **Emma**, **Isabella**, and **Olivia**. Of the names in our study, only **Elizabeth** remains on the list of top 20 female names for babies in 2012, at rank 10. Even **Maria** is not being maintained in the naming of 21st-century babies.