**Et billede, der indeholder Grafik, Font/skrifttype, rød, design

Automatisk genereret beskrivelseProject Plan**

**Master’s Thesis**

Data-Driven Decision Making for Sustainable Building Development

**Student:** Nicolaj Hans Nielsen

**Study no.:** s184335

**Supervisors:** Henrik Madsen and Justinas Smertinas

**External Partner:** Center Danmark

**Start:** 01/09-2023

**Deadline:** 20/01-2024

This document contains the proposed project plan for the thesis specified above.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Phases** | **Preliminary** | **Data Processing and Initial Modelling** | **Research** | **Implementation** | **Evaluation** | **Ultimate** |
| **Time** | Start: 01/09-2023  End: 14/09-2023  2 weeks. | Start: 14/09-2023  End: 30/09-2023  2 weeks. | Start: 30/9-2023  End: 30/11-2023  2 months. | Start: 15/10-2023  End: 22/12-2023  2 months. | Start: 17/11-2023  End: 22/12-2023  1 month. | Start: 02/01-2023  End: 14/01-2023  2 weeks. |
| **Milestones** | Formulate project plan and learning objectives.  Prepare the report with relevant sections in Overleaf.  Acquire sample data and get familiar with the type of data used in the field.  Get familiar with core concepts and literature on signature modelling. | Collect and preprocess more comprehensive data sets.  Formulate an initial model for the energy signature of a building.  Formulate an initial presentation about the project. | Presentation in Nationalbanken on 05/09-2023.  Write literature reviews.  Start writing the theory section.  Make small implementations for proof of concept in a Bayesian setting. | End-to-end implementation of the three modules; energy signature modelling, statistical clustering, and the decision support framework.  Construct a well-functioning codebase.  Finish the theory section. | Evaluate the results of each of the three modules.  Finish the empirical section.  Present progress, models, and performance at Nationalbanken. | Finish conclusion and abstract.  Revisit all sections in the report.  Final presentation at Nationalbanken before thesis defense. |
| **Activities** | Formulate learning objectives and finish a plan.  Set up GitHub and other systems to ensure a stable codebase.  Set up Overleaf to facilitate a fast start on the writing phase.  Acquire sample data.  Find papers, read abstracts, and start writing small sections to collect thoughts. | Acquire and work with larger data volumes and preprocess the data, including managing missing values.  A deep dive directly into formulating and testing models on data and documenting observations. These learnings enable me to understand model limitations, the models' interpretability, and the data's intricacies.  Prepare an initial project presentation at the ClimateCluster group at Danmark’s Nationalbank | Select and Read Relevant Papers on Energy Signature Modelling.  Select and Read Relevant Papers on Statistical Clustering.  Select and Read Relevant Papers and Reports on Climate Policy Scenarios and Existing Climate Navigation Tools.  Begin on the theory section.  Make Small Implementations for Proof of Concept in the Bayesian Setting. | The models and methods described during the research period should be implemented and tested on data.  The implementations are sequential by nature for the models, so the goal is to get a working model and then make it more advanced.  This phase will overlap with the research phase as models might not fit houses when working on larger and diverse data volumes.  The decision support framework can materialize in different formats, depending on inputs from Danmarks’s Nationalbank and stakeholders. | Compare the performance of each module. The focus is always on locating sources of uncertainty, e.g., uncertainty in input data and wrong energy labels, to make adequate and robust decisions in the last module.  Revise the theory section.  Write the empirical section.  Prepare presentations at Nationalbanken.  Find appropriate metrics to assess performance for each of the modules.  Work on graphical representation of model results to convey results. | Write concluding remarks and abstract.  Revise and read the report thoroughly and keep only the relevant sections in the main report.  Work on final presentations and the defense. |

**A screenshot of a project

Description automatically generated**

**Learning Objectives**

The following are the learning objectives for the project:

* Describe and motivate the need for data-driven methods in building performance estimation with focus on sustainable building development.
* Describe the fundamentals of energy signature modeling, highlighting its importance in data-driven characterizing the energy profile of buildings.
* Describe the advantages of employing Bayesian regressions for modeling energy signatures in building performance estimation. Elaborate on the practical implementation aspects, including the choice of sampling methods and their relevance in probabilistic programming languages like Stan.
* Implement several Bayesian regression in a probabilistic programming language, analyze model results, and evaluate bias-variance tradeoffs and interpretability.
* Explain how statistical clustering methods can be used to group houses and introduce learning approaches that utilize information from the subset of buildings with existing energy certificates.
* Implement a methodology to cluster houses based on their energy performance.
* Discuss similarities and discrepancies between model-assigned energy performance and the performance listed in the energy certificates. Address model uncertainty, data uncertainty, and errors in the performance certificates[[1]](#footnote-1).
* Compile possible economic and climate policy scenarios for the building industry and explain how compliance can be demonstrated. Introduce existing climate navigation tools like CREEM[[2]](#footnote-2) and how they can be employed.
* Implement a probabilistic decision support framework to suggest necessary enhancement initiatives to comply with climate policy and align with climate goals.
* Assess the effectiveness and robustness of the decision support framework and discuss the potential impact for sustainable building development.

1. The Danish Energy Agency. *Status for Energimærkningsordning for Bygninger*; Technical report; Energistyrelsen (The Danish Energy Agency): Copenhagen, Denmark, 2018. [↑](#footnote-ref-1)
2. https://www.crrem.eu [↑](#footnote-ref-2)