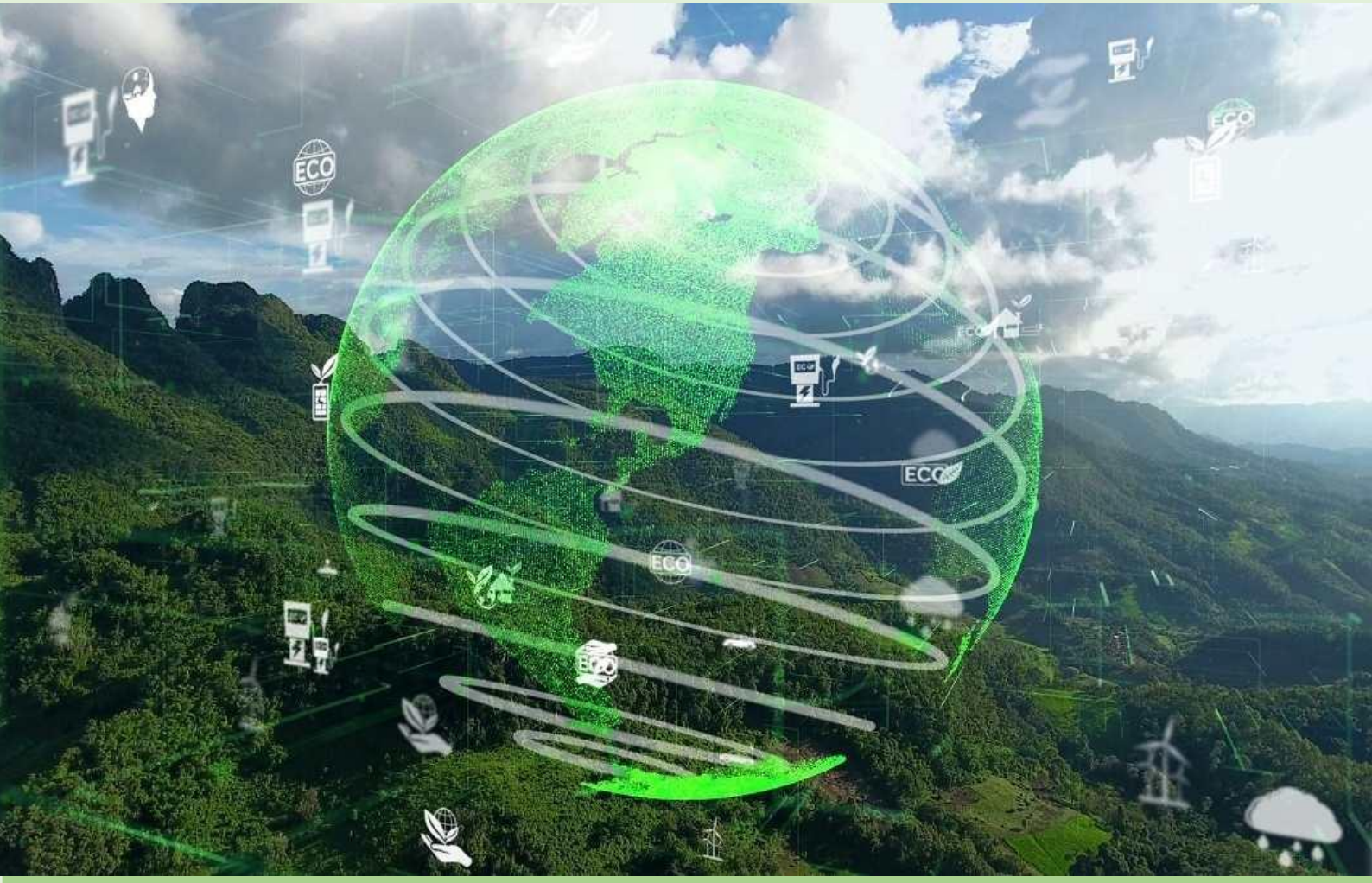
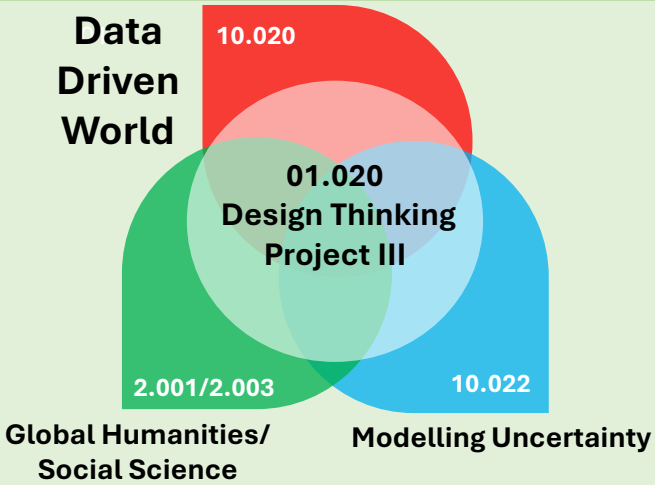
SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN | Science, Mathematics and Technology

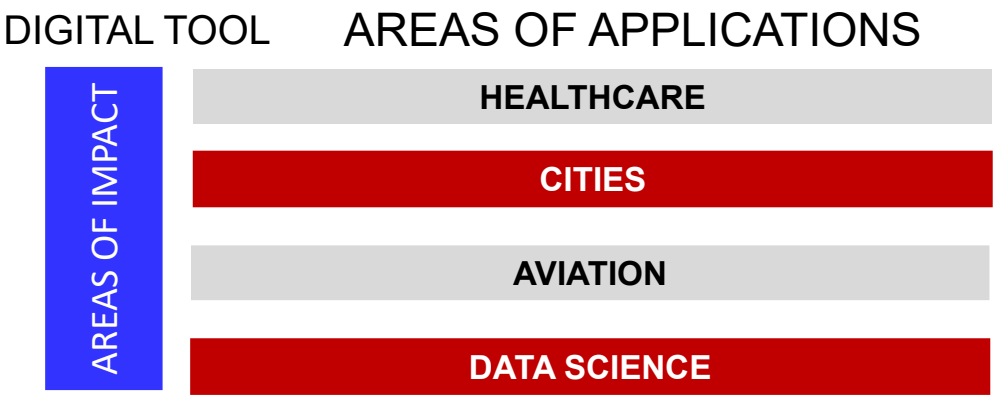
Modelling for Sustainability-Related Innovation and Solution using Excel and Python Programming Language



01.020 Design Thinking Project III
Design Brief



Progressive Learning



FACULTY TEAM

Bong Eng Ying	Project Lead
Fredy Tantri	Data Driven World
Sun Zeyu	Modelling Uncertainty
Zhao Zhenxing	Global Humanities
Pang Yang Huei	Social Science

ACKNOWLEDGMENT OF SUPPORT

Dario Poletti	Head of Cluster (SMT)
Lin Wuzhong	Office of Education
Rina Rin Binte Sodar	Office of SMT

1. Design Thinking Project Title

"Modelling for Sustainability-Related Innovation and Solution using Excel and Python Programming Language"

2. Design Thinking Project Learning Objectives

By the end of the project, students should be able to:

10.020 Data Driven World (DDW)	**10.022 Modelling Uncertainty (MU)**	**2.001 Global Humanities (GH)/ 2.003 Social Science (SS)**
1. Use plots and visualization to understand data. 2. Transform data for multiple linear regression. 3. Build multiple linear regression model. 4. Evaluate machine learning model using metrics evaluation.	1. Develop and evaluate simple probabilistic models for a variety of situations. 2. Examine data and use tools to visualize data and uncover relationships. 3. Build regression models and estimate their parameters	Apply self-reflection honed by the course when: 1. Identifying and analyzing the specific community of people you are hoping to serve and 2. Assessing the social implications of your design solution in terms of the ideals of equity, equality, and justice.

3. Design Thinking Project Grading

Each task will contribute 100% to the respective subject's DTP III (2D) Project component.

Subject	Tasks	Submission
2.001 Global Humanities/2.003 Social Science	Task 1	Oral Presentation
10.022 Modelling Uncertainty	Task 2	Report
10.020 Data Driven World	Task 3	Video and Jupyter Notebook

4. Design Thinking Project Overview

The Design Thinking Project III is a transdisciplinary project across three subjects, and this project is being led by the elective 10.020 Data Driven World. In this project, students are expected to cohesively apply knowledge learnt in 10.020 Data Driven World, 10.022 Modelling Uncertainty, and 2.001 Global Humanities or 2.003 Social Science.

In this project, students are given a specific scenario to an open-ended problem to design (discover, define, develop and delivery phases, from the UK Design Council Double Diamond design thinking framework) a solution by applying technical domain knowledge, skills, methods, and attitudes and behaviours (competencies) associated with design.

5. Design Thinking Project Scenario

In this design thinking scenario, the United Nations and World Health Organisation are forming a common taskforce together with industry leaders. As part of this fictitious taskforce, you will be exploring this **sustainability-related problem** in multidimensional ways and **building machine learning solution**, through what you have learned in Data-Driven World, Modelling Uncertainty, and Global Humanities/Social Sciece.

Sustainability has a very broad definition. It is the practice of maintaining processes, resources, and systems in a way that ensures long-term balance between environmental health, economic prosperity, and social well-being. In this project, you can propose your own sustainability-related topic or choose one of the topics below:

- Energy Consumption
- Climate Change
- Waste Management
- Carbon Footprint
- Agriculture
- Renewable Energy
- Air Quality
- Water Conservation and Management
- Biodiversity
- Culture
- Infrastructures (housing, transportation, cities)
- Public health

6. Design Thinking Project Tasks

UNDERSTANDING THE CHALLENGE & FRAMING THE TASK

Task 1 (Global Humanities/Social Science):

Lead-in

Your group should have a brainstorming session to canvass a series of real-world scenarios which are relevant to your STEM segments. You may use AI as your intellectual companion or other database empirical search methods (e.g. J-stor, EBSCO, Google Scholar) to further fine-tune your selections. Each student group is expected to select one viable scenario based on group consensus. Each group will have a unique scenario, no two groups will have similar scenarios.

The Next Step

Having selected the scenario, each group then refer to and engage with '**Pedagogical Scaffolding for Students' Presentation**' (refer to DTP-HASS grading criteria). At this main stage, it is essential to select the social science text/theorist to be used as a critical lens to critique your scenario. You may also use AI to test your ideas. To demonstrate your deep understanding of your selected theorist, judicious selection of short relevant quotes (page number indicated) is immensely useful.

Final Outcome

Please produce a set of PowerPoint slides for your presentation.

Make a short (10-minute) oral presentation during Week 12 in your recitations. All team members are required to present.
Refer to DTP-Hass Presentation Rubric for grading criteria.

N.B. A good presentation will refer to and engage with ideas found in the assigned and supplementary readings in a synthesizing way. The more deeply you engage, the stronger your work will be.

MODEL THE TECHNICAL DESIGN SOLUTION & EMBODIMENT OF THE TECHNICAL DESIGN

Tasks 2 (Modelling Uncertainty) and Task 3 (Data Driven World):

In the following description, Task 2 refers to Modelling Uncertainty deliverables, while Task 3 refers to Data Driven World deliverables.

Your goal here is to explore sustainability problem through the topics listed in the Section 5 above using *Multiple Linear Regression*, done in **both** *Excel* and *Python*. This task is open-ended. You are free to find and define a problem (apply the *discovery* and *define* phases first, from the UK Design Council Double Diamond Design Thinking Framework, introduced in 3.007 Design Thinking and Innovation) of your interest related to Sustainability. This project is restricted to use only linear regression model, not other models such as logistic regression.

In completing this task, it is important that you treat all these data within the frame of the cultural and social realities that characterise any given context. In other words, you must start with a clear understanding of the community that you are dealing with, keeping in mind that all the quantifiable data are the result of non-quantifiable elements (religious beliefs, cultural positions, traditional ways of doing things etc.). Additionally, you can make use of socio-economic indicators, geographical variables, etc. to add more dimensions to your analysis.

As a general guide, you may need to undertake the following actions:

- Identify a problem regarding sustainability that interests you, and which you want to solve using Multiple Linear Regression **(please check with your instructors first on whether the problem makes sense)**.
- Apply Design Thinking in scoping your problem statement (Define phase in the first diamond).
    - Brainstorm different ideas for the problem statement that you went through with your team and how you finally scope it down (*define phase*) it to the selected chosen problem statement. Read some brainstorming techniques[1].
    - Create *personas* to illustrate the pain points of the users in framing the problem statement. Read how you can create personas for data analytics tasks[2].
- Write your problem statement using "How might we" statement[3]. Since this is a Machine Learning task, your statement might be something like: 'how might we predict…'
- Find a dataset that is relevant to the problem you are solving. Ideally, the data should cover a diverse range of regions, time periods, or categories to ensure the model captures meaningful patterns. Aim for at least 30 distinct data points or samples in each category to support reliable analysis—though more data is generally better.". You are

[1] https://www.sessionlab.com/blog/brainstorming-techniques/
[2] https://towardsdatascience.com/how-data-driven-companies-build-customer-personas-c1559d82a1d4
[3] https://conceptboard.com/blog/how-might-we-template/#How_Might_We_questions

free to choose which datasets to use and from multiple resources. Many useful datasets can be found in, for instance, https://www.worldbank.org, https://data.cdc.gov/, https://data.gov/, https://www.kaggle.com/datasets, etc.

- At this juncture, you will need to approach the deliverable for Modelling Uncertainty and Data Driven World deliverables differently.
  - **Task 2** (Modelling Uncertainty) guidelines:
    - Identify the response (dependent) variable, and research for appropriate predictor variables – some irrelevant predictor variables may be dropped from your dataset, some predictor variables may require you to look up more data.
    - Build a multiple linear regression model, and measure the accuracy of your model, using an appropriately chosen metric (hint: r^2 is *not* a good metric for this task).
    - If needed, improve your model by incorporating other predictors, and/or removing existing ones.
    - Discuss your data sets, model, accuracy, what metrics you used to judge the accuracy, as well as any insights, predictions or recommendations you can draw from your model.
    - For this task, **you may only use *Excel*,** and not Python.
  - **Task 3 (**Data Driven World) guidelines:
    - You must use Python with Jupyter notebook to do this task.
    - Use charts to visualize, summarize and understand your data.
    - Create training and test data sets.
    - Build linear regression model.
    - Choose an appropriate metric to evaluate your model (you may use the same metric as the one used in Modelling Uncertainty part).
    - Improve your model.
    - The following technical/tool constraint applies: you are **NOT** allowed to use Neural Networks or other Machine Learning models from any library. You **must use Python** and Jupyter Notebook.
    - You must use the functions that you created in your problem sets. If you require other additional functions not done in the problem sets, you can consider using those functions provided by scikit-learn. Do consult with your instructor first.

7. Design Thinking Project Deliverables

Important Timeline

Week 5 – 7	Literature research, brainstorm on design options, and frame the task.
Week 9 – 11	Finalize your design and prepare for the presentations and report.
Week 12 (DTP week)	Task 1 GH/SS presentation (Cohort 2 session) Task 2 MU report (Friday, 8 Aug, 11 pm) Task 3 DDW video presentation and report (Thursday, 7 Aug, 9 am)
Week 13	**DTP III Showcase at Campus Centre** (15 Aug, 2 – 4 pm) Shortlisted projects will be invited to present. Prizes will be awarded to the top 3 projects.

You can ask and post your questions on Ed Discussion (Category 2D) platform about this project. Alternatively, you can approach your instructor for consultation.

Task 1 (Global Humanities/Social Science):
Produce a 10-min oral presentation to demonstrate the ability to self-reflect and practice accountability for your engineering project. Discuss your slides in a short 10-minute oral presentation during **Week 12 in your recitation section.** All team members are required to present. Submit your presentation in either PDF or PPT (if video is used) format on eDimension **under '2025 May Term - 01.020: Design Thinking Project III'** course.

Tasks 2 (Modelling Uncertainty):
Submit your work through eDimension at the 01.020 Design Thinking Project III page. The submission should consist of a PDF report limited to 5 pages, and an implementation of the final multiple linear regression model in *Excel*. The *Excel* spreadsheet needs to be annotated, interactive (i.e. contains formulas, so that changing the data input results in different output), and contain the raw datasets used. The report is due on Friday of week 12, by 11:00pm, on eDimension at 01.020 Design Thinking Project III page.

Task 3 (Data Driven World):
Submit your work for through ***Vocareum Assignment*** and ***eDimension*** submission before **9AM Thursday, 7th August 2025**. The submission should use the template provided in Vocareum and must contain the following:

- 7-minutes (max) video presentation. This is to be submitted on eDimension. The video should emphasize on general overview of your work: clearly explain your problem statement, how the data is prepared, analysis and notable insights from the data, how you build the model, model evaluation results, and conclusions, without going too much detail about the code line by line. The video cannot be fast forwarded.
- All dataset files.
- In addition, the Jupyter notebook should contain the following (refer to the given template in Vocareum for more details):
    - Names of the group members, and a short sentence summarizing each member's contribution.
    - Problem Statement
        - Background description of the problem.
        - User Persona.
        - Problem Statement using "how might we ..." statement.
    - Data
        - Link to data sources.
        - Discussion on why these data were used
        - Steps to clean the data and their Python codes.
        - Analyse the data using various statistical measures and meaningful visualizations.
        - Discussion on the predictors or features.
    - Model
        - Discussion on building the linear regression model and their Python codes.
        - Any other relevant mathematical details.
    - Model Evaluation

- Discussion on the chosen metric to evaluate your final model, how that model compares with your earlier attempts, any mathematical details, and all resources consulted.
- Code and results of the metric on your test data set.

- Discussion of the result
  - Discussion of your code and steps taken to improve the accuracy of your model

7.1 Peer Contribution

The Design Thinking Project III is a team collaborative effort where all members are expected to contribute equally towards the project using the Design Thinking Process (Discover, Define, Develop, and Deliver phases, *UK Design Council's Double Diamond Design Thinking Framework*).

In your Modelling Uncertainty report (Task 2), include a Peer Contribution section describing each member's contributions to the overall project, covering all three tasks (Task 1, 2 and 3).

This section is mandatory.

8. Design Thinking Project Budget

No budget is allocated for this project.

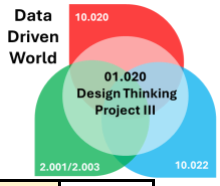
9. Design Thinking Project Rubrics

Rubrics provide a set of expectations for the students and serve as a guide for accomplishing the design thinking project

01.020 Design Thinking Project III

10.020 Data Driven World – Rubrics for Video and Jupyter Notebook

Criteria (Ability)	Design Skills	Design deliverables (Expectations for Student Score)			
		Below Exp. (0-1)	Meets Exp. (2-3)	Exceeds Exp. (4-5)	Score
Problem Statement (Define the Challenge)	- Problem scoping. - Definition of the challenge. - Formulate the design opportunity	The problem statement is incoherent.	The problem statement is lacking one or more of the following: clarity in its definition, relevance, and the design opportunity.	The problem statement clearly defines the problem, its relevance, and the design opportunity.	
Dataset Choice and Data Preparation (Develop a solution)	- Analysing & interpretating quantitative data - Identifying patterns & relationships in data. - Making well-informed decision. - Constructing visual representations of data.	Students did not choose an appropriate dataset. The number of data is lacking. The dataset is not cleaned up. The proportion of the training and the test set are not appropriate. Data is not described nor visualized.	Students chose an appropriate dataset. Students may not do the necessary clean up or preparation and just use the data as it is. Students may not split the data in the right proportion. Data are visualized.	Students chose an appropriate dataset with a suitable amount of data. Students did the necessary clean up and preparation. Students split the data for training and test in a right proportion. Description of the dataset is clear and the sources are properly cited. Data are visualized with meaningful insights.	
Model and Evaluation (Model the solution)	Identifying the functionalities of the model. - Building the model. - Modelling data.	Students do not build the model properly nor use appropriate metrics. No discussion or analysis is found.	Students build the model properly and use appropriate metrics to evaluate the model. The discussion and analysis on the model and the metrics are lacking.	Students build the model properly and use appropriate metrics to evaluate the model. Students discuss the outcome and model evaluation in great depth.	
Improvements (Iteration of solution)	- Iterating. - Enhancing the solution. - Conforming with technical requirements.	Students do not attempt to improve the model's accuracy.	Students show modest improvement to the accuracy of the model through a better data transformation	Students are able to show significant improvements in the accuracy of the model through improved model building and/or data transformation. Students are able to discuss why the model performance is improved.	
Durability of code (Sustainability)	- Communicating the code. - Applying modularity. - Documenting the process.	Students do not write code in a modular way. The code is not documented or properly commented. The code is difficult to understand.	Students write code in a modular way. The code is documented and commented.	Students write code in a modular way. The code is properly documented and commented. The code can be easily understood by other programmers.	
				Total Score (Max 25 pts)	
Web app (BONUS) Total score for the bonus component (maximum of 6 points) to contribute to a maximum of 31 total points.					
Criteria	Design Skills	Below Exp. (0)	Meets Exp. (1)	Exceeds Exp (2)	Score



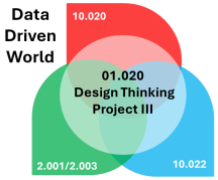
Functionality (Prototype the solution)	- Digital prototyping. - Adding functions to prototype - User testing.	Web app only has informational contents with no prediction function. Web app did not run in Vocareum.	Web app has a function to make prediction using the ML model developed in the design task. Web app run in Vocareum.	Web app has a function to make prediction using the ML model developed in the design task. The function has an appropriate data as its input. The predicted output solves the particular problem that the group is trying to solve. Web app run in Vocareum.	
UI/UX (Prototyping the solution)	- Designing user experience - Designing user interaction.	The input and output are not appropriate. The web app is difficult to use.	The prediction function in the Web app has an appropriate input and output elements.	The prediction function in the Web app has an appropriate input and output elements. It is intuitive and user-friendly. Its input data is something that user can easily enter. The output provides insight to the particular problem and easy to digest.	
Coding Style (Prototyping the solution)	- Documenting the process. - Optimizing functionalities.	The code is not written in a modular way. Many code repetition. Documentation is difficult to understand or very little. There is no README file. Students do not use Streamlit framework but other frameworks that they learn outside of the course.	The code is written in a modular way and contains documentation on how to deploy and use the app. Contains a README file. Students use Streamlit framework as taught through the 1D Mini Project and make use of the template given.	The code is written in a modular way and contains documentation on how to deploy and use the app. It follows PEP8 convention. Documentation is well written for every functions, classes and modules. Contains a README that gives all the necessary information to understand the project structure and to run the web app. Students use Streamlit framework as taught through the 1D Mini Project.	
				Total Score after Bonus (Max 31 pts)	

01.020 Design Thinking Project III

10.022 Modelling Uncertainty – Rubrics for Report

Criteria (Ability)	Design Skills	Design deliverables (Expectations for Student Score)			
		Below Exp. (0-2)	Meets Exp. (3-4)	Exceeds Exp. (5)	Score
Dataset Acquisition and Data Preparation (Develop Solutions)	- Analysing & interpretating quantitative data -Identifying patterns & relationships in data. - Making well-informed decision. - Constructing visual representations of data.	Students did not choose an appropriate dataset. The number of data is lacking. The dataset is not cleaned up. Data is not described nor visualized.	Students chose an appropriate dataset. Students may not do the necessary clean up or preparation and just use the data as it is. Data are visualized.	Students chose an appropriate dataset with a suitable amount of data. Students did the necessary clean up and preparation. description of the dataset is clear and the sources are properly cited. Data are visualized and analyzed to give superior understanding.	
Model (Model the Solution)	- Identifying the functionalities of the model. - Selecting the correct model. - Developing metrics.	Students do not use an appropriate model nor appropriate metrics. No discussion or analysis is found.	Students use an appropriate model for the data and appropriate metrics to evaluate the model. The discussion and analysis on the model and the metrics are lacking.	Students use an appropriate model for the data and appropriate metrics to evaluate the model. Students discuss and analyse the model and the metrics in great depth.	
Improvements (Iteration of the solution)	- Optimizing a model.	Students do not attempt to improve the model's accuracy.	Students show modest improvement to the accuracy of the model through some modification to the parameters or the dataset.	Students show significant improvements in the accuracy of the model by using various fine tuning or better data preparation. Students are able to discuss the improvement gained.	
				Total Score (Max 15 pts)	


Data
Driven
World
10.020
01.020
Design Thinking
Project III
2.001/2.003
10.022



01.020 Design Thinking Project III

2.001 Global Humanities / 2.003 Social Science – Rubrics for Presentation

Criterion (Weightage)	Design Skills	Presentation Deliverables (Expectations for Student Score)			
		Below Exp. (0-2)	Meets Exp. (3-4)	Exceeds Exp. (5)	Score
Understanding of Social and Cultural Context (5 points)	Qualitative Research Cultural Sensitivity	Little or no research has been done. Sources of information are poorly chosen and/or are unreferenced. Facts are absent or inaccurate. Information about community needs is unspecific and inappropriate.	Research has been done sincerely, but information conveyed is somewhat common- place, general, or vague. Information about community needs is specific and mostly appropriate.	Extensive research has been done. information is thorough and accurate, is taken from valid/vetted sources, with references as appropriate. Information about community needs is very specific, well-thought through, and extremely appropriate.	
Application of Social Scientific Concepts (5 points)	Conceptual Thinking Critical Thinking	Little or no reference to social scientific concepts. Concepts are applied randomly, inappropriately, and/or without justification.	Some sincere reference is made to social scientific concepts, but their application is sometimes inaccurate, inappropriate, or unclear, or made with limited justification.	Social scientific concepts are described accurately and applied appropriately to the details of the scenario. Approach is very clearly justified.	
Logic and Analysis (5 points)	Sense Making Decision Making Problem Solving	Links between components of DTP project are weak, unclear. Does not consider the connections between the social and technical dimensions of the design solution.	Links between DTP components are expressed but not extensively. Considers connection between social and technical dimensions of the design, but these are general or vague.	Draws clear links to other dimensions of the DTP project. Connections between the social and technical dimensions of the design solution are well-reasoned and clearly expressed.	
Presentation Style and Delivery (5 points)	Communication Creative Thinking	Oral presentation was disorganized, unclear, and unconvincing. Accompanying materials are dull, boring, or uninspired.	Oral presentation was organized and clear. All group members contributed to the presentation. Presentation was good, but accompanying materials were poor (or vice versa). Poster/slides were generic.	Oral presentation is very well organized and clear. All group members contributed equally. Presentation and accompanying visual materials were both very good. Design of poster/slides was unique, compelling, and visually appealing.	
				Total Score (Max 20 pts)	