Disclaimer

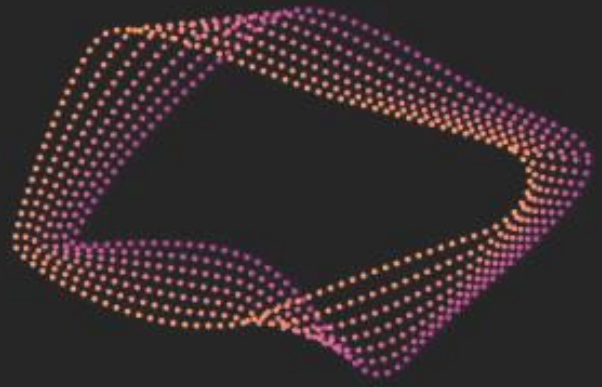
The various initiatives as listed in this presentation are **only applicable to public officers serving the Singapore Government**.

Content is updated and accurate as of 6th June 2018

Click here to view the video recording
or visit GovTech home page> Resources Folder > locate "11. Tableau Public Sector Day (Post Event Resources)"

GOVTECH SINGAPORE

Growing Data Science & AI Capabilities Across WOG



GOVTECH
SINGAPORE

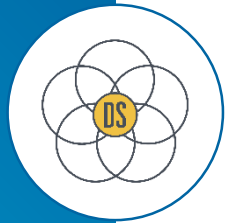
Liu Feng-Yuan, Director
Data Science and Artificial Intelligence Division

Objective

- This presentation is to highlight the strategic intent and plans on **broadening data science and artificial intelligence capabilities across Whole-of-Government (WOG)**.

- Various initiatives in:
  - Frameworks
  - Training roadmap
  - Satellite teams
  - Community

Overview



Broad Strategies

GOVTECH SINGAPORE

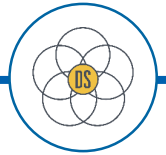
Broad Strategies

- Digitalisation will enable the Government to serve the public better, and to enable policies and services that are better designed

- Important to step up digital and data capabilities in our public officers. Data science training targets to achieve this.

4 Key Pillars Of WOG Capability Development

GOVTECH SINGAPORE



Framework

1. Establish **competency framework** and **training roadmap**.

Training

2. **Develop, curate and deliver training** options; **Aggregate training demand** from agency.

Satellite

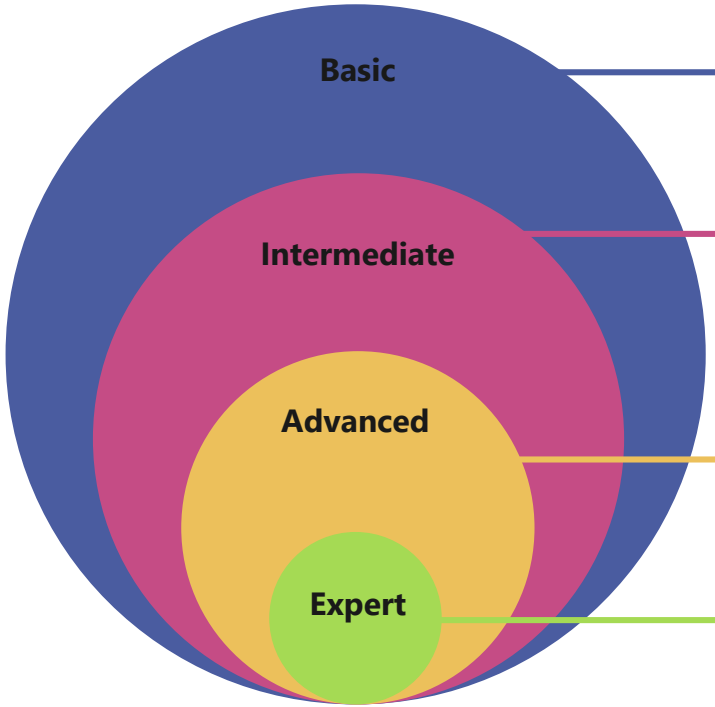
3. Assist agency in **setting up a Data Science team**; **Forward-deployment** of Data Scientists.

Community

4. **Engage enthusiasts and practitioners across WOG** through the **Data Science Community**.

GOVTECH SINGAPORE

Grow Broad and Deep capabilities



Basic Data Literate Officers

Power Users

Ministry/ Agency Core Teams

Cap Centre Data Scientists

Broad-based: Full-time job <u>is not</u> analytics work

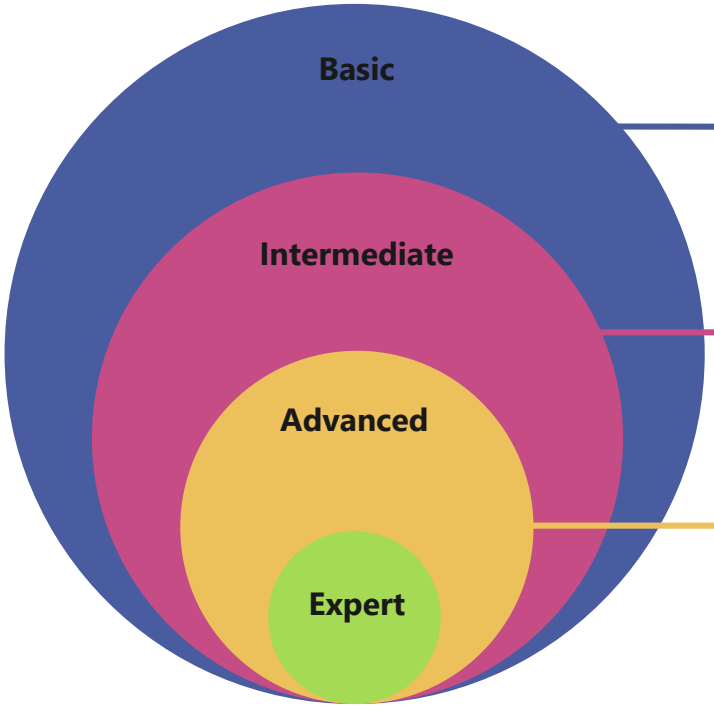
- Public officers with **basic data literacy** and conceptual understanding.
- Data literate public officers with **some technical hands-on** ability.

Deeper Capabilities: Full-time job <u>is</u> analytics work

- Data Science Practitioners whose full-time job role is analytics work, **embedded within the agencies with domain-specific skills**
- Multi-disciplinary experts housed at **GovTech** with broad range of expertise to solve complex, cross-agency problems, and always at the forefront of applied data science techniques, tools.

GOVTECH SINGAPORE

Map out relevant training for various Proficiency Tiers and Job Roles



Job Role	Actionables	Training required
Manager	Drive culture, originate projects	Understand concepts/myths, apply to business problems
Policy / Ops Analyst	Data exploration & modelling	Data visualization, data cleaning, simple techniques
Data Scientist	Full range of techniques	Advanced Machine Learning, e.g. deep learning

Support Agencies At Different Maturity Stages

Agency's maturity along the Data Science journey

#1 

#2 

#3 

COMMUNITY

- Regardless of maturity, to **engage enthusiasts and practitioners** in the agencies to be part of the wider WOG Data Science community

AGENCY TAKEOFF

- **Exploration & discovery** of projects
- **Alignment with staff and management** on Data Science potential through sharing, workshop sessions.

EXECUTION & SUPPORT

- GovTech to **help execute on** specific agency **projects.**
- **Train data literate public officers**, and a **pool of power users** to support the data science projects

SATELLITE DEVELOPMENT

- **Level up agency's own practitioners** through apprenticeships, project collaborations.
- **Forward-deploy data scientists on-site** to scope and deliver data science projects

Workplace community

DS Connect meet-ups

Primer sessions

Project scoping workshops

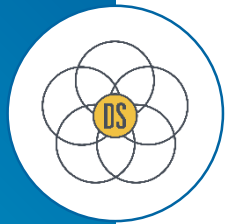
Partnerships with GovTech

Ad-hoc consultancy

Retainer / Satellite teams

GOVTECH SINGAPORE

Framework



1. Establish **competency framework** and **training roadmap**.

Assist agencies in developing training plans

Training should be relevant to job role

Structured way to choose relevant courses based on officer's job role

GOVTECH SINGAPORE

Competency Framework

Outlines required competencies and performable tasks
Assists HR in hiring and identifying training needs

Competency Groups

Data Governance

Project Scoping / Management

Mathematics

Data Integration

Data Systems Engineering

Data Exploration & Analysis

Data Visualisation & Comms

Software Design / Development

Competency Categories

- Policy and Process
- Data Handling & Ethics
- Problem Statement Formulation
- Project and Stakeholder Management
- Algebra
- Probability and Statistics
- Data Collection
- Data Quality
- Data Systems Architecture
- Data Pipeline Setup
- Exploration and Analysis
- Machine Learning
- Data Visualisation Principles
- Data Visualisation Development
- Programming Concept & Construct
- Systems Integration
- Software Architecture

GOVTECH SINGAPORE

Defining A Data Literate Public Officer

Data Quality

- **Understand** the various data types (e.g. number, date, string), underlying representation in common locally stored data formats (e.g. CSV, Excel).
- **Use** common features of office productivity tools (e.g. Excel) to manipulate and clean up the data.
- **Understand** the impact of data cleansing (e.g. removal of duplicate entries, removal of incomplete entries) to the overall analysis.

Proficiency	Foundation Level I
Role	Public Officer (Data Literate)
Competency Category	Competencies
Policy and Process	Understand basic considerations of data governance policies.
Data Handling & Ethics	Understand the need to exercise sensitivity and protect as appropriate the confidentiality of stakeholder's data and information. Understand the need for objectivity in the data analysis without undue pre-conceived biasness. Understand the possibilities of misrepresentations of analysis and insights due to distortion of facts and figures.
Problem Statement Formulation	Understand some of the possibilities and limitations of what data can and cannot do for you.
Project and Stakeholder Management	---
Algebra	Understand and apply basic concepts of algebra (e.g. coordinate geometry, functions and graphs) to analysis work.
Probability and Statistics	Understand broad concepts and terminologies of probability and statistics. Calculate and interpret significance of statistical measures (mean, median, mode, standard deviation, correlation) from data sets.
Data Collection	Exercise sensitivity in the collection process for data to ensure clarity and consistency for downstream analysis. Use common office productivity tools (e.g. Excel) to collate data. Create the associated data dictionary for the extracted data set.
Data Quality	Understand the various data types (e.g. number, date, string), underlying representation in common locally stored data formats (e.g. CSV, Excel). Use common features of office productivity tools (e.g. Excel) to manipulate and clean up the data. Understand the impact of data cleansing (e.g. removal of duplicate entries, removal of incomplete entries) to the overall analysis.
Exploration and Analysis	Use common features / in-built formula of office productivity tools (e.g. PivotTable in Excel) to obtain insights from the data (e.g. mean, median, mode).
Machine Learning	Relate some use cases where machine learning (e.g. clustering, regression, classification) has been used to solve specific problems.
Data Visualisation Principles	Understand some considerations of best practices in presenting various data types and information.
Data Visualisation Development	Use common visualisation charts (e.g. Bar, Line, Histogram, PivotChart) for analysis. Validate accuracy and clarify of information presented.
Communication and Presentation	Explain discoveries and insights with the help of visualisations (e.g. charts, tables, dashboards). Explain findings related to specific analytical questions or problem statements, and the team's conclusions and recommendations. Explain limitations of the available data and the impact on the overall analysis.
Programming Concept & Construct	---

Defining different types of Advanced/Expert Data Scientists

Tracks / Roles	Quantitative Strategy 	Computational Methods 	Data Engineering 	Data Visualisation 
Description	Quantitative consultant at the core.	Computer scientist at the core.	Software engineer at the core.	Software developer at the core.
Skilled in ...	Problem formulation, statistics, data governance, translating analyses into visuals & actionable insights for key stakeholders	Computer algorithms, machine learning, statistics.	Software, network and systems engineering.	Software development, data visualisation development and principles.
Focus on ...	Enhancing government policies and operations through strategic data science projects with government agencies.	Algorithms relating to natural language processing, machine learning, deep learning or graph theory to solve complex problems.	Engineering data pipelines for analysis work and architecting of distributed, big data systems.	Developing and integrating data visualisations into software/apps on a variety of platforms (e.g. web, mobile, apps).

Training Roadmap

Builds upon competency framework
Lists relevant training courses given job roles & required competencies

Challenges faced

- ✗ Too many courses out there
- ✗ Unfamiliar with tech jargon and ambiguous course titles
- ✗ Difficulty in identifying appropriate courses to address specialized training needs
- ✗ Courses may not be of quality

Our Training Roadmap Can Help

- ✓ Allow HR / supervisor to easily develop training plan for their officer
- ✓ Choose course from menu
- ✓ How to navigate menu?
  - Archetype Job role of officer
  - Determines what skills needed
  - Per skill, 5 courses to choose
  - Star courses that we curated or had consistently good feedback

GOVTECH SINGAPORE

Mapping required skills to job role

Example Archetype	Job role	Training	Actionables
IT Manager	Leading a team/division doing tech related operations.	Understand overall ecosystem of Data Science / Big Data tools (such as Hadoop, Tableau, R, Python) to provide oversight and supervision for the tech team.	Drive change in the IT infrastructure so that Proof-of-Concept analytics projects can transit into deployment.
Business Analyst	Embedded within an IT team to manage vendors, contracts and bridge between users and the tech aspects.	Understand overall ecosystem of Data Science / Big Data tools; Straightforward data extraction and transformation skills.	Carry proof-of-concepts to production and deployment; Execute straightforward data extraction and data manipulation requests to facilitate subsequent analytics work
Policy/Ops Analyst	**Put up policy/ops recommendations. Typically requires simulation of policy/ops changes, cost-benefit analysis etc.**	**Hands-on learning. Such as data cleaning; simple techniques like clustering, regression; Data visualization skills, e.g. building dashboards**	**Perform simple analysis such as merging different data sets, cluster analysis to find archetypes, and presenting insights using impactful visualisations**

Different types of training available

Primer

Broad awareness of use-cases, concepts.

Improves data literacy and encourage data-driven mindset & culture

Highlights practical applications of AI / Data Science

--- For Testdrive sessions ---

Try out functions and features of Data Science-related tools

0.5 – 1 day

Onboarding

Know enough to start and know where to deep-dive further.

Recommended as the first course for officers to take (no prior background required)

Cover foundational skills, broad-based knowledge and build awareness

Prepare participants to embark on enhancement courses

Basic / Intermediate: 2-3 days
Advanced: 3-5 months

Enhancement

Hone specific skills for immediate, effective application.

Develop an officer further along a particular functional track

Equip practical, hands-on and ready-to-apply skills back at work

Validate / Accredit skills through certifications and/or project work

2 – 5 days per course

Training

2. Develop, curate and deliver training options; **Aggregate training demand** from agency.

a) Primer Talks

b) Starting A Data Science Project

c) Data Science For Management

d) Data Science Masterclasses

GOVTECH SINGAPORE

We Conduct Sessions to Help Government Agencies Through Their Data Science Transformation Journey

Primers

- One primer explains the possibilities of data science with actual public sector use cases
- Another primer explains how data science works, concepts of some techniques in a layman way

Project Scoping Workshops

- We act as bridge, know both public policies/ops context, and data science techniques
- Sharpen scope of project to something doable within 3 to 6 months
- Develop project pipeline for agency

Lunch Time Talk – Data Science & AI: A Primer

- Lunch-time talks that are informal, and easily digestible (no intimidating math)
- Helps public officers appreciate the use of data and intuition behind some data science techniques

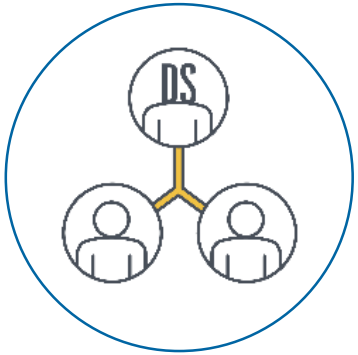
Starting A Data Science Project

- Onboarding course for Intermediate level (power users).
- Covers fundamentals, project scoping, data preparation and analysis with Python, data visualisation and communication with Tableau.

Data Science For Management

- Onboarding course for Senior Management.
- Covers applications of Data Science and AI across government, primer on Data Science and AI to debunk the myths, how to scope a good project and how to build up a team.
- Approach CSC or GovTech for a customised run of this.

Data Science Masterclass



- Short length workshop, course, seminar sessions for Advanced, Expert proficiency levels.
- Intro to niche, specialised tools, training on tips & tricks, best practices, dialogues with gurus.
- E.g. Workshops on Spark, Deep Learning

Capability Building

3. Assist government agencies in **starting data science projects**, and **building data science capabilities**

a) Project Scoping Workshop

b) Data Science Apprenticeship

c) Satellite Teams

GOVTECH SINGAPORE

Workshops to get agencies started on projects

- Help government agencies scope data science projects
- Agencies typically start with broad high-level goals
- We help to define a clearer scope, ensure that relevant data is available, and a proof-of-concept can be done within 3-6 months
- Outcome of workshop is to help **build a pipeline of projects** and **prioritize** the list of projects

Components Of Project Scoping Workshops

Workshop Preparation

Participants guided through project scoping worksheets **to gain a better understanding** of the potential projects.

Workshop Session #1

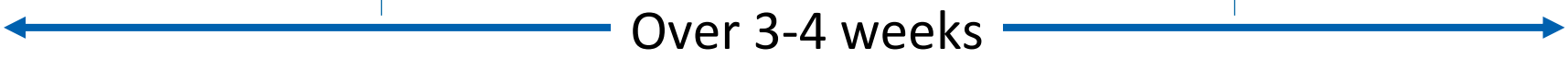
Participants guided to gain **a better understanding of Data Science concepts** and methodologies, and to **define their problem statements.**

Review of Project Scoping Worksheets

Participants to **refine their problem statements** based on inputs from the Workshop Session #1.

Workshop Session #2

Participants to **present their problem statements for critic** and **assess for feasibility of project.**



Data Science Apprenticeship

- 6-9 months attachment to GovTech for an immersive learning experience, working on GovTech and/or agency projects.
- Jointly chart out learning milestones between agency, apprentice and GovTech.

Only applicable to public officers serving the Singapore Government

GovTech Partnerships

- **Ad-hoc consultancy**
  - GovTech data scientists work on agency's particular project
- **Retainer**
  - Agency engages GovTech for X man-months
  - GovTech deploys relevant officers to work on range of projects or initiatives
- **Satellite teams**
  - GovTech posts data scientists into agency
  - Assess suitability of agencies:
    - Strong project pipeline
    - Strong management support

Community

4. Engage enthusiasts and practitioners across WOG through the **Data Science Community**.

a) [WOG] Data Science @ Workplace

b) Data Science Connect Portal

c) Data Science Connect Meet-Up

GOVTECH SINGAPORE

We build WOG Data Science Community

[WOG] Data Science @ Workplace

Data Science Connect Meetup

- Meet up every 2-3 mths
- Sharing knowledge and experiences

Data Viz User Group

- Comprise of practitioners
- Conducted the first Data Science Clinic at Connect Meetup #6

Data Science Connect Meet-Up

Format of a Data Science Connect Meet-Up session

Agency sharing on how they use analytics in their work, their transformation journey	
Tools / Products sharing (related to Data, Data Science & AI)	
Technical Sharing sessions for **Enthusiasts**	Technical Sharing sessions for **Practitioners**

Agencies are forthcoming in volunteering to share at Data Science Connect!

Data Science Connect Portal (Intranet)

Competency Fmk v2.0

Data Science Playbook for Project Scoping

Data Science Playbook for Project Scoping

Articles and Videos

GOVTECH SINGAPORE

Thank you!

GOVTECH SINGAPORE