

## Useful Excel Data Validation of Qualtrics Respondent Data Prior to Analysis

John R. Hamilton <sup>1,\*</sup>  
Stephen Maxwell <sup>2</sup>  
Singwhat Tee <sup>3</sup>

\*Corresponding author

<sup>1</sup> Adjunct Chair Professor, James Cook University, Cairns, Qld., Australia, john.hamilton@jcu.edu.au

<sup>2</sup> James Cook University, Cairns, Australia, stephen.maxwell@my.jcu.edu.au

<sup>3</sup> James Cook University, Cairns, Australia, singwhat.tee@jcu.edu.au

### ABSTRACT

This study reports on useful tracking components of the Qualtrics survey instrument, and their application to higher degree research and survey validity. Qualtrics captures the unique survey responses from an individual as coming from a specific device or computer with its own unique internet protocol (IP) address. This IP address is expressed as a numeric in 4 blocks of 3 digits. This numeric provides information that separately locates each individual respondent's online survey completion as created from a specific, unique, Internet-connected digital device or computer. Thus, a Qualtrics *uniqueness-of-position* survey-location-check is achieved for each individual respondent. Qualtrics adds a day, date, and time of survey response stamp. Applied in conjunction with digital device or computer IPv4 address one can detect any multiple survey data sets created from one digital device or computer. Qualtrics longitude and latitude data can precision-map each survey respondent's digital device or computer *geographical location*. This specific geographical data can location-compares each respondent's survey digital device or computer position. Hence, *representativeness* to an overall targeted population is possible – and sometimes even an area-by-area comparison. Used collectively, these three digital device or computer survey checks help measure (or gauge) and *authenticate uniqueness*, and so further *validate survey representativeness* against a targeted population. Also, higher degree researchers, and supervisors, can check validity, uniqueness, and representativeness of each individual respondent's survey, and can readily eliminate any respondent survey duplication(s) – thereby *improving final data set quality*. One benefit of qualified data sets is facilitation of AI systems to make real time operational recommendations. However, in some countries, this raises governance concerns regarding the privacy dilemma such as university ethics respondent anonymity requirements.

**Keywords:** Digital survey, data capture, positional location, 3d Excel map, geographical location, representativeness, validity, reliability.

### INTRODUCTION

Surveying is trending towards an online data collection model (Perera et al., 2013). This approach adds near-real time digital connections towards generative AI inspired business environments. Today, many firms, their workforces, and their business operations, are engaging with smart, dynamic, real-time business and implementing AI related solutions. Such real-time digital intelligence shifts prod surveying towards new methods of internet-shared, data accumulation connectivities, and near-real-time AI assisted analysis (Wang et al., 2023).

The nature of 'work' continues to change from physical manufacturing, assembling, sales, and distribution environments towards automation, robotics, and autonomous machinery. and more towards virtual, and/or digital services, and/or collective intelligences (Drucker, 2023). This promotes a knowledge-based need to better understand firm-related, ever-changing marketplaces, and to better match these towards the inherent dynamic change behaviors of the firm's consumers.

This inclusive generative AI can extend throughout most aspects of a firm's business, and/or its operational activities (Kanbach et al., 2024; Gupta, 2024). Such generative AI business activities help coalesce and network knowledge, theory, and empirical work practices towards new intelligence(s) solutions. Prasad Agrawal (2023) indicates these generative AI changes apply in different degrees and forms to a multitude of public and corporate disciplines including business, education administration, sociology physiology, intelligences, economics, finance, foreign investment, farming, mining, and even into transnational corporate sciences. Arguably, inclusive generative AI will lead to public and private sector operational refocus from current internal management systems towards the development of in-house systems software development with a focused commercial practical application. Survey-supported studies can help move the effects of these changes into the social and economic wellbeing of those who are affected by such structural workplace changes.

Within today's AI-supported management, firm, or societies environments, analytical and scientific survey-supported studies can help understand where and how to pursue productive operational change, refinement, or optimization implications. Such studies often involve real-time, capture-specific, digital, informative, data-capture processes and near real-time survey data-capture information assessments (Sekaran & Bougie, 2013; Hair et al., 2019). Data-capture can be digital and possibly involve

automation, machine, robotic, autonomous digital device data collation, or physical workforce-based data collection, or both (Zarouali et al., 2024).

This study adds insight into distribution and validity aspects of online, and real-time digital surveying (Evans and Mathur, 2018). It suggests how researchers or firms with no prior relationship to respondents/recipients, can externally monitor, and assess, the reliability of their respondent data set – and determine specifically if this small specific respondent cohort can suitably represent the targeted population.

### DIGITAL SURVEYING

Digital surveys may be specific and literature-developed, or machine activity generated. They can digitally merge workforce responsive actions with automatic machine actions. Digital surveys offer low cost, and robust data collection, plus near real-time assessment decisions and response actions (Johansen & Wedderkopp, 2010). This information can then apply to operational systems to mitigate effects of transformational change with the ever-increasing influence of AI supported systems, and a resultant move towards a digital workforce.

Machine-based surveys can provide digital data specifically around operational quality control tests, machine throughput volumes, visual shape/size/fit positions, machine accuracy/precision, process bottlenecks, multi-batch integrations, and the like. This data is fused as a way of integrating imperfect raw data and for capturing reliable, valuable, and accurate information (Meng et al., 2020). This data can also be AI assessed and interpreted for management and for the firm. It can be used to further assist in the modification of the workplace environment – specifically towards delivering an improved net-operational system.

Ongoing digital workforce surveys can provide digital management and firm data that may lead towards further enhancing human contributions throughout the firm's operational systems (Colbert et al., 2016). They can also be indirect, digital device-based, integrated data collection systems including (1) email, (2) LinkedIn or other social media connections, (3) computer keystrokes, (4) web page activities, (5) imagery, or even (6) online distributed documents. This negates barriers of distance and time on the collection of data within an organization, enabling real time analysis that when integrated with AI systems, results in firm focused real time recommendations that can coalesce towards operational optimization and increased competitiveness. These survey-discovered human adjustment options such as virtual reality support can coalesce and focus towards helping to lower negative workforce barriers, such as stress, burnout, and poor workplace actions. They can also be applied towards raising workforce enablers, such as fit-for-purpose individual roles operating with strong management support (Adhyaru & Kemp, 2022). This enables implementation of focused welfare and workplace support systems, ones capable of meeting workforce needs of AI mediated workplace structural changes.

Taken and used together, firm-specific machine and workforce survey data, likely offer a blended solution potentially capable of optimizing such operational systems. Externally, these combined optimized operational systems can also be assessed against workforce acceptance and against the expectations of societies. Thus, across firms digital surveying also holds both internal and external work environment considerations. However, digital workforce surveys sometimes create disadvantages related to correct interpretation, rivalry, boredom, repetitiveness, and limited interest. Thus, digital surveying is extremely diverse and differentiated depending on its purpose and perceived benefits. This study focuses on workforce or human digital surveys and their validity.

### Quantitative Digital Surveying

Workforce or human quantitative digital surveys can support social science studies and their behavioral flows from motives (intentions to act) to consumptive choices (actions), and through to user gratification (outcomes) (Mäntymäki & Riemer, 2014). In addition, they can be applied within psychology, business, industry, societal, or governmental situations. They can assess educative learning systems, software development(s), digital leadership, marketing channels, website benchmarking, body corporate risks, statistical systems, and numerous other research situations. Such informative digital data is collected as quantitative Likert scale (1 = strongly disagree to 5 = strongly agree) responses, qualitative open-ended recorded responses to broad overarching questions, scaled demographics grouping responses, or as reassessed secondary data.

Current survey companies, such as Salesforce.com, suggest quantitative digital surveys offer real time responses to questions, repeatable comparative data collection surveys, and tracking of consumer trend shifts (Lee, XXXX). Such surveying approaches allows a firm to strategically benefit, and rapidly apply its findings, and/or data accumulations, towards generating positive business content, and towards generating further business competitiveness.

Today quantitative digital surveying has trended towards an online data collection model. Many alternative approaches are now available including (1) SurveyMonkey (cloud-based software; used by business for operational and marketplace insights; captures product, services, and workforce experience; multiple additive program extension capabilities; uses one-off business survey of multiple questionnaire item types; has respondent customization possibilities with mobile, tablet, PC connections), (2) SoGoSurvey (enterprise-level software; multiple templates, demographics, and question texts; can be manipulated with email or social media connectivities), (3) Hubspot (qualitative/quantitative consumer dashboard; feedback from email or website survey tools; understanding of consumers habits across time), (4) Google Forms (free online software/questionnaire tool; drag

and drop interface and customizable design; offers real-time response info and charts; has Google interconnections), (5) Survey Planet (simple online questionnaire tool; suite of custom themes and questions; interactive layout; multi-language options), (6) Typeform (user interface to build visual surveys; unlimited questions/answers; data export options; custom-designed themes; readily-made templates), and (7) Qualtrics (cloud-based online survey software toolkit; extends questionnaire boundaries; with SPSS analyses survey data, workforce experiences and behavioral psychology information; multiple question templates and display logic; email triggers, logic branching, real-time web reports; graphical/export options).

This study specifically focuses on the survey instrument Qualtrics (<https://www.qualtrics.com/>). It identifies multiple individual survey uniqueness checks and balances. It shows how these coalesce towards gauging the suitability of overall data capture as representative of the population across a desired data capture domain. Data capture population representativeness is particularly important when considering higher degree research studies that demand precision and sampling data validity, whilst also overcoming potential privacy issues.

## METHODOLOGY

Qualtrics, as a digital survey instrument, offers useful continual digital tracking components. These present particular application to researchers – particularly higher degree researchers, supervisors, and other firm or societies investigators that require high levels of data precision. This study offers approaches to data cleaning, uniqueness, geographical-location-sorting, validity, and representativeness, and so offers a strong data checking mechanism that can extend to a gauging of reliability, prior to establishing each final Qualtrics data set, and prior to conducting the study's analysis phase.

Qualtrics data cleaning, as an initial data preparatory check, is a quick procedure. First, the Qualtrics data is imported into SPSS, and a case ID added. The mean, SD, kurtosis, skew, and range for each survey item is assessed for item spread, uniformity, and suitability. Then SPSS missing values assessment is conducted across the quantitative data. Respondent cases with missing values of over 10% are eliminated, and the missing values are replaced with the appropriate average value. Hair et al. 2019 suggest missing values up to 20% can sometimes be accepted, but this needs careful consideration against data set quality, number of constructs, and sample size. Then a SPSS 'missing values replacement' is completed, and a Little's 'missing completely at random' (MCAR) test is checked for a  $p > 0.05$  significance level result. This suggests any missing data is missing completely at random, and that no pattern exists in the missing data, and so missing value replacement is deemed acceptable (Cunningham 2006; Hair et al. 2019) and is conducted. Resultant remaining Likert scale missing values replacement-numbers contain decimals. These are removed in SPSS by collectively EM recoding each into whole numbers as the same variables (0-1.49999 = 1, 1.5 -2.49999 = 2, ..., 4.5 to 5.5 = 5). A 'missing values' check then conforms no missing values remain. The data set with other qualitative and demographic responses is rejoined using the suite of IDNo's. Hence, the quantitative data is now a 'correct no missing values final data set' and ready for the next stage of statistical preparedness.

To perform a preliminary data capture assessment process, the Qualtrics survey respondent's digitally captured information is investigated for uniqueness, location-sorting, validity, and representativeness. This digitally captured information (IDNo, StartDate, EndDate, Status, IPAddress, Progress, Duration (seconds), Finished, RecordedDate, Latitude, Longitude) is block-extracted to Excel from the above 'correct no missing values final data set.' Three further preliminary investigatory assessments are now considered before the data analysis phase (Computergaga.com, SimoneSezIT.com).

## ANALYSIS AND RESULTS

### Study 1: Uniqueness of Position

Qualtrics captures the unique survey responses from an individual as coming from a specific digital device or computer with its own unique internet protocol version 4 (IPv4) address. The IPv4 address is the most widely format. It is a 32-bit number address, written as four 8-bit fields that allows any digital device or computer to perform specific addressing operations, and so solve relevant addressing problems. It is expressed as a 12 digit, 4 block numeric, formatted as: xxx.xxx.xxx.xxx, with each 'xxx' block having a numerical value from 0 to 255 (256 possible values in each block).

An IPv4 address is typically a numeric world-of-networking identification format. For example, consider IPv4 addresses such as: 10.0.0.177, or 119.157.132.52, or 192.168.1.1. Here, the first block (10, or 119, or 192) represents the network or subnetwork to which the host/computer is connected. The second block (0, or 157, or 168) identifies the subnet within the broader network and with its own set of digital devices and services. The third block (0, or 132, or 1) uniquely identifies each digital device on the network and ensures data the digital device receives is as intended. The fourth block (177, or 52, or 1) uniquely identifies the host or service on the digital device and directs correct applications and services to the specific digital device or computer.

Some IPv4 numerics are not allowed as they are (1) default network addresses (0.0.0.0) that are non-connectable (non-routable), (2) loop hole addresses (127.0.0.1) used by computer to identify itself, (3) prior and automatically assigned addresses (169.254.0.1 to 169.254.254.254) where a digital device remains unsuccessful in receiving any valid IPv4, or digital device IPv4 address from its client/server Dynamic Host Configuration Protocol source, (4) a special address that can send messages to every computer on the network (255.255.255.255).

The four-block numeric information locates the survey's individual responder as positioned at a specific, unique, internet-connected digital device or computer that connects to the internet with its own unique IPv4 address. The IPv4 address consists of the first 3 blocks provide the network positional ID address and the fourth IP block is the specific host/computer address on this network.

The Excel Table 1 Qualtrics survey check of each individual's full group of responses (termed one individual's 'case' of survey responses) exposes the unique IPv4 address. This can be sorted from smallest to largest. It is readily checkable for the unique IPv4 numeric from just one (and only one) survey respondent digital device or computer. Hence, every survey respondent's individual complete suite of survey responses is aligned to just one Internet-connected IPv4 address. This differentiation exposes if the individual respondent made multiple submissions (or if multiple survey respondents 'hot-desked' and used the same digital device or computer). This can trigger a researcher consideration and/or elimination process for incorrect, duplicate submission attempts. Several other survey programs (including Survey Monkey) also offer this IPv4 data capture facility, but to date this author has rarely seen fellow researchers attempt to utilize IPv4 addressing as any form of a preliminary data check.

Table 1a: Excel Individual Respondent Data - Sorting by IP Address

FileHomeInsertDrawPage LayoutFormulasDataReviewViewAutomateHelpAcrobatPower PivotComments																			
Clipboard		Font			Alignment			Number		Styles		Cells		Formulas		Analysis		Add-ins	
Paste	Format Painter	Calibri	11	A A	Text	Wrap Text	General	\$ %	Conditional Formatting	Table	Cell Styles	Insert	Delete	Format	Sum	Sort & Filter	Find & Select	Analyze Data	Add-ins
		B I U				Merge & Center												Create PDF	Create and Share Link
F2																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1																			
2		IDNo	StartDate	EndDate	Status	IPAddress	Progress	Duration(secs)	Finished	RecordedDate	Responded	Responded	Responded	LastName	FirstName	ExtReference	Latitude	Longitude	
3		72	13/09/2023 1:33	13/09/2023 1:49	0	101.53.234.3	100	932	1	13/09/2023 1:49	R_1k0qbjj5916sGG						31.5826	74.3276	
4		279	6/03/2024 22:56	6/03/2024 23:11	0	103.130.156.34	100	923	1	6/03/2024 23:11	R_4pff811XYhtCSj						31.5826	74.3276	
5		351	3/04/2024 19:43	4/04/2024 1:32	0	103.134.3.26	100	20918	1	4/04/2024 1:32	R_4WGs65XVmOVI6nt						31.5511	74.3437	
6		107	15/11/2023 14:09	15/11/2023 14:36	0	103.137.24.0	100	1610	1	15/11/2023 14:36	R_rfChnWWf7mk0uQh						36.0087	74.1241	
7		70	12/09/2023 22:44	12/09/2023 22:56	0	103.149.241.137	100	685	1	12/09/2023 22:56	R_ssVfWNLWHdmj5zb						33.6054	73.0354	
8		175	5/01/2024 23:39	5/01/2024 23:47	0	103.150.239.201	100	466	1	5/01/2024 23:47	R_47OmtOhNsvnl6KK						31.5826	74.3276	
9		74	13/09/2023 3:57	13/09/2023 4:06	0	103.150.43.245	100	554	1	13/09/2023 4:06	R_NW4EBguQFM0XbEZ						31.558	74.3587	
10		111	20/11/2023 15:39	20/11/2023 15:45	0	103.151.42.62	100	390	1	20/11/2023 15:45	R_3hg2Bdy6E1kva0O						31.558	74.3587	
11		52	28/08/2023 4:21	28/08/2023 7:04	0	103.151.43.217	100	9803	1	28/08/2023 7:04	R_2zOVQacDcZr3mBw						31.5826	74.3276	
12		9	15/07/2023 18:42	15/07/2023 18:45	0	103.152.43.39	100	166	1	15/07/2023 18:45	R_3im4vkXdwVPrDW9						31.4495	73.6958	
13		49	23/08/2023 11:33	23/08/2023 11:38	0	103.163.254.110	100	326	1	23/08/2023 11:38	R_8j1CRbzyWrgAFj						33.6054	73.0354	
14		126	26/11/2023 18:26	26/11/2023 18:36	0	103.163.255.134	100	619	1	26/11/2023 18:36	R_Qn3PkvMDbuHvrrr						33.6054	73.0354	
15		201	4/02/2024 0:53	4/02/2024 4:27	0	103.169.64.18	100	12819	1	4/02/2024 4:27	R_4pQB5mObvUklFgt						33.6054	73.0354	
16		204	4/02/2024 23:30	4/02/2024 23:34	0	103.170.179.185	100	262	1	4/02/2024 23:34	R_1Y5fwd5ZhyOOxT						33.6054	73.0354	
17		45	22/08/2023 18:40	22/08/2023 23:03	0	103.174.4.26	100	15782	1	22/08/2023 23:03	R_3kCgalJlYLa9zaC						33.6054	73.0354	

#### Study 1: Uniqueness of Position: Day, Date, and Time Addendum

Qualtrics provides day, date, and time information, and when used in conjunction with a digital device or computer IPv4 address, any individual respondent making multiple submissions from any one digital device or computer (such as a 'new' submission per day) can be further interrogated for possible removal.

#### Study 1: Uniqueness of Position: Excel Validity Check

When interrogating an IPv4 address a 'classless inter-domain routing' notation is followed. This facilitates a more flexible allocation of address space against both a network address, and the subnet within the address, allowing each IPv4 address to differentiate, and remain simple to interpret. Thus, checks for survey respondent uniqueness can be quickly assessed in Excel by highlighting all the Excel spreadsheet's data set, and using the sort and filter functions to custom sort by Table 1's 'IPAddress.' Each respondent's survey data set can be checked for uniqueness-of-position as shown in Table 1. If an identical address occurs, then day, date, and time information, plus a direct comparison of these specific identical address cases, can further indicate whether a second person likely used the same digital device or computer, or whether this duplicate IPv4 addressed case needs removal prior to final data analysis.

#### Study 2: Geographical Location Sorting

Qualtrics captures both the latitude and longitude of each digital device or computer from which the survey respondent's individual suite of survey responses from a computer can be positionally located. This specific data capture point generally should be unique for each survey respondent's individual suite of survey responses – as each individual digital device or computer tends to be a specific individual's operational or task-specific workstation, and so can be digitally positioned a specific longitude or latitude (except when travelling, and possibly when using a mobile device or laptop).

The IP address delivers another form of digital device location, but a precise unique computer address. Shown below the Table 1b latitude and longitude for each digital device or computer. As shown in Figure 1, These 3D or 2D map and offer another close proximity to locate the engaged individual's case of survey responses - as a unique point that is pinned to a specific

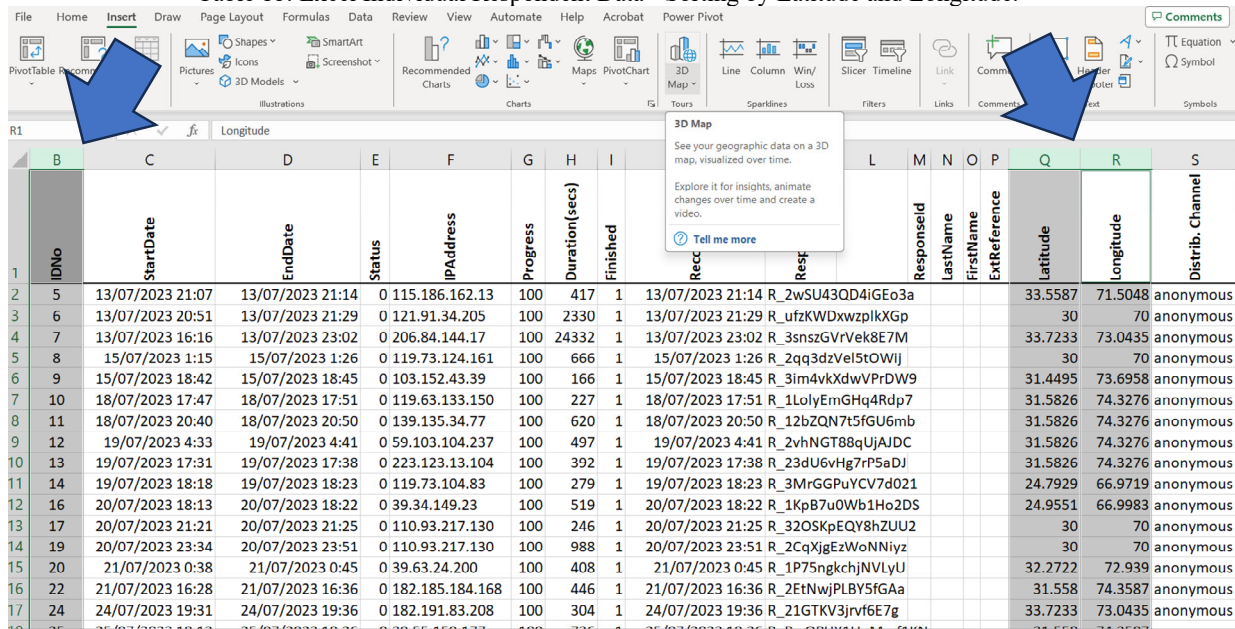
location on a 'World 3D Map.' This 3D World Map can also be rotated in all directions - to check global response locations. It can enlarge and 'zoom-in' to check any chosen specific differentiation from other nearby locality response positions.

### Study 2: Geographical Location Sorting: Day, Date, and Time Addendum

Further longitude and latitude plus date, day, and time information again helps delineate a computer's use as a single individual's survey response input, or as a digital device that adds multiple surveys to a study, and so restricts representation of the targeted population.

Figure 1a shows a 3D Map of Pakistan and the respondents to a Pakistan Software Houses (firms) survey. The concentrations of respondents closely match the concentration zones of these major software houses. Figure 1b expands this study's 395 respondent group as a 2D World Map. This 3D or 2D Map approach allows searches for respondents beyond Pakistan. This study shows external contributors appear in the US, Sweden, Netherlands, UK, France, Germany, Turkey, Saudi Arabia (2), and UAE (2). These geographical locations are also where specific individual major Pakistan Software Houses operate overseas subsidiary corporate software developer facilities. Consideration is then given as to whether these eleven survey respondents should (or should not) be eliminated.

Table 1b: Excel Individual Respondent Data - Sorting by Latitude and Longitude.



IDNo	StartDate	EndDate	Status	IPAddress	Progress	Duration(secs)	Finished	Recd	ResponseId	LastName	FirstReference	Latitude	Longitude	Distrib. Channel
5	13/07/2023 21:07	13/07/2023 21:14	0	115.186.162.13	100	417	1	13/07/2023 21:14	R_2wSU43QD4iGEo3a			33.5587	71.5048	anonymous
6	13/07/2023 20:51	13/07/2023 21:29	0	121.91.34.205	100	2330	1	13/07/2023 21:29	R_ufzKWDXwzplkXGp			30	70	anonymous
7	13/07/2023 16:16	13/07/2023 23:02	0	206.84.144.17	100	24332	1	13/07/2023 23:02	R_3snszGVrVek8E7M			33.7233	73.0435	anonymous
8	15/07/2023 1:15	15/07/2023 1:26	0	119.73.124.161	100	666	1	15/07/2023 1:26	R_2qq3dzVel5tOWij			30	70	anonymous
9	15/07/2023 18:42	15/07/2023 18:45	0	103.152.43.39	100	166	1	15/07/2023 18:45	R_3im4vkXdwVPrDW9			31.4495	73.6958	anonymous
10	18/07/2023 17:47	18/07/2023 17:51	0	119.63.133.150	100	227	1	18/07/2023 17:51	R_1LolyEmGHq4Rdp7			31.5826	74.3276	anonymous
11	18/07/2023 20:40	18/07/2023 20:50	0	139.135.34.77	100	620	1	18/07/2023 20:50	R_12bzQN7t5fGU6mb			31.5826	74.3276	anonymous
12	19/07/2023 4:33	19/07/2023 4:41	0	59.103.104.237	100	497	1	19/07/2023 4:41	R_2vhNGT88qUjAIDC			31.5826	74.3276	anonymous
13	19/07/2023 17:31	19/07/2023 17:38	0	223.123.13.104	100	392	1	19/07/2023 17:38	R_23dU6vHg7rP5aDJ			31.5826	74.3276	anonymous
14	19/07/2023 18:18	19/07/2023 18:23	0	119.73.104.83	100	279	1	19/07/2023 18:23	R_3MrGGPuYCV7d021			24.7929	66.9719	anonymous
16	20/07/2023 18:13	20/07/2023 18:22	0	39.34.149.23	100	519	1	20/07/2023 18:22	R_1KpB7u0Wb1Ho2DS			24.9551	66.9983	anonymous
17	20/07/2023 21:21	20/07/2023 21:25	0	110.93.217.130	100	246	1	20/07/2023 21:25	R_32OSkpEQY8hZUU2			30	70	anonymous
19	20/07/2023 23:34	20/07/2023 23:51	0	110.93.217.130	100	988	1	20/07/2023 23:51	R_2CqXjgEzWoNNiyz			30	70	anonymous
20	21/07/2023 0:38	21/07/2023 0:45	0	39.63.24.200	100	408	1	21/07/2023 0:45	R_1P75ngkchjNVLYU			32.2722	72.939	anonymous
22	21/07/2023 16:28	21/07/2023 16:36	0	182.185.184.168	100	446	1	21/07/2023 16:36	R_2EtNwjPLBY5fGAA			31.558	74.3587	anonymous
24	24/07/2023 19:31	24/07/2023 19:36	0	182.191.83.208	100	304	1	24/07/2023 19:36	R_21GTKV3jrvf6E7g			33.7233	73.0435	anonymous

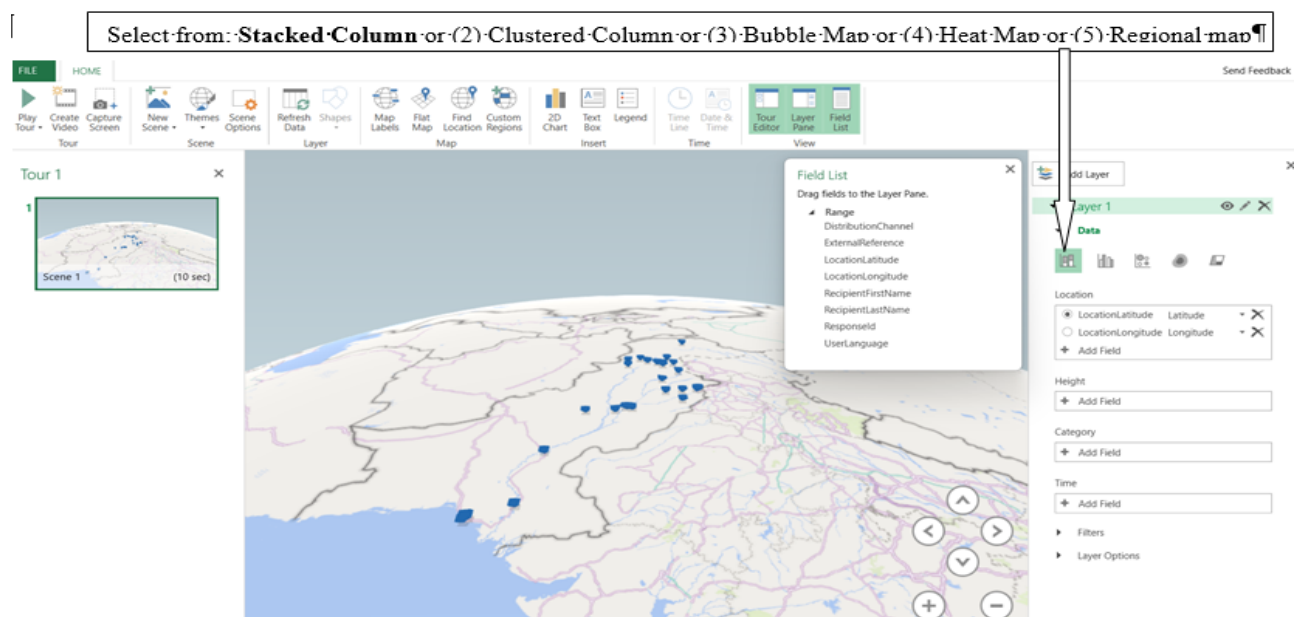


Figure 1a: 3D World Map of Respondent Global Survey Locations with Visual Global Validity Check



### Study 2: Geographical Location Sorting: Excel Positional and Density Check

With the 'IDNo, StartDate, EndDate, Status, IPAddress, Progress, Duration (seconds), Finished, RecordedDate, Latitude, Longitude extracted to Excel, the 'CTRL' key is depressed and held to highlight the IDNo, the latitude, and the longitude columns (headings included) and their data only. Next > select the Insert function (this activates access the 3D Maps icon). Next > select the 3D Maps icon > select (+) New Tour > select Open New Scene > select World map (this yields a rotatable World map) > click on location box to the right of the map > select Add Field > click on Longitude (add longitude field to location box) > click on Latitude (add latitude field to location box) > above this location box > select from 5 data icons, the second from right (heat icon) > view World map and drag and rotate to areas to localities where most of respondents are likely located. This process is summarized in the Figure 1 sequence below. Now select 2D Map and zoom-out to see part or all of the World and the survey respondents' localities (refer Figure 2). These 3D and 2D world maps are downloadable, complete with each survey's latitude and longitude points. Each specific location can be visually located in conjunction with excel maps major city locations, and/or even a specific contour map geographical terrain.



Figure 1b: 3D Map of Respondent Global Survey Locations with Respondent Clusters Outside Pakistan

Engaging the Heat Map icon allows a relative intensity of respondents to be heat positioned on a 3D Map and/or a 2D Map. Figure 2's, 2D regional survey respondent heat map approach allows multiple relative respondent locality distribution checks to be instantly mind-mapped. Further, this allows in-depth comparative understanding of survey respondent intensity sub-sets within or around a localized population, or within the targeted overall survey population domain. This can further validate the representativeness of the respondents within, and across, the targeted overall survey population domain.



Figure 2: 2D Map of Respondent Regional Survey Location Heat Map with Regional Density Checks

Engaging the Figure 3, 2D map of respondent global survey localities as a size intensity (bubble) map allows visualization of the relative location density of respondents. This can present as a 2D and/or 3D Map. It also helps differentiate and clarify the positional clustering of respondents in differing parts of a chosen region. This allows in-depth understanding of stronger survey respondent regional-, or locational-clustering, relative to the overall density locations across the targeted population. This in-turn further validates the representativeness of the respondents within, and across, the targeted overall survey population domain.



Figure 3: 2D Map of Respondent Global Survey Location Intensity Map with Global Bubble No. Position Checks

## DISCUSSION

Researchers, or firms, or societies, with no prior relationship to respondents/recipients, can externally monitor, and assess, the reliability of their respondent data set. Following the processes above they can deliver a correct no missing values data set, a uniqueness of position regarding each survey respondent, a geographical-location-sorting that helps visually understand the survey and its representativeness, and the net assessment of the quality of the data and its overall validity prior to the researcher's or firm's specific analysis processes. Generative AI can also digitally assess this data capture pre-analysis phase, and in near real-time, validate the data set, or suggest correction sand/or help implement each AI-programmed delivery of the 'correct no missing values final data set.'

### Implications of Research

#### *Theoretical Implications*

For researchers, and/or their supervisors, this initial excel investigatory approach provides study-validity in terms of a respondent's uniqueness of presence. Study one's 'uniqueness of position' and its Addendum deliver unique IPv4 addressing that pinpoints each individual respondent as completing the survey at a specific digital device or computer, and at a specific time. AI tools can also real-time capture this information thereby enabling near instantaneous data capture and near real-time case inclusions into any operational model's or system's' investigatory or analysis contributions. This study shows Qualtrics adds to 'data capture theory' by extracting digital survey information into concise useful data that is potentially useful for analysis, and 'grounded theory' by including additional preliminary digital data capture analysis, and iterative comparison and by delivering purposeful validation and reliability assessments.

#### *Practical Implications*

A 'classless inter-domain routing' notation is followed when deploying IPv4, and when interrogating an IP address. As discussed in the analysis and results section above, this approach facilitates a more flexible allocation of address space against both a network address, and the subnet within the address, allowing each IP address to differentiate, and remain simple to interpret. Furthermore, this data analysis allows targeting of survey hotspots, where a number of responses are received from the same location. This information allows for abnormal respondent clusters to be identified, improving the oversight and integrity of the data collected and increasing the integrity of the survey results.

For understanding, and general information, Study 2's 'geographical location sorting' provides a rapid visual assessment for relevant workforce members. It allows for an overall physical interpretation process regarding the spread, or reach, or clustering-density of the survey respondents. This translates towards decision making regarding whether the survey does capture its overall targeted market. AI tools can continually Qualtrics-extract and perform this physical assessment task in near real time. They can then re-assess a new individual respondent's data set for inclusion into relevant parameters that may frame further change inclusions across any relevant operational process or system.

There are many more options available when using 3D and 2D Maps. This information can be found at YouTube sites including: Computergaga.com and SimoneSezIT.com. These sights expand the visual deliverables and offers many different

trace possibilities. These can be presented in video formats.

### **Future Research**

#### *Measurement Aspects*

Survey sampling is non-random and purposeful, and it selects ‘fit-for-purpose’ or ‘expert’ individuals as respondents (Boddy, 2016). Surveying follows a grounded theory approach where data collected is capable of simultaneous and iterative comparison and offers purposeful ideas or theory advancements (Glaser & Strauss, 1967, 1998, and 2017; Bowen, 2008).

Reliability of data capture can be investigated across AI-supported management environments, or firm environments, or societies environments. For example, in this study’s Pakistan software developer firm investigation reliability is tracked by the comparative degree of the survey respondent localities matches against the global localities of major registered Pakistan software developer houses. Similarly, reliability comparisons may also expand into other AI-supported management environments, or firm environments, or societies environments involving digital measurement studies. In such cases, if a relevant targeted population size can be determined, then a respondent population percentage may be determined and gauged for significance and representativeness. This reliability comparison across localities and populations also exposes an assessment option regarding the degree of variability between differing location zones within the overall location spread across the complete survey data collection.

A future Qualtrics study within an IPv6 country like France or Germany, and another Qualtrics study in a 50% IPv4 and 50% IPV6 country like the USA should also be attempted. This then tests if Qualtrics consistently operates in the same manner as it does within IPv4 countries like Australia or Pakistan.

#### *Theoretical Aspects*

Theory articulates connectivities between data and theory and so frames an essential component of research. It predicts/explains (Glaser & Strauss, 1967, 1998, 2017). It shows how and/or why phenomena occur (Corley & Gioia, 2011). It systematically orders ideas, concepts, models, and purpose towards explaining, interpreting and/or shaping practices. (Garrison, 2000). It helps construct relationships and boundaries (Weber, 2003). Theory also brings argumentation and discussion about theoretical mechanisms of an investigated statistical relationship (Elmer, 2023).

Uniqueness theory (Snyder, 1992; Snyder & Fromkin, 1977) is where a difference to others is positioned. Validation enables, confirms, and supports process-initiated actions or developments (Abosag, et al., 2020), but validation theory is largely absent from the literature. This study suggests there is a case for a combined uniqueness and validation theory, one that typically ‘delivers differentiated position regarding a consistent, documentable, process or system.’

#### *Management Aspects*

Across the firm’s digital generative AI deployments and its strategic management operational solutions well-mapped digital survey data collections can be usefully engaged when considering an industrial item, and/or a product, and/or a marketplace purchasing behavior. This assessment and decision making likely focuses on networked operational constructs and their digital survey response interpretations. These can further AI assess in conjunction with survey-identifiable marketplace needs at either a national, or regional, or city, or rural town level. Here, digitally captured survey information can plot intelligent solutions that best support possible strategic management intelligent operational processes and their re-targeting, and/or the intelligent changing of marketplace solutions. Such intelligent digital solutions can benefit management and/or positively modify its generative AI processes.

## **CONCLUSION**

This digitally captured full data set capture of survey information allows researchers, supervisors, firms, or societies to check and assess their respondent group for each individual’s uniqueness-of-position, geographical-location-sorting, validity, representativeness, and possibly to then frame an estimate of their reliability towards suitably representing a population. This quick procedural check follows a simple path. First, a case IDNo. is added, and the data set version 1 is saved. A data cleaning process first conducts a missing values assessment across the quantitative data. Cases with missing values over 10% are considered for elimination. A missing values replacement SPSS solution is then corrected back to whole numbers (no decimals left in data set) and saved as the ‘cleaned’ data set version 2. Data sets versions 1 and 2 are appropriately remerged ready for the Excel validation processes described above. This validation process substantially prevents an individual respondent making multiple submissions and ensures researchers and supervisors work with optimal data sets improving the integrity and reliability of the data enabling more qualified decisions improving operations and competitiveness.

## **LIMITATIONS**

### ***Qualtrics: A Cautionary Limitation Consideration***

However, this also creates a potential ethics dilemma, as it can in some instances, it may raise location privacy issues and university respondent anonymity claims against the researcher.

The IP versions can differ around the world depending on country. Most countries deploy IPv4 addresses which are 32 bits long. However, some countries including France, Germany and India mainly use Google and IPv6, and countries including the



USA, Brazil and Japan are around 50% IPv6 deployment. These IPv6 addresses are stronger and are 128bits long. At the present time, we do not know if Qualtrics behaves the same way and if it records the above information in the same way under different Internet Protocol Domains.

A researcher can specifically turn-off this Qualtrics location option in the 'Survey Termination' section by checking-off 'Anonymize Response,' and so remove all identifiable individual addressing.

## REFERENCES

- Abosag, I., Ramadan, Z. B., Baker, T., & Jin, Z. (2020). Customers' need for uniqueness theory versus brand congruence theory: The impact on satisfaction with social network sites. *Journal of Business Research*, 117, 862-872. <https://doi.org/10.1016/j.jbusres.2019.03.016>
- Boddy, C. R., (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal*, 19(4), 426-432. <https://doi.org/10.1108/QMR-06-2016-0053>
- Bowen, G., (2008). Grounded theory and sensitizing concepts. *International Journal of Qualitative Methods*, 5(3), 12-23. <https://doi.org/10.1177/160940690600500304>
- Colbert, A., Yee, N., & George, G. (2016). The digital workforce and the workplace of the future. *Academy of Management Journal*, 59(3), 731-739. <https://doi.org/10.5465/amj.2016.4003>
- Adhyaru, J. S., & Kemp, C. (2022). Virtual reality as a tool to promote wellbeing in the workplace. *Digital Health*, 8, 20552076221084473, 1-12. <https://doi.org/10.1177/20552076221084473>
- Computergaga.com. (2024). Use 3D maps in Excel - create beautiful map charts. Retrieved from <https://www.youtube.com/watch?v=P--qP4mfxEg>. (accessed 12 May 2024).
- Corley, K. G., & Gioia, D. A., (2011). Building theory about theory building: What constitutes a theoretical contribution? *Academy of Management Review*, 36(1), 12-32. <https://doi.org/10.5465/amr.2009.0486>
- Cunningham, E. (2008). *A practical guide to structure equation modeling using AMOS*, Streams, Statsline: Melbourne, Australia.
- Drucker, P. (2023). *Landmarks of tomorrow: a report on the new post modern world*. Oxfordshire, UK: Routledge. <https://doi.org/10.4324/9781003423461>
- Elmer, T. (2023). Computational social science is growing up: Why puberty consists of embracing measurement validation, theory development, and open science practices. *EPJ Data Science*, 12(1), 58. <https://doi.org/10.1140/epjds/s13688-023-00434-1>
- Evans, J. R., & Mathur, A. (2018). The value of online surveys: A look back and a look ahead. *Internet Research*, 28(4), 854-887. <https://doi.org/10.1108/IntR-03-2018-0089>
- Garrison, R., (2000). Theoretical challenges for distance education in the 21st century: A shift from structural to transactional issues. *International Review of Research in Open and Distributed Learning*, 1(1), 1-17. <https://doi.org/10.19173/irrodl.v1i1.2>
- Glaser, B. G., & Strauss, A. L., 1967, 1998, 2017. *The Discovery of Grounded theory*. Berne, Switzerland: Huber.
- Gupta, V. (2024). An empirical evaluation of a generative artificial intelligence technology adoption model from entrepreneurs' perspectives. *Systems*, 12(3), 103. <https://doi.org/10.3390/systems12030103>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson R. E. (2019). *Multivariate data analysis, (8th ed.)*, Essex, UK: Pearson Education.
- Johansen, B., & Wedderkopp, N. (2010). Comparison between data obtained through real-time data capture by SMS and a retrospective telephone interview. *Chiropractic & Osteopathy*, 18, 1-7. <https://doi.org/10.1186/1746-1340-18-10>
- Kanbach, D. K., Heiduk, L., Blueher, G., Schreiter, M., & Lahmann, A. (2024). The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. *Review of Managerial Science*, 18(4), 1189-1220. <https://doi.org/10.1007/s11846-023-00696-z>
- Mäntymäki, M., & Riemer, K. (2014). Digital natives in social virtual worlds: A multi-method study of gratifications and social influences in Habbo Hotel. *International Journal of Information Management*, 34(2), 210-220. <https://doi.org/10.1016/j.ijinfomgt.2013.12.010>
- Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A survey on machine learning for data fusion. *Information Fusion*, 57, 115-129. <https://doi.org/10.1016/j.inffus.2019.12.001>
- Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. (2013). Context aware computing for the internet of things: A survey. *IEEE Communications Surveys & Tutorials*, 16(1), 414-454. <https://doi.org/10.1109/SURV.2013.042313.00197>
- Prasad Agrawal, K. (2023). Towards adoption of generative AI in organizational settings. *Journal of Computer Information Systems*, 63, 1-16. <https://doi.org/10.1080/08874417.2023.2240744>
- Sekaran, U., & Bougie, R. (2013). *Research methods for business (6th ed.)*, Chichester, W. Sussex UK: John Wiley & Sons.
- SimoneSezIT.com. (2024). How to use the excel maps function: Excel maps tutorial. Retrieved from <https://www.youtube.com/watch?v=1G1PXwa-AW0> (accessed 12 May 2024).
- Snyder, C.R. (1992). Product scarcity by need for uniqueness interaction: A consumer catch-carousel? *Basic and Applied Social Psychology*, 13, 9-24. [https://doi.org/10.1207/s15324834basps1301\\_3](https://doi.org/10.1207/s15324834basps1301_3)
- Snyder, C.R., & Fromkin H.L. (1977). Abnormality as a positive characteristic: The development and validation of a scale measuring need for uniqueness *Journal of Abnormal Psychology*, 86(5), 518-527. <https://doi.org/10.1037/0021-843X.86.5.518>

- Wang, Y., Pan, Y., Yan, M., Su, Z., & Luan, T. H. (2023). A survey on ChatGPT: AI-generated contents, challenges, and solutions. *IEEE Open Journal of the Computer Society*, arXiv:2305.18339. <https://doi.org/10.48550/arXiv.2305.18339>.
- Weber, R., (2003). Editor's comments: The reflexive researcher. *MIS Quarterly*, 27(4), v-xiv. <https://doi.org/10.2307/30036546>
- Zarouali, B., Araujo, T., Ohme, J., & de Vreese, C. (2024). Comparing chatbots and online surveys for (longitudinal) data collection: an investigation of response characteristics, data quality, and user evaluation. *Communication Methods and Measures*, 18(1), 72-91. <https://doi.org/10.1080/19312458.2022.2156489>