WIH3001 – DATA SCIENCE PROJECT

SEMESTER 1, 2022/2023

---

3 credit hours

Pre-Requisite: Pass all Faculty and Program Core courses except for Industrial Training.

Learning Outcomes:

1. Define the problem background
2. Determine the objective of the project
3. Identify suitable solution approach for the stated problem.
4. Review literatures relevant to the stated problem
5. Conduct data gathering using suitable techniques
6. Develop a prototype of the proposed solution
7. Write a project report.

Synopsis of Course Content

This course covers the following research activities including problem and objectives identification, literature review, data collection, prototype development, report writing and project presentation.

Assessment Method

Continuous Assessment 100%

Individual project

Supervisor – 60%, Panel – 40%

Close supervision (minimum 5 meetings)

Panels (3 person)

WIH3001 – Data Science Project Timeline

Week	Suggested Steps (Guideline)
Week 1 17-10-2022	*Briefing* **Appointment of Supervisor** • Identify the project supervisor (max 3 / lecturer)
Week 2 24-10-2022	**Define Domain / Topic of Interest** • Project Idea Generation facilitate by Supervisor
Week 3 31-10-2022	**Business / Problem Understanding** • Ask relevant questions • State project objectives • Identify required data
Week 4 7-11-2022	**Data Acquisition (Find & Get Data)** • Gather or scrape the data needed for the project • Project proposal (10%) due for supervisor to check.
Week 5 14-11-2022	**Data Preparation** • Clean the data, fix inconsistencies in data & handle missing values
Week 6 21-11-2022	**Data Preparation** • Continue data cleaning, data tidying, etc.
Week 7 28-11-2022	**Exploratory Data Analysis** • Summarize the data main characteristics • Plot the data to see patterns or anomalies
5-12-2022	**Semester Break**
Week 8 12-12-2022	**Analysis / Modelling** • Approach based on the nature of problem to be tackled.
Week 9 19-12-2022	**Analysis / Modelling** • Approach based on the nature of problem to be tackled.
Week 10 26-12-2022	**Evaluation & Interpretation** • Validating the analysis or the model • Sense making results, insights and findings.
Week 11 2-1-2023	**Deployment** • Creating the data product / application
Week 12	**Deployment**

9-1-2023	• Creating the data product / application
Week 13 16-1-2023	**Communicating Results** • Presentation and Report Preparation
Week 14 23-1-2023	**Project Presentation** (schedule will be given) **Report Submission** (template for report will be provided)

Assessment / Marking Score Sheet (Supervisor)

A. Data Science Project Proposal (10%)	1	2	3	4	5
1) Title of Project					
2) Introduction / background					
3) Problem statement					
4) Objectives					
5) DS Methodology					
6) References					
TOTAL	/ 30				

B. Data Science Project Report (20%)	1	2	3	4	5
1) Abstract					
2) Introduction					
3) Literature Review / Previous Research (minimum 10 references)					
4) Problem Statement					
5) Model or Methods					
6) Results					
7) Discussion					
8) Conclusion, limitation, and future directions					

TOTAL	/ 40

C. Data Product Presentation (20%)	1	2	3	4	5
1) Suitable catchy name for the apps and the availability of link to apps (i.e. shinyapps.io) in the report / presentation slides & Presentation host (i.e., if R on Rpubs and done in RStudio Presenter or Slidify)					
2) Demonstration of the data product					
3) Associated supporting documentation (user manual)					
4) Functionality of data product, ease of use and attractiveness					
5) Identification of potential stakeholders & usefulness of the product to them.					
6) Informative and engaging slides materials					
7) Incorporate interesting data story with visualization					
8) Ability to answer and defend during Q & A session.					
TOTAL	/ 40				

D. Supervisee Conduct (10%)	1	2	3	4	5
1) Student interest and contribution of idea in the project					
2) Effort and resourcefulness					
3) Attitude					
4) Attendance					
5) Follow the checklist (instructions) given for the project					
TOTAL	/ 25				

Assessment / Marking Score Sheet (Panel)

Data Science Project Presentation / Pitching (40%)	1	2	3	4	5
1) Introduction & Problem Statement					
2) Objectives					
3) Data Science Methodology					
4) Analysis & Modelling Work					
5) Tools and coding					
6) Demonstration of the apps - Functionality of apps, usefulness etc.					
7) Interesting data story with visualization					
8) Ability to answer and defend during Q & A session.					
TOTAL	/ 40				

Outline for Data Science Project Report

- Conference paper format (Samples to be given)
8 -10 pages (including references) due Week 15 (30 January 2023)

Abstract (summary of the report)

Introduction – a gentle introduction to the subject matter and written in easily understandable format.

Literature Review

Data Section - Include written descriptions of data and follow with relevant spreadsheets or variables in the data source.

Data Science Project Methodology Section — Contains research methods and data sources used in the report.

Analysis / Modelling Section — explain what we analysed. Include any charts here.

Result Section followed by a discussion section. Describe the results of your analysis in the result section. The discussion section is where you rely on the power of narrative to enable numbers to communicate the thesis to your readers.

Conclusion, Limitation & Future direction: Concluding the overall report, outcomes of the report, application of the research in future. Identify limitation(s) to your work, and suggest appropriate future directions.

References and acknowledgement section. **Appendix** is optional