

SIM DSA 2025

INTERNAL

PROJECT

OBJECTIVES

- DEVELOP STRONG TECHNICAL CAPABILITY BY EXECUTING AN END-TO-END DATA SCIENCE PROJECT.
- DEMONSTRATE READINESS FOR INDUSTRY PROJECT THROUGH TEAMWORK, CLEAR COMMUNICATION, AND THE ABILITY TO JUSTIFY YOUR MODEL CHOICES, RESULTS, AND THE ABILITY TO TRANSLATE FINDINGS THAT BRINGS VALUE.

TIMELINE

- 18 NOV (TUE) - PROJECT COMMENCEMENT
- 25 NOV, 27 NOV - LAST TWO LESSONS
- 20 NOV, 2 DEC - MANDATORY NETWORKING EVENTS
- 4 DEC, 9 DEC - VENUE AVAILABLE, OPTIONAL
- **10 DEC 2359 (WED)** - SUBMISSION OF PRESENTATION SLIDE DECK AND CODE SCRIPTS ON GOOGLE DRIVE
- **11 DEC (THU)** - FINAL PRESENTATION (8 MINUTES MAX PER GROUP)

YOUR TASK

YOU ARE REQUIRED TO PLAN, DEVELOP, AND PRESENT AN END-TO-END DATA SCIENCE PROJECT USING YOUR ALLOCATED DATASET.

YOUR WORKFLOW MUST INCLUDE:

1. DATASET EXPLORATION & UNDERSTANDING

- Perform exploratory data analysis (EDA) to understand the structure, features, patterns, and challenges of the data.

2. PROBLEM DEFINITION & SCOPING

- Based on your EDA, define a clear problem statement (e.g., classification, regression, image recognition, anomaly detection).
- Justify why you chose this problem and what success looks like.

YOUR TASK

3. MODEL DEVELOPMENT

- Build a complete ML or deep learning pipeline (data cleaning, feature engineering, modeling, evaluation).
- Use appropriate metrics to evaluate model performance and iterate to improve it.

4. INSIGHTS & INTERPRETATION

- Explain your findings, model results, key drivers, limitations, and trade-offs.

5. FINAL PRESENTATION

- Deliver a polished, structured presentation that communicates your workflow, decisions, results, and key takeaways.

SKILLSETS

- YOU MUST DEMONSTRATE AT LEAST 3 OF THESE EXPECTED SKILLSETS IN YOUR CODE WORKFLOW AND PROJECT PRESENTATION!
- EXPECTED SKILLSETS:
 - PYTHON/SQL
 - DATA VISUALISATION
 - MACHINE LEARNING AND DEEP LEARNING
 - GENAI AND RAG (OPTIONAL BUT PREFERRED)
 - AI AGENTS (OPTIONAL)

GROUPINGS

- PLEASE REFER TO THIS TABLE FOR THE GROUPINGS.

DSA INTERNAL PROJECTS 2025					
GROUP 1	GROUP 2	GROUP 3	GROUP 4	GROUP 5	GROUP 6
Sammi Chua Ying Jie	Shaista Muskan	Tanishqa Sachin Akude	Janelle Ho Rui Yee	Leng Ji Herng	Pay Meow Lin Zara
San Sok Heng	Tay Jing Lin, Cheryl	Ang Lija	Uzma Husna Binte Nordin	Siah Ching Loong	Roshan Ashok Ramchandani
Mitch Lian Jun Yu	Thaddeus Ng Kai Qi	Nadon Panwong	Avneet Kaur	Sivaraman Leika	Venson Tok Zing Yung
Dhirana Sundaram	Tan De Xun	Long Xingxu	Muthurajan Koshinamrit	Lim Rong Jun	Joel Yeo
Ramkumar Sakthi Maheshwari	Yamini Karlmarx	Ryan Tan	Andrew Jiang Wenyao	Harneet Kaur Pannu	Jessica Noellyn Japara
			Muhamad Raid Ramlee		Kiefer Russel Sulijanto

- NOMINATE A LEADER FOR YOUR GROUP AND CREATE A GROUP CHAT.

DATASET

- ACCESS THIS DRIVE LINK TO ACCESS THE DATASETS.
- LEADER TO TEXT JULIANA THE TEAM'S TOP 3 DATASET CHOICES. IT WILL BE ASSIGNED FCFS BEFORE GOING HOME.

-  Group []: Brain Tumor Images (Ching Xi)
-  Group []: Car Prices (Ching Xi)
-  Group []: Credit Card Transactions (Juliana)
-  Group []: Deepfake Images (Ching Xi)
-  Group []: Dog Emotion Images (Juliana)
-  Group []: Garbage Images (Michelle)
-  Group []: USA Real Estate (Michelle)

SUBMISSION

- SUBMIT YOUR WELL-DOCUMENTED CODE SCRIPTS, SLIDE DECK, DASHBOARDS (IF ANY) IN THE SAME FOLDER ON 10 DEC (WED) 2359.
- FEEL FREE TO MESSAGE YOUR MENTOR FOR ANY QUESTIONS OR PROBLEMS.

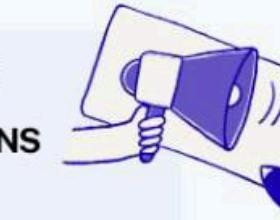
ALL THE BEST!

A FIRESIDE CHAT BY SIM DAC X SCS4YOUTH

DEMYSTIFYING *Data Careers*

JOIN US FOR AN EXCLUSIVE EVENT IN COLLABORATION WITH THE
SINGAPORE COMPUTER SOCIETY (SCS)!

- PANEL TALK WITH DATA PROFESSIONALS FROM TOP ROLES!
- UNDERSTAND WHAT EACH DATA CAREER PATH REALLY MEANS
- LEARN WHICH SKILLS AND TOOLS MAKE YOU STAND OUT
- BUILD CONNECTIONS WITH MENTORS AND THE DATA COMMUNITY



20TH NOV 2025 (THURSDAY) 1830-2130 SIM B.2.17

Meet our panelists from SCS!

Ruven



Data Analytics
Solutions Engineer
@IBM

Yuxin Seow



Aland Data Delivery
Management
@Temus

Daniel Carlier



Manager, Strategy &
Consulting -
Data & AI @ Accenture

Joey Chua



Principal Data
Scientist at the
National University
Health System
(NUHS)

Community partner



Co-Hosted by



The Un-Programmed Mind: Thinking Beyond the Algorithm



Zhi Yi
Chang

Product lead @AsiaVerify
Ex TikTok Product
Marketing intern



Benjamin
Goh

Talent Acquisition
Specialist Cloud
computing &
Infrastructure @ TikTok



Prasanna
Loganathar

Web3, blockchain tech
& distributed systems
Entrepreneur,
Ex Googler

6:30 – 7:20 PM	Networking Dinner	7:20 – 8:20 PM	Fireside Chat: How to Stand out & Get Hired	8:20 – 9:45 PM	Interactive Workshops: Elevator Pitch & Resume	9:45 – 10:00 PM	Personal Branding: AI Resume Builder & Digital Business Cards
----------------	-------------------	----------------	---	----------------	--	-----------------	---

2nd December, Tuesday, 6:30 PM to 10PM

Location: SIM Global Education

Join Us