

RESEARCH QUESTIONS – BORKUM SURVEY

Here some inspiration for possible questions to initiate the discussion and selection process of research questions for your project.

Natural Language Processing (NLP)

1. **Topic Modeling:** What are the main topics that emerge from the open-ended responses about why visitors travel to Borkum?
2. **Keyword Extraction:** What are the most frequent keywords in the responses that describe visitors' understanding of sustainability?
3. **Text Classification:** Can we classify the text responses into categories such as positive or negative sentiments, complaints, praises, or suggestions?
4. **Etc.**

Clustering (Hierarchical, k-means)

1. **Visitor Segmentation:** Can we identify distinct clusters of visitors based on their travel habits, reasons for visiting, perceived health, demographic profiles? What are the distinct groups of visitors based on their connection to the island and eco-stress levels?
2. **Segmentation by Sustainability Behaviors:** How do different clusters of tourists vary in terms of their sustainable behaviors and attitudes during their stay? How do different clusters of tourists vary in terms of their sustainable practices at home?
3. **Segmentation by Knowledge, Intention and Behavior:** Can we find clusters of tourists with varying knowledge, intention and behavior regarding water consumption, plastics, and/or wadden sea?
4. **Etc.**

Classification / Prediction

1. **Predictive Modeling:** Can we predict a visitor's eco stress level based on their responses about sustainability knowledge, intention, or behavior?
2. **Visitor prediction:** Can we predict whether a visitor is a frequent visitor based on their demographics and travel preferences?
3. **Sustainable practice prediction:** Can we classify visitors based on their sustainability practices during their stay? Can we predict the impact of environmental awareness on visitors' behaviors and attitudes?
4. **Classification regarding sustainability:** Can we classify visitors regarding Closenes to Nature based on their survey responses regarding sustainability?
5. **Classification regarding environmental activism:** Can we classify visitors regarding their Environmental Activism based on other survey responses regarding sustainability practices or others?
6. **Classification regarding perceived health:** Can we classify visitors regarding their perceived Health based on other survey responses?
7. **Prediction of Sustainability Index:** Can we develop a model to predict a sustainability score of a visitor based on their detailed behaviors and attitudes as reported in the survey?
8. **Classification of Visitor Types:** Can we classify visitors to a visitor type (define value ranges of selected responses) and predict the type based on other variables?
9. **Etc.**