



Assignment 1: Blog Posts 1-7

REFLECTIONS ON QUANTITATIVE RESEARCH

Lan Zhang2005977 | Applied Quantitative Business Research | 25th November

words count: 2921

contents

Blog 1 Alice In Wonderland.....	3
Blog 2 Get Your “Quantitative Brain”!(word count 462).....	6
Blog 3 The 1936 Literary Digest Poll(word count 491).....	9
Blog 4 The Ethical Dilemma of Netflix(word count: 395).....	13
Blog5 The More Revenue Generated By Arcades, The More Computer Doctorates Awarded In The US?.....	16
Blog6 \$1 Billion Gambling.....	20
Blog7 One Giant Leap For A Man.....	23

List of figures

Figure 1: Alice in Wonderland (Source: Disney/Sportsphoto Ltd/Allstar).....	3
Figure 2 The Questionnaire of My optional Course, The Ancient Chinese, in 2019.....	7
Figure 3 The Population of The US from 1936 to 2100.....	11
Figure 4 Tyler Vigen: Spurious Correlations (Resource: Tyler Vigen.com).....	16
Figure 5 Tyler Vigen: Spurious Correlations (Resource: Tyler Vigen.com).....	17
Figure 6 Tyler Vigen: Spurious Correlations (Resource: Tyler Vigen.com).....	18
Figure 7 Tyler Vigen: Spurious Correlations (Resource: Tyler Vigen.com).....	18
Figure 8: Attendance at Disney Theme Park Locations in 2019 (Resource: Statista).....	22

Blog 1 Alice in Wonderland (words count 366)



Figure 1: Alice in Wonderland (Source: Disney/Sportsphoto Ltd/Allstar)¹

This is my first time to blog, for my DBA programme of study.

This is also my first time to have an opportunity to acquire an insight into the quantitative research, by the incoming module, the applied quantitative research.

A fairy tale Alice's adventure in wonderland suddenly occur to me. Alice dropped into the rabbit hole and gulped down a bottle of water. Magically She was scaled down and could see a whole new world which had been invisible to her before.

¹ Guardian available at: <https://www.theguardian.com/film/2010/feb/23/alice-in-wonderland-disney-cinemas> (Accessed 10 November 2020).

I am grateful that the long-coveted adventure of Alice will possibly fall on me in the next several months. Would this module be like the bottle of magic drink to open my eyes and lead me to an adventurous, exciting and unknown world of research?

In my past research, qualitative method is the one I used most, which try to examine people's experiences by using a wide range of methods like the interviews, observations, discussions, etc. It is a flexible and interpretive approach to do the research.² But The Ipsos 2016 Perils of Perception survey³ which focus on how wrong the public is about the current news let me begin to doubt if the perception is truthful to the reality.

Maybe it's the time to drink the magic water to see the world in another perspective, a more accurate way -- quantitative method. The quantitative research is so strange to me full of bewildering figures and equations which can successively scare most of the liberal art students away. But Savage told us "Statistics is the science of making decisions under uncertainty."⁴ The foundations of statistics. Courier Corporation. The use of data only cannot make you a quantitative researcher. It's only when you are going beyond the data and you can explain or predict something unknown that you are going to get access to the statistics.

I'd like to think statistics as the science of changing my mind, as well as the way I see my research. how to wear the two pairs of glasses at the same time and integrate the new

² Hennink, M, Hutter. & Bailey, A. (2020) *Qualitative Research Methods*. SAGE Publications Limited.

³ Ipsos (2016) *perceptions are not reality: what the world gets wrong*. Available at: <https://www.ipsos.com/ipsos-mori/en-uk/perceptions-are-not-reality-what-world-gets-wrong> (Accessed: 15 November 2020).

⁴ Savage, L. J. (1972) *The Foundations of Statistics*. Courier Corporation.

quantitative approach into my previous study is my biggest challenge. I am looking forward to starting my new journey in the wonderland.

Blog 2 Get Your “Quantitative Brain”!(words count 462)



When I got to the topic to discuss the biggest challenge of using quantitative research in my organization, I was saddened and at a loss. As a lecturer working in a university targeting the ethnical minorities in China, I have a very slim chance to utilize quantitative research in my work. Unfortunately, most of the teachers in my universities seldom adopt the quantitative method in our teaching.

As regards to me, the biggest challenge is that I don't have a “quantitative brain”. I am accustomed to making a judgement in teaching by my observation, real-time interviews and the assessments. Quantitative research is the last way I will resort to.

Neither did I have “quantitative eyes”. Sometimes even when I was so close to the

quantitative method, some valuable data I collected will be overlooked and some stupid mistakes would be made. After learning the quantitative method introduced in this module, I got to realize that the data like their genders, grades, nationalities, preferences to the individual sessions, the skills realized are all valuable data to be analyzed but were ignored previously.

For example, I was in charge of a one-semester optional course, the ancient Chinese last year and most of the students were ethnical minorities like Tibetan and Uyghurs. To improve the quality of the class and prompt ethnical students' understanding of ancient Chinese, I designed a rough questionnaire to get feedback from the students attending the class.

Questionnaire

1 Are you satisfied with the optional course, the ancient Chinese? If not, please list the reasons.

2 which session in the module do you like best and why?

3 Will you recommend this module to your classmates? If not, please list the reasons.

4 Are you satisfied with the teacher's performance? If not, please list the reasons.

5 What is your advice for the teacher of this module?

6 Are you satisfied with the amount of the work given by the teacher, too much, medium, or too little?

Figure 2 The Questionnaire of My Optional Course, The Ancient Chinese, in 2019.

I have to confess that the questionnaire is a total failure especially after knowing how to design the questionnaire. What I have collected is the descriptive comment instead of

the data used for analysis. The validity and reliability of the questionnaire can only be checked by the “quantitative eyes”.

Firstly, the Likert scale should have been used to promote the efficiency of the data.⁵

Secondly, The questions were far from being clear and simplified. What’s worse, most of the questions are so generalized that a lot of useful variables got lost.

Thirdly, A huge amount of precious variables was lost in the design, like gender, nationality, age, the skills realized, etc. all theses data can tell me much more than the test performance. Only with accurate feedback could I adjust my module effectively. The effective feedback is one of the seven rules for informative teaching concluded by David Nicol and Debra MacFarlane-Dick.⁶ By utilizing the data analysis, teachers can study their teaching in more elaborate ways.⁷

“Education is a tree shaking a tree, a cloud to promote a cloud, a soul awakens another soul”, the famous line from Karl Theodor Jaspers might take on a new meaning in this computerized era. the quantitative method will “shake” and “promote” the whole education.

⁵ Nemoto, T & Beglar, D. (2014) *Likert-scale Questionnaires*. In *JALT 2013 Conference Proceedings*. pp. 1-8.

⁶ Nicol, D. J. & Macfarlane-Dick, D. (2006) ‘Formative assessment and self-regulated learning: A model and seven principles of good feedback practice.’ *Studies in higher education*, 31(2), pp. 199-218.

⁷ 7 Castro, F, Vellido, A, Nebot, A, & Mugica, F. (2007) *Applying data mining techniques to e-learning problems*. In *Evolution of Teaching and Learning Paradigms in Intelligent Environment* pp. 183-221.

Blog 3 The 1936 Literary Digest Poll (words count 491)



The Literary Digest was one of the most respected magazines in the 1930s and had accurately predicted the winners of the last four presidential elections. But its reputation was ruined by the fateful polling of 1936 presidential election. For that election, the Literary Digest predicted that Landon would defeat Roosevelt by 57% of the vote to 43%. The actual result was on the contrary. This tragic polling accident, leading to the falling of the Literary Digest, tells us why sampling could be regarded as the foundation of robust quantitative research.

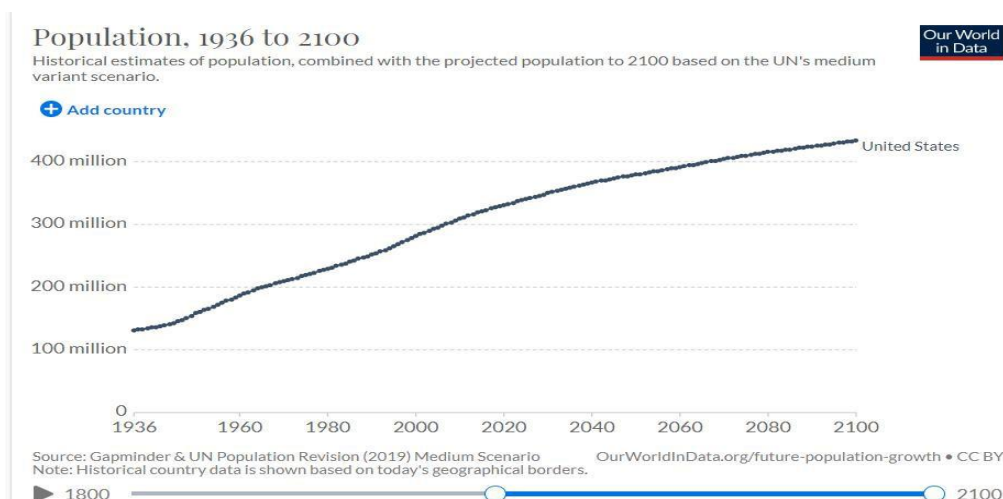
The population is the object of study, which are usually too large for researchers to have access to. So we collect data from a sub-set, named Sample.⁸ “we can use a sample of data to estimate what’s happening in a larger population to which we don’t have

⁸ Field, A. (2013) *Discovering statistics using IBM SPSS statistics*. sage.

access.”⁹ Therefore, the sample that fair-represents the population lays the foundation for convincing quantitative research.

All quantitative analyses require trustworthy data to make logical decisions.¹⁰ To be exact, the appropriate sample size, sample procedure and high rate of participation will lay a robust foundation for the reliability and validity of the result and vice versa.

Andy Field believes that there isn't any consensual benchmark for sample size. “The simplest rule of thumb is that the bigger the sample size, the better.”¹¹ Sociologist Gay notes that in the descriptive research, the sample must account for 10% of the population.¹² Is the sample size of 1936 polling too small? The 1936 Literary Digest poll was one of the most extensive and costly polls ever conducted, with a sample size of around 2.4 million people, which account for over 20 per cents of the population of US in 1936.



⁹Field, A. (2013) *Discovering statistics using IBM SPSS statistics*. sage.pp.147.

¹⁰ Waters, D. & Waters, C. D. J. (2008) *Quantitative methods for business*. Pearson Education.

¹¹ Field, A. (2013) *Discovering statistics using IBM SPSS statistics*. sage.pp.618.

¹² Gay, L. R. Mills, G. E. & Airasian, P. W. (1992). *Educational research: Competencies for analysis and application*.

The fact that the historically big sample size can't prevent the failure of the prediction tells us the sampling procedure is the top of the priorities. Because sampling is innately inaccurate, the biggest challenge is to identify a sample giving a great representation of the population.

The Literary Digest choose to send letters to voters which were taken from the telephone directories, club membership lists. In 1936 when a depression stroke the country, people with telephone and club membership mostly belonged to middle-and-upper class. polling to the public is disastrously narrowed to the middle-and -upper class. The misleading result comes from the difference between the population and sampling. Apparently, convenience sampling is not a wise choice for this polling, which is the leading reason for the failure. How to run the sampling procedure play a fundamental role in this polling.

Last but not the least, only 2.4 million in 10 million responded to the mail survey. The participation rate is too low to be ignored. Baruch suggested in 1999 sampling need a high response rate in the study to have dependable, valid, and reliable results.¹⁴

The missing statistics analysis should also be carried out to avoid sampling errors.

¹³ Our world in data (2020) the population from 1936 to 2100. Available at: <https://ourworldindata.org/search?q=population++US> (Accessed: 15 November 2020)

¹⁴ Baruch, Y. (1999) Response rate in academic studies-A comparative analysis. *Human relations*, 52(4), pp. 421-438.

Blog 4 The Ethical Dilemma of Netflix (words count: 395)

The image shows the Netflix logo, which consists of the word "NETFLIX" in a bold, red, sans-serif font. The letters are slightly tilted and have a 3D effect, appearing to float above a black background.

It is cruel and uncomfortable to find that the Netflix users, maybe most of us, are the dolls being manipulated by the platform. Unfortunately, it's likely the case.

Netflix, claiming itself “world’s leading internet TV service”, lay its business model on the foundation of subscription, instead of the traditional profit model, advertising.¹⁵ It collects data from its subscribers, without any negotiation with advertising agencies, to boost customers retention, get satisfactory click ratings and long-term honesty.¹⁶

¹⁵ Gavilán, J, C. et al (2013) *Actores y nuevos modelos de negocio*. Universitat Oberta de Catalunya.

¹⁶ Gomez-Urbe, C. A.& Hunt, N. (2015). The Netflix Recommend System: Algorithms, Business Value, And Innovation. *ACM Transactions on Management Information Systems (TMIS)*, 6(4), pp. 1-19.

On the Netflix tech blog, machine learning is proudly introduced as “historically impacts many exciting areas throughout the company, which powers the recommendation algorithms.”¹⁷

Moral problems appear when the company’s interest in conflict with users’ interest. Xiao et al’s study in 2015 concluded that the recommendation systems, now pervasive on websites, can be constructed to produce a platform- friendly result rather than a customer-friendly one.¹⁸ A controversial hot show in Netflix, “13 reasons why,” aimed at the teen audience, which glorify the misconduct in school, was commented by Reed Hastings, Netflix co-founder and chairman. “it has been enormously popular and successful,” she said, “It is controversial. But nobody has to watch it.”¹⁹ In other words, no matter the show is beneficial to the teens or not, as long as it can bring profits, Netflix will say yes to it.

In Netflix, how to use the data of the subscribers is inclined to be out of control. Sad to say, the collected data is frequently put to use beyond its supposed purpose and many people would regard such behaviour as an invasion to their privacy.²⁰ Netflix recommended the content only when the algorithmic rules think it will guarantee the click rates and with autoplaying function, users even lose the right to refuse what

¹⁷ Netflix research (2020) machine learning. Available at: <https://research.netflix.com/research-area/machine-learning> (accessed: 24 November 2020).

¹⁸ Xiao, B. & Benbasat, I. (2015) Designing warning messages for detecting biased online product recommendations: An empirical investigation. *Information Systems Research*, 26(4), pp. 793-811.

¹⁹ Deadline (2018) Netflix CEO Reed Hastings defends ‘13 reasons why’. Available at: <https://deadline.com/2018/06/parents-television-council-raises-concerns-13-reasons-why-netflix-shareholders-meeting-1202404918/> (Accessed: 24 November 2020).

²⁰ Pascalev, M. (2017) Privacy exchanges: restoring consent in privacy self-management. *Ethics and Information Technology*, 19(1), pp. 39-48.

Netflix provide them. Even worse, the platform targeted customers with advertisements by collecting data.

Who can protect the interest of the users? The data science and design team should take individuals' right to choose whether and which of their data and facts are shared into consideration. After all, it has been proved that with the data of zip code, birth date and gender, people can be re-identified, with eighty-seven accuracy.²¹

This company has a superior power to control what we will consume. If it is not adequately supporting you to cancel the subscription, it should at least sound the alarm to reflect on the use of big data.

²¹ Gumbus, A. & Grodzinsky, F. (2016) Era of big data: Danger of discrimination. *ACM SIGCAS Computers and Society*, 45(3), pp. 118-125.

Blog5 The More Revenue Generated By Arcades, The More Computer Doctorates Awarded In The US?(word count 418)

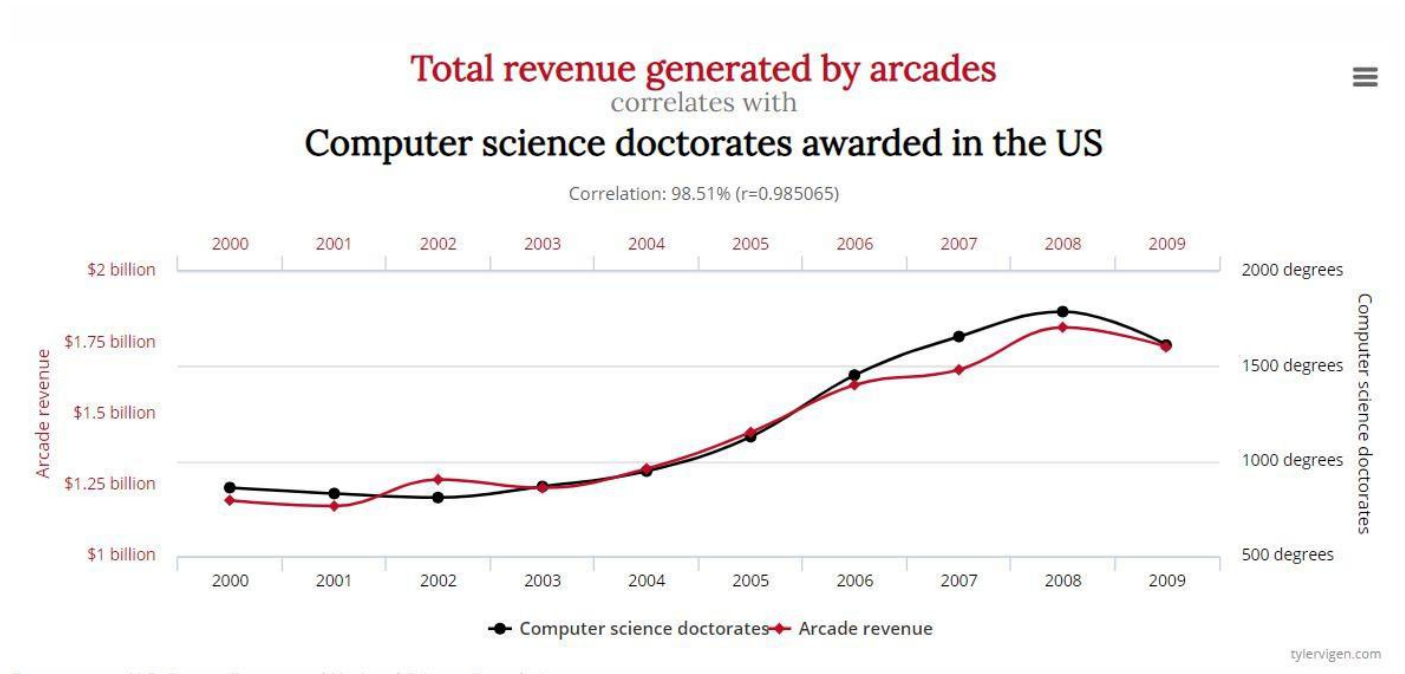


Figure 4 Tyler Vigen: Spurious Correlations (Resource: TylerVigen.com)²²

A book *spurious correlation* written by Tyler Vigen invoke the thinking about the reliance of the significance test. The figure above taken from his website illustrate our point to a degree, the significance test is the detective, not the judge.

According to Anderson who was *the editor in chief of* Wired, “Correlation is enough.”²³

²² Tyler Vigen. (2020) spurious correlation. Available at: <http://www.tylervigen.com/spurious-correlations> (Accessed: 25 November 2020).

²³ Chris Anderson. (2000) 'The End of Theory: The Data Deluge Makes The Scientific Method Obsolete'. Wired[blog] Available at: <https://www.wired.com/2008/06/pb-theory/> (Accessed: 25 November 2020).

.05, a figure was given birth by Ronald Fisher 90 years ago, in his book *statistical methods for researchers*, has become a maxim observed by most researchers. However, in real research, too many statistically significant findings, under .05, are proved false. This magic figure could turn any stones into gold. A high school teacher from Greece published his paper “Why Most Published Research Findings Are False” fire a battle on it. “It can be proven,” he writes, “that most claimed research findings are false.”²⁴

Tyler Vigen posted a wide variety of figures on his website to show the absurdity of believing the significance test only. Like...

Cheese consumption affect the number of people killed by their bed sheets?

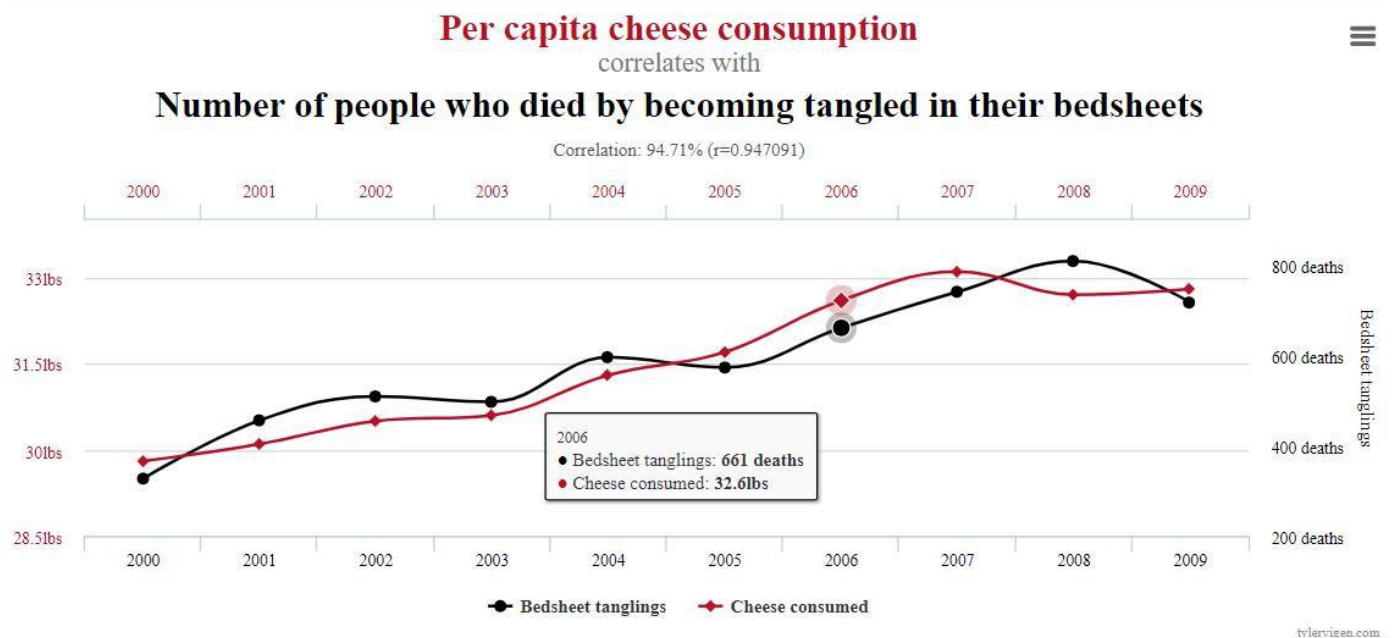


Figure 5 Tyler Vigen: spurious correlations (resource: Tyler Vigen.com)²⁵

²⁴ Ioannidis, J, P. (2005) Why most published research findings are false. *PLOS medicine*, 2(8), pp. 124.

²⁵ Tyler Vigen. (2020) spurious correlation. Available at: <http://www.tylervigen.com/spurious-correlations> (Accessed: 25 November 2020).

Or Nicolas Cage related to the number of drowning people?

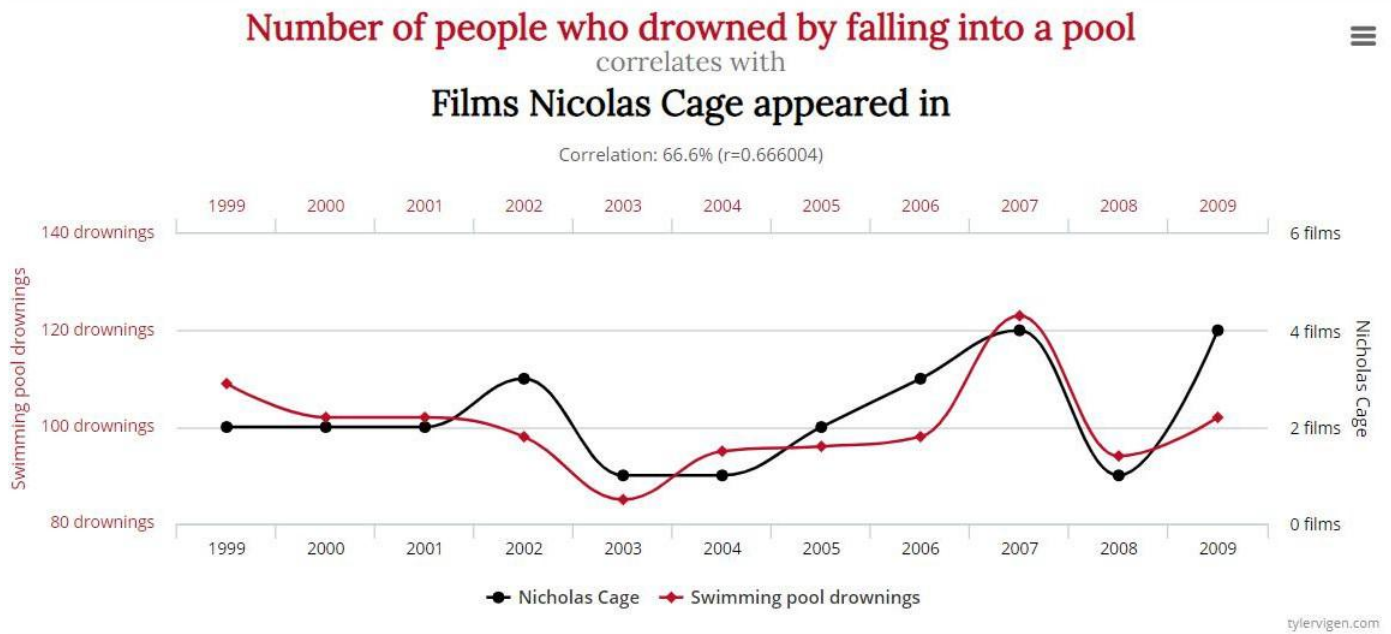


Figure 6 Tyler Vigen: Spurious Correlations (Resource: TylerVigen.com)²⁶

Most sadly, margarine consumption was involved in the divorce rate.

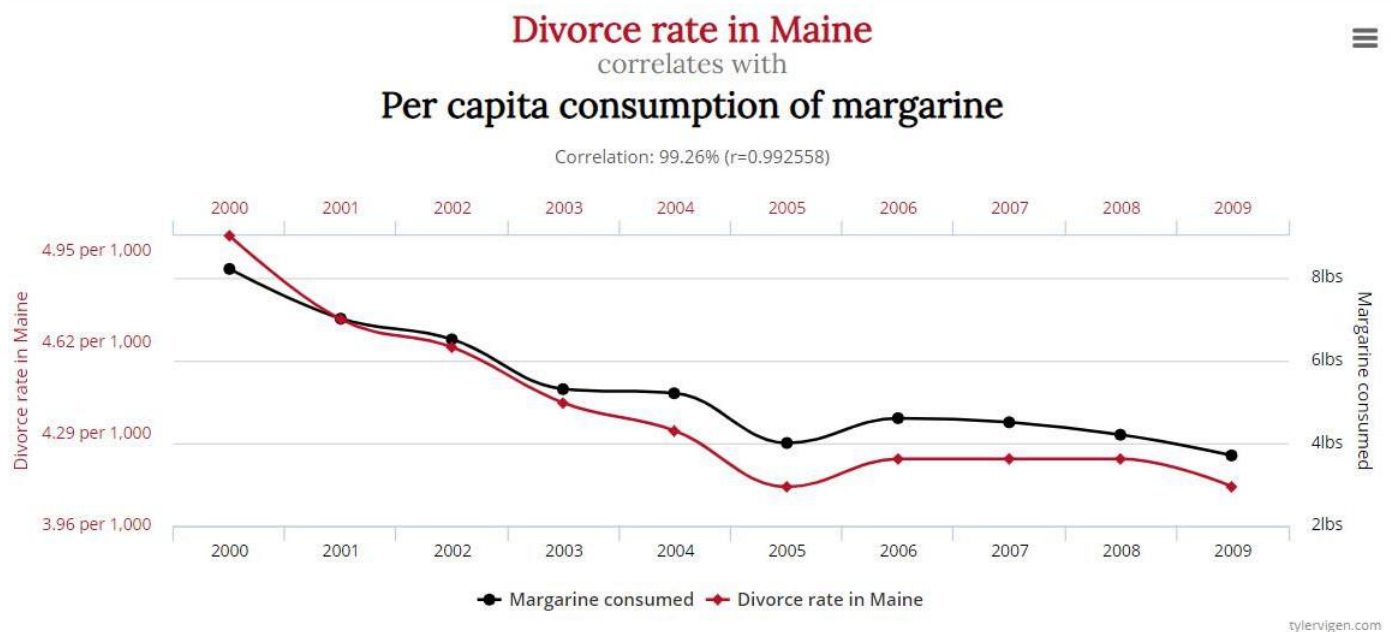


Figure 7 Tyler Vigen: Spurious Correlations (Resource: Tyler Vigen.com)²⁷

²⁶ Tyler Vigen. (2020) Spurious Correlation. Available at: <http://www.tylervigen.com/spurious-correlations> (Accessed: 25 November 2020).

²⁷ Tyler Vigen. (2020) Spurious Correlation. Available at: <http://www.tylervigen.com/spurious-correlations> (Accessed: 25 November 2020).

Cristian S. Claudel has proved the big database cover arbitrary correlations, only out of the size, not the essence, of data. He warned the danger of allowing the scientific methods being substituted by the pursuit of correlations.²⁸

What is worse, the p-value can be manipulated by people. Uri Simonsohn, a researcher who works in Pennsylvania University, named it “p-hacking”.²⁹ Ellenberg vividly recreated the setting. “Did we control for age? Did we control for the weather outside? Did we control for age and the weather outside? Give yourself license to tweak and shade the statistical tests you carry out on your results, and you can often get that .06 down to a .04”.³⁰

It's time for the researchers to go back to the definition of the P-value. The conventional benchmark .05 which means 1 out of 20. it shows us even the result is significant, the null hypothesis is possibly proved true in the odds of 1/20, which could be realized by repeated experiments. Even Fisher himself knew that a significant result doesn't equal the truth. He emphasized, “A scientific fact should be regarded as experimentally established only if a properly designed experiment rarely fails to give this level of significance”³¹

²⁸ Calude, C. S. & Longo, G. (2017) The Deluge of Spurious Correlations in Big Data. *Foundations of Science*, 22(3), pp.595-612.

²⁹ Simmons, J. P. Nelson, L. D. & Simonsohn, U. (2013) Life After P-hacking. Meeting Society of Personality Social Psychology, New Orleans, LA.

³⁰ .Ellenberg, J. (2015) *How Not To Be Wrong: The Power of Mathematical Thinking*. Penguin.

³¹ Fisher, I. (1926) A statistical relation between unemployment and price changes. *Int'l Lab. Rev.* 13, pp. 785.

Blog6 \$1 Billion Gambling (words count: 399)



Disneyland has invested 1 billion on its machine learning and big data process. The MagicBand introduced several years ago is the typical production out of the investment, which once again illuminates the magic power of big data analysis.³²

From the official website of Disneyworld, the MagicBand is described as a wrist band with a compound function of the tickets, hotel keys and reservation.³³ It can get a huge

³² Disney's 1 Billion Bet on Wrist Band. *Wired* [blog] Available at: <https://www.wired.com/2015/03/disney-magicband/> (Accessed: 25 November 2020).

³³ Disneyland available at: <https://www.disneyworld.co.uk/plan/my-disney-experience/bands-cards/> (Accessed: 24 November 2020).

amount of information about you and your family, including when and where you are, what you have consumed in the restaurant, how much money you have spent, what are the money used for, etc. undoubtedly, with all these data, Disneyland can afford the customer with a better service, like more convenient transportation, dinners without lineups.³⁴

However, I still seriously doubted all these eating and drinking information worth 1 billion, until a new word, behavioural analytics, popped into my eyes.

Behavioural analytics is a kind of data analytics concentrating on providing ideas into the actions of people. it is used to optimize to realize specific business outcomes.³⁵ That is to say, Disneyland has dug out the additional value of these data collected via MagicBand. They make use of it to generalize the behavioural pattern to discover the consumption habit of the guests as well as their preference and taste. Based on the information, Disneyland can save the labour cost by allocating the character to the appropriate place and cut the expenses greatly by accurate estimation of the stock. What's more, they can even target their guest with advertisement and other pushing content for profit, which will revolutionarily change the way the customers interact with them.

³⁴ Christopher Palmeri. (2014) Disney Bets 1 Billion to Track Disney Visitors. Bloomberg[blog] Available at: <https://www.bloomberg.com/news/articles/2014-03-07/disney-bets-1-billion-on-technology-to-track-theme-park-visitors> (Accessed: 24 November 2020).

³⁵ Wikipedia (2020) Behavioural Analytics. Available at: https://en.wikipedia.org/wiki/Behavioral_analytics (Accessed: 24 November 2020).

MagicBand is used only at Walt Disney World Resort in Florida named Magic Kingdom. The data below is much more convincing to prove that MgcBand is a great help for attracting the visitors and increase the accommodation of the guests.

Attendance at Disney theme park locations in 2019, by park (in millions)

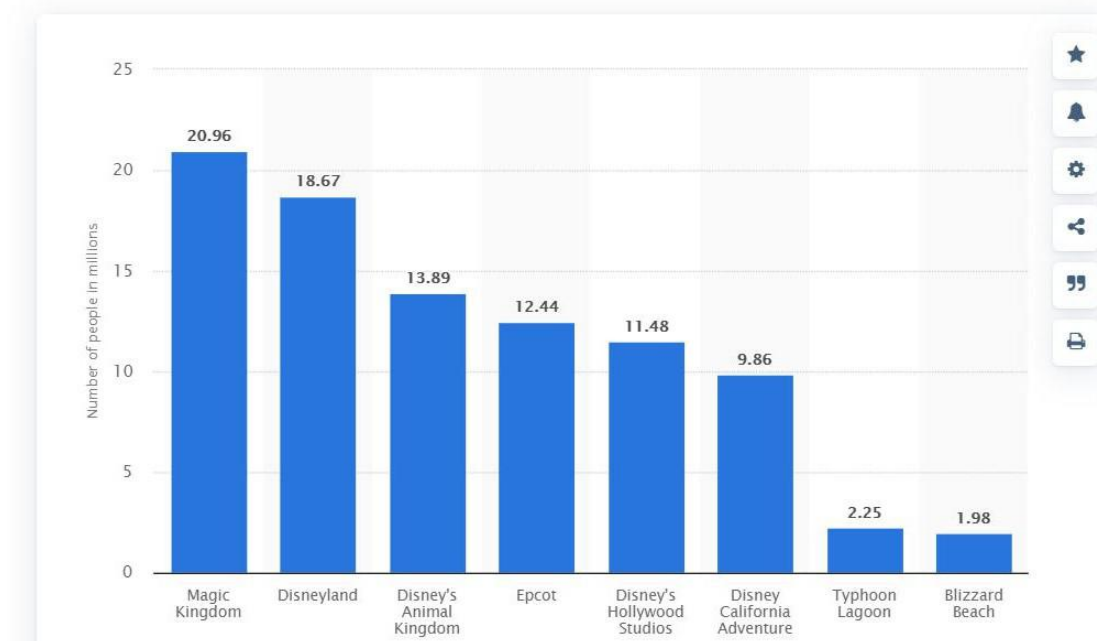


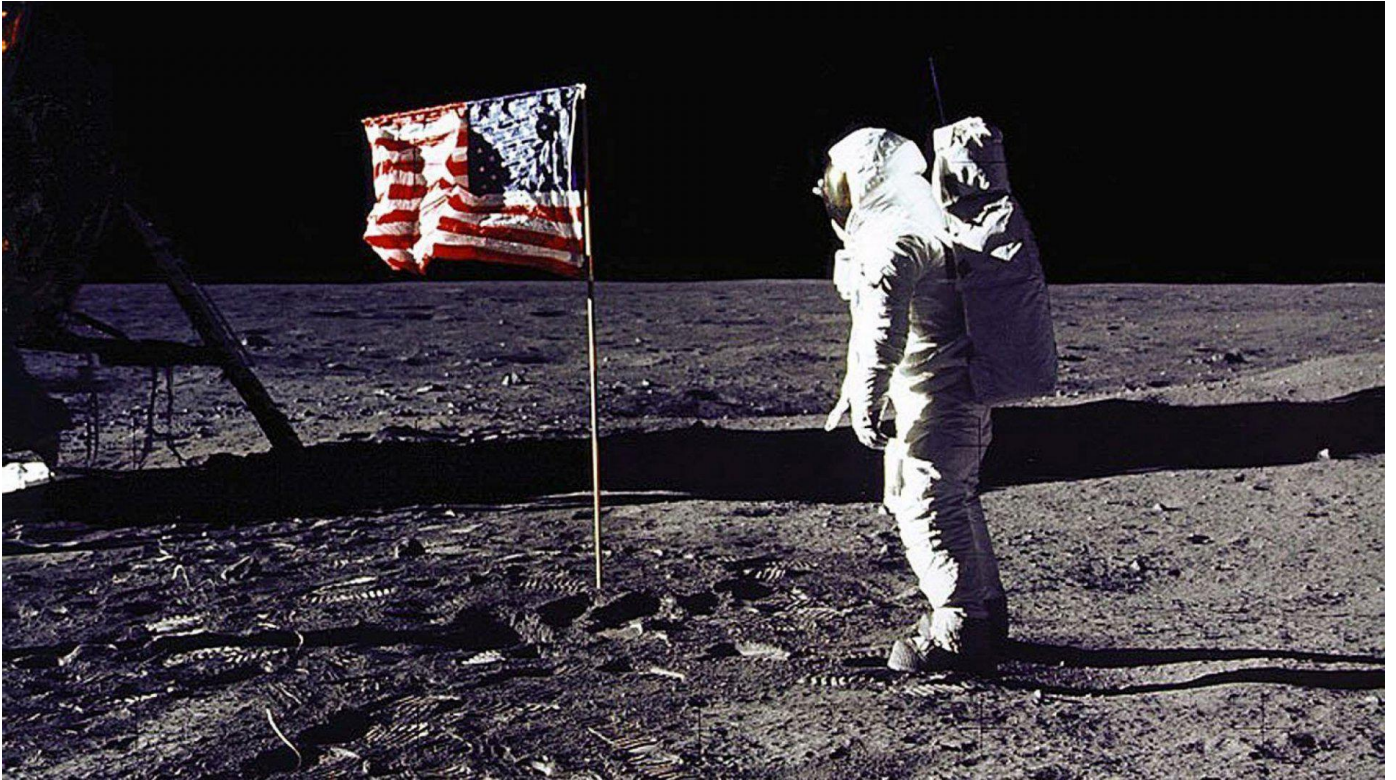
Figure 8: attendance at Disney Theme park locations in 2019 (resource: statista, 2019)³⁶

Unlike the other big companies, FaceBook, Amazon, Netflix, who utilizing the data to acquire insight into customers' behavioural pattern, as a traditional enterprise, Disney's huge investment into the big data is of epoch-making significance.

The potential of these data is unlimited because you never know the extreme of the utilization of data. Compared with the possible profit they will bring, I believe 1 billion dollars is a richly- rewarded investment ever.

³⁶ Statista (2019) Attendance at Disney Theme Park Locations in 2019. Available at: <https://www.statista.com/statistics/227490/disney-world-visitors-orlando-usa/> (Accessed: 24 November 2020).

Blog7 One Giant Leap For A Man (words count: 390)



When I got to the last entry of my blog, I realized that the module, both scary and enchanting, finally came to an end. I was extremely surprised when I found myself unwilling to part with it. I have spent a six-week exciting journey with Zach, the musician who knows nothing about statistics. (A role in Field's *an adventure in statistics*.) *I take him as a metaphor that encourages all the novice struggling with statistics like me.*

I am wondering if he has found “the reality prism”, as I am longing to know if I could apply the quantitative method in my DBA research convincingly. But there is one thing I

am quite sure, this module has opened the door of a wonderland and I can't wait to enjoy the scenery inside.

Before this module, the qualitative method is the tool I often used and now I know I got a new option. The two methods are so different but not against each other. The quantitative-qualitative debate is commonly described as subjectivity versus objectivity, quantification versus generalization.³⁷ After watching a lot of Youtube videos, I found a better way to understand it. The quantitative method works well with deductive research while qualitative method works well with inductive research. Field noted that the two methods are complementary, not competing.³⁸

With its inherent nature of objectivity, "Quantitative methods form a broad range of numerical approaches for analyzing and solving problems."³⁹ This language of mathematics and numbers can not only give a clear measurement but also could be used in calculations. So in my research, the combination of the two methods is most possibly applied to strengthen my points.

I am working on an essay about how religion will affect the economic growth on city-level in China, to resonate with the idea from Weber in the book *The Protestant Ethic and the Spirit of Capitalism*. To seek the correlation of two or more variables, the quantitative method can display its skills to full. After literature reviewing, identifying

³⁷ Hanson, B. (2008) Wither Qualitative/Quantitative?: Grounds for Methodological Convergence. *Quality & Quantity*, 42(1),pp. 97-111.

³⁸ Field, A. (2013) *Discovering Statistics Using IBM SPSS Statistics*. Sage.

³⁹ Waters, D., & Waters, C. D. J. (2008) *Quantitative Methods for Business*. Pearson Education.

the independent and dependent variables, I am going to start the next critical phase, collecting data. I don't know if I could succeed or not for my first trial of the quantitative method, but I am confident that I am on my way to integrating and enriching my research by this language of mathematics and numbers.