

SUPPLEMENTARY METHODOLOGICAL DOCUMENTATION

This file provides supplementary methodological documentation prepared to support the evaluation and verification of the manuscript entitled **“How Destination Attractiveness, Transportation Accessibility, and Infrastructure Shape Sustainable Halal Tourism Business Development: Evidence from a Multi-Regional Survey.”**

Specifically, this document contains the materials requested by the journal, including:

- A. Request No. 4: Survey Protocol / Data Collection Procedure
This section explains the procedures used in collecting the research data, including the research location, research period, respondent groups, sampling technique, and questionnaire distribution process.
- B. Request No. 6: Coding Scheme
This section describes the coding system used for the questionnaire responses, including the scoring of the Likert-scale items and the interpretation of each response category.
- C. Request No. 7: Data Transformation Description
This section explains the procedures applied to prepare the data for analysis, including data checking, coding, cleaning, tabulation, and preparation of the dataset for statistical modelling.
- D. Request No. 9: Formulae Used for Modelling
This section presents the statistical and modelling formulae used in the study, particularly those related to Structural Equation Modeling (SEM), measurement models, structural models, direct effects, indirect effects, total effects, and model identification.

This documentation is intended to provide transparent information on how the data were collected, coded, processed, and analysed, thereby enabling reviewers and editors to assess the methodological consistency and reproducibility of the study.

A. Survey Protocol and Data Collection Procedure

This study employed a quantitative survey design to examine the relationships among Tourist Attractiveness (ToAt), Tourist Transportation (ToTr), Tourist Infrastructure (ToIn), and Sustainable Tourism Business Development (STBD) in the context of world halal tourist destinations. The study was conducted using a structured questionnaire as the main research instrument, supported by limited interviews and documentation studies.

Research Location

The research locus covered halal tourist destinations in three major regions: **Asia, Europe, and America**. These regions were selected because they represent different geographical, cultural, and tourism system contexts. Halal tourism destinations in Asia generally operate in countries with Muslim-majority populations and relatively established halal ecosystems, whereas destinations in Europe and America operate mainly in non-Muslim-majority contexts, creating different dynamics in providing Muslim-friendly services, attractions, transportation, and tourism infrastructure.

Research Period

The research was carried out during the February 2025– December 2025 period. The research process consisted of four main stages:

1. Preliminary stage, including field orientation, preliminary study, development of the research proposal, preparation of the research instrument, proposal seminar, and research permission.
2. Data collection stage, including instrument testing, instrument revision, questionnaire distribution, and questionnaire retrieval.
3. Data processing stage, including data tabulation, statistical calculation, data interpretation, and hypothesis testing.
4. Discussion and conclusion stage, including discussion of findings, formulation of conclusions, and final report writing.

Respondents

The respondents consisted of stakeholders related to world halal tourism destinations. The study involved **385 respondents** from six respondent groups:

No.	Respondent Group	Sample
1	Tourism investors	9
2	Representatives of tourism ministries or state institutions implementing halal tourism policies	7
3	World halal tourism destination managers	6
4	Visitors who had visited halal tourism destinations within the last five years	334
5	Toll road managers related to halal tourism destinations	5
6	Bus drivers who had transported tourists to halal tourism areas	24
	Total	385

The target respondents were selected because they represent different stakeholder perspectives in halal tourism, including investors, government or policy representatives, destination managers, tourists, transport-related managers, and transport service providers.

Sampling Technique

The sampling procedure was determined based on the requirements of **Structural Equation Modeling (SEM)**. The sample was selected using predetermined criteria formulated by the researcher. The sampling process included determining the research area, identifying relevant respondent groups, and determining the minimum sample size using the **Maximin Method**, with a 95% confidence level, $\alpha = 5\%$, and a bound of error of 10%.

The steps undertaken in determining the research sample were as follows:

1. First, the sampling area or research object was determined. The research focused on respondents related to halal tourism destinations and stakeholders involved in halal tourism business development.
2. Second, the minimum sample size was determined using the **Maximin Method**, with a confidence level of 95% or $\alpha = 5\%$, and a **Bound of Error = 10%**, which refers to the acceptable level of statistical error in estimating the sample size. This method was applied under the assumption that the population condition was relatively homogeneous. The formula used was developed by Cochran (2010), as follows:

$$n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} \quad n_0 = \left(\frac{Z \frac{1}{2} \alpha}{2BE} \right)^2$$

The decision rules are as follows:

$$\begin{aligned} \text{If } n_0 \leq \alpha N, \text{ then} \quad & n = n_0 \\ \text{If } n_0 > \alpha N, \text{ then} \quad & n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} \end{aligned}$$

Notes:

n_0	=	Sample size or observation unit based on statistical calculation
n	=	Required sample size or observation unit
N	=	Population size or total observation units
$Z\frac{1}{2}\alpha$	=	Z-table value, where the curve area is $-\frac{1}{2}\alpha = 0,5 - 0,025 = 0,475 \approx 1,96$ based on the Z-table)

The initial sample size was calculated as follows:

$$n_0 = \left(\frac{Z \frac{1}{2} \alpha}{2BE} \right)^2 \quad n_0 = \left(\frac{1,96}{2 \times 5\%} \right)^2$$

$$n_0 = 384,16; \quad \alpha N = 0,05 \times 3,735 = 186,75$$

then the corrected sample size formula was applied:

$$n_0 > 0,05N \text{ or } 384,16 > 186,75$$

$$\text{then,} \quad n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} \quad \rightarrow \quad n = \frac{384,16}{1 + \frac{384,16 - 1}{3.735}}$$

Therefore, based on a population size of 3,735, the minimum sample size would be approximately 348 respondents. However, in the research calculation presented in the dissertation, the sample size used for visitors was 334 respondents, based on the proportional allocation table of visitor data over the last five years. Thus, after the sampling process, the minimum unit of analysis for visitors in this study was determined to be 334 respondents.

3. The proportion of observation units was determined. The sample of 334 respondents was proportionally allocated using the following formula:

$$n_i = \frac{N_i}{N} \times n$$

Notes:

n_i	=	Sample size for each observation unit
N	=	Population size of each observation unit
N_i	=	Total population size
N	=	Total sample size

For example, if one group consisted of 59 people, then the number of respondents representing that group was calculated as follows: $n_i = \frac{59}{2.505} \times 334 = 7,86 \approx 8$

Thus, the minimum number of respondents representing that group was 8 people. The proportional allocation of visitor respondents over the last five years is summarized in Table 1.

Table 1.
Determination of the Number of Visitors in the Last Five Years

Visited Country / Destination	Year	Number of Departures	Number of Visitors	Number of Representatives
Hongkong Shenzhen Macau	2015	2	50	4
Sydney Brisbane Gold Coast		3	90	8
New Zealand		2	50	4
Eropa Barat		4	125	11
Eropa Timur		4	150	13
Russian Scandinavia		1	30	3
Turkey		2	50	4
Jepang		2	40	4
Beijing Shanghari HKG	2016	2	50	5
HKG Shenzhen Macau		2	50	4
Jepang		2	25	2
Scandinavia		1	25	2
Eropa Barat		8	200	18
Eropa Timur		3	75	7
Turkey		2	50	4
Maroko Spain Portugal		2	50	5
SIN Malaysia, Thailand Overland	2017	2	45	4
BKK Pattaya		2	60	6
Russian, Scandinavia		2	60	6
Maroko, Spanyol, Portugal		2	60	6
Eropa Barat		5	175	16
Eropa Timur		4	100	9
Turkey		2	60	6
USA, France, Canada		2	50	5
Korea		2	50	5
Thailand	2018	2	105	10

Jepang		1	25	2
Korea		3	60	6
HKG Shenzhen Macau		2	60	6
Eropa Barat		8	200	18
Eropa Timur		5	125	12
Turkey		2	50	5
Eropa Barat	2019	7	175	16
Eropa Timur		5	125	12
Turkey		2	50	5
Jepang		1	25	2
Korea		3	90	8
Eropa Barat	2020	1	25	2
Eropa Barat		1	25	2
Turkey		3	75	7
Turkey	2021	8	160	15
Eropa Barat		1	28	2
Eropa Timur		1	16	1
Turkey	2022	1	78	7
Vietnam		1	38	4
Canada		1	16	1
USA		1	15	1
Eropa Barat		9	270	25
Eropa Timur & Barat London		1	62	6
Umrah plus Turkey	2023	1	31	3
Turkeys plus Int'l Meeting		1	25	2
West Europe Edutrip		1	11	1
Total			3.735	334

Based on this calculation, the target sample of visitor respondents was determined to be 334 people. Furthermore, the study also included other stakeholder groups, namely tourism investors, representatives of tourism ministries or state institutions implementing halal tourism policies, world halal tourism destination managers, toll road managers related to halal tourism destinations, and bus drivers who had transported tourists to halal tourism areas. The detailed population and sample size of the study are presented in Table 2.

Table 2.
Population and Sample Size of the Study

Unit of Analysis	Total		Percentage (%)
	Population	Sample	
Tourism investors, as representatives of investors	833	9	1,08
Representatives of tourism ministries or state institutions implementing halal tourism policies, represented by ambassadors or staff based in Indonesia	72	7	9,72
World halal tourism destination managers	294	6	2,04

Unit of Analysis	Total		Percentage (%)
	Population	Sample	
People who had visited halal tourism destinations within the last five years, as representatives of visitors	3.735	334	8,94
Toll road managers related to world halal tourism destinations	107	5	4,67
Bus drivers who had transported tourists to halal tourism areas	2.796	24	0,86
Total	7.837	385	27,31

Source: Primary Data Recapitulation, Author, 2025.

Data Collection Technique

Data were collected using both **indirect communication** and **direct communication** techniques. Indirect communication was conducted through a structured questionnaire, while direct communication was conducted through limited interviews and documentation studies. The questionnaire served as the primary instrument for collecting respondents' perceptions of the research variables.

Data Sources

The study used both **primary data** and **secondary data**. Primary data were collected directly from respondents through the survey, limited interviews, and questionnaire responses. Secondary data were obtained from available publications, books, reports, electronic sources, and websites relevant to halal tourism, tourist attractiveness, transportation, infrastructure, and sustainable tourism business development.

B. Data Transformation Description

The data transformation process was conducted to ensure that the survey data were complete, consistent, valid, and ready for statistical analysis using Structural Equation Modeling (SEM). The transformation process included data screening, coding, cleaning, tabulation, composite score preparation, and data formatting for LISREL 10.2.

1. Data Screening

After the questionnaires were collected, all responses were screened to ensure that they met the criteria for analysis. The screening process included checking whether the respondents belonged to the target respondent groups and whether the questionnaires were completed properly.

Responses were considered eligible for analysis if they met the following criteria:

No. Screening Criterion

- 1 The respondent belonged to one of the predetermined respondent groups.
- 2 The questionnaire was completed with valid responses.
- 3 The respondent had sufficient knowledge or experience related to halal tourism destinations.
- 4 For visitor respondents, the respondent had visited halal tourism destinations within the last five years.
- 5 The response pattern did not indicate careless or inconsistent answering.

Incomplete or invalid questionnaires were excluded from the final dataset.

2. Data Coding

All questionnaire responses were coded numerically based on a five-point Likert scale. Each response option was assigned a numerical score to allow quantitative analysis.

Response Category	Code	Interpretation
Strongly Disagree	1	Very Low / Very Poor
Disagree	2	Low / Poor
Neutral	3	Moderate
Agree	4	High / Good
Strongly Agree	5	Very High / Very Good

All questionnaire items were designed as positive or favorable statements. Therefore, higher scores indicate a stronger or more positive respondent perception toward the measured construct.

3. Reverse Coding

Since all questionnaire statements were formulated as positive statements, no reverse coding was applied. Each item was interpreted in the same direction, where a higher score reflects a more favorable assessment of the variable being measured.

Thus, the interpretation of the scoring was as follows:

Higher Score=More Positive Assessment

Lower Score=Less Positive Assessment

This scoring approach was used consistently across all variables, namely Tourist Attractiveness, Tourist Transportation, Tourist Infrastructure, and Sustainable Tourism Business Development.

4. Data Cleaning

Data cleaning was conducted to identify and correct possible data entry errors, missing values, inconsistent responses, and out-of-range values. The cleaning process included:

No.	Data Cleaning Step
1	Checking whether all responses were within the valid range of 1 to 5.
2	Identifying missing responses or blank answers.
3	Checking duplicate entries.
4	Reviewing inconsistent response patterns.
5	Ensuring that each respondent record had complete demographic and questionnaire data.
6	Removing responses that did not meet the inclusion criteria.

Any value outside the valid range of 1–5 was treated as an input error and corrected by referring to the original questionnaire, where possible. If correction was not possible, the response was excluded from the analysis.

5. Data Tabulation

After coding and cleaning, the data were tabulated into a structured dataset. Each row represented one respondent, while each column represented a questionnaire item, respondent characteristic, or research variable.

The dataset structure was prepared as follows:

Data Component	Description
Respondent ID	Unique anonymous code assigned to each respondent
Respondent Group	Category of respondent, such as visitor, investor, tourism manager, government representative, toll road manager, or bus driver
Demographic Variables	Gender, education level, employment status, work sector, and other relevant characteristics
ToAt Items	Indicators measuring Tourist Attractiveness
ToTr Items	Indicators measuring Tourist Transportation
ToIn Items	Indicators measuring Tourist Infrastructure
STBD Items	Indicators measuring Sustainable Tourism Business Development

No personally identifiable information was included in the final dataset.

6. Composite Score Preparation

For descriptive analysis, the responses for each variable were summarized using item scores and aggregate scores. Composite scores were prepared by calculating the mean score of the indicators belonging to each construct.

The general formula used was:

$$\bar{X}_j = \frac{\sum_{i=1}^k X_{ij}}{k}$$

Where:

Symbol Description

$\bar{X}_j$	Composite mean score for construct j
X_{ij}	Score of item i in construct j
k	Number of items or indicators in construct j

For example, the composite score of Tourist Attractiveness was calculated from all indicators measuring the Tourist Attractiveness construct.

7. Preparation for SEM Analysis

The cleaned and coded dataset was prepared for SEM analysis using LISREL 10.2. The data were arranged according to the latent variables and their indicators.

The main latent variables were:

Latent Variable	Abbreviation	Description
Tourist Attractiveness	ToAt	Attractiveness of halal tourism destinations
Tourist Transportation	ToTr	Accessibility, mobility, and transport support for Muslim tourists
Tourist Infrastructure	ToIn	Physical facilities and supporting services for halal tourism
Sustainable Tourism Business Development	STBD	Sustainable development of halal tourism businesses

Each latent variable was measured through its respective observed indicators. The prepared dataset was then used to estimate the measurement model and structural model.

8. Treatment of Missing Data

Missing data were checked before analysis. Questionnaires with substantial missing values were excluded from the final dataset. For minor missing values, the response was reviewed against the original questionnaire record where available. If the missing response could not be verified, the case was excluded from the relevant analysis to maintain data accuracy.

9. Data Standardization

No substantive transformation that changed the meaning of the original responses was applied. The original Likert-scale scores from 1 to 5 were retained for analysis. The data were treated as ordinal responses suitable for descriptive analysis and as observed indicators in the SEM procedure.

Where required by the statistical software, the dataset was formatted according to LISREL input requirements, including variable naming, item ordering, and data matrix preparation.

10. Final Analytical Dataset

After screening, coding, cleaning, and tabulation, the final dataset consisted of **385 valid respondents**. These respondents represented six stakeholder groups involved in halal tourism destinations:

No.	Respondent Group	Final Sample
1	Tourism investors	9
2	Representatives of tourism ministries or state institutions implementing halal tourism policies	7
3	World halal tourism destination managers	6
4	Visitors who had visited halal tourism destinations within the last five years	334
5	Toll road managers related to world halal tourism destinations	5
6	Bus drivers who had transported tourists to halal tourism areas	24
	Total	385

The final dataset was used for descriptive statistics, validity and reliability testing, measurement model estimation, structural model estimation, and hypothesis testing.

D. Formulae Used for Modelling

This study employed **Structural Equation Modeling (SEM)** using **LISREL 10.2** to examine the causal relationships among Tourist Attractiveness (ToAt), Tourist Transportation (ToTr), Tourist Infrastructure (ToIn), and Sustainable Tourism Business Development (STBD). The dissertation states that the study used a quantitative approach with SEM-LISREL to analyse causal relationships among tourist attractiveness, tourist transportation, tourist infrastructure, and sustainable tourism business development.

1. Structural Model Formula

The structural model was used to estimate the causal relationships among latent variables. In general, the structural model can be written as follows:

$$\eta = \Gamma\xi + B\eta + \zeta$$

Where:

Symbol	Description
η	Endogenous latent variable
ξ	Exogenous latent variable
Γ	Path coefficient matrix from exogenous latent variables to endogenous latent variables
B	Path coefficient matrix among endogenous latent variables
ζ	Structural error term or disturbance

In this study, the structural relationships were formulated based on the hypothesised model involving ToAt, ToTr, ToIn, and STBD. The model tested direct, simultaneous, and indirect relationships among these variables, including the mediating role of ToIn. The hypothesis diagram also shows the relationships among ToAt, ToTr, ToIn, and STBD, including ToAt with ToTr, ToAt $\rightarrow$ ToIn, ToTr $\rightarrow$ ToIn, ToIn $\rightarrow$ STBD, and simultaneous effects on STBD.

The main structural equations can be written as follows:

$$\text{ToTr} = \gamma_1 \text{ToAt} + \zeta_1$$

$$\text{ToIn} = \gamma_2 \text{ToAt} + \gamma_3 \text{ToTr} + \zeta_2$$

$$\text{STBD} = \gamma_4 \text{ToAt} + \gamma_5 \text{ToTr} + \beta_1 \text{ToIn} + \zeta_3$$

Where:

Variable	Description
ToAt	Tourist Attractiveness
ToTr	Tourist Transportation

Variable	Description
ToIn	Tourist Infrastructure
STBD	Sustainable Tourism Business Development
$\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$	Direct path coefficients
β_1	Path coefficient from ToIn to STBD
$\zeta_1, \zeta_2, \zeta_3$	Error terms in the structural equations

Thus, the final model tested whether tourist attractiveness and tourist transportation influence tourist infrastructure and whether tourist attractiveness, tourist transportation, and tourist infrastructure jointly influence sustainable tourism business development. This is consistent with the model description in the dissertation, which explains that ToAt, ToTr, and ToIn simultaneously explain variation in STBD.

2. Measurement Model Formula

The measurement model was used to estimate the relationships between latent variables and their observed indicators. The dissertation explains that the measurement model represents the relationship between manifest variables and latent variables.

For exogenous constructs, the measurement model can be expressed as:

$$X = \Lambda_x \xi + \delta X$$

For endogenous constructs, the measurement model can be expressed as:

$$Y = \Lambda_y \eta + \varepsilon$$

Where:

Symbol	Description
X	Observed indicators of exogenous latent variables
Y	Observed indicators of endogenous latent variables
Λ_x	Factor loading matrix for exogenous indicators
Λ_y	Factor loading matrix for endogenous indicators
ξ	Exogenous latent variable
η	Endogenous latent variable
δ	Measurement error for exogenous indicators
ε	Measurement error for endogenous indicators

In the simplified form for each construct, the measurement equations can be written as:

$$ToAt_i = \lambda_i ToAt + \delta_i$$

$$ToTri = \lambda_i ToTr + \delta_i$$

$$ToIni = \lambda_i ToIn + \epsilon_i$$

$$STBDi = \lambda_i STBD + \epsilon_i$$

Where i represents each observed indicator of the corresponding latent construct.

3. Direct Effect Formula

The direct effect refers to the path coefficient between one latent variable and another latent variable. For example:

$$\text{Direct Effect} = \gamma_{xy}$$

Examples in this study include:

$$ToAt \rightarrow ToTr$$

$$ToAt \rightarrow ToIn$$

$$ToTr \rightarrow ToIn$$

$$ToIn \rightarrow STBD$$

$$ToAt \rightarrow STBD$$

$$ToTr \rightarrow STBD$$

4. Indirect Effect Formula

The indirect effect was used to examine the mediating role of Tourist Infrastructure (ToIn). The indirect effect is calculated by multiplying the path coefficient from the independent variable to the mediator by the path coefficient from the mediator to the dependent variable.

For the indirect effect of ToAt on STBD through ToIn:

$$\text{Indirect Effect } ToAt \rightarrow STBD = (ToAt \rightarrow ToIn) \times (ToIn \rightarrow STBD)$$

or:

$$IEToAt = \gamma_2 \times \beta_1$$

For the indirect effect of ToTr on STBD through ToIn:

$$\text{Indirect Effect ToTr} \rightarrow \text{STBD} = (\text{ToTr} \rightarrow \text{ToIn}) \times (\text{ToIn} \rightarrow \text{STBD})$$

or:

$$\text{IEToTr} = \gamma_3 \times \beta_1$$

The dissertation explains that ToIn functions as an important mediator in converting the potential of attractiveness and transportation into sustainable tourism business development.

5. Total Effect Formula

The total effect is the sum of the direct effect and indirect effect.

For ToAt on STBD:

$$\text{Total Effect ToAt} = \text{Direct Effect ToAt} \rightarrow \text{STBD} + \text{Indirect Effect ToAt} \rightarrow \text{ToIn} \rightarrow \text{STBD}$$

or:

$$\text{TEToAt} = \gamma_4 + (\gamma_2 \times \beta_1)$$

For ToTr on STBD:

$$\text{Total Effect ToTr} = \text{Direct Effect ToTr} \rightarrow \text{STBD} + \text{Indirect Effect ToTr} \rightarrow \text{ToIn} \rightarrow \text{STBD}$$

or:

$$\text{TEToTr} = \gamma_5 + (\gamma_3 \times \beta_1)$$

6. Coefficient of Determination

The coefficient of determination was used to explain the proportion of variance in an endogenous variable that can be explained by its predictor variables.

$$R^2 = 1 - \frac{\text{Var}(\zeta)}{\text{Var}(Y)}$$

Where:

Symbol	Description
R^2	Coefficient of determination

Symbol	Description
Var(ζ)	Variance of the residual or error term
Var(Y)	Total variance of the endogenous variable

In this study, R^2 was used to assess how much variation in ToTr, ToIn, and STBD could be explained by the predictor variables in the SEM model.

7. Model Identification Formula

Model identification was assessed using the degree of freedom formula:

$$df = \frac{1}{2}(p + q)(p + q + 1) - t$$

Where:

Symbol	Description
df	Degree of freedom
p	Number of exogenous manifest variables
q	Number of endogenous manifest variables
t	Number of estimated parameters

A model is considered **over-identified** when:

$$df > 0$$

The dissertation states that the model had a positive degree of freedom, meaning that the parameters in the model could be estimated using the collected data.