What's in a Name?

Paper written by Sonja van Kerkhoff,
Data collected (figure 1) by Ning Shan.
Paper for the Media Technology masters course
Creative Research, LIACS, University of Leiden, 2006.
Lecturers: Bas Haring + Maarten Lamers.

The assignment was to work in a small group to research correlatons seeking to see if this investigation could reveal new insights – in short, engage in creative research.

Our original topic for this project was to look at correlations between surnames and occupations, using the Leiden Yellow Pages as our source of information. However since two (then) members of my group could not read Dutch and I was the only member of the group able to write the paper, we changed the question to:

Are there correlations between surnames and occupations from a random sample of 1111 Chinese residents living in the city of Beijing?

Abstract:
An outline on the background of family names is China presented, followed by some points on the demography of Beijing which also gives some general impressions about employment there. This can only be superficial because the main thrust of this course is on creative research as a topic in itself: the details of a particular society are not at issue.
This is followed by a discussion of the data collected by Ning Shan. The data is divided into 7 job categories and 11 categories of family names. 1111 names were collected, showing in general that surnames derived from the names of provinces and cities show correlations with the job categories. Finally the paper finishes with some general thoughts and problems on the topic of creative research in relation to writing a scientific paper in 3 weeks.

Keywords: creative, research, surnames, occupation, Chinese, correlations, intuition, meanings of surnames

1. Chinese Family Names

A Chinese name is written with the surname first and the given name second. For instance, Ning is Ning Shan's surname.
Today, there are over 700 different Chinese family names, but as few as twenty surnames cover a majority of Chinese people. The variety in Chinese names therefore depends greatly on given names rather than family names. There is a colloquial expression "the hundred surnames" (bǎi xìng) used in Chinese to mean "the people" or "commoners."
The expression is probably derived from a document recording 100 family names from around 1000 A.D. There is a selection of one hundred Chinese surnames here: http:// en.wikipedia.org/ wiki/List_of_common_Chinese_surnames.[1]
Another unnamed source from wikipedia speaks of there being over 3,625 family names across China, and says that 75 cent of all Chinese have one of the 45 most popular surnames. A paper by YUAN Yi-Da, ZHANG Cheng, and YANG Huan-Ming, states that of "*The 100 common surnames, less than 5% of the total number of Chinese surnames, are connected with more than 85% of the population.*"[2]
It is clear from the above that the data in Ning Shan's list of 11 of the most popular types of names is sufficient to give an initial answer to our question, although as we will see it is not a sufficiently large or rigorous sample to say what correlations are statistically significant, or what the reasons for the correlations may be.
The names he chose are categorized under the following types:

1. *Provinces*: Family names derived from the name of a province. *Chen*, for example, is the name of a province, but was initially recorded from about 1040 BC as the name of a marquis. Hundreds of years later it was adopted by all members of that province as their surname. Another surname is *Wu*, the name of a small state in China. After it was destroyed by Yue in 340 B.C, its citizens used this name as their family name.

2. *Cities*: *Pan* and *Zhao* are examples of surnames derived from names of cities.

3. *Small Towns: Yan* is one surname derived from the name of a small feudal town.

4. *Residence*: The surname *Mon*, for example is derived from the name of a residence. The descendents of the

---

1 Accessed 11 February 2006.
2. YUAN Yi-Da, ZHANG Cheng, YANG Huan-Ming, *Population Genetics of Chinese Surnames in* Inheritance Stability of Surnames and Regional Consanguinity of Population, (Institute of Genetics, Chinese Academy of Sciences, Beijing 100101, China)

official in charge of sacrifices on mount Mon, then took the name Mon as their surname.

5. Title: (as a form of memorial, like the word "The Honorable"): *Tong* and *Yuan* are surnames that refer to titles or are honorifics for the deceased. *Tong* comes from a grandson of Emperor Huang (China's first emperor, circa, 200 B.C.). *Yuan* was initially used as a word to honour someone, and later adopted by descendants of that person as their surname.

6. Birth Rank: Meng for example, was originally the title of the eldest child in a family.

7. Official position: Zhen is a tool used for making ceramic. Later the officials in charge of making ceramic goods were called by the name *Zhen*. Later this name was adopted by descendents of these officials as their surname. *Qian* means money in Chinese, so it seems likely to derive from the officer who controlled finances.

8. Skill: Tu means butcher in Chinese, from the famous butcher, Kuai Tu, from circa 600 - 200 B.C. *Bu* is derived from the word wizard, the person who made potions for protection, and such like.

9. Minority culture: The surname *Murong* for example, is also the name of a clan of XianBei nationality, an ancient Chinese minority culture. In the period of Three Kingdoms, the head of XianBei nationality, Tohuba set up a state in the north of China.

10. In Memory of: For example the surname W*u* comes from a title given to the emperor of the State of Song (circa 600-200 B.C.), given the posthumously.

11. Assigned by Royalty: An example is the surname *Ling*, which was assigned by the Emperor of the Chou Dynasty (before 600 B.C.) to a descendent of the defeated previous emperor, as a form of protection.
Ning Shan and I discussed how we could put these surnames into separate semantic categories, such as names meaning animal names or names meaning a particular skill or rank or location. Ning Shan, drawing on his knowledge of Chinese culture, decided on the above 11 categories, as those that would be most likely to produce correlations. I cannot read Chinese, and have only visited Beijing twice and Ning Shan's English is limited.

My role in this project was to see what relationships I can find, from the data I was given. So I have to assume that the distinctions made by Ning Shan for the various categories are meaningful. This does not exclude the possibility that a different categorization of names, or a different collection of surnames might produce weaker or stronger correlations.

2. Demography: Beijing in particular

China's overall population is 1.3 billion. The population of Beijing Municipality, defined as the total number of people who reside in Beijing for 6 months or more per year, was 15.38 million in 2005. 11,870 million people in Beijing Municipality had permanent residence status (Beijing *hukou*) and the remainder were on temporary residence permits.[3]

Over 95% of Beijing's residents belong to the Han Chinese majority. Smaller populations consisting of members of the Manchu, Hui, and Mongol ethnic groups also call the city home. Beijing is a major transportation hub, with dozens of railways, roads and express ways entering and leaving it in all directions. Beijing is recognized as the political, educational, and cultural centre of China. The population of Beijing's urban core (city proper) is around 7.5 million and it is from this population that Ning Shan took his data.

3. Methodology for collecting the surnames

Ning Shan searched widely for surnames in Chinese using Google and Sina (a Chinese search engine), and also looked for places where names are cited, such as the names of teachers on the websites of Beijing educational institutes, the names of police officers in reports and disputes, and newspaper articles, as well as searching in the yellow pages.

Ning Shan looked for surnames for seven job categories: 1: lawyers, 2: chefs, 3: teachers, 4: police officers or officials, 5: family owned businesses or entrepreneurs, 6: actors, and 7: (taxi) drivers. All of these jobs require some level of training or skill.
Lawyers and Teachers require 4-5 years of training and a similar level of education. The entrepreneur category are family businesses, mainly in IT or real estate. Drivers and chefs (cooks) need a similar level of training and both involve some manual labour.

---

3 http://www.bjstats.gov.cn/tjyl/tjgb/200601240023.htm
  Accessed 10 February 2006.

4. The correlations shown by the data

	Provinces	Cities	Small towns	Residence	Title	Birth rank	Official position	Skills	Minorities	In memory of	Assigned by royalty	Total
Lawyer	35	30	2	4	6	3	9	3	3	4	24	128
	Yang 9	Liu 12					Li 6				Zhang 11	
	Xu 8	Zhao 8					others 3				Wang 9	
	Chen 8	Ma 5									others 4	
	others 10	others 5										
Chef	42	29	3	5	9	5	15	4	1	2	15	108
	Chen 10	Liu 10					Li 9				Wang 8	
	Guo 6	Zhao 6					others 6				Zhang 5	
	others 26	others 13									others 2	
Teacher	65	31	5	9	13	5	20	4	2	2	34	190
	Zheng 10	Liu 11			Zhong 5		Li 10				Wang 20	
	Wu 7	Zhou 4			others 8		others 10				Zhang 8	
	Xu 7	others 15									others 6	
	others 41											
Police	26	28	3	5	7	3	18	6	2	2	25	125
	Xu 5	Liu 8	Nie 2				Li 11				Zhang 12	
	Chen 5	Zhao 6					others 7				Wang 5	
	Yang 3	Yu 6									others 8	
	others 13	others 8										
Enterpriser	56	26	5	7	9	3	15	3	5	7	23	160
	Huang 15	Zhou 8			Sun 3		Li 7				Zhang 12	
	Yang 9	Liu 8			others 6		Zhu 3				Wang 2	
	Wu 8	others 10					others 5				others 9	
	others 24											
Actor	68	40	7	11	12	7	26	9	8	5	40	233
	Chen 15	Liu 13			Hu 4	Meng 3	Li 14				Zhang 12	
	Yang 11	Zhao 10			Sun 3	others 4	Shi 3				Wang 7	
	Wu 10	Shao 5			others 5		others 9				Deng 5	
	Xu 10	others 12									others 16	
	Bao 5											
	others 27											
Driver	35	27	9	5	3	4	28	9	4	4	39	167
	Chen 9	Zhao 10	Pan 3				Li 15				Zhang 17	
	Shen 7	Fan 6	others 6				others 13				Wang 12	
	Yang 5	others 11									others 10	
	Xu 5											
	others 9											

Figure 1, surnames of Beijing residents collected by Ning Shan, January 2006.

To say anything about this data, I have to assume that the numbers in the category "others" means that no individual surname of that type, in that same job category, scored more than the lowest amount of specified names in the same category.

I can only make general qualitative observations about the above data because I do not know how representative the example is and I do not have the statistical skills to say whether apparent patterns are statistically significant. The number of people for each job category is not consistent, and not representative of the size of that occupation in the Beijing population, but this does not matter if we look only down the columns, at the order of frequency of the names.

We may suspect, for instance, that teachers are over-represented because there was a good source of teachers' names, but it is still remarkable, looking down the first column that *Zheng* and *Wu* are the most common names for teachers, but are lower in the list or missing entirely in the other occupational categories. As already mentioned this only gives us a qualitative indication of whether true correlations are to be found.

Here are some observations from the data. In the first column (names that derive from provincial names) it is noticeable that the name *Guo* only appears in the chefs category and that the meaning of the surname indicates a province. Perhaps this particular province has a strong

cullinary culture?

The surname *Zheng* is popular amongst teachers: that is, a name that derives from a particular area of China now relates to a socio-economic class in Beijing.

Similarly the surname *Huang* referred originally to particular province, and now is the most common among entrepreneurs, but this does not appear as a major surname for any other occupational category. This might be because current immigrants from that area are attracted to the entrepreneurial sector, or it might be because this name is over-represented in a longer-standing socio-economic class in Beijing, because of some pattern of migration in the past.

The surname *Chen* appears as a common name for every category except for teachers and entrepreneurs.

Xu also appears to a common name for all the categories, except for chefs and entrepreneurs.

Shen appears to be a rather common name for drivers, but this name is not mentioned in any of the other categories. Perhaps people originating from a particular area of the country are now found in a lower skills category in Beijing?

On the other hand, in the column of city-related surnames, the name *Liu* appears in all the categories except for drivers.

The city name *Zhou*, like Huang in the provinces

column, only appears in the category of entrepreneurs. Perhaps entrepreneurs in Beijing are migrating disproportionately from a particular province or city, or that the entrepreneurial class in Beijing once originated from that area.

In the column for names that relate to titles, *Zhong* is the only surname to stand out and it is in the category for teachers. There is little data here, but it is possible that it is meaningful that *Zhong* appears for teachers and not for lawyers. Perhaps this is an indication that the teaching profession has more continuity with traditional society, as an honoured occupation handed from father to son, whereas the occupation of lawyer has been newly created with the adoption of modern systems of law.

The names derived from official positions only show that *Li* as a consistently common name. Likewise in the column for names assigned by royalty: the most common names are most common in all job categories. This is the expected result if the commonness of a surname in the population is the only factor determining its commonness in a particular job category.

The data shows that the names that relate to province and city names do seem to correlate with some job categories. However a larger sample might produce differing results for the columns that have such low numbers that no pattern is visible. We can at least say that statistic research on occupation-name correlations would yield results of interest, but we cannot say that only the provincial and city names, or only these job categories, would be of interest.

5. Problems with writing a scientific paper on a creative research topic in 3 weeks

I understood the main purpose of this course as being to experience various approaches to creative research so that we are better equipped to develop more creative or interesting topics for our own projects. However the fact that we had to write a paper along scientific lines about our creative research question severely restricted the possibilities to events that could be documented within 3 weeks. I am not convinced that it is very creative to spend a lot of time recording data (even though I realise this is import for research in general), because we are not skilled enough in statistics to be able to draw strong conclusions and we do not have the resources (for 2 individuals) to collect enough data to show statistical significance. We have also not had the time to take the really interesting step, which is to look for the present or historical patterns of name-forming, migration, social classes and family occupational traditions that lie behind the correlations.

I will finish this paper with some suggestions on the topic of creative research and some approaches that could have been taken to stretch the imagination, and with an indication of some things I did learn from the assignment as it was set.

If the assignment had been to write about methodologies for one or more questions, we could have been more adventurous in the questions we could ask. For example, we could have asked if elderly women with traditional chin tattoos (a middle eastern custom) wore more or less make-up, or if the type or amount of make-up worn by women had a relation to how they used the internet, or whether the amount of plants in a home is affected by the number of computers, and so on. It would be too much leg work to actually collect the data on such questions, but I would have found it more interesting to say, choose 4 or 5 questions and then write out a plan of action and resources required to actually research the topic. For instance: is there a correlation between human height and the fear of heights: the research approach would have to include not only how to find participants, and how to measure them, but also creative ways of measuring their fear. Apart from asking for subjective evaluations, one could test what greater effort people are willing to go to (for instance, by taking a longer route) in order to avoid a height-challenging situation. This question could include approaches for making a novel object, so people would be inclined to climb 'height challenging' stairs.

The hour-long interview with the astronomer, Vincent Icke, conducted by Joris Slob, Zimo Xu and myself on 16 january 2006 was insightful on the topic of creative research. To start with, his office was full of contrasts: a bicycle, pictures of stars lying on the floor, and a book "The Science of Harry Potter – How magic really works" by Roger Highfield[4] on his desk. He commented that it was a great misunderstanding to think that exact science is methodical. It is the history of the development of science that is what is methodical. The *a posteriori* (methodical) approach being the average approach taken in the sciences, but that the interesting stuff is what is on the periphery of the known – in the terra incognita. This is not a denial of the importance of skill or methodology, but of the importance of intuition. He gave the example of when we are with close friends: a situation where we would tend use our intuition to detect a subtly changed state in a person. This 'attentiveness' (opmerkzaamheid) is how, he said we should treat the world around us. We need to be so intimate with our area of interest or research that we are able to use our intuition to 'feel' the subject. So when you notice something, the noticing is like a small act of creation.

---

4 Roger Highfield, science editor of *The Daily Telegraph,* London, carried out research at Oxford University and the Institute Lane Langevin, Grenoble, where he became the first to bounce a neutron off a soap bubble. He has co-authored three other books: *Frontiers of Complexity*, *The Private Lives of Albert Einstein* and *The Arrow of Time.* With the BBC, he has organized several mass experiments, dubbed Megalab, which have attracted the participation of hundreds of thousands of people. He has also contributed to *Esquire* magazine.