

Quantitative Research Techniques

June 2009

Ray Poynter, The Future Place

<http://thefutureplace.typepad.com/>

Introduction

This report sets out the findings from a survey about quantitative techniques conducted by Ray Poynter in June 2009. The invitations were distributed via social network groups used by market researchers.

The main driver for this survey was a project I am currently engaged in with a North American university looking at their online learning course. The key textbook for the course is one of the most internationally respected books on the market. However, when I looked at the analysis section I was concerned that the techniques the textbook was saying were the most common were actually, in some case quite rare.

Amongst the major sections in the textbook were:

- Chi-square
- Kolmogorov-Smirnov
- Mann-Whitney U Test
- ANOVA
- MANOVA
- Correlation
- Partial Correlation
- Regression
- Discriminant Analysis
- Logit Analysis
- Factor Analysis
- Cluster Analysis
- Multidimensional Scaling
- Conjoint Analysis

Whilst some of these techniques are, in my opinion, very commonly used, others are, in my experience, rarely used after leaving university.

The aim of the survey was to establish which techniques are in common use and which are somewhat rarer. In classic 'wisdom of crowds' fashion the survey asks the respondents to express their views about what was happening in the industry, as opposed to how often they personally used these techniques.

The survey collected the views of 138 researchers, from 35 countries, with all the fieldwork being conducted in June 2009.

Analysis

The analysis has divided the quantitative techniques into three bands, these being the techniques used often, those used occasionally, and those only used rarely or by specialists.

Used Often

Of the 22 techniques researched, just 7 were rated by 50% or more of the respondents as *Used really often* or *Quite often used*. In descending order of being rated as common they were:

Technique	Often
Correlation	78%
t-tests	71%
Regression	60%
Factor Analysis	56%
Perceptual Mapping	52%
Cluster Analysis	52%
z-tests	51%

Looking at the responses I suspect that it is possible that some respondents might not be aware that z-tests are used by many cross-tabulation programs (the others using t-tests).

The implication for agencies is that if they are quantitative specialists they really need to be offering all of these.

Used Occasionally

Of the remaining 15 techniques 6 were rated as *Used occasionally*, in descending order of popular use:

Technique	At least occasionally
Conjoint Analysis	73%
Chi-square	69%
Correspondence Analysis	69%
Discriminant Analysis	60%
Discrete Choice Modelling (DCM)	53%
ANOVA	51%

One thing that jumps off the page here is how low in the ranking ANOVA is. Most market research textbooks which have statistical techniques section describe Chi-square as ubiquitous, and ANOVA to be fundamental to multivariate analysis. However, as I suspected, both of these are occasional techniques.

The implication for most quant departments is that these techniques are relatively specialised, for example Conjoint Analysis and DCM. Therefore, new staff and clients may be familiar with them.

Used Rarely

Nine of the techniques were rated as “Really rare or specialised”, “Almost never used”, or “Don’t know”. Ranked in terms of how unlikely people are to use them are these nine:

Technique	Rare/Special/DK
Archetype Analysis	90%
Ridge Regression	89%
Structural Equation Modelling	75%
Logit Analysis	74%
TURF	70%
MaxDiff Scaling	63%
Partial Correlation	56%
Multidimensional Scaling	55%
CHAID (or similar eg CART, Decision Tree)	55%

The first thing that struck me was that I use three of these nine frequently, which might mean I am specialised, of perhaps out of step!

One implication for agencies from this finding is that if they are offering any of the techniques on this list they need to ensure clients are familiar with the technique.

The impact of Region

Using the region coding I looked at whether there were differences in the techniques that were seen as ‘often’ used. This produced the table below:

	Europe	North America	Other	Significance
Base	48	45	45	
ANOVA	23%	27%	29%	
Archetype	4%	0%	4%	
CHAID	21%	22%	11%	
Chi-square	46%	47%	40%	
Cluster	50%	42%	56%	
Conjoint	46%	51%	40%	
Correlation	71%	62%	89%	Other sig more than Europe/North A.
Correspondence	42%	24%	64%	Other sig more than North A.
DCM	27%	40%	31%	
Factor Analysis	44%	51%	67%	Other sig more than Europe
Logit Analysis	6%	18%	11%	
MaxDiff	15%	27%	7%	North A. sig more than other
Multidim'l Scaling	17%	16%	40%	Other sig more than Europe/North A.
Partial Correlation	8%	9%	27%	Other sig more than Europe/North A.
Regression	46%	56%	71%	Other sig more than Europe
Ridge Regression	2%	0%	2%	
SEM	6%	11%	7%	
TURF	8%	22%	4%	North A. sig more than other
t-tests	60%	76%	69%	
z-tests	40%	53%	56%	
Discriminant Analysis	15%	24%	22%	
Perceptual Mapping	38%	31%	69%	Other sig more than Europe/North A.

Please note significance testing has been used to highlight differences that are larger. Given the convenience nature of the sample, it does not mean that the probability of these findings being replicated by the population can be realistically estimated. But, for the record, the figures were tested with a confidence level of 95%, with a H_0 that the proportions were equal.

The table suggests the following:

1. The Other region tends to rank more techniques as being used 'Often'
2. Other tends to score higher than Europe/North America on Correlation, Correspondence Analysis, Factor Analysis, Multidimensional Scaling, Partial Correlation, Regression, and Perceptual Mapping.
3. North America tended to score higher than Other on MaxDiff and TURF.
4. Europe was not the highest on any technique.

The implications of this analysis are somewhat muddy. Although Other appears stronger than say Europe, the differences between agencies and likely to be bigger than the differences between regions.

Techniques missed by this research!

At the end of the questionnaire I asked an open-ended question about which techniques I had missed from the questionnaire. There were many suggestions for possible techniques, but I think that most of them would have ended up in the used rarely segment.

The most frequent suggestion was for Latent Class, with a strong mention for pricing techniques.

Overall Conclusions and Closing Thoughts

I started this project because I was concerned that academic textbooks were out of sync with what researchers were actually using. This assumption has been borne out by the research. However, some techniques such as Discriminant Analysis were more common than I had thought, whilst other, most markedly MaxDiff, were less common than I thought.

I think there are two lessons for research providers in this research, and one for research buyers.

1. If you are a provider of quant research, you should probably have reasonable competence in all of the techniques classified in this research as 'Often'
2. If you are providing research that falls into the 'Occasional' or 'Rare' category you probably ought to take extra steps to ensure your clients are aware of the limitations and implications of the research, as it might be unsafe to assume that buyers are well versed with these techniques.
3. If you are a research buyer, you might want to check that your providers of quant research have at least considered the 'Often' category, and probably the 'Occasional' category in terms of your projects.

The final point to make is that fashions change. New techniques are always being developed, whilst others go into decline. This sort of research about research should probably be carried out every few years.

Postscript on Qualitative Research

Several people have contacted me to ask me if there are plans to conduct a similar piece of research into qual techniques. There is now! The qual project will probably happen in the autumn (as we call it in the UK, or fall in the US, or spring in the Sothern Hemisphere).

Appendix

Sample

The data were collected between June 15th and June 26th.

The invitations were distributed via Facebook groups such as The Big List of Market Researchers [<http://www.facebook.com/groups.php?id=637847066#/group.php?gid=18095664464>] and via MRSpace [<http://mrspace.ning.com/>].

138 people completed the survey, from 35 countries. The table below shows the countries, the responses per country, and the way that the countries have been allocated to the regions North America, Europe, and Other.

Country	Europe	North America	Other	Grand Total
Argentina			1	1
Australia			9	9
Belgium	3			3
Bolivia			1	1
Brazil			1	1
Canada		6		6
Chile			1	1
China			1	1
Croatia	1			1
Denmark	1			1
Egypt			2	2
Europe	1			1
France	4			4
Germany	2			2
Greece	2			2
Hong Kong			3	3
India			5	5
Indonesia			5	5
Israel			1	1
Morocco			1	1
Netherlands	3			3
New Zealand			1	1
Pakistan			1	1
Peru			1	1
Singapore			5	5
Slovakia	1			1
Slovenia	1			1
South Africa			2	2
Spain	2			2
Sweden	2			2
Taiwan			2	2
Thailand			2	2
Turkey	2			2
UK	23			23
United States		39		39
Grand Total	48	45	45	138

Note the countries were identified by IP address, which is about 95% accurate.

Summary of Responses

The table below shows the summarised results:

Techniques	Base	Used really often	Quite often used	Used occasionally	Really rare or specialised	Almost never used	DK
ANOVA	130	11%	17%	23%	18%	20%	11%
Archetype Analysis	129	1%	2%	7%	13%	36%	41%
CHAID	129	4%	16%	26%	26%	19%	11%
Chi-square	131	26%	21%	22%	11%	15%	5%
Cluster Analysis	132	21%	30%	33%	10%	5%	1%
Conjoint Analysis	132	21%	27%	25%	17%	9%	1%
Correlation	131	46%	32%	16%	2%	2%	2%
Correspondence Analysis	131	15%	31%	23%	12%	14%	5%
Discrete Choice	131	10%	24%	18%	19%	18%	11%
Discriminant Analysis	122	7%	16%	37%	16%	17%	7%
Factor Analysis	133	27%	29%	31%	8%	5%	1%
Logit Analysis	129	4%	9%	14%	20%	29%	24%
MaxDiff Scaling	131	8%	9%	20%	11%	30%	23%
Multidimensional Scaling	129	10%	16%	19%	22%	23%	9%
Partial Correlation	129	6%	9%	29%	16%	22%	18%
Perceptual Mapping	121	24%	28%	22%	9%	12%	5%
Regression	131	33%	27%	24%	10%	4%	2%
Ridge Regression	128	1%	1%	9%	13%	36%	40%
SEM	130	4%	5%	16%	25%	31%	19%
TURF	129	3%	9%	18%	14%	26%	30%
t-tests	132	50%	21%	15%	7%	4%	3%
z-tests	133	28%	23%	16%	11%	17%	6%

Verbatims

The following verbatim related to techniques that could, perhaps should, have been included on the questionnaire:

- A lot of these are attitude battery techniques, if you avoid attitude batteries for other methods the techniques are less used. Hierarchical bayes is also worth mentioning separate from conjoint.
- AFC
- Cronbach alpha
- All the pricing stuff like PSM, Gabor granger, not really stat but prob relevant
- ancova
- Basic response frequencies (descriptives) with statistical testing.
- CAPM, Gordon Growth Model
- cronbach alpha; only fun survey i've done in years
- Data fusion
- descriptive analysis
- ECONOMETRIC TIME SERIES FORECASTING
- frequency analysis
- From a strictly data processing side, I would add something as simple as WEIGHTING the data, and while talking to all these people I would also wonder what the split is between PAPER/PENCIL surveys, CATI and ONLINE data collection.

- Genetic algorithms, prediction markets
- GLM (nee MANOVA), logistic regression,
- I use Van Westendorp Pricing Analysis quite often...love it!
- jaccard analysis;
- Latent Class Analysis
- Latent class regression, multiple linear regression etc
- latent class regression, shapley-value analysis
- Maybe Latent class segmentation. Quite popular in Europe these days.
- Multinomial Logit Regression ???
- neural networks
- New methods: agent based models, operations research (such as Genetic Algorithms), time series forecasting/modeling (very important!!!), bayesian belief systems, etc.
- Non-parametric techniques
- Other Segmentation techniques such as Latent Class, penalty analysis
- panel (longitudinal/static or non-static)- would be "really rare or specialised"
- Partial Least Squares; Propensity Weighting
- Penalty analysis
- Segmentation techniques that don't involve cluster analysis.
- Survival Analysis (Kaplan-Meier), Econometric Methods (mainly regression)
- Targeted Boot Strapping (TABOO), PLS (Partial Least Squares),
- time series analysis
- Time series Analysis
- tradeoff,