Journal of Management Info

Journal Homepage: http://readersinsight.net/JMI

Research Article

Univariate and multivariate normality in data cleaning and bias analysis: A case study on mobile recycling framework

Taher Ben Yahya*

Faculty of Management, Universiti Teknologi Malaysia (UTM), Skudai Johor, 81310, Malaysia

*Corresponding Author email: taher.binyahya@gmail.com

ABSTRACT

The survey questions and results serve as the primary tools for testing hypotheses. However, the accuracy and quality of the final data analysis results depend on the accuracy of the generated data. Therefore, data normalization and cleaning are vital steps in the data processing phase of any analysis. This study focuses on new methods and techniques currently employed for data normalization, with a particular emphasis on univariate normality and multivariate normality analysis, as well as an examination of common-method bias (CMB), especially about data used in PLS-SEM analysis. Researchers should consider data screening methods during the survey design process and select appropriate data normality techniques based on theoretical principles. This paper provides a comprehensive guideline for conducting data normalization and measuring common method bias. Additionally, the paper highlights outdated methods that should be avoided.

Keywords: *Data Cleaning; Research Design; Data Quality; Normality of Data; Multivariate Normality; Common Method Bias*

Submitted: 19 October 2023 *Revised: 12 December 2023* *Accepted: 17 December 2023*

1. INTRODUCTION

Several strategies can be employed to collect data, such as experiments, surveys, case studies, action research, grounded theory, ethnography, and archival research. Some of these strategies align with the deductive approach, while others are more aligned with the inductive approach. According to Saunders et al. (2019), surveys are typically associated with the deductive approach as they seek to answer questions about the who, what, where, and how many/much aspects of the research subject. Surveys are widely used in business and management research due to their ability to collect large amounts of data using questionnaires distributed to a sizable population. They are cost-effective, considered reliable by participants, and easy to interpret and understand (Saunders et al., 2019). Moreover, surveys and questionnaires have been widely used as a primary research method for data collection from end-users in recycling studies (Cao & Liu, 2019; Chan & Chan, 2008; Khan et al., 2019; Kianpour et al., 2017; Nikabadi & Hajihoseinali, 2018; Park & Ha, 2014; Rosenthal, 2018; Yin et al., 2014; Zhang et al., 2020).

There are two main types of questionnaires: self-administered and interviewer-administered. In self-administered questionnaires, respondents are required to answer a

https://doi.org/10.31580/jmi.v10i4.2903

set of questions without the assistance of an interviewer. On the other hand, interviewer-administered questionnaires involve structured face-to-face or telephone conversations between an interviewer and an interviewee (Maylor et al., 2016). In this study, the self-administered questionnaire was chosen as the most suitable method to achieve the research objective. This method offers convenience to prospective respondents in terms of time and location and is cost-effective for the researcher (Bell et al., 2018). Additionally, a high response rate can be achieved through strategies such as constructing a precise questionnaire, targeting specific customer groups during working or studying hours, collecting responses upon completion, and conducting follow-up sessions to capture respondents' attention (Saunders et al., 2019).

However, the process of normalizing and cleaning data significantly impacts the accuracy and quality of data analysis outcomes. Understanding the data and its attributes before implementing any normalization or cleansing methodologies is essential. Therefore, this paper discusses new methods that should be employed in data cleaning and normalization, with a focus on univariate normality and multivariate normality analysis, as well as common-method bias (CMB) analysis. It also highlights outdated methods that should be avoided, as they are still commonly used by many researchers in various studies.

2. FRAMEWORK USED FOR MEASUREMENT

This paper utilized a conceptual framework proposed by Ben Yahya et al. (2022, 2023), as depicted in Fig. 1. In their analysis, Ben Yahya et al. (2022, 2023) examined factors from the Theory of Planned Behavior (TPB), four additional factors from the IBM (Information, Belief, and Motivation) model, and two additional proposed factors (data security and perceived benefits). The incorporation of these factors was expected to enhance mobile phone consumers' intention towards recycling. The framework comprised nine latent variables. The analysis of the data collected for this study, which included 601 responses obtained through a self-administered online survey in the UAE, was conducted using the partial least squares structural equation modeling (PLS-SEM) procedure. The data supporting the findings of this study are available at https://data.mendeley.com/datasets/n4ypn8sp9m/1.

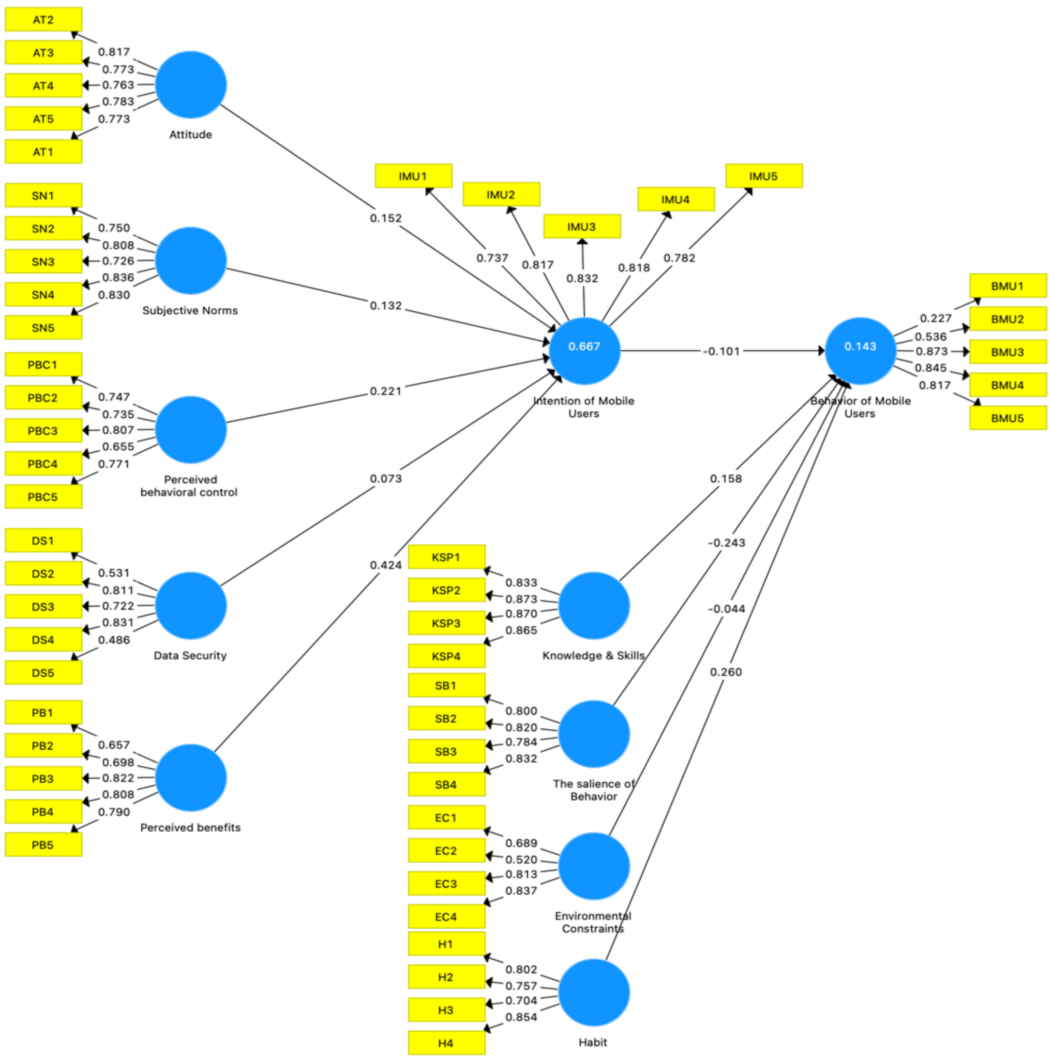

Fig. 1. The Conceptual Model by Ben Yahya et al. (2022, 2023)

3. DATA SCREENING

Data screening is essential to ensure the usability, reliability, and validity of testing causal theories (Field, 2013). It involves examining the collected data to ensure its correctness, cleanliness, and readiness for statistical analysis. This section discusses and illustrates the preliminary steps of data screening, including methods for addressing missing data and outliers. However, this paper will focus on providing detailed information about data normality methods and CMB analysis.

3.1. Missing Data

Missing data is a common occurrence during data collection (Hair Jr et al., 2021). It arises when respondents fail to adequately answer one or more survey items. In this study, partial responses were observed as some participants did not complete the entire survey. To mitigate the issue of incomplete data, online surveys were designed with specific questions designated as mandatory. The present study does not delve into the methods used to address missing data, as extensive research and established techniques are available in the literature.

3.2. Outliers

An outlier, as defined by Hair Jr et al. (2021), is a response that is either too extreme for a specific question or excessively deviates from the norm across all questions. To address outliers, they first need to be identified. Outliers are typically identified for observable variables (or indicators) using z-scores calculated through a descriptive function in software such as SPSS. In large sample sizes (>100), it is expected to have a few participants with z-scores exceeding 3, and consideration might be given to extending the rule to $Z > 4$. A score of ±4.0 is considered an extreme outlier for a large sample size, as the statistical analysis should accurately reflect the majority of the data and not be heavily influenced by a few outliers (Pituch & Stevens, 2015; Stevens, 2012).

3.3. Normality of Data

3.3.1. Univariate normality

The purpose of univariate normality testing is to determine whether the sampled data is derived from a normally distributed population. While PLS-SEM analysis generally does not require the data to be univariate and normally distributed, it is a significant mistake that many scholars ignore the analysis of data normality when using PLS-SEM. It is crucial to check and ensure that the data does not deviate significantly from a normal distribution, and, importantly, scholars need to assess multivariate normality when employing PLS-SEM. (Kline, 2016).

Scholars commonly employ the Kolmogorov-Smirnov and Shapiro-Wilk tests to assess the normality of the data, with a significance level (p-value) greater than 0.05 considered as indicative of normality. Software such as SPSS is often utilized for conducting these tests. Additionally, researchers may calculate the skewness and kurtosis values for the collected data, ensuring that the skewness falls within the range of -1 to +1, and the kurtosis value does not exceed +2 or fall below -2. However, it is important to note that both methods assess data normality rather than multivariate normality.

To provide further clarification, let's consider an example using a dataset applied within the framework illustrated in Fig. 1. The results of the Kolmogorov-Smirnov and Shapiro-Wilk tests for univariate normality show that all variables exhibited significant p-values below 0.05, as illustrated in Table 1.

Table 1. The Kolmogorov-Smirnova and Shapiro-Wilk tests

Construct	item	Kolmogorov-Smirnova			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Attitude (AT)	AQ1	0.272	601	<.001	0.755	601	<.001
	AQ2	0.264	601	<.001	0.799	601	<.001
	AQ3	0.248	601	<.001	0.807	601	<.001
	AQ4	0.283	601	<.001	0.764	601	<.001
	AQ5	0.276	601	<.001	0.76	601	<.001
Subjective Norms (SN)	SNQ1	0.257	601	<.001	0.873	601	<.001
	SNQ2	0.246	601	<.001	0.874	601	<.001
	SNQ3	0.251	601	<.001	0.815	601	<.001
	SNQ4	0.225	601	<.001	0.882	601	<.001
	SNQ5	0.231	601	<.001	0.883	601	<.001

Construct	item	Kolmogorov-Smirnova			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Perceived behavioural control (PBC)	PBCQ1	0.244	601	<.001	0.866	601	<.001
	PBCQ2	0.214	601	<.001	0.898	601	<.001
	PBCQ3	0.234	601	<.001	0.891	601	<.001
	PBCQ4	0.267	601	<.001	0.833	601	<.001
	PBCQ5	0.251	601	<.001	0.874	601	<.001
Data Security (DS)	DSQ1	0.252	601	<.001	0.852	601	<.001
	DSQ2	0.254	601	<.001	0.809	601	<.001
	DSQ3	0.26	601	<.001	0.81	601	<.001
	DSQ4	0.244	601	<.001	0.803	601	<.001
	DSQ5	0.218	601	<.001	0.897	601	<.001
Perceived benefits (PB)	PBQ1	0.232	601	<.001	0.836	601	<.001
	PBQ2	0.241	601	<.001	0.857	601	<.001
	PBQ3	0.258	601	<.001	0.811	601	<.001
	PBQ4	0.251	601	<.001	0.801	601	<.001
	PBQ5	0.243	601	<.001	0.793	601	<.001
Intention of Mobile Users (IMU)	IQ1	0.247	601	<.001	0.838	601	<.001
	IQ2	0.259	601	<.001	0.855	601	<.001
	IQ3	0.274	601	<.001	0.842	601	<.001
	IQ4	0.275	601	<.001	0.827	601	<.001
	IQ5	0.277	601	<.001	0.802	601	<.001
Knowledge & Skills to Perform the Behaviour (KSP)	KSQ1	0.233	601	<.001	0.888	601	<.001
	KSQ2	0.209	601	<.001	0.904	601	<.001
	KSQ3	0.223	601	<.001	0.901	601	<.001
	KSQ4	0.208	601	<.001	0.906	601	<.001
The salience of Behaviour (SB)	SBQ1	0.267	601	<.001	0.822	601	<.001
	SBQ2	0.272	601	<.001	0.84	601	<.001
	SBQ3	0.252	601	<.001	0.852	601	<.001
	SBQ4	0.25	601	<.001	0.83	601	<.001
Environmental Constraints (EC)	ECQ1	0.215	601	<.001	0.887	601	<.001
	ECQ2	0.223	601	<.001	0.901	601	<.001
	ECQ3	0.181	601	<.001	0.912	601	<.001
	ECQ4	0.185	601	<.001	0.913	601	<.001
Habit (H)	HQ1	0.213	601	<.001	0.899	601	<.001
	HQ2	0.204	601	<.001	0.901	601	<.001
	HQ3	0.247	601	<.001	0.883	601	<.001
	HQ4	0.188	601	<.001	0.909	601	<.001
Behaviour of Mobile Users (BMU)	BQ1	0.15	601	<.001	0.905	601	<.001
	BQ2	0.167	601	<.001	0.907	601	<.001
	BQ3	0.259	601	<.001	0.82	601	<.001
	BQ4	0.332	601	<.001	0.75	601	<.001
	BQ5	0.319	601	<.001	0.76	601	<.001

Secondly, the skewness and kurtosis for the collected data are excellent, as the skewness value ranged between -1 and +1, and the kurtosis value in the range of +2 and -2. Table 2 summarises the results by using SPSS software.

Scholars calculate both methods to prove data normality, even while using PLS-SEM. However, as explained before, PLS-SEM analysis does not require the data to be univariate and normally distributed, and this is one of the PLS-SEM advantages. However, it is mandatory to calculate the multivariate normality.

Table 2. Skewness and Kurtosis results using SPSS

Construct	item	N	Mean	Std. Deviation	Skewness		Kurtosis	
		Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Attitude (AT)	AQ1	601	4.2396	0.93764	-1.431	0.1	1.973	0.199
	AQ2	601	4.1398	0.88531	-1.072	0.1	0.973	0.199
	AQ3	601	4.1115	0.94475	-0.997	0.1	0.424	0.199
	AQ4	601	4.2978	0.82632	-1.242	0.1	1.561	0.199
	AQ5	601	4.2812	0.85389	-1.347	0.1	1.993	0.199
Subjective Norms (SN)	SNQ1	601	3.7488	0.99084	-0.625	0.1	-0.081	0.199
	SNQ2	601	3.7521	1.01	-0.637	0.1	-0.052	0.199
	SNQ3	601	4.1082	0.86984	-0.989	0.1	1.106	0.199
	SNQ4	601	3.7205	1.02065	-0.514	0.1	-0.319	0.199
	SNQ5	601	3.7088	1.00668	-0.474	0.1	-0.404	0.199
Perceived behavioural control (PBC)	PBCQ1	601	3.7221	1.11997	-0.572	0.1	-0.656	0.199
	PBCQ2	601	3.4676	1.15009	-0.362	0.1	-0.76	0.199
	PBCQ3	601	3.5474	1.10973	-0.447	0.1	-0.631	0.199
	PBCQ4	601	3.9933	0.9345	-0.921	0.1	0.658	0.199
	PBCQ5	601	3.7221	1.05245	-0.649	0.1	-0.193	0.199
Data Security (DS)	DSQ1	601	3.8419	1.08165	-0.8	0.1	-0.063	0.199
	DSQ2	601	4.0849	0.96667	-1.092	0.1	0.889	0.199
	DSQ3	601	4.0300	1.02426	-1.115	0.1	0.846	0.199
	DSQ4	601	4.1198	0.95164	-1.126	0.1	1.056	0.199
	DSQ5	601	3.5324	1.058	-0.429	0.1	-0.39	0.199
Perceived benefits (PB)	PBQ1	601	4.0116	0.97033	-0.869	0.1	0.355	0.199
	PBQ2	601	3.8802	1.00281	-0.753	0.1	0.112	0.199
	PBQ3	601	4.1115	0.87705	-1.037	0.1	1.176	0.199
	PBQ4	601	4.1631	0.84855	-1.073	0.1	1.382	0.199
	PBQ5	601	4.198	0.84994	-1.109	0.1	1.335	0.199
Intention of Mobile Users (IMU)	IQ1	601	4	0.94868	-0.87	0.1	0.452	0.199
	IQ2	601	3.8735	0.97501	-0.795	0.1	0.342	0.199
	IQ3	601	3.9301	0.92472	-0.863	0.1	0.666	0.199
	IQ4	601	4.0183	0.86679	-0.913	0.1	1.017	0.199
	IQ5	601	4.0965	0.84696	-1.11	0.1	1.825	0.199
Knowledge & Skills to Perform the Behaviour (KSP)	KSQ1	601	3.609	1.08713	-0.465	0.1	-0.588	0.199
	KSQ2	601	3.4093	1.11752	-0.296	0.1	-0.73	0.199
	KSQ3	601	3.4276	1.10537	-0.392	0.1	-0.584	0.199
	KSQ4	601	3.3677	1.12675	-0.296	0.1	-0.723	0.199
The salience of Behaviour (SB)	SBQ1	601	4.0383	0.93463	-0.986	0.1	0.731	0.199
	SBQ2	601	3.9617	0.92387	-0.827	0.1	0.379	0.199
	SBQ3	601	3.9418	0.92102	-0.681	0.1	0.004	0.199
	SBQ4	601	4.0283	0.93854	-0.942	0.1	0.715	0.199
Environmental Constraints (EC)	ECQ1	601	3.3394	1.30177	-0.275	0.1	-1.142	0.199
	ECQ2	601	3.3993	1.07561	-0.395	0.1	-0.475	0.199
	ECQ3	601	3.1947	1.13006	-0.075	0.1	-0.855	0.199
	ECQ4	601	3.2063	1.08044	-0.139	0.1	-0.677	0.199
Habit (H)	HQ1	601	3.5058	1.09714	-0.342	0.1	-0.689	0.199
	HQ2	601	3.4626	1.1412	-0.353	0.1	-0.701	0.199
	HQ3	601	3.6672	0.98948	-0.527	0.1	-0.181	0.199
	HQ4	601	3.2762	1.1805	-0.164	0.1	-0.922	0.199
Behaviour of Mobile Users (BMU)	BQ1	601	2.8785	1.30137	0.117	0.1	-1.028	0.199
	BQ2	601	2.6489	1.17397	0.207	0.1	-0.784	0.199
	BQ3	601	2.1131	1.21813	0.804	0.1	-0.388	0.199
	BQ4	601	1.8536	1.13219	1.12	0.1	0.235	0.199
	BQ5	601	1.9251	1.19974	1.068	0.1	0.011	0.199

3.3.2. Multivariate normality

When employing PLS-SEM, multivariate normality testing is the measurement that should be prioritized above all others. The multivariate skewness should be ≤±3, and the multivariate kurtosis should be ≤±20 (Ramayah et al., 2018). To start the analysis firstly by using SmartPLS software so we need to generate the latent variables for the 601 answers and upload this data to a web application URL "https://webpower.psychstat.org/models/kurtosis/" developed by Cain et al. (2017). Which will generate a result. Summarised in Table 3.

The calculated results in Table 3 show that the collected data was not multivariate normal, Mardia's multivariate skewness (β = 20.16, $p < 0.01$) and Mardia's multivariate kurtosis (β = 205.82, $p < 0.01$). Thus, for such a situation and following the suggestions of Ramayah et al. (2018), the researcher should report the path coefficients, the standard errors, t-values and p-values for the structural model using a 10,000-sample re-sample bootstrapping procedure using SmartPLS software.

Table 3. Test of Multivariate Normality

Variables	Skewness	SE_skew	Z_skew	Kurtosis	SE_kurt	Z_kurt
AT	-1.063	0.100	-10.665	1.593	0.199	8.005
BMU	0.804	0.100	8.066	0.035	0.199	0.176
DS	-1.001	0.100	-10.048	1.490	0.199	7.488
EC	-0.091	0.100	-0.911	-0.822	0.199	-4.130
H	-0.079	0.100	-0.791	-0.536	0.199	-2.693
IMU	-0.688	0.100	-6.907	0.716	0.199	3.598
KS	-0.334	0.100	-3.355	-0.503	0.199	-2.527
PBC	-0.383	0.100	-3.839	0.073	0.199	0.365
PB	-0.767	0.100	-7.700	1.159	0.199	5.824
SN	-0.558	0.100	-5.596	0.522	0.199	2.624
SB	-0.747	0.100	-7.498	0.827	0.199	4.155
Mardia's multivariate skewness and kurtosis						
	β	**Z**	***p*-value**			
Skewness	20.16	2019.09	0.000			
Kurtosis	205.82	45.54	0.000			

3.4. Common-Method Bias (CMB)

Method biases pose a significant challenge as they represent one of the primary causes of measurement error, which jeopardizes the validity of conclusions regarding relationships between measures. Common method bias (CMB), also known as common method variance (CMV), is a potential source of systematic error that can arise when survey respondents or research participants are influenced by the measurement method itself, rather than the underlying construct being assessed. This error is recognized to have both a random and systematic component (Podsakoff et al., 2003). One common issue in data collection for surveys is the use of a single-source questionnaire to measure all variables in the study, particularly if the researcher fails to provide evidence of construct validity for the questionnaire measures or include independent variables unrelated to the questionnaire's core (Conway & Lance, 2010).

The presence of CMB often leads to ambiguous conclusions, which can impact the advancement of theoretical frameworks and practical applications. It has the potential to result in either an overestimation or underestimation of associations among variables, thus undermining the accuracy and significance of research outcomes (Craighead et al., 2011). Many research studies rely on a single source of data, which can give rise to the common issue associated with single-source data. However, two primary methodologies exist to address this issue: the marker variable technique and the full collinearity technique.

3.4.1. The marker variable technique

The marker variable technique (Lindell & Whitney, 2001) is a statistical method employed to control for common method bias (CMB). Many researchers and editors consider the marker variable approach as a gold standard for addressing CMB. Therefore, researchers should carefully consider pre-data collection procedures and conditions before collecting the final data. The marker variable technique involves incorporating a variable that lacks theoretical association with the constructs under examination but is likely to be influenced by CMB. Subsequently, the marker variable is subjected to correlation analysis with the other variables under investigation, and the marker variable coefficient is derived based on its correlations with the other variables. The marker variable technique is a straightforward and user-friendly approach to mitigate the effects of CMV/CMB. However, it is important to acknowledge that the marker variable technique is not without limitations. The selection of the marker variable requires careful consideration, and interpretation of the analysis outcomes should be approached cautiously.

To measure the marker variable technique, it is essential to plan for it before data collection by including additional variables that are unrelated to any variables in the conceptual model. In PLS-SEM, the marker variables are connected to all the dependent variables in the framework (Fig. 2). After running the new model in PLS-SEM, two things should be investigated. Firstly, the R^2 value should be examined. If the R^2 value increases by 10% or more compared to the original model without the marker variable, it indicates a significant problem with the collected data, suggesting the presence of CMB. Secondly, the Beta value (β) should be checked. If the β values were initially significant without the marker variable but become insignificant after adding the marker variables, it indicates that the independent variables were affected by the marker variables. In other words, the data exhibits CMB. If such a scenario arises, the researcher should report the model, including the marker variables.

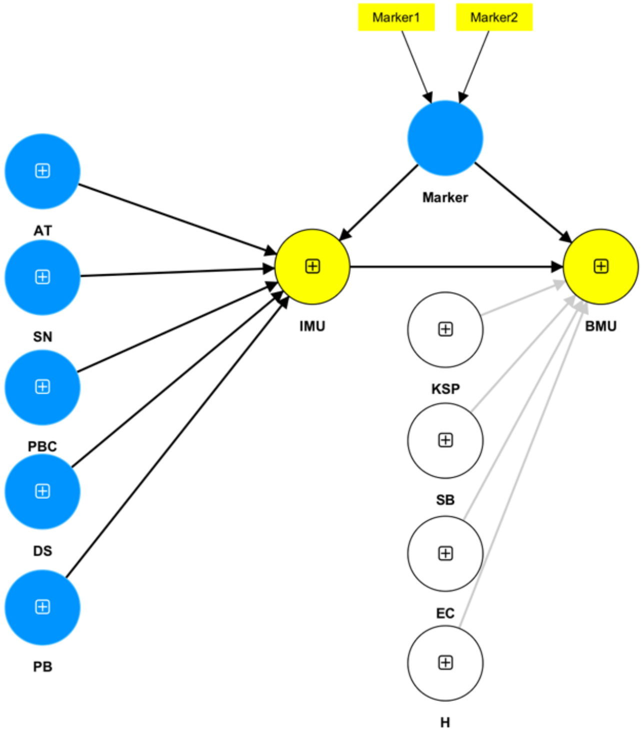

Fig. 2. The Conceptual Model with marker variable

3.4.2. The full collinearity technique

In some cases, researchers may not have the opportunity to employ the marker variable technique and may instead opt to use the Harman single-factor test to assess the extent of CMB. The Harman single-factor test assumes that if a significant amount of CMB exists, an exploratory factor analysis of the data will result in a single factor that explains the majority of the covariance between the independent and dependent constructs. Table 4 presents the results of the CMB analysis based on the single-factor test. The findings indicate that the first factor accounts for only 29.81% of the total variance. This suggests that the variation associated with CMB does not significantly impact the results, as it is less than the threshold value of 50% for variance (Podsakoff et al., 2003).

Table 4. Common-Method Bias Result (CMB) - Harman One-Factor

Total Variance Explained						
Component	Total	Initial Eigenvalues		Extraction Sums of Squared Loadings		
		% of Variance	Cumulative%	Total	% of Variance	Cumulative%
1	15.205	29.813	29.813	15.205	29.813	29.813
2	4.903	9.614	39.427			
3	3.175	6.225	45.652			
4	1.779	3.489	49.141			
5	1.66	3.256	52.397			
...	...	...	...			
...	...	...	...			
...	...	...	...			
...	...	...	...			
...	...	...	...			
51	0.179	0.35	100			

In contrast, there is an alternative method proposed by Kock (2015) that suggests testing for full collinearity to calculate common method bias (CMB). Ramayah et al. (2018) support this approach as the correct method, indicating that correlation and single-factor tests are no longer considered acceptable. To perform the full collinearity test, the latent variable data extracted from SmartPLS is loaded into SPSS, and all variables are regressed against a common variable (e.g., a random variable) in this technique. If the variance inflation factor (VIF) is ≤ 5 and the tolerance level is 0.2 or higher, it suggests that there is no bias from single-source data (Hair et al., 2011; Wong, 2019). Table 5 presents the results of the analysis, indicating VIF values that are less than 5. This suggests that single-source bias is not a significant problem with the sample data used in this paper.

Table 5. Common-Method Bias Result (CMB) - Full Collinearity Testing

(Constant)	**Collinearity Statistics**	
	Tolerance	**VIF**
Attitude	0.419	2.388
Behaviour of Mobile Phone Users	0.813	1.230
Data Security	0.557	1.794
Environmental Constraints	0.754	1.326
Habit	0.406	2.465
Intention of Mobile Phone Users	0.290	3.445
Knowledge & Skills	0.381	2.622
Perceived Behavioural Control	0.447	2.236
Perceived Benefits	0.342	2.923
Subjective Norms	0.400	2.501
The Salience of Behaviour	0.415	2.412

4. CONCLUSION

In this paper, we have emphasized the importance of data normalization and cleaning in the data analysis process. We have provided an overview of various new methods that can be utilized for data normalization, with a particular focus on univariate normality and multivariate normality analysis. Additionally, we have cautioned against the use of outdated methods that are no longer considered acceptable in the field. Furthermore, we have presented a case study to illustrate the practical application of these methods in a real-world data analysis project. The findings of the case study demonstrate that employing these new methods for data normalization and cleaning can yield significant improvements in the quality of data analysis results. We believe that the methods discussed in this paper have the potential to enhance the quality of data analysis in various research and business contexts. We encourage researchers and practitioners to adopt these methods in order to improve the accuracy and reliability of their data analysis outcomes, while also avoiding the utilization of outdated and less effective approaches.

Author Contributions:

The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Funding:

This research received no external funding.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement:

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Acknowledgments:

The author declares that he has no known competing financial interests or personal relationships that could influence the work reported in this paper.

Conflicts of Interest:

No Conflicts of Interest.

Reference:

Bell, E., Bryman, A., & Harley, B. (2018). *Business Research Methods.* Oxford university press.

Ben Yahya, T., Jamal, N. M., Sundarakani, B., & Omain, S. Z. (2022). The Potential Determinants for Smartphone Recycling Behaviour Sustainability in Uae. *Sustainability, 14*(4), 2282. https://www.mdpi.com/2071-1050/14/4/2282

Ben Yahya, T., Jamal, N. M., Sundarakani, B., & Omain, S. Z. (2023). The Effects of Data Security and Perceived Benefits on Mobile Phone Recycling Behaviour and the Recycling Intention Mediation Role. *Recycling, 8*(1), 27. https://www.mdpi.com/2313-4321/8/1/27

Cain, M. K., Zhang, Z., & Yuan, K.-H. (2017). Univariate and Multivariate Skewness and Kurtosis for Measuring Nonnormality: Prevalence, Influence and Estimation. *Behavior research methods, 49*(5), 1716-1735.

Cao, X., & Liu, C. (2019, 12–14 April 2019; IOP Publishing: Bristol, UK). Research on Customers' Willingness to Participate in Express Package Recycling. IOP Conference Series: Earth and Environmental Science, Kuala Lumpur, Malaysia.

Chan, F. T. S., & Chan, H. K. (2008). A Survey on Reverse Logistics System of Mobile Phone Industry in Hong Kong [Article]. *Management Decision, 46*(5-6), 702-708. https://doi.org/10.1108/00251740810873464

Conway, J. M., & Lance, C. E. (2010, 2010/09/01). What Reviewers Should Expect from Authors Regarding Common Method Bias in Organizational Research. *Journal of Business and Psychology, 25*(3), 325-334. https://doi.org/10.1007/s10869-010-9181-6

Craighead, C. W., Ketchen, D. J., Dunn, K. S., & Hult, G. T. M. (2011). Addressing Common Method Variance: Guidelines for Survey Research on Information Technology, Operations, and Supply Chain Management. *IEEE transactions on engineering management, 58*(3), 578-588.

Field, A. (2013). *Discovering Statistics Using Ibm Spss Statistics*. sage.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011, 2011/04/01). Pls-Sem: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice, 19*(2), 139-152. https://doi.org/10.2753/MTP1069-6679190202

Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2021). *A Primer on Partial Least Squares Structural Equation Modeling (Pls-Sem)* (Third ed.). Sage publications.

Khan, F., Ahmed, W., & Najmi, A. (2019, 2019/03/01/). Understanding Consumers' Behavior Intentions Towards Dealing with the Plastic Waste: Perspective of a Developing Country. *Resources, Conservation and Recycling, 142*, 49-58. https://doi.org/https://doi.org/10.1016/j.resconrec.2018.11.020

Kianpour, K., Jusoh, A., Mardani, A., Streimikiene, D., Cavallaro, F., Nor, K. M., & Zavadskas, E. (2017). Factors Influencing Consumers' Intention to Return the End of Life Electronic Products through Reverse Supply Chain Management for Reuse, Repair and Recycling. *Sustainability, 9*(9), 1657.

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (3rd ed ed.). Guilford publications.

Kock, N. (2015). Common Method Bias in Pls-Sem: A Full Collinearity Assessment Approach. *International Journal of e-Collaboration (ijec), 11*(4), 1-10.

Lindell, M. K., & Whitney, D. J. (2001). Accounting for Common Method Variance in Cross-Sectional Research Designs. *Journal of applied psychology, 86*(1), 114.

Maylor, H., Blackmon, K., & Huemann, M. (2016). *Researching Business and Management*. Macmillan International Higher Education.

Nikabadi, M. S., & Hajihoseinali, A. (2018, Apr-Jun). E-Waste Recycling System in Closed Loop Supply Chain [Article]. *International Journal of System Dynamics Applications, 7*(2), 55-80. https://doi.org/10.4018/ijsda.2018040104

Park, J., & Ha, S. (2014). Understanding Consumer Recycling Behavior: Combining the Theory of Planned Behavior and the Norm Activation Model. *Family and Consumer Sciences Research Journal, 42*(3), 278-291.

Pituch, K. A., & Stevens, J. P. (2015). *Applied Multivariate Statistics for the Social Sciences: Analyses with Sas and Ibm's Spss*. Routledge.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of applied psychology, 88*(5), 879.

Ramayah, T., Cheah, J., Chuah, F., Ting, H., & Memon, M. A. (2018). *Partial Least Squares Structural Equation Modeling (Pls-Sem) Using Smartpls 3.0*. Pearson.

Rosenthal, S. (2018). Procedural Information and Behavioral Control: Longitudinal Analysis of the Intention-Behavior Gap in the Context of Recycling. *Recycling, 3*(1), 5.

Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research Methods for Business Students* (8th Edition ed.). Pearson education.

Stevens, J. P. (2012). *Applied Multivariate Statistics for the Social Sciences*. Routledge.

Wong, K. K.-K. (2019). *Mastering Partial Least Squares Structural Equation Modeling (Pls-Sem) with Smartpls in 38 Hours*. IUniverse.

Yin, J. F., Gao, Y. N., & Xu, H. (2014, Feb 15). Survey and Analysis of Consumers' Behaviour of Waste Mobile Phone Recycling in China. *Journal of Cleaner Production, 65*, 517-525. https://doi.org/10.1016/j.jclepro.2013.10.006

Zhang, Y., Wu, S., & Rasheed, M. I. (2020). Conscientiousness and Smartphone Recycling Intention: The Moderating Effect of Risk Perception [Article]. *Waste Management, 101*, 116-125. https://doi.org/10.1016/j.wasman.2019.09.040