LET'S FACE IT: USING CHERNOFF FACES TO PORTRAY SOCIAL MEDIA BRAND IMAGE

Anthony Chan*, Leyland F. Pitt**, Deon Nel*

Abstract

The age of social media sees that brands are consciously utilizing social media to reach existing customers, acquire new ones, establish credibility, maintain reputation, or simply become part of the conversation. Those who manage brands need to understand the strategic importance of their visibility, the sentiment toward them, and the passion with which they are discussed in the most popular social media relative to competitors. This study describes a source of data of brand visibility in social media, and then presents a simple yet powerful graphical tool for portraying this information. This permits, it is contended, a means of quickly assimilating and understanding this information. The managerial implications of the approach are discussed, its limitations are acknowledged, and avenues for future research are identified.

Keywords: Social Media, Chernoff Faces Brand Image, Brand Visibility, Brand Sentiment, Social Media Strategy

** Division of Industrial Marketing, eCommerce and Logistics, Luleå tekniska universitet, Fakturaservice, FE 227SE-833 83 STRÖMSUND, Sweden*
Phone: +1 778 782 5129
Email: anthonyc@sfu.ca
*** Segal Graduate School of Business, Simon Fraser University, 500 Granville Street, Vancouver, BC, Canada V6C 1W6*
Phone: +1 778 782 7712
Email: lpitt@sfu.ca
**** Professor of Marketing, Department of Marketing Management, Faculty of Economic and Management Sciences, University of Pretoria, Pretoria, South Africa*
Email: deon.nel@up.ac.za

1. Introduction

Social media are now as influential, if not more so than, conventional media and the impact this is having on brands is considerable. Strong existing brands have been threatened – for example, "United Breaks Guitars" is a protest song written by Canadian musician Dave Carroll chronicling his experience with United Airlines, which quickly became viral and was on Time magazine's top 10 viral videos of 2009. Old, declining brands have experienced Lazarus-like rejuvenation, thanks in no small part to well-managed social media campaigns. Procter and Gamble's Old Spice mens' toiletries range enjoyed spectacular revival during 2010 with more than 120 000 followers on Twitter, devoted fans on Facebook, and many millions of downloads of the ads on YouTube. New brands have even been created in social media: Journalist Justin Halpern's tweeting of the sayings of his crusty 78-year old father on Twitter attracted more than a million followers within a few weeks. This enabled him to publish a bestselling book (Halpern, 2010) and also became the basis of a television sitcom starring William Shatner as Halpern senior.

These examples highlights a critical need for brand managers and scholars alike to gain an insight into what is being said about brands in social media and how such discourse affects consumer sentiments. Since social media are multi-dimensional and attempts to understand them require tracking different measures simultaneously, what is the best way to portray this data? We will address these issues in this research article.

Our paper will first provide a brief overview of the various forms of social media platforms. Second, we discuss the use of Social Mention, a data collection tool on brand visibility data in social media. Then we detail a study of the relative positing of some competing IT brands according to their assessment in Social Mention. We then explain the use of a graphic technique called Chernoff Faces (Chernoff, 1971; 1973) which simultaneously portray the chosen brands on a number of significant dimensions so that their relative positioning can be contrasted with each other. The conclusion of this paper discusses the limitations of this study and provides a list of implications for brand managers. Areas of future research are identified.

2. Social Media: An Overview

Kaplan and Haenlein (2010) define Social Media as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content." Social media tools are typically designed to be highly accessible (easy to get to and easy to use) and scalable (can handle a large number of users) publishing techniques (Brogan, 2010; Zarella, 2010). Such tools utilize the Internet and other related technologies to transform broadcast monologues (one to many) into social media dialogues (many to many). They support the democratization information by transforming individuals from mere content consumers into content producers. Social media are sometimes referred to as user-generated content (UGC) or consumer-generated media (CGM) (Berthon et al., 2008). When consumers create ads of brands that they either love or hate, they are referred to as Consumer Generated Advertising (CGA) (Berthon et al., 2008).

Social media have altered the way business is conducted and how consumers interact with businesses and brands. If managed effectively, organizations can now become part of customer conversations in addition to reaching and interacting with them online. The advent of social media enables brands to utilize social media to better serve existing customers, acquire new ones, and establish or maintain reputation and credibility.

Social media tools are enabled by the emergence of a class of Internet technologies commonly known as "Web 2.0". "Web 2.0" features technologies that provide enhanced interactivity between users and a website. Prior to "Web 2.0", websites are more or less static pages which allow users to view and retrieve information.

Social Mention and (Un)Green IT Brands

Social Mention is a social media search and analysis tool that aggregates user generated content from a number of different social media sources (such as MySpace, Twitter, Facebook, YouTube, and Digg) into a single stream of information. Social Mention provides users with a real-time tool to keep track of what people are saying about a particular issue and measure the frequencies of comments made. Issues can be related to a product or a service, a brand or an organization's reputation, or any topic that is of interest to users. Social Mention monitors over 100 social media sources constantly and provides a point-in-time search and analysis service that can include daily social media alerts. Social Mention reports on a number of metrics for each of the brands chosen for this study. The metrics used are described in table 1.

For illustration purposes, we have chosen "green" and less "green" information technology (IT) brands from a report by a joint research project called GreenFactor by Strategic Oxygen and Cohn & Wolfe. GreenFactor's (2008) research study surveyed 11,740 IT professionals of small, medium, and large companies in 13 countries who had involvement in IT purchase decisions. The objective of the study is to "illuminate 'green' marketing opportunities and further 'green'-focused research on a global scale" (GreenFactor, 2008). "Green" is defined as "efficient power consumption, recyclable/reusable packaging, recycling offers for older equipment, use of non-toxic materials, or making investments in future green concepts such as alternative materials" (GreenFactor, 2008 downloaded November 29th 2009 from http://www.greenfactorstudy.com/).

To illustrate the portrayal of brands in social media using Chernoff Faces, we chose to contrast the top three brands, which respondents see as "most green" (Hewlett Packard, Dell, and IBM) in the GreenFactor survey, with the bottom three brands, which respondents see as "least green" (Fujitsu, SAP, and NEC). The clustering of the most "green" and the least "green" is a useful technique to assist the research in discriminating between the contrasting Chernoff Faces (Johnson and Wichern, 2007).

3. Methodology

In order to illustrate our approach to using data from Social Mention to portray IT brands, we gathered data on the 6 IT brands identified above in the GreenFactor study by entering their brand names into the Social Mention website, and having the website calculate the metrics (shown in table 1) for each brand in all the social media Social Mention reports on. Then a contingency table was created with the 6 IT brands as columns, and the Social Mention metrics (Strength, Positive/Negative Mentions, Passion, Reach, Unique Authors and Relative Frequency) as rows. The contingency table was then used as data input for the construction of Chernoff Faces using the statistical package Stata.

Table 2 shows a summary of the scores for each of the six brands on each Social Mention dimensions. Interpreting relatively complex tables such as Table 2 can be challenging. The reader not only wishes to determine where a specific brand is performing well or poorly, but would also want to discern how a brand stacked up against other brands on a specific indicator. The ability to create a picture of a human face in which the brands and the criteria are portrayed would make this easier.

Chernoff Faces

Over the years various graphic display techniques including pie-charts, histograms and scatter diagrams have been used to portray statistical data (Beniger and Robyn 1978; Zelazny 1972). From the 1980's onward, the accessibility of user-friendly software and relatively inexpensive graphics plotters and printers, as well as other media-producing devices with which to create these displays, greatly expedited the task of researchers and managers in communicating numeric information. Unfortunately the ability of many of these displays to depict multi-dimensional data was severely constrained, particularly when a basis for generalizing and communicating relationships was desired. Some researchers explored icons as a way of displaying multivariate data (Everitt, 1978); Cleveland, 1985). This variety of icons for representing multivariate data included tools such as Fourier blobs (Fienberg, 1979), glyphs (Anderson, 1969) and faces (Chernoff, 1973). These offered novel ways of presenting intricate data by means of straightforward, interpretable pictures.

Unlike most graphs, icons are not designed to communicate absolute numerical information. They are intended for recognizing clusters of similar variables and are useful for sorting or organizing, and especially comparing, variables that differ in many respects. While some researchers argued the use of icons to be subjective and ad hoc, cognitive science research on multi-attributable visual processing, has shown that people can accurately categorize multivariate data based on appropriate visual cues (Garner, 1974; Spoehr and Lehmkuhle 1982). The human face (or a simpler representation of it) is one of the most effective graphical icons for visually clustering multivariate data, particularly for long-term memory processing. Wang (1978) describes a number of papers on applications of faces to multivariate data, while Wilkenson (1982) showed that faces can be more effective than many other icons for similarity comparisons.

Chernoff (1973), a statistician, originally proposed the facial technique. It is helpful first, in that widely divergent facial features are shown, each of which can be associated with a different variable. Second, most people are able to discern correctly between faces with different features. In Chernoff's (1973) opinion, "People grow up studying and reacting to faces all of the time. Small and barely measurable differences are easily detected and evoke emotional reactions from a long catalogue buried in the memory" (p.362). He later went on to say (Chernoff, 1978): "I believe that we learn very early to study and react to real faces. We perceive the face as a gestalt and our built-in computer is quick to pick out the relevant information and to filter out the noise when looking at a limited number of faces" (p.1).

The Chernoff Faces procedure has been incorporated into many statistics and statistical graphics packages. Essentially the procedure involves the assignment of variables in the data set to the features of a face. It is both relatively cogent and flexible and can be tailored to suit the prerequisites of almost any data set, and the technique has been applied in a wide range of disciplines and field. Chernoff (1973) describes its use in such diverse fields as the study of fossil data, in geology. Apaiwongse (1995) uses the approach to detect perceptions among market drivers toward environmental protection policies, while more recently Raciborski (2009) applies Chernoff Faces to a portrayal of public utility data. In marketing specifically, Huff, Mahajan and Black (1981) used faces to illustrate progressions of business failure and success, and Golden and Sirdesai (1992) displayed consumers' perceptions of multi-dimensional, multi-object attributes (brand and retail image impressions) using Chernoff Faces. Nel, Pitt and Webb (1994) illustrated and compared service quality scores from a large-sample customer satisfaction study using the procedure. Nowadays, rudimentary Chernoff Faces can even be constructed using simple commands in spreadsheets such as Excel (Hunt, 2004).

Generating Chernoff Faces

Despite the advances in computer statistical and graphic processing capabilities that had occurred in the previous twenty years, Nel, Pitt and Webb (1994) still lamented the difficulty of generating Chernoff Faces. By the mid-1990s, relatively few user-friendly software packages offered control over the drawing and interpretation of cartoon faces. This despite the fact that Chernoff's original methods had been improved by superior algorithms that addressed concerns about unequal facial sizes and the handling of extreme data values (Flury, 1980; Schüpbach, 1987; Friendly, 1991). To a large extent, these problems have been overcome recently by Raciborski's (2009) published syntax for easy use with the high end statistical and data analysis package by StataCorp (2009) that generates Chernoff Faces to "detect patterns, clusters, outliers, and temporal trends".

This study uses the Chernoff command in StataCorp (2009) as described by Raciborski (2009). The command offers wide-ranging control over the allocation of facial features that represents selected independent variables. Eighteen facial features can be programmed to construct individual face graphs for each observation. Detailed facial expressions can be drawn by controlling, for example eye size, pupil size, eye position, eye brows, mouth curvature, hair line, hair density and nose to name but a few. This allows for a greater variety of different looking faces that offers the researcher a rich source of generating Chernoff Faces that makes the clustering of multivariate easier to interpret. Raciborski (2009) illustrates enhanced clustering abilities by generating twenty vastly different looking faces using only eight of the available programmable facial features with public utility data for demonstration purposes (Johnson and Wichern, 2007).

The variables used in this study were allocated facial features using the Raciborsky (2009) Chernoff command syntax. Our procedure for doing this is summarized in table 3.

As can be seen from figure 1, the ideal IT brand in terms of social media presence as measure by Social Mention, would have a fat face, dense hair, bushy eyebrows, large round eyes with large round pupils, a large broad nose and a broadly smiling mouth. The least preferred face for a brand to have would be the minimum values face shown in figure 1.

We then used the Social Mention data for the six IT brands shown in table 2 to construct the Chernoff Faces presented in figure 2. It is obvious that the faces of the 6 brands are quite different, and that their differences are shown quite clearly in the figure. While the brand Dell is not perfect on all the attributes, it does present a favorable picture overall. The nice fat face = high brand strength; large eyes = high positive sentiment, but large round pupils = high negative sentiment as well; reasonable smile (not as good as Fujitsu, for example) = good passion; thin eyebrows = quite low on reach; fattish nose = reasonably high unique authors; and, dense hair = high relative frequency. The NEC brand has the least ideal face – thin, with small oval eyes, and a downturned smile. However its face does have small round pupils, relatively bushy eyebrows and a bigger nose. There do not appear to be any major or obvious facial differences between those brands highly rated in the Greenfactor study (Dell, HP and IBM) and those rated low (NEC, SAP and Fujitsu).

Limitations of this Study

In this paper we outline an approach for simultaneously mapping 6 competing IT brands using Chernoff Faces, based on their scores on the dimensions of Social Mention. Our intention is to illustrate the use of a technique by selecting few brands within a limited number of social media. We do not claim this to be a definitive study of the positioning of all IT brands in social media. Obviously it would be relatively simple to get the data on other brands from Social Mention. The pictures presented here could have looked very differently had different, or more brands been used, or if a tool other than Social Mention, with different metric had been chosen.

A study such as this provides more of a snapshot in time than an ultimate set of results. Social Mention is a dynamic tool, so that the scores obtained, and thus the Chernoff Faces, are for brands up to that particular point in time. Our study provides more of a snapshot in time than an ultimate set of results. The Chernoff Faces presented here represents the scores obtained by Social Mention, a dynamic tool, up to that particular point in time. It is very likely that if the data had been collected at an earlier or later time, dramatically different faces would have appeared. Social media, by nature, is such that their content evolves constantly.

The allocation of Social Mention criteria to facial features is in a sense always arbitrary and will depend on the allocator's personal preferences. For example, a smile can signal a lot, and could be perceived by many to be the most important feature, with possible negative interpretation as a result. For example, in Figure 2, the IBM brand performs relatively well on most criteria except Passion. If Passion is regarded as the most important criterion, then IBM is not doing well, and that can be expected; however, if it is not, or is equally important, then IBM might be perceived negatively to a disproportionate extent. This method is not sensitive enough to single out the effects of a characteristic such as brand greenness. Furthermore, we have chosen to use the assessments of IT brand greenness from a single research study. There are many more sources (such as Greenpeace) that evaluate and report on environmental performance of IT brands. Lastly, the accuracy and reliability of Social Mention as a tool for measure social media discourse is assumed.

Managerial Implications

A number of managerial implications are apparent. A brand manager of a technology company needs to clearly define its social media strategy. While social media is a reasonably new phenomenon, there are many competitors who are actively seeking to shape the conversation and how their brands are portrayed. Developing a strong capability and a focused strategy may be a source of competitive advantage. Tools like Social Mention allow brand managers to keep track of the characteristics of the dimensions which they deem as critical to their company's success. Since social media is a continuously changing stream of collective consciousness, communication strategies need to be refined on a regular basis to ensure their effectiveness.

As mentioned in our earlier discussion, visualizing complex data sets is a challenge for brand managers. The use of Chernoff faces provides a valuable tool for the simultaneous portrayals of brands in multidimensional space. This comes in handy when brand managers need to monitor social media content and the visibility of brands they regard as competitors or benchmarks.

Finally, brand managers will be able to monitor the effectiveness of social media strategies by using data from tools like Social Mention to track their brands' positions over time, and changes in Chernoff Faces, to measure the effectiveness of a particular strategy against competing brands. Raciborski (2009) provides an excellent example, using influenza data, to illustrate this.

Our exploratory study opens up the stream of opportunities for further research. First, trustworthiness and reliability of data by providers such as Social Mention needs to be confirmed independently. This can be done by working directly with these services to gain a better understanding of the methodologies. Second, findings based on secondary data research could be combined with primary data collection in the markets of the brands in question. Chernoff Faces are as amenable to showing the results of survey studies (cf. Nel, Pitt and Web, 1994) as they are to showing the results of secondary data form sources like Social Mention.

Conclusion

The management of brands in an age of social media is not only more difficult, but it is also even more critical than it has been previously. Consumers post videos about proprietary brands on YouTube. They like, friend, and unfriend them on Facebook, and declare their love for their favorite brands to their friends, while just as easily vilifying the brands they hate. Their tweets about brands on Twitter reach potentially massive audiences at the speed of light. The proliferation of consumer generated content and its rapid diffusion wrestles much of the control over messages away from brand managers (Deighton & Kornfeld, 2007). While it is almost impossible to control every aspect of a brand in Social Media, Brand managers should at least make an attempt to monitor what is being said. We suggest that Chernoff Faces can be a simple yet powerful tool in the brand manager's toolbox that will allow them to assimilate complex information quickly, and to track the ongoing conversations that create this information over time.

Table 1. Social Mention Metrics: Descriptions

Metric	Definition	How Calculated
Strength	The likelihood that your brand is being discussed in social media	Phrase mention within the last 24 hours divided by the number of total possible mentions
Sentiment	The ratio of generally positive mentions to the number of generally negative mentions	Number of Generally Positive Mentions/ Number of Generally Negative Mentions This measure can also be gauged in absolute terms by counting the number of positive mentions, the number of neutral mentions and the number of negative mentions*
Passion	A measure of the likelihood of individuals talking about your brand in social media will do so repeatedly	A small number of users mentioning a brand repeatedly will give a high passion score. A large number of individuals talking about your brand, but only infrequently per individual, will give a low passion score
Reach	A measure of the range of influence	Ratio of the number of unique individuals talking about your brand as a % of the number of total possible mentions
Unique Authors	An indicator or the number of authors messaging about a brand	The number of unique authors on a brand within a specific time period
Frequency	The frequency with which mentions of a brand appear	Measured in minutes or seconds. For our purposes, this indicator is reverse-scored; e.g. a brand being mentioned every 30 seconds vs a brand mentioned every 60 seconds would score 60 and the second brand 30. We term this Relative Frequency.

**For this study we counted positive and negative mentions separately rather than simply use the ratio of positives to negatives*

Table 2. IT Brands and Social Mention Scores

IT Brand	Strength	Positive Sentiment	Negative Sentiment	Passion	Reach	Unique Authors	Relative Frequency
SAP	22	114	32	32	24	387	212
NEC	12	96	13	35	22	380	210
Fujitsu	15	82	7	42	20	335	30
IBM	25	107	21	31	22	384	235
HP	17	80	21	39	20	318	90
Dell	25	141	39	38	21	358	240
Totals	116	620	133	217	129	2162	1017

Table 3. IT Brands and Social Mention Scores - Explanation of Facial Features

Social Mention Mention	Facial Feature Allocation
Strength	Facial line - The fatter the face the higher the brand strength
Positive Sentiment	Eye Size - The larger the eye size the higher the positive sentiment
Negative Sentiment	Pupil Size - The larger the pupil size the higher the negative sentiment
Passion	Mouth - The higher the passion the greater the curvature of the smile
Reach	Eye Brows - The larger the reach the bushier the eye brows
Unique Authors	Nose - The more unique authors the larger the nose
Relative Frequency	Hair Density - Greater relative frequency results in higher hair density

As reference points, two extreme faces were generated using the actual minimum and maximum values from the dataset. These are shown in Figure 1.

Figure 1. Maximum and Minimum Values Chernoff Faces

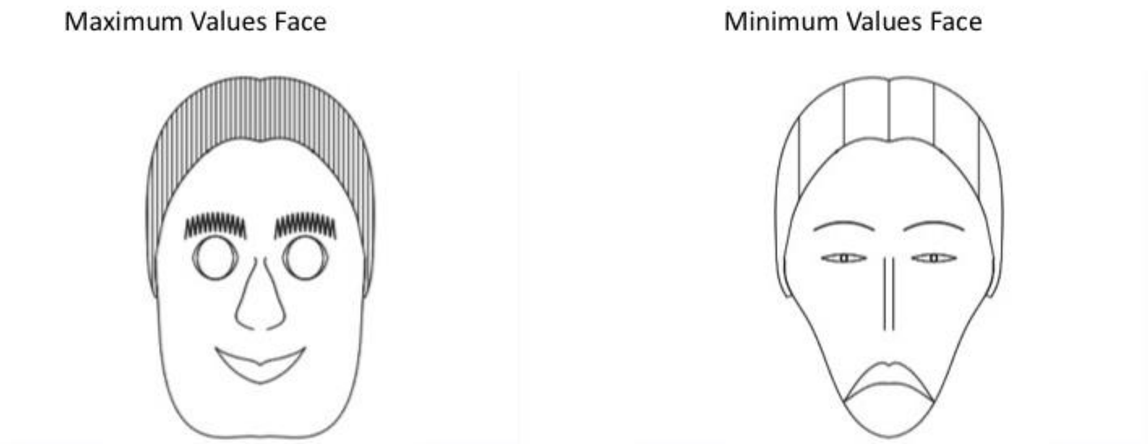
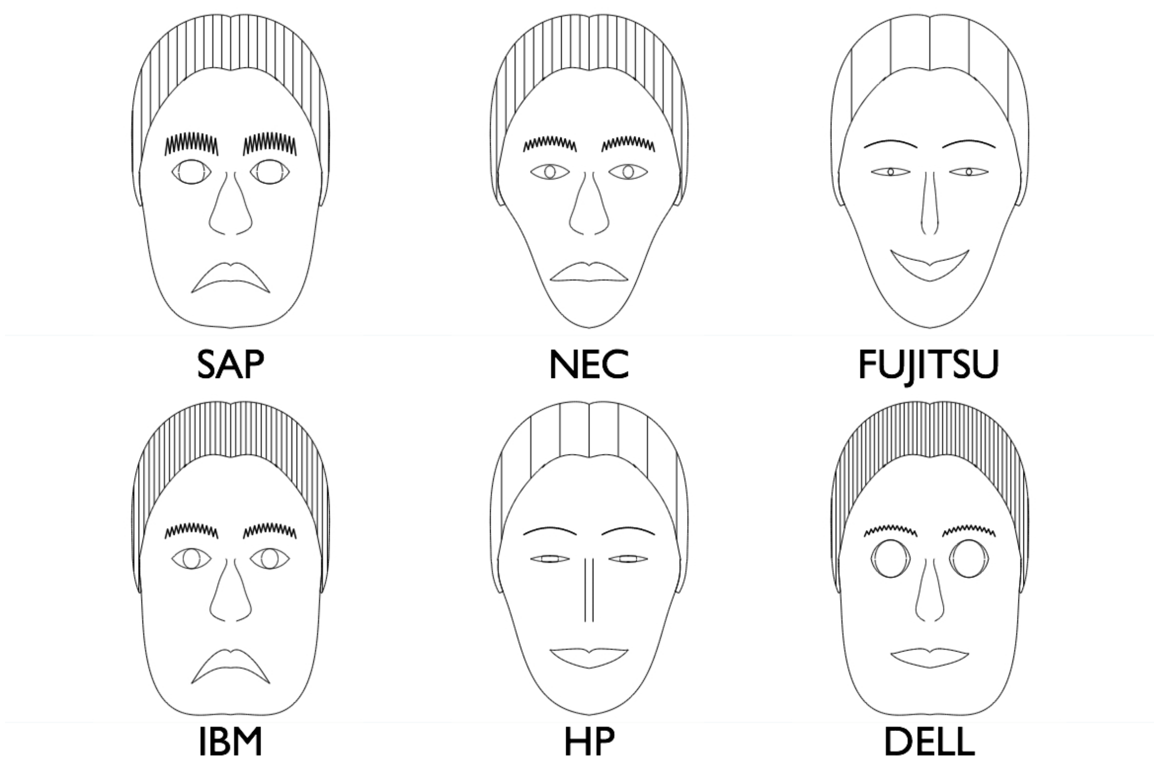


Figure 2. Chernoff Faces of 6 IT Brands, Based on Social Mention Characteristics



References

1. Anderson, E. (1969). A Semigraphical Method for the Analysis of Complex Problems. Technometrics, 2, (3), pp. 387-39.
2. Apaiwongse, T.S. (1995). Facial Display of Environmental Policy Uncertainty. Journal of Business and Psychology, 10, (1), pp. 65-74.
3. Beniger, J. R. and Robyn, D. L. (1978). Quantitative Graphics in Statistics: A Brief History. The American Statistician, 32, (February), pp. 1-10.
4. Berthon, P.R., Pitt, L.F., and Campbell, C. (2008). Ad Lib: When Customers Create the Ad. California Management Review, 50, (4), pp. 6-30.
5. Brogan, C. (2010). Social Media 101: Tactics and Tips to Develop Your Business Online, Hoboken, NJ: John Wiley and Sons.
6. Chernoff, H. (1971). The use of faces to represent points in n-dimensional space graphically. Technical Report 71, Department of Statistics, Stanford University.
7. Chernoff, H., (1973). The use of faces to represent points in k-dimensional space graphically. Journal of American Statistical Association, 68, pp. 361-368.
8. Chernoff, H. (1978). Graphical Representation as a Discipline. In Wang, Peter C. (Eds.), Graphical Representation of Multivariate Data. pp. 1-11. New York: Academic Press.
9. Cleveland, W. S. (1985). The Elements of Graphic Data, Monterey, CA: Wadsworth Advanced Books.
10. Deighton, J., & Kornfeld, L. (2007). Digital Interactivity: Unanticipated Consequences for Markets, Marketing, and Consumers, Harvard Business School Working Paper, pp. 08-017.
11. Everitt, B. (1978). Graphical Techniques for Multivariate Data, London: Heinemann Educational Books.
12. Fienberg, S. E, (1979). Graphical Methods in Statistics. The American Statistician, 33, (November), pp. 165-178.
13. Flury, B. (1980). Construction of the asymmetrical face to represent multivariate data graphically. Technical Report 3, Institute of Mathematical Statistics and Actuarial Science, Bern, Switzerland: Bern University.
14. Friendly, M. (1991). Faces: Faces display of multivariate data. Department of Psychology, York University. http://math.york.ca/SCS/sasmac/faces.html.
15. Garner, W. R. (1974). The Processing of Information and Structure. Hillsdale, NJ: Lawrence Erlbaum.
16. Golden, L.L., & Sirdesai, M. (1992). Chernoff faces: A Useful technique for Comparative Image Analysis and Representation. Advances In Consumer Research, 19, pp. 123-128.
17. Halpern, J. (2010). Shit My Dad Says. New York, NY: It Books.
18. Huff, D. L., Mahajan, V. & Black, W. C. (1981). Facial Representation of Multivariate Data. Journal of Marketing, 45, (Fall), pp. 53-59.
19. Hunt, N. (2004). Chernoff Faces in Microsoft Excel, Teaching Statistics, 26, (3), pp. 75-77.
20. Johnson, R. A., & Wichern, D. W. (2007). Applied Multivariate Statistical Analysis. 6th ed. Upper Saddle River, NJ: Prentice Hall.
21. Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! the challenges and opportunities of social media. Business Horizons, 53(1), 59-68.
22. Nel, D., Pitt, L., & Webb, T. (1994). Using Chernoff Faces to portray service quality data. Journal of Marketing Management, 10, pp. 247-255.
23. Raciborski, R. (2009). Graphical representation of multivariate data using Chernoff faces. The Stata Journal, 9, (3), pp. 374–387.
24. Spoehr, K. T. & Lehmkuhle, S. W. (1982). Visual Information Processing. San Francisco: W. H. Freeman.
25. StataCorp., (2009). Stata Statistical Software: Release 11. College Station, TX: StataCorp LP.
26. Wang, P. C. C. (Ed.). (1978). Graphical Representation of Multivariate Data. New York, NY: Academic Press.
27. Wilkenson, L. (1982). An Experimental Evaluation of Multivariate Graphical Point Representation. In: Human Factors in Computer Systems Proceedings, Gaithers-brug, MD, pp. 202-209.
28. Zelazny, G. (1972). Choosing and Using Charts, New York, NY: Video Arts.
29. Zarella, D. (2010). The Social Media Marketing Book. North Sebastopol, CA: O'Reilly Media.