DRAFT

Department of Education

Preparing for the Analytics Platform of 2020

13th December 2019

BUSINESS ANALYTICS ADVISORY

Final Report Draft- Version 1.0
Released

Australian Government
Department of Education

sas
THE POWER TO KNOW.

Contents

1. Executive Summary 5
   - The path so far and working out the next steps 6
   - Key Priorities 9-19
   - Major Trends 20
   - High Level Development Milestones 21
2. Engagement Approach 22-26
3. Analytics within the Department 27
   - The Analytics Maturity Curve 29-32
   - Current and Future Positioning within Department 33-34
4. Early Childhood and Childcare 35-48
5. Schools and Youth 49-54
6. Higher Education, Research and International 55-59
7. NEEB 60-63
8. Appendices 64-73

Document Information

Revision History

Version	Author	Date	Summary of change
Version 0.1 Draft	[redacted]	27 November 2019	Initial draft.
Version Preliminary Report (Draft) 1.0	[redacted]	5th December 2019	Feedback from the Department
FINAL Report Draft (V1.0)	[redacted]	12th December	Feedback following presentation 10/12/19
FINAL Report	[redacted]	13th December	Minor

Source Documentation

The following source documents were referenced in creating this document.

Data Management Strategy 2018-2021	Department of Education	November 2018	
Data Management Strategy – Appendices	Department of Education	November 2018	
Data Management Strategy Implementation Plan	Department of Education	November 2018	
NEEB – Health Check	PWC	March 2019	
NEEB Program Review of Recommendations	PWC	October 2019	
Highlights from the Data Capability Flash Survey	Department of Education	June2019	
Strategic Analysis (VI Pilot Project)	SAS	June 2018	
Corporate Plan 2018-19	Department of Education	2018	
NEEB Insights & Capabilities Final Report	Department of Education	January 2018	

Document Distribution

Name	Title	Organisation
[redacted]		Department of Education
[redacted]		Department of Education
[redacted]		SAS

Executive Summary

Executive Summary

Background & Context

The path so far:

The Department of Education and Training embarked on developing and delivering a Data Management Strategy 2018-2021 ("the strategy") to support a stronger Australian education system. In response to the issue of dispersed education data, the department has created a Chief Data Officer role, adopted a data management strategy and developed a National Education Evidence Base (NEEB) reform agenda, focused on enabling greater use and availability of data to support improved education outcomes. Funding of $14.3 million was provided to the Department through Data Integration Partnership for Australia (DIPA) for NEEB, and the Department has also committed extra funds.

Aligned to the overall Corporate Plan, there are key factors and implications linked to changes in technology.

It is acknowledged that technology and computing have changed significantly and rapidly in recent times and the use of data, automation and artificial intelligence is growing.

The Department is implementing its Technology Strategy and the associated transformation is underway to transform the department's technology services and systems. The strategy sets a direction for providing technology solutions that are modular, less complex, easier to maintain and more responsive in addressing departmental business needs.

This report identifies where a number of branches can benefit from enterprise based technology – specifically in the area of using data and analytics to support research, policy making, reporting and identifying fraud.

As part of this review, NEEB has also been discussed. However PwC has been involved in reviewing NEEB specifically, being aware of this SAS Advisory has aimed not to replicate the work done by PwC, but to augment it as appropriate.

The Business Analytics Advisory team at SAS was asked to produce this report to help the Department plan for the increased demand for analytics.

This review will largely focus on the use of SAS and how it will evolve into the analytics platform for the future.

Executive Summary

Working out the next steps:

The Department has the requirement to understand how the analytics and the data sharing journey will develop into the future and what might be required to continue on this journey to:

- Provide a first-class service in the provision and democratisation of data (and insights) to a range of stakeholders, additionally generate greater value from such data assets
- Improve operational efficiency and increase agility and innovation
- Increase service offerings and move deeper into the area of advanced analytics
- Continue to advance levels of maturity in analytics and the use of data

A key focus of this exercise is the **Analytical Platform** and how it should evolve going forward. The focus of this was in terms of its application across:

- Data Analytics across the Department, and across the
- National Education Evidence Base – (NEEB)

Improving policy making and decision making is at the core of the Department and this will be achieved through the on-going task of improving accessibility to quality data and the application of analytical methods. With the proliferation of data, the emergence of relatively inexpensive and abundant computing power as well as advanced analytical algorithms, it is of interest to the Department to understand where and how they can continue on the journey and deliver a range of data driven services.

This exercise focuses on a short review of current state in terms of current service delivery. However, the main focus is looking into the future in terms of how the Department can continue to maximise the evolving analytical platforms available

.

Definition of "Analytics" across the Department

Across the Department there is a range of different analytical capabilities associated with different clusters. This covers a focus on consolidating and interrogating raw data, reporting / BI, advanced analytical modelling, as well as the delivery of insights and research through the application of advanced statistical techniques.

Basic Analytics

- Data Consolidation
- Standard Reporting
- Ad Hoc Analysis & Reporting

Intermediate

Producing new insights that reveal the relationship between data elements (e.g. the cause and effect, survey design and analysis, hypotheses testing)

Visualisation, exploratory and confirmatory data analysis

Intermediate

- Statistical Analysis
- Trend & Exception Analysis
- Advanced Visualisation
- Research and quantitative studies

Advanced Analytics:

<u>Predictive Analytics / Anomaly Detection and advanced linkage</u>

Determining the relative importance and impact of drivers that indicate fraudulent behaviour.

Advanced Analytics

- Forecasting
- Predictive Analysis
- Optimisation

Analytics within Department of Education

Within the Department, and within the context of this review and roadmap development, current state "analytics" covers a wide range of activities, from the manipulation and integration of data sources, to the subsequent analysis of such data, providing meaningful and actionable information in the form of business intelligence and reporting.

Key Priorities

Key Priorities for Department of Education

The following *general* observations can be seen as key priorities to be addressed to deliver value to the Department in line with the ongoing analytical and data strategy including the implementation and modernisation of the SAS platform (over the next 12 -18 months):

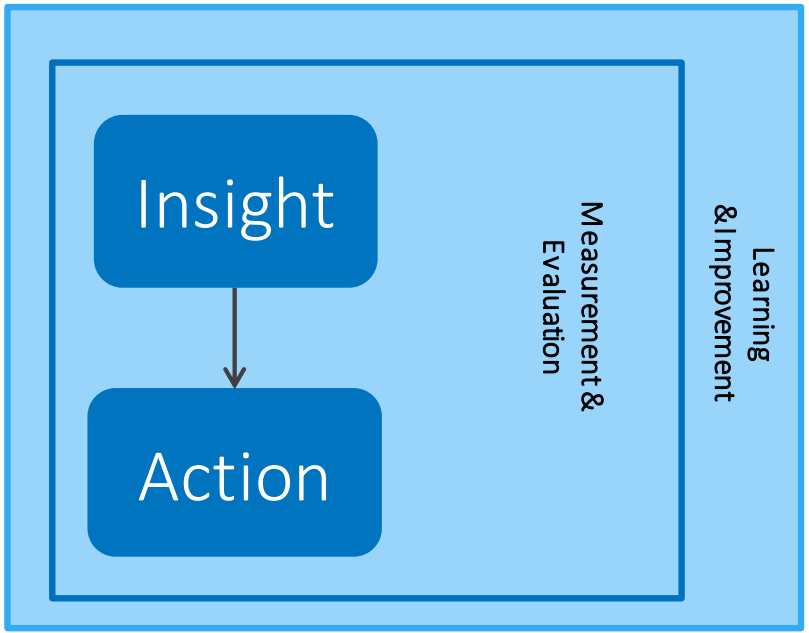
1. Adopt an "Insight to Action" way of thinking

Implement a program of work covering innovative prototyping projects to evaluate and assess capabilities (not just with the SAS Analytics Platform) and developing advanced analytics capabilities, but also the general principle of insight driving decision and policy making (and subsequent implementation).

Align with change management, training and adoption plans for migration (this will be multi-speed based on current competencies across clusters).

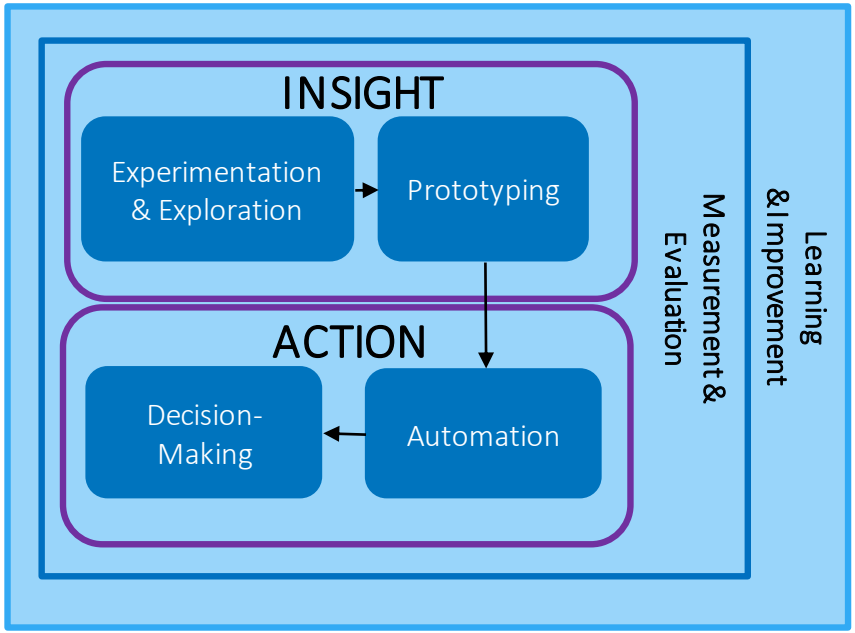
See Appendices for further clarification

Priority One: Continue to adopt an "Insight to Action" way of thinking



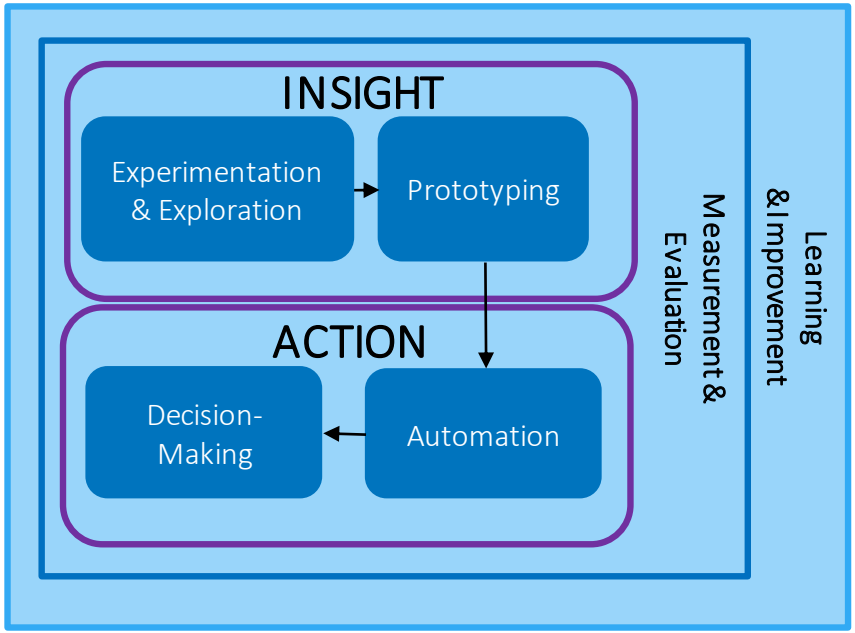
- Insight without action destroys value
- Action without insight is guesswork
- Outcomes without measurement prevents learning and improvements
- Measurement without learning creates stagnation

Priority One: Continue to adopt an "Insight to Action" way of thinking



- Prototyping reduces risk and drives agility
- Insight with action creates the potential for value
- Measurement drives self-awareness, focus and prioritisation
- Learning enables growth and improvement

Priority One: Continue to adopt an "Insight to Action" way of thinking



- Prototyping reduces risk and drives agility
- Insight with action creates the potential for value
- Measurement drives self-awareness, focus and prioritisation
- Learning enables growth and improvement

Key Priorities

Key Priorities for Department of Education

2. Not every branch or cluster needs to be "Levels 4/5"

Based on current analytical capability, continue the development of analytics teams, increasing their maturity to a desired and realistic level, promoting and showcasing success as it occurs.

Particular branches (e.g. HERI cluster) given the nature of their services do not require to be "Analytically Innovative" if data manipulation and reporting are core focus areas of delivery, then sensible aspirations need to evolve around delivery in a timely, repeatable and cost effective manner.

Adoption of new technologies such as the SAS Analytics Platform (and associated up to date functionality) will enable such branches to increase and optimise their level of capability.

Maturity will drive automation of data handling, analytics, decision-making and action; duplication and manual processes are eliminated, making sure the most efficient approaches to deliver the required capability are followed.

Key Priorities

Key Priorities for Department of Education

3. Continue to evolve and optimise technology architecture

Identify weaknesses in the IT administration and support of IT that enables the full spectrum of reporting and analytics required to deliver services, ensuring sufficient capacity for teams to deliver outcomes efficiently.

A diligent approach to the management of IT infrastructure, encouraging innovation (through on-going technological improvements) will be critical elements of delivering analytical assets efficiently. This even relates to individuals having up to date laptops.

The continual exploration both within existing technology and additional emerging solutions can continue to offer progressive disruption and potential productive efficiencies offered by innovative

Key Priorities

Key Priorities for Department of Education

4. Targeted promotion and awareness of NEEB

With further additions to the program – the SAS Analytics Platform becoming available, there will be focus on the use of analytics (and platform), focusing on the ability to share and explore data to drive meaningful information and insights and to generate further hypotheses to support policy making.

Early adopters could be individuals who are beginning their analytics journey, or, are low maturity in terms using data and analytics.

Encourage "exploration" and exploratory data analysis in the first instance and promote how the NEEB infrastructure can support hypotheses. Once an interesting opportunity has been found, teams need to find the answer to the identified opportunity. Exploration focuses on clarifying the "known unknowns" and is often driven by specific requirements, NEEB is a perfect opportunity to practice this approach

Key Priorities

Key Priorities for Department of Education

5. Review (and restate) respective service catalogues for each cluster (as required),

Given the ambition of the Department and funding to modernise and increase the capacity of the use of data and technology it is key that the Department maximises the offerings that can be achieved for both internal and external stakeholders.

Focusing on the vision and priorities of each branch, continue to review the needs of stakeholders and their satisfaction with current outputs to maintain a relevant, complete and up to date catalogue of services.

With the advent of a modern analytics platform it is important to review existing service catalogue and service levels, identifying potential enhancements, including the use of advanced analytics can be put forward.

An example framework is shown in *Appendix 5*

Key Priorities

Key Priorities for Department of Education

6. Specific expansion of capabilities within Compliance Intelligence and Analysis

Investigation of additional and incremental approaches across the fraud lifecycle.

The opportunity for integrating analytics into the Compliance team's activity is immense – some great progress has already been made and the team continues to develop high quality processes and solutions.

Further opportunities with technology and solutions exist to increase their level of competency even further Additionally, review opportunities to extend detection of other types of fraud.

This project did not investigate in detail the team's depth of services, but the use of Analytics in fraud "detection and prevention" is far reaching and includes:

- Risk assessment and monitoring
- Assessing if and how data and analytics can be used to combat fraud ("fraud insights")
- On-going effectiveness of fraud prevention and potential new patterns ("fraud insights")
- Creation of continuous protection capabilities to protected against known patterns of fraud ("fraud scenarios")
- Management and tuning of continuous protection capabilities to minimise false positives, but maximise protection ("scenario tuning")
- Supporting investigators with triage and execution of cases ("intelligence")
- Management reporting to provide oversight and governance of the entire process and results

Key Priorities

Key Priorities for Department of Education (continued)

7. Continuous development of competencies and capacity to support delivery of current and future services

Given the significant change being observed currently and the rate of change expected in the future, there will be a strong requirement to continually support individuals with up to date training and skills programs. As the need to further expand the service catalogues (priority 5) and deliver high quality services, coupled with the fast changing technology enablers, a range of parallel activities will be required:

- Training in relevant technology and tools
- Training and skills transference in terms of analytical methods and application - not just the technology
- Operating model review(s) within clusters
- Assessing the balancing internal and external resourcing levels to meet demand for services

Key Priorities

Key Priorities for Department of Education (continued)

8. Drive the continual execution of the change journey and deliver sustainable change to the Department to gain value from analytics and the analytics platform

Maintain a range of specific objectives for the Change Management (and team) to drive successful execution, specifically:

- Continue to educate leaders on the importance of managing change so they can create sustainable and manageable sponsorship across the department
- Enable team members to understand the journey to arrive at desired future state
- Create a "demand" for the change at all levels – will influence employees as to why they should do something differently and work in new ways
- Identify quick-wins to release business value early and create demonstrable value and build confidence
- Transition employees and work processes to the new operations / platform in shorter but manageable timeframes
- Measure of the change journey – will indicate benefits re realised and impact as appropriately manages
- Make sure employees are given opportunities to build their changes skills through training, knowledge transfer (even across siloed clusters) activities and coaching

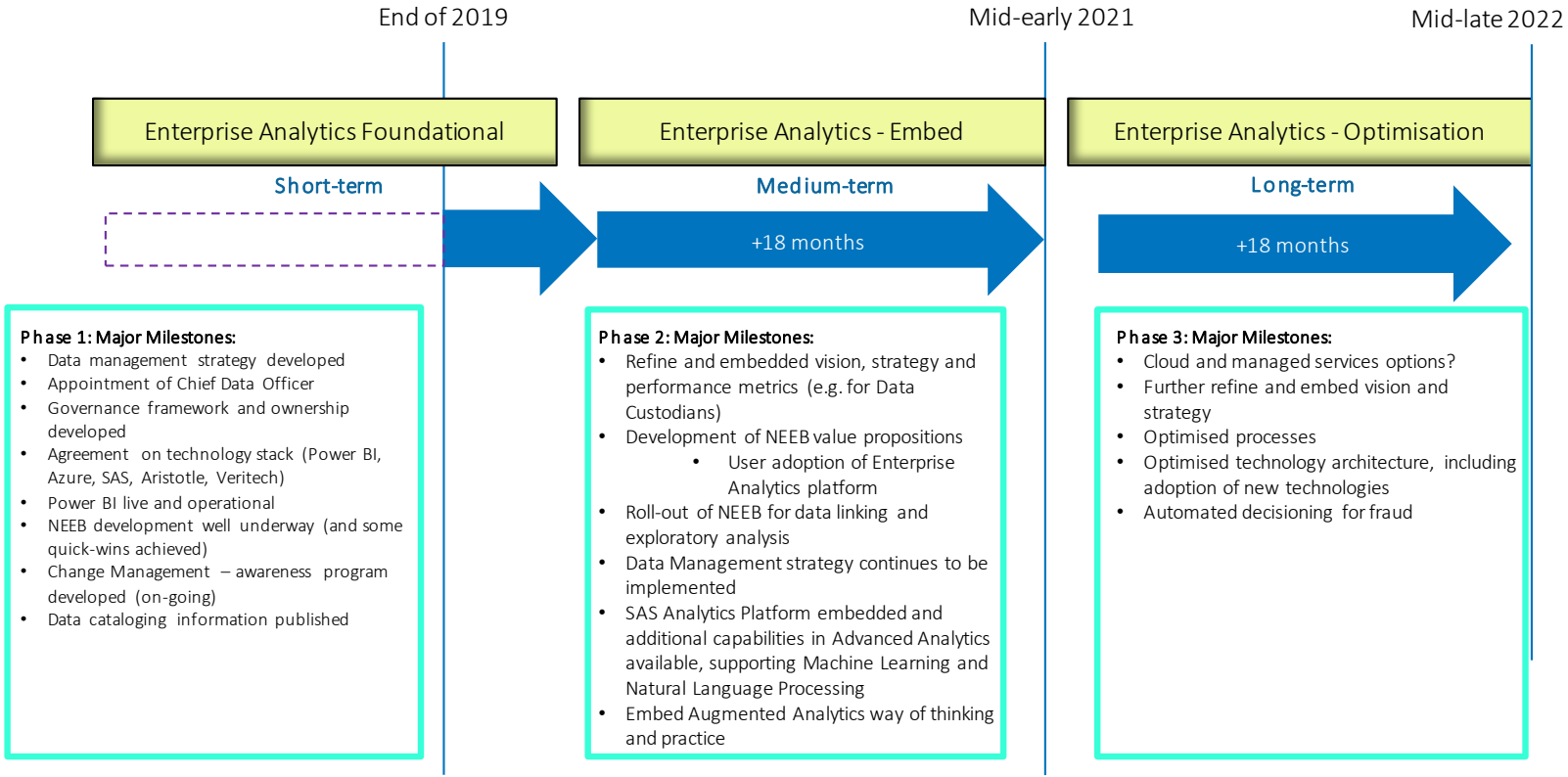
See Appendix 6

Major Trends in Reporting and Analytics Today

The major trends in enterprise reporting and analytics affecting government, including education:

Artificial Intelligence and machine learning	•Automation of traditional roles •(e.g. image and speech recognition)
Internet of things	•Sensor-driven real time decision automation •(e.g. asset management, building maintenance and safety)
Pervasive advanced analytics	•Operational integration of analytics •(e.g. embedded analytics processes and task management)
Big data	•Ever increasing data volumes, variety and velocity •(e.g. notes and digitisation, camera video streams)
Data governance and stewardship	•Data governance and stewardship as dedicated roles •(e.g. Chief Data Officer role)
Cloud flexibility	•Capacity on demand •(e.g. cloud based solutions)
Enterprise strategy	•Business led enterprise reporting and analytics strategy •(e.g. Data Management strategy)
Analytics on demand	•Services on demand •(e.g. Results as a Service)

High Level Development Roadmap Milestones



Engagement Approach

Overview of Approach

SAS Advisory Project Approach – Roadmap Development

Step 1.Project Preparation	Step 2. Information Gathering 1	Step 3. Information Gathering 2	Step 4. Initial feedback of draft report	Step 5. Initial feedback of draft report
Meet with CDO and support to understand scope and priorities, engagement approach and outcomes required of the project	SAS to adapt/prepare relevant information request templates, workshop content	Further individual one-on-ones identifying requirements and collection of further insights as required	Draft report with findings, recommendations and roadmap for discussion	Preliminary report and recommendations presentation to the Department Executive and Reporting / Analytics teams
SAS, with the Department project team identify who to meet and adapt engagement approach to focus on the agreed priorities required of the project	Small group workshops and "interview" sessions selected with Directors, Managers and Analysts to understand current state, challenges and priorities	Further consolidate findings, develop draft report with recommendations and roadmap	Finalise feedback and incorporate into Preliminary report	Production and delivery of final report
	Consolidate findings, identify follow-up sessions with selective individuals			

Note: No survey's were used due to the need to ask a variety of different questions to different teams. Additionally, probing further information at a face-to-face level is extremely valuable. It is also understand that many teams have been the subject of several self-completion questionnaires in recent times so there is an element of "survey fatigue".

Branches included in the Exercise

Person	Role	Section	Branch	Group	Cluster
[redacted]			Data Analytics Branch	Strategy and Data Group	Higher Edu, Research & International
[redacted]			Data Analytics Branch	Strategy and Data Group	Higher Edu, Research & International
[redacted]			Data Analytics Branch	Strategy and Data Group	Higher Edu, Research & International
[redacted]			Market Analysis & Data	Strategy and Data Group	Higher Edu, Research & International
[redacted]			Market Analysis & Data	Strategy and Data Group	Higher Edu, Research & International
[redacted]			Payments and Collection	Funding and Data Collection	Schools
[redacted]			Assessment and Analysis	National Policy and Data Reform	Schools
[redacted]			Assessment and Analysis	National Policy and Data Reform	Schools
[redacted]			Assessment and Analysis	National Policy and Data Reform	Schools
[redacted]			Modelling, Data and Finance	Payment Integrity	Early Childhood and Child Care
[redacted]			Modelling, Data and Finance	Payment Integrity	Early Childhood and Child Care
[redacted]			ECCC Network	Early Learning Policy and Network	Early Childhood and Child Care
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy

Branches included in the Exercise (continued)

Person	Role	Section	Branch	Group	Cluster
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy
[redacted]			Digital and Business Transformation	Finance Technology and Business Services	Corporate Strategy

Analytics for the Future

Focus of this Review

Use of Analytics

Use of analytics (and what will be possible)

- Within clusters - current skills & capabilities
- Future platform requirements & benefits
- Roadmap and maturity levels
- Change requirements

- *User base(s)*
- *Adoption*
- *"Change"*
- *On-boarding*
- *Training*

NEEB

Bringing data together from across clusters

- High level review
  - Status
  - Benefits
  - Challenges

- ❑ Open analytics capability
- ❑ Insight driven decisions and policy support
- ❑ Measure outcomes

SAS Analytics Platform spans both

Analytics within the Department

Analytics by Cluster

High-Level Review

Early Childhood & Child Care

Cluster that focuses largely on compliance and intelligence. Large team of skilled data scientists; manipulation and analysis of large volume data sets

Services:

- Cleansing and manipulation of data sets for input to advanced analytics and further investigative procedures (e.g. state network compliance)
- Provision and delivery of predictive models and outlier and anomaly detection through advanced analytics, including multivariate methods, to identify fraudulent behaviours
- Provide dashboards and reports through VA budgeting and summary reports

Schools & Youth

Highly skilled team of researchers who drive policy through longitudinal studies and application of advanced research techniques.
Additional resources supplying data into researchers

Services:

Production and provision of data

- Use of Enterprise Guide to manipulate and provide data
- Response to a range of ad-hoc data requests

Research & Analytics

- Fact checking
- Data manipulation
- Research for driving policy

Higher Education & International

Provider of data to internal and external clients (SAS datasets and Excel/CSV formats). Use Power BI, SAS , SQL- deliver high profile annual reports

Services:

- Publication of 2 annual reports
- Provide response to ad-hoc requests (1000 per year) for data and information
- Very focused data centric roles

Corporate Strategy

Focus was on NEEB – the data portal to allow access to a large number of integrated data sets allowing users to select data for ad-hoc linkage and analysis to support policy exploration. Power BI and SAS core platforms are / will be available to enable users extract value from data

Services:

- Highly catalogued repository of data
- Allows sharing of data
- Enables insight generation through easy to access and qualified data sources
- Allows self-serve analytics and bespoke data wrangling

Note: There maybe other areas delivering analytical insights that have not been captured in this review, and pockets of capability that exist in each of the clusters that are not reflected on this page

The Analytics Maturity Curve

Different stages and levels for different clusters

Analytical Maturity

The Department has the requirement to understand how the analytics (and data sharing journey) will develop into the future and what might be required to continue on this journey to:

- Provide a first-class service in the provision and democratisation of data and insights to the range of stakeholders
- Improve operational efficiency and increase agility and innovation
- Increase service offerings and move deeper into the area of advanced analytics

A key focus of this exercise is the development and expansion of the **Analytics Platform** and how it will evolve going forward. The focus being in terms of:

- Data Analytics across clusters
- National Education Evidence Base (NEEB)

Improving policy making and decision making is at the core of the Department's analytics strategy, and this will be achieved through the on-going task of improving accessibility to quality data and the application of analytical methods. With the proliferation of data, the emergence of relatively inexpensive and abundant computing power as well as advanced analytical algorithms it is of interest to the Department to understand where and how they can get on the journey and continue to deliver a range of data driven services.

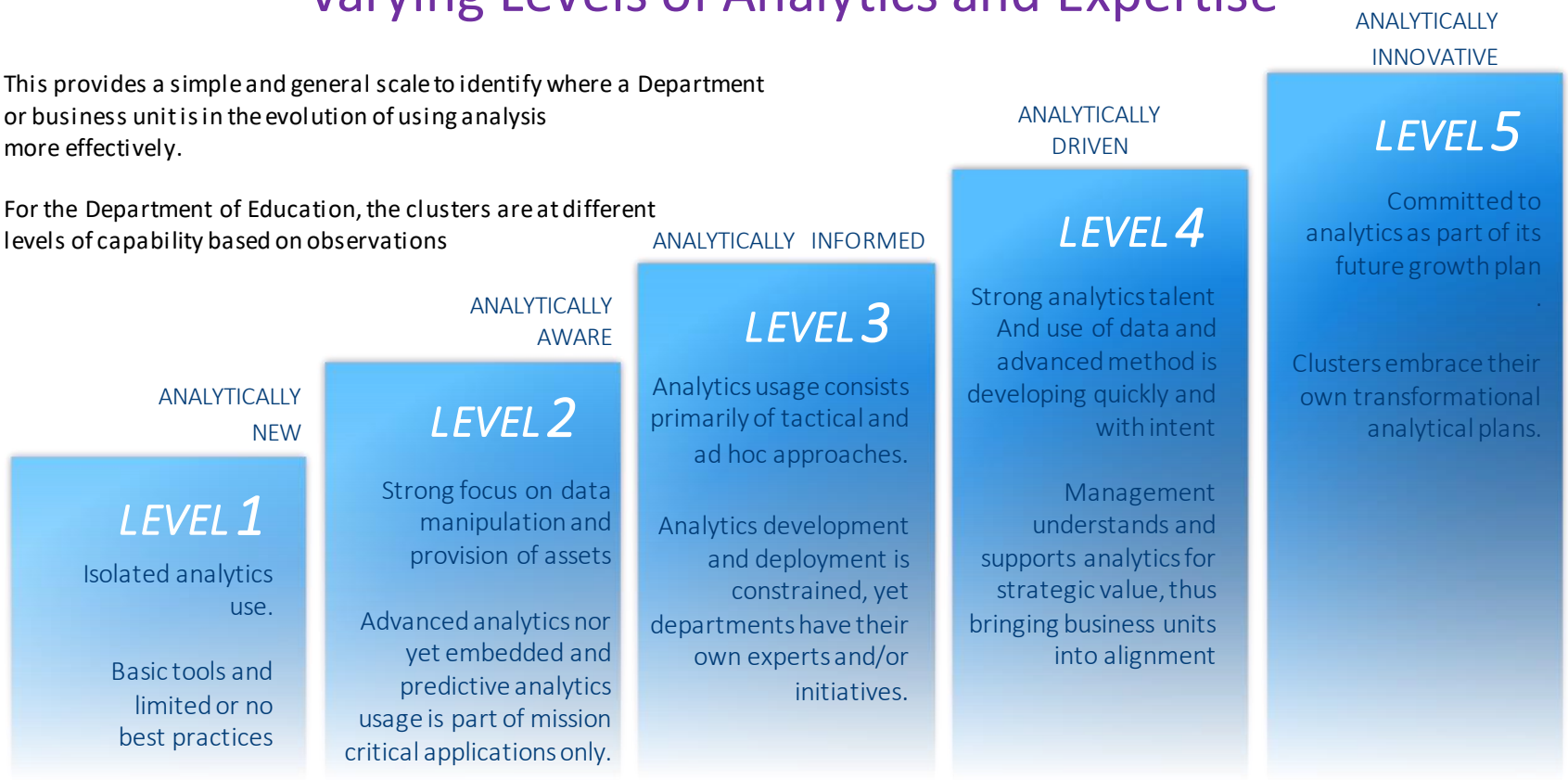
This exercise focuses on a short review of current state in terms of current service delivery. However, the main focus is looking into the future in terms of how the Department can continue to maximise the evolving analytical platforms available

.

Varying Levels of Analytics and Expertise

This provides a simple and general scale to identify where a Department or business unit is in the evolution of using analysis more effectively.

For the Department of Education, the clusters are at different levels of capability based on observations



Maturity is variable across clusters, but also within clusters

Varying Levels of Analytics and Expertise

Each cluster needs to be considered separately in terms of current service offerings and the analytical capabilities required to deliver against such services

Different service offerings are plotted on the bespoke maturity chart on the following page with clusters aligned to current state.

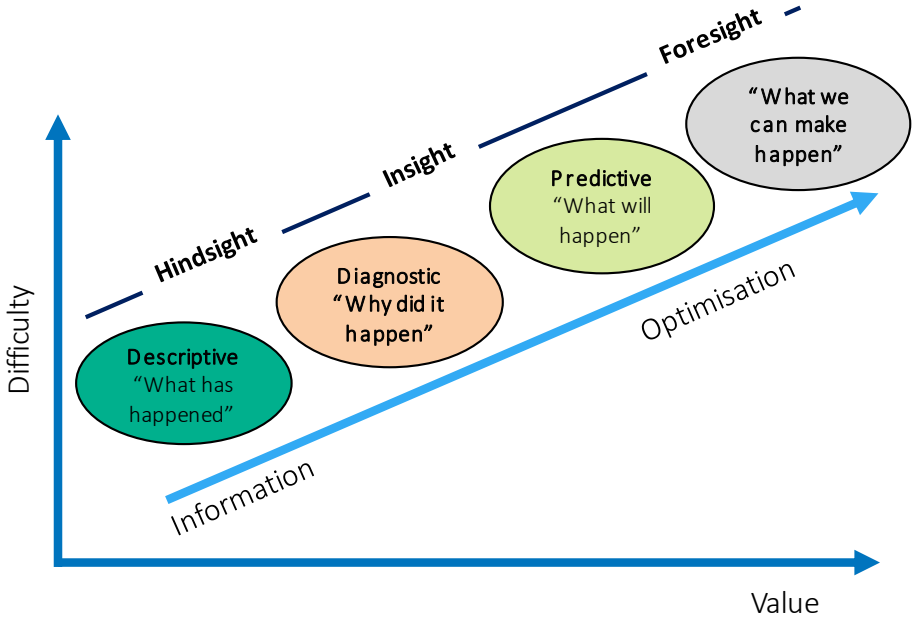
The ability to move up the analytical maturity curve is variable across clusters and the level achievable will be based on the following criteria:

- **Future need, relevance and requirement to adapt analytics and the use of data beyond current status quo**
- **Current level of analytical maturity – how much further can a branch or section be expected to go?**
- **Current skill sets and capability – the nature of current work**
- **'Ambition' to move from current "business as usual" to a more advanced level**

The Analytics Maturity – Continuum of Capabilities

Further detail can be found
In the Appendix 2

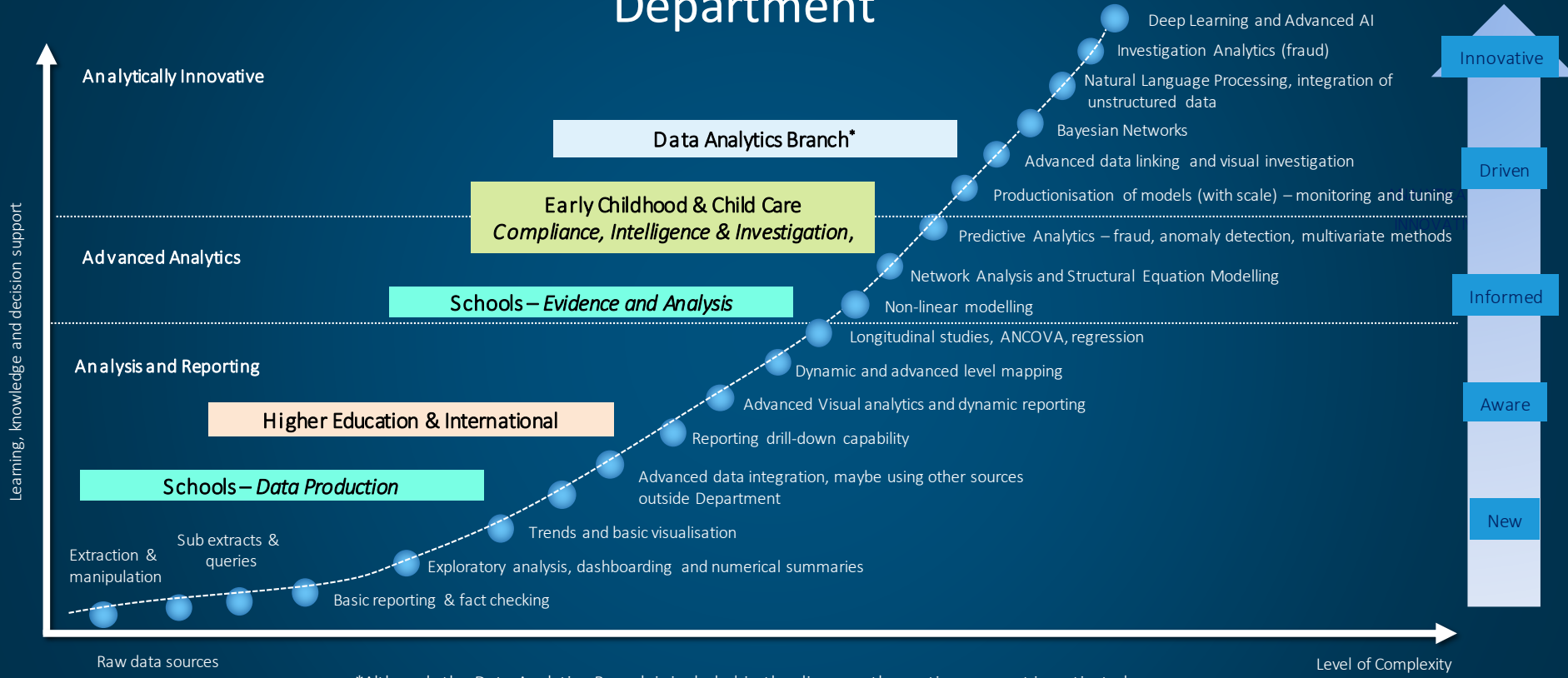
Analytics	Description
Descriptive	Descriptive Analytics is the examination of data or content, usually manually performed, to answer the question "What happened?" (or What is happening?), characterised by traditional business intelligence (BI) and visualisations such as pie charts, bar charts, line graphs, tables, or generated narratives.
Diagnostic	Diagnostic Analytics is a form of advanced analytics which examines data or content to answer the question "Why did it happen?", and is characterised by techniques such as drill-down, data discovery, data mining and correlations.
Predictive	Predictive Analytics is a form of advanced analytics which examines data or content to answer the question "What is going to happen?" or more precisely, "What is likely to happen?", and is characterised by techniques such as regression analysis, forecasting, multivariate statistics, pattern matching, predictive modelling, and forecasting.
Prescriptive Analytics	Prescriptive Analytics is a form of advanced analytics which examines data or content to answer the question "What should be done?" or "What can we do to make _____ happen?", and is characterised by techniques such as graph analysis, simulation, complex event processing, neural networks, recommendation engines, heuristics, and machine learning



Source: www.gartner.com

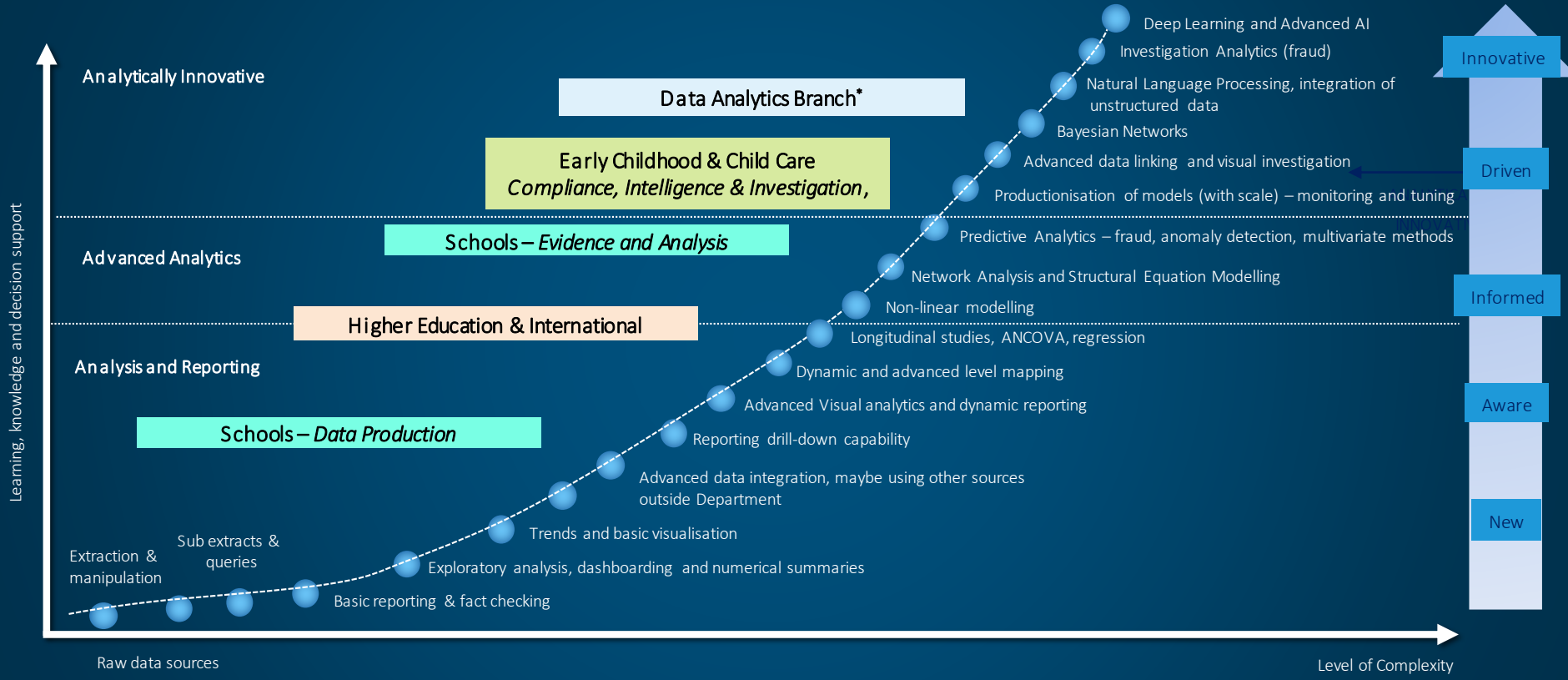
The following page shows key branches and their current position

The **Current** Analytics Maturity Continuum for the Department



*Although the Data Analytics Branch is included in the diagram, the section was not investigated as most of their work is done using the ABS data lab and therefore a separate SAS platform

The **Future** Analytics Maturity Continuum for the Department



Early Childhood and Child Care Cluster

Analytics and Fraud

Vision:

The Compliance Intelligence and Analysis team requires a vision that encapsulates them being an analytics-led function that combines advanced analytical techniques, intelligence products, and domain knowledge to proactively detect and prevent fraud.

They play a critical role in the fraud control with the aim of **proactive** and **sophisticated prevention** detection and response mechanisms to manage the risks of fraud and corruption.

Analytics – where is it used for fraud:

In order to assess where the Department currently is and where its potentially needs to be, it is important to understand the use of Analytics (all four categories) in fraud "detection and prevention". This includes usage across:

- Risk assessment and monitoring
- Assessing if and how data and analytics can be used to combat fraud ("fraud insights")
- On-going effectiveness of fraud prevention and potential new patterns ("fraud insights")
- Creation of continuous protection capabilities to protected against known patterns of fraud ("fraud scenarios")
- Management and tuning of continuous protection capabilities to minimise false positives, but maximise protection ("scenario tuning")
- Supporting investigators with triage and execution of cases ("intelligence")
- Management reporting to provide oversight and governance of the entire process and results

.

Early Childhood and Childcare Review

Compliance Intelligence and Analysis

Priorities and Activities

- A relatively large team (has contracted slightly), with a number of skilled data scientists with strong analytical capabilities
- Delivery of predictive models and outlier detection rules through some fairly advanced analytics to identify fraudulent behaviours
- High volumes of unit level data are analysed; stored procedures can deliver unit level data

Observed Challenges

- There have been serious performance problems with the SAS platform and the application of VA. SAS VIYA/VA has not performed well historically (though some improvements are now being observed).
- There are some manual processes for cleansing and manipulating large volumes of data - a lack of automation and scheduling is still prevalent

Positive Messages

- This team has highly skilled analytics and data science personnel focusing on applying analytical methods to look for insights, provide reporting and apply techniques such as business rules, anomaly detection and statistical modelling / machine learning for fraud detection and prevention
- Further movement into the advance analytics space is definitely where this front-running team needs to be and this team given the type of work with tangible benefits as outcomes is in a position to showcase advanced analytics to the Department

Summary Recommendations

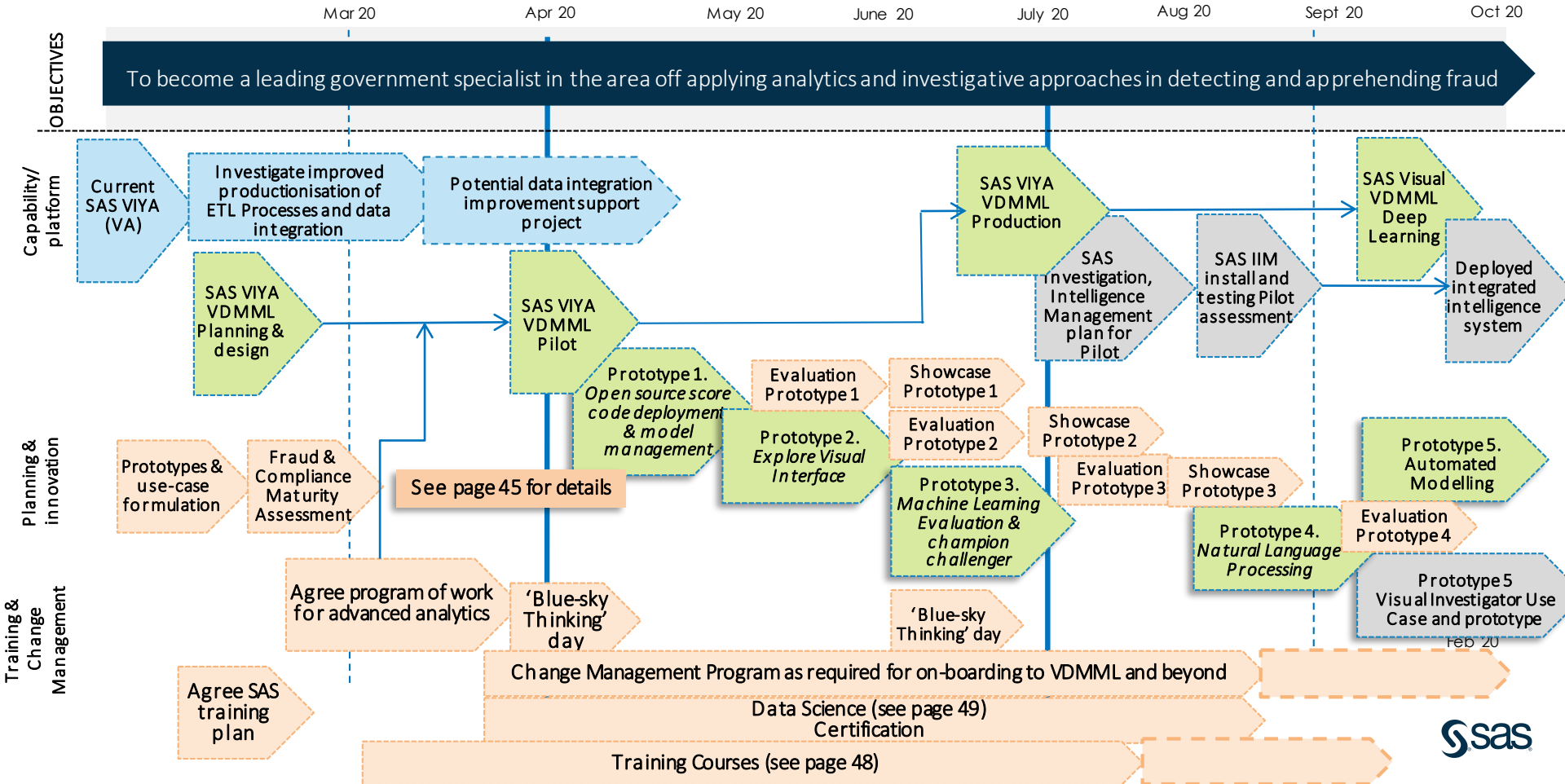
- Prototype and test SAS VIYA for a number of purposes including:
  - Productionisation of predictive models (currently built in R) through VIYA Model Management capability
  - Test and evaluate more advanced machine learning and NLP methods, use VIYA for model comparison
- Perform a detailed and more specific fraud control maturity assessment for the Early Childhood cluster – investigate further opportunities?

Plan for Delivery

Compliance, Intelligence and Analysis

Recommended platform enhancements plan 2020

Compliance, Intelligence & Analysis



Detail for Early Childhood & Childcare
Compliance Intelligence and Analysis

Capability & Platform

The key aspect of this is to make sure Early Childhood & Child Care (specifically the groups closer to gaining a high level of maturity regarding analytics) are able to use the latest version of Viya as well as develop VDMML capability across a range of agreed 'prototype' initiatives.

The demand for advanced analytics is likely to grow quickly, indicating optimal platform support, efficiencies in data manipulation and strong governance (of analytical assets) will be required

Specific projects and Prototypes will be tightly defined during early 2020 enabling quick-wins and on-going learning to be achieved

Recommended Components first 6 months

- SAS Visual Statistics
- SAS Visual Data Mining and Machine Learning
- SAS Visual Forecasting
- SAS Visual Text Analytics
- "SAS Natural Language Studio"

Recommended components beyond 6/12 months

Additional VDMML functionality as it becomes available:

- Model Building in Model Studio
- Fully guided/automated end to end modeling for business users and data scientists, using common automation framework
- Bagging, boosting, looping within a pipeline
- Further integration with Text Analytics
- Interactive Deep Learning
- Narrative summary, interpretability and reporting automatically
- Two-stage/incremental modelling

SAS IIM (intelligence and investigation management)

Planning & Innovation:

Further, detailed maturity assessment of Fraud

Fraud data and analytics requires not only capabilities and tools for base-line analytics – but also know-how in detailed fraud patterns and investigation processes. This requires more skills and capability than a 'typical' data and analytics.

SAS Advisory can assist in offering a focused and specific assessment of the following areas (amongst others):

- Fraud Management
- Domain knowledge and integration
- Analytical Capability
- Data & Data Engineering
- Insights (fraud)
- Investigative analytics
- Governance and monitoring
- Tools, platforms and infrastructure review

See page 35 for further details

Blue-Sky Thinking Days

Sessions structured to push the boundaries of innovation and identify where advanced technology and analytics can drive

Potential Prototypes for Childhood & Childcare

Open source code deployment

- Integrate models and open source code through SAS VIYA

Explore Visual Interface capabilities

- Variable selection (regression models)
- Validation data based tree pruning
- Insert models directly into Model Studio projects.

Machine Learning evaluation & champion challenger

- Machine Learning methods and compare, implement through using Model Management
- Adopt champion challenger test with competing models
- (*See Appendix 4)*

Natural Language Processing

- Natural Language Processing investigating use of unstructured data and how it might indicate potential fraud)

Automated Modelling

- Fully guided/automated end to end modeling for business users and data scientists, using automation framework

A Product Perspective – Summary Roadmap

Information is subject to change.

Recent Releases

- SAS Visual Statistics (SAS Viya and SAS 9)
- SAS Visual Data Mining and Machine Learning
- SAS Visual Forecasting
- SAS Visual Text Analytics
- SAS Econometrics
- SAS Optimization
- SAS Advanced Analytics Suite on SAS 9

Q3 2019

- SAS Machine Learning on SAS Analytics Cloud
- SAS Machine Learning + Forecasting + Econometrics on SAS Analytics Cloud

Q4 2019

- SAS Visual Statistics
- SAS Visual Data Mining and Machine Learning
- SAS Visual Forecasting
- SAS Visual Text Analytics
- SAS Econometrics
- SAS Optimization
- SAS IML
- "SAS Natural Language Studio"

Q1 | Q2 2020

- SAS Visual Statistics
- SAS Visual Data Mining and Machine Learning
- SAS Visual Forecasting
- SAS Visual Text Analytics
- SAS Econometrics
- SAS Optimization
- SAS IML
- "SAS Natural Language Studio"

Q3 | Q4 2020+

- SAS Advanced Analytics Suite on SAS 9

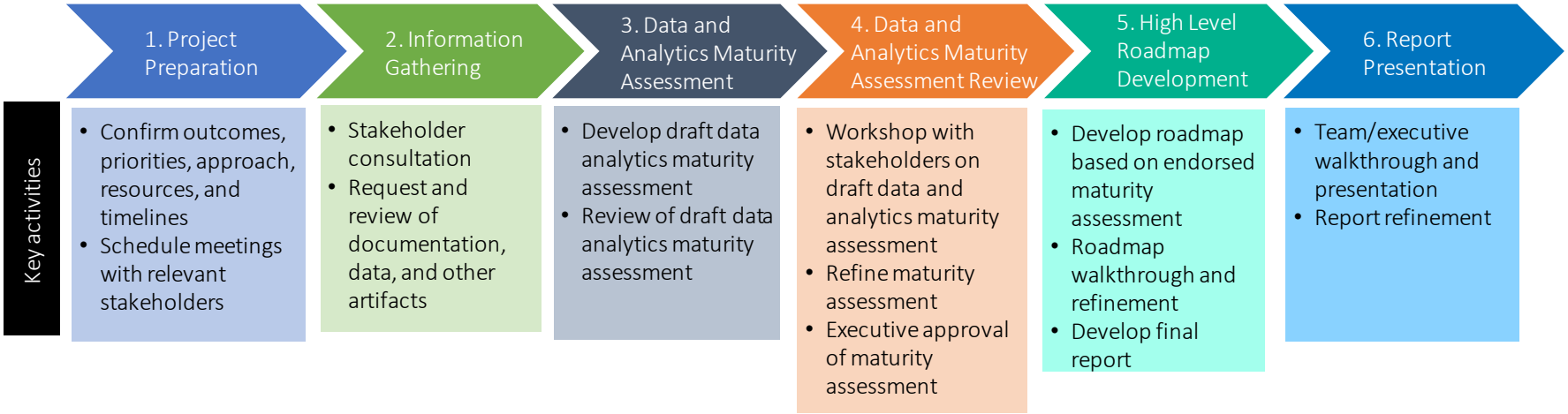
Further Details

Fraud & Control Maturity Assessment – What it is?

Recommendation: Maturity Assessment Specific to Fraud

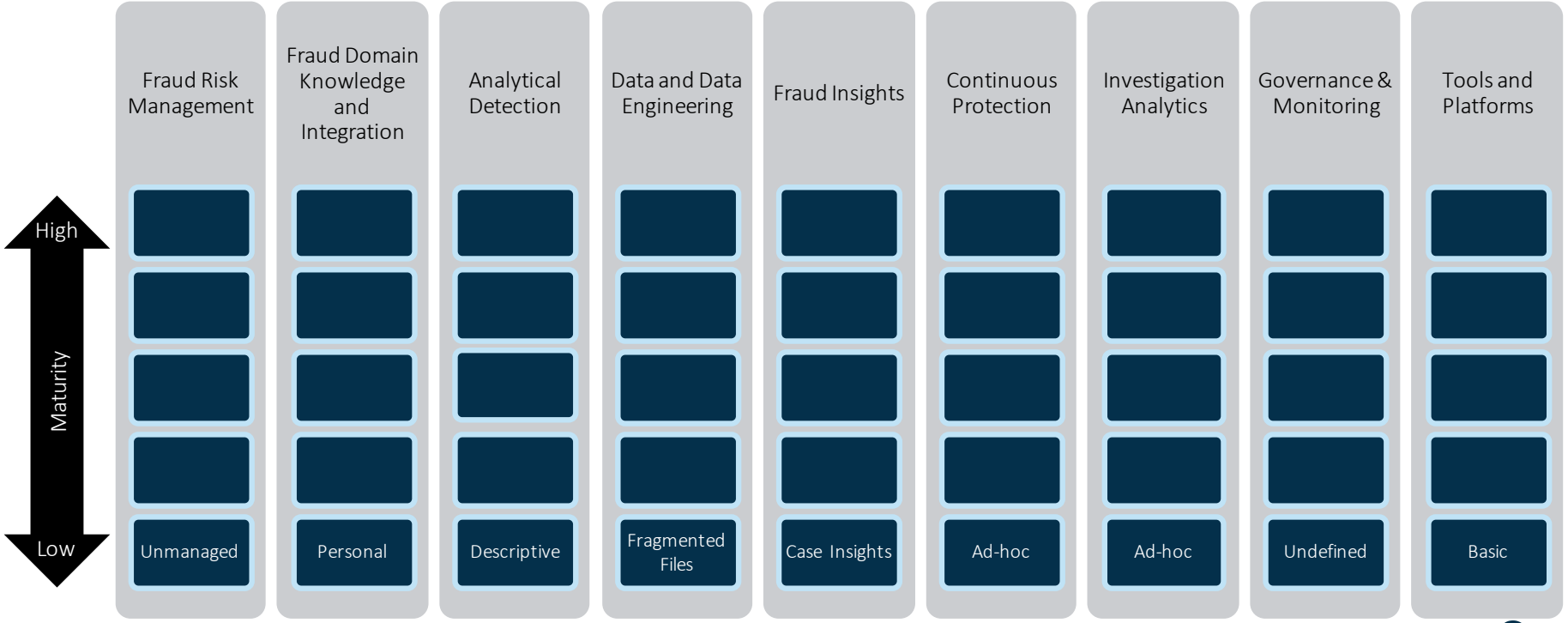
The aim of such a specific and detailed project include:

- Analysis and maturity assessment of the data analytics capability of the Compliance Intelligence and Analysis team
- Development of a data analytics roadmap based on the data analytics maturity assessment.
- A final report that includes the maturity assessment, best practice guidance on fraud data analytics, and a data analytics roadmap that supports the strategic objectives of the business.
- Consultation with stakeholders as recommended by the Department, and additional stakeholder consultation as required to inform the project deliverables.



Areas of Assessment for Fraud Analytics

The objective to become more proactive in fraud detection has been front of mind for this assessment. In this context, assessment areas have been found to be relatively immature, however this maturity assessment should not be viewed in isolation, and balanced against the high level themes and issues identified earlier in this report.



Further Details

Training Opportunities

Recommended SAS Training to Support Early Childhood & Childcare

Training Course	Description
Supervised Machine Learning Procedures using SAS VIYA in SAS Studio	This course combines data exploration, visualization, data preparation, feature engineering, sampling and partitioning, model training, scoring, and assessment. It covers a variety of statistical, data mining, and machine learning techniques performed in a scalable and in-memory execution environment. The course provides theoretical foundation and hands-on experience with SAS Visual Data Mining and Machine Learning through SAS Studio, a user interface for SAS programming. The course includes predictive modeling techniques such as linear and logistic regression, decision tree and ensemble of trees (forest and gradient boosting), neural networks, support vector machine, and factorization machine.
SAS Visual Analytics in SAS VIYA: Interactive Model Building	This course introduces SAS Visual Statistics for building predictive models in an interactive, exploratory way. Exploratory model fitting is a critical step in modeling big data. This course is appropriate for users of SAS Visual Analytics in SAS Viya 8.3
SAS Visual Data Mining and Machine Learning on SAS VIYA; Interactive Machine Learning	This course provides a theoretical foundation for SAS Visual Data Mining and Machine Learning, as well as hands-on experience using the tool through the SAS Visual Analytics interface. The course uses an interactive approach to teach you visualization, model assessment, and model deployment while introducing you to a variety of machine learning techniques

A matrix of training courses by cluster/branch is given in Appendix 7

Enrich Data Literacy

- Custom training program
- Blend of soft skill & technical skill courses
- Coaching and Mentoring
- Six months to complete
- Certification & Badges to achieve along the way

Data Science Fast Track

- **Module 1: Data Preparation for Analytics**
  - Programming on SAS Viya
  - Self-Service Data Preparation in SAS Viya
- **Module 2: Statistical Modeling**
  - Introduction to Statistical Concepts
  - Crafting compelling (and true) Data Stories
  - SAS Visual Statistics in SAS Viya: Interactive Model Building
- **Module 3: Machine Learning & Model Implementation**
  - SAS Visual Data Mining & Machine Learning in SAS Viya: interactive Machine Learning
  - Machine Learning using SAS Viya
  - Managing Models and Building Decisions in SAS Viya
  - Practice Exam / Exam + Certification
- **Module 4: Communication with Analytics Stakeholders**
  - Leading with Analytics
  - Communicating Technical Findings with a non-technical

Schools & Youth

Data Production

Schools Evidence and Analysis

Assessment and Analysis *

Vision:

The Schools Cluster, specifically the Assessment and Analysis branch provide design and analyse high quality longitudinal research studies for Schools. On-going, this area aims to continue to develop advanced analytical practices to maximise knowledge and drive effective policy

Analytics:

The primary focus of activity within this Section is around the following services:

- Delivery and design of research initiatives to drive and influence policy
- Large users of aggregated , school level data , additional manipulation done by the team
- Use of some advanced analytics – traditional statistical analysis
- Use of visualisation tool (Power BI)
- 'Fact checking' activity
- The Data Production Section provides a range of curated data sets as well as ad-hoc requests for data and information

.

* Additional conversations took place between SAS Advisory and the Data Production Team

Schools
Assessment & Analysis Branch

Priorities and Activities

- A team well versed in research, design, execution and evaluation techniques to drive policy for Schools
- High users of SAS, coding, datasets and analytical procedures are all used heavily
- Merging and aggregation of datasets is prevalent (for longitudinal studies) as is the use of additional data sources and assets (e.g. ABS Data lab)
- Use Power BI as opposed to SAS VA for visualisation

Observed Challenges

- SAS platform generally works well within A & A – any performance issues maybe due to laptops
- Old version of SAS VA has meant VA has been "left-behind" and Power BI now tool of choice in visualisation
- Limited geo-spatial analysis and mapping
- Some schools data will be going into NEEB e.g. school location from administrative data sources, survey data to be included in the future

Positive Messages

- No problems with the SAS Platform within Assessment and Analytics.
- Creativity with data is high – use of additional complimentary data sets is occurring
- Skilled and experienced research and analytically driven team – data journalistic capabilities are also well established – i.e. explanation of output
- Further advancement into Level 4 is possible and the way to maximise strong capabilities that currently exist

Summary Recommendations

- Demonstrate new functionality within SAS VA.
- Demonstrate SAS Visual Statistics (and down the line VDMML)
- Further use of multivariate techniques such as decision trees and random forests
- Longitudinal data analysis with discrete & continuous response training
- Explore SAS capability in Structural Equation Modelling
- Explore opportunities in Fraud
- Explore opportunities in forecasting opportunities and extrapolation methods

Schools

Data Production Branch

Priorities and Activities

- A team delivering curated schools data to the Assessment and Analytics team as well as responding to other ad-hoc data requests
- Provide counts and summary information to teams

Positive Messages

- Efficient repeatable processes in place to deliver data
- Use of Power BI will occur during early part of 2020
- Ambition is there to provide additional services and explore additional software and tools.

Observed Challenges

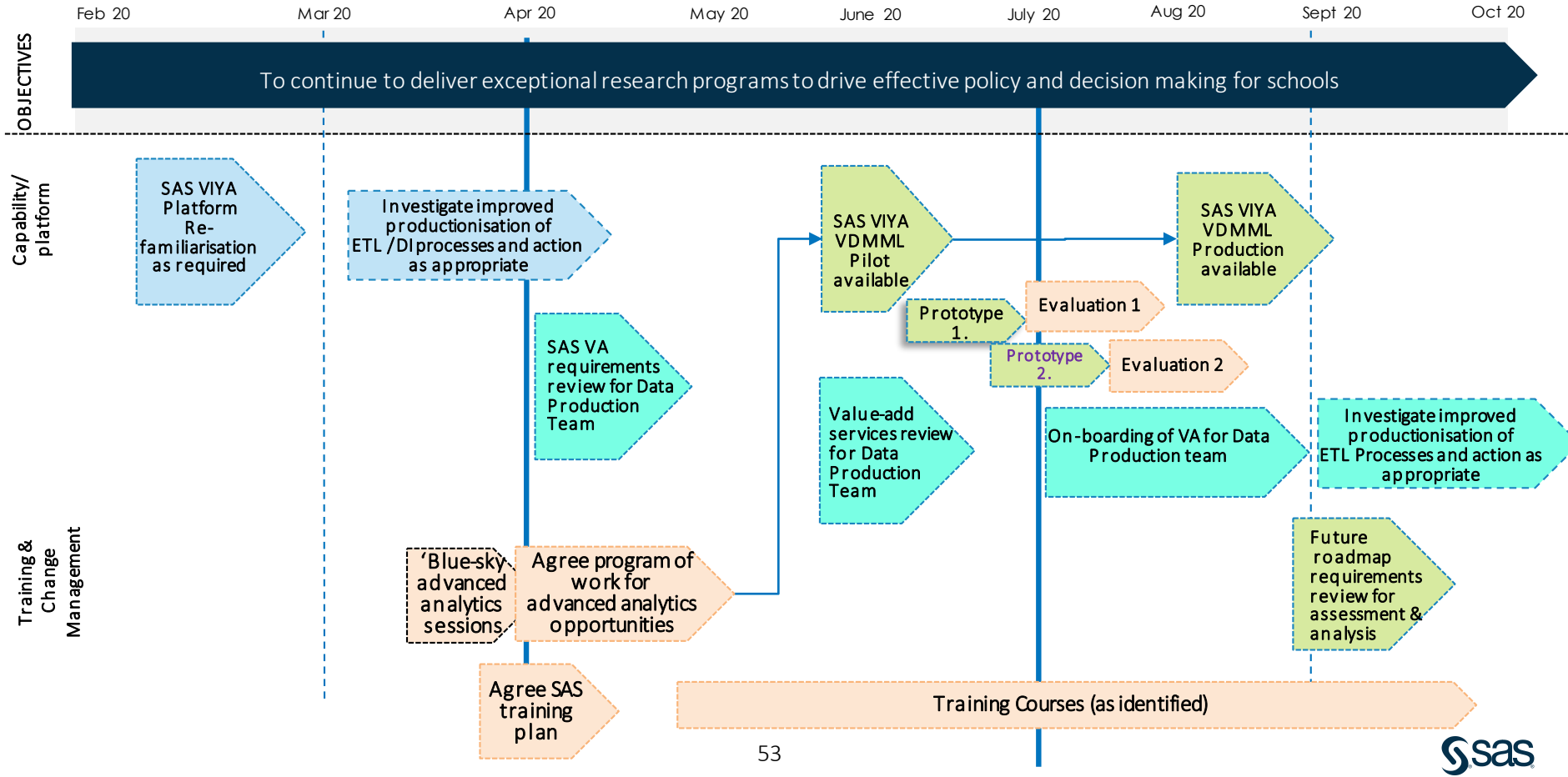
- SAS Grid has been somewhat unstable
- Do not use SAS VA as yet

Summary Recommendations

- Plan for additional service value – one such idea was a simple "Audit Service" which delivers a simple report of dataset content along with curated datasets
- Begin exposure and awareness to SAS VIYA and VA capability – plan for the future – there is a desire to investigate in the future (around mid-20202)

Recommended platform enhancements plan 2020

Schools



Detail for Schools & Youth
Schools Evidence & Analysis / Data Production

Capability & Platform
- Re-familiarisation with SAS VIYA (VA) capability
- Given new capabilities of platform or additional requirements within cluster – review current data manipulation and integration practices to identify if any additional efficiencies can be made using up to date tool set (applies to both Assessment and Analysis and Data Production)
- Given proposed Prototype for SAS VDMML (as proposed for Compliance, Intelligence & Analysis) it is recommended that Schools Evidence and Analysis is the *next group* to prototype the platform
- The will support additional capability and allow for the Evidence & Analysis branch to move up the Maturity Curve and develop into applying advanced analytics across the Schools domain

Recommended Components first 12 months
- Enhanced Visualisation and mapping in VA
- SAS Visual Data Mining and Machine Learning
- SAS Visual Forecasting

- **Recommended components beyond 12 months**
- This is harder to suggest – recommend a review following the VDMML Pilot availability and initial prototype activity, aligned with additional SAS VIYA roadmap entities will enable further recommendations at this point in the future

Planning & Innovation:
- For both branches – Payments & Collection and Assessment & Analytics, further in depth reviews of current practices in relation to data engineering and manipulation is recommended to identify where further efficiencies may be made
- If such opportunities are clear and acted upon ,this will, in turn, free up analytical resources to investigate how advanced methods and techniques can be embedded into their respective 'tool-kits'

Blue-Sky Thinking Days

Sessions structured to push the boundaries of innovation and identify where advanced technology and analytics can drive

Higher Education, Research and International

University Statistics Section

Analytics

Vision:

The University and Statistics Section within the Higher Education Cluster aims to be seen as a team focused on providing data and information to the highest quality, in the form of tables, files and statistics to end-users. This section is extremely knowledgeable in terms of the nuances associated with the data sources it works with.

Analytics:

The primary focus of activity within this Section is largely around the following services:

- Current users of SAS VA - provision of data centric solutions and providing a quality and efficient service to internal and external users
- Production of curated datasets (through Enterprise Guide, Base SAS and programming skills)
- Being responsive to a large number of ad-hoc requests for information
- Manipulation of SAS datasets and subsequent storage and cataloguing of assets

.

Higher Education, Research and International

University Statistics Section

Priorities and Activities

- A team well versed in derivation and provision of repeatable data assets
- Application of a rigorous set of data quality rules to check the data that flows into this team prior to further curation
- Filtering of datasets is a key operation performed by this team – knowledge of data intricacies and definitions is well understood

Observed Challenges

- Things appear to run relatively smoothly within this team in terms of day-to-day delivery, managing ad-hoc requests and documenting and storing data
- Perhaps an over reliance on coding at present - modification of existing Base SAS programs for future execution in SAS Viya is proposed
- Resistant to significant change and don't think NEEB will something they utilise

Positive Messages

- The SAS Platform is considered stable
- Data volumes are constant and not problematical in terms of manipulation and storage
- Processes in place to react to ad-hoc requests are well established and appear to work well
- There is no issue with key-man dependencies and reliable and accessible documentation is in place
- Team is happy with Power BI

Summary Recommendations

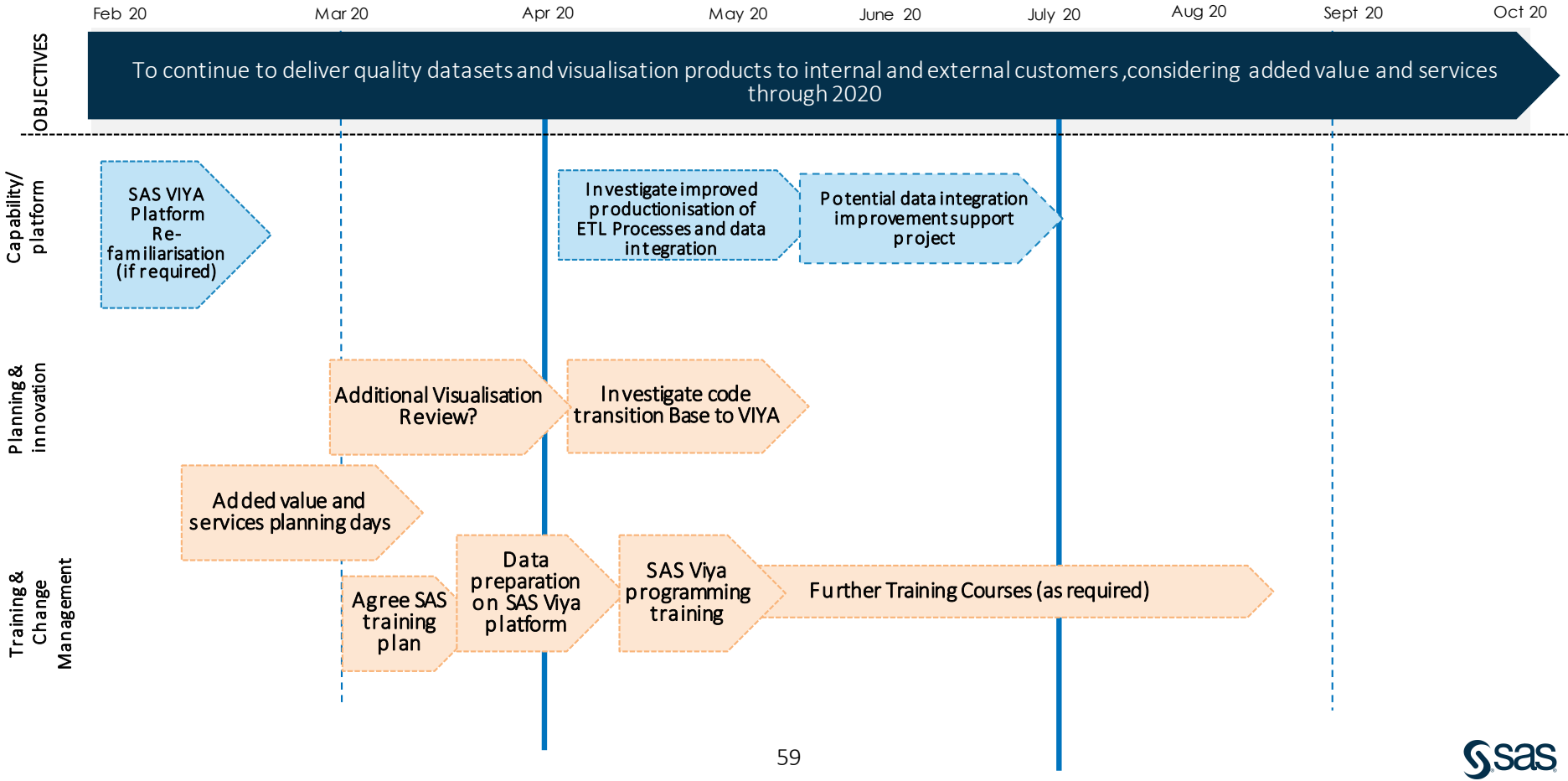
- Consider added value services that the branch can deliver, e.g. further advanced visualisation and analysis (compare Power BI and VA capability – not to replace Power BI, but enhance it)
- Additional collateral to be provided with dataset extracts (exceptions, significant changes from previous version, etc.)
- Further embrace SAS VIYA for a number of purposes including:
- Explore advanced features of data preparation, analytics, exploratory analysis and report creation in SAS VA

Plan for Delivery

University Statistics Section

Recommended platform enhancements plan 2020

University Section



National Education Evidence Base (NEEB)

NEEB & The Analytics Platform

Vision:

NEEB is a secure data integration platform that uses micro-services to bring together data assets in a flexible way and makes them available to staff through data analytics and visualisation solutions. This will answer multi-dimensional policy questions drawing on an ability to see individual journeys through Australia's education system, as well as enable policy and program areas to understand whole of education issues more clearly

Analytics:

SAS has been selected as the Analytics Platform to enable and support all users of NEEB to analyse data, over and above BI, reporting and dashboards. The SAS Platform will enable users to initially apply exploratory and confirmatory analysis techniques.

Key focus will be on:

- Delivery of accessible range of shared data
- The on-boarding and familiarisation program coupled with change management to drive users to the portal
- The use of additional visualisation and mapping (over and above Power BI capability)
- "Approachable Analytics" - develop capabilities for current maturity levels 1 and 2 to deliver additional insights and value
- The use of some advanced analytics , generally around traditional statistical analysis, hypotheses testing, exploratory data analysis etc.

.

NEEB
General Comments

Priorities and Activities

- Establish a data integration platform
- Increase data Departmental data literacy and capability building
- Change management, stakeholder engagement

Observed Challenges

- One of the biggest challenges will be the adoption of NEEB and driving users to the portal.
- Promotion of access to data and communication of what's available – articulation of benefits it can deliver
- Further adoption for use of the Analytics Platform will also be a priority once the data is accessible
- User requirements will be on-going, quick-wins and evidence will be required (showcasing of success) to drive teams to the portal
- NEEB's end-state is difficult to define and will be fluid for at least 2-3 years

Positive Messages

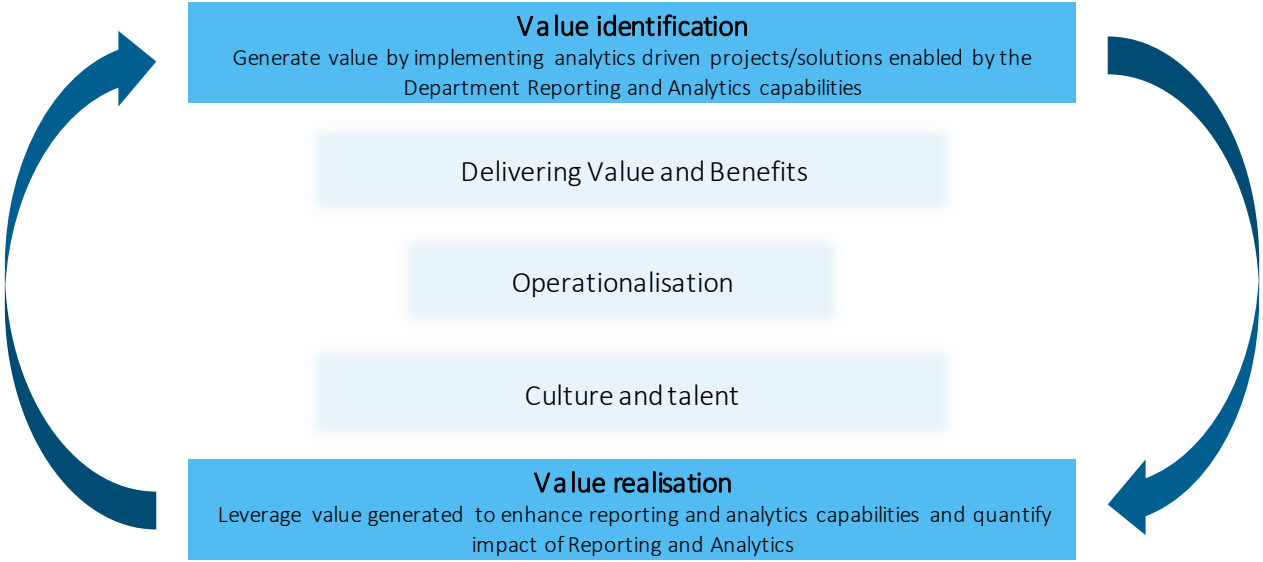
- Positive indication of initial 'quick-wins'
- Development of Information Services is underway – this will help define what can be achieved using the data
- Data sharing agreement and process flow finalised
- Stewardship manual is available
- Data management fundamentals being put in place: data catalogue, data quality framework, investment in training, data stewards manual
- Power BI has gained a footing in the department with demonstrated benefits
- Investment in training, and user forums established

Summary Recommendations

- Identify early-adopters and build capability and on-board teams
- Demonstrate initial functionality of SAS Platform to prospective users – identify potential Education program for Analytics (technology and general analytics)
- Invite prospective users to discuss specific use-cases and identify how data sources can support delivery of outcomes
- Follow-up with delegates from the initial SAS Day to identify potential opportunities and use-cases
- Establish process for feeding back success and findings

NEEB & the Analytics Platform

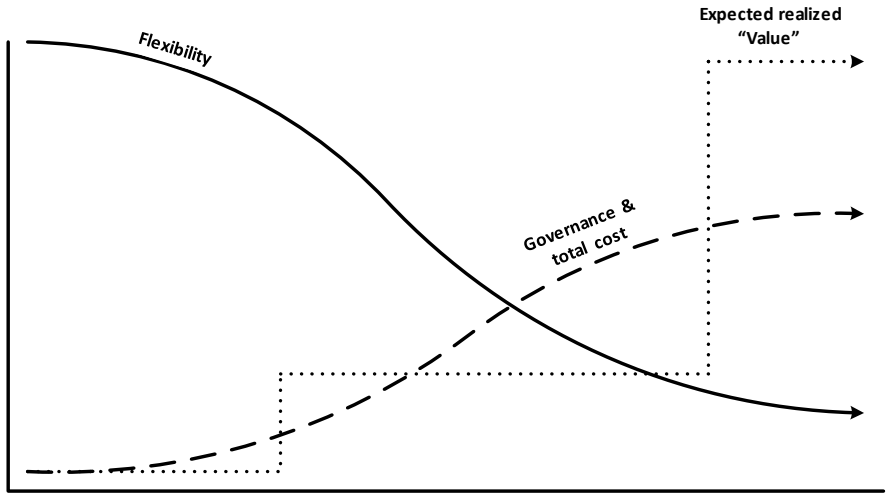
A strong framework for adoption of the Analytics Platform aligned to NEEB will enable the Department to rapidly generate value from the initial reporting and exploratory analytics capability, enabling further investment in order to fund further reporting and analytics investment needs.



Appendices

Supporting Material

Appendix 1. “Insight to Action”



Not all analysis needs go through all these stages.
For example, simply answering a challenging question might stop after exploration. By contrast, deploying a fraud model and integrating it with an organisation’s Decision support system will necessarily go right through to automation.

Overview

In moving from an idea to an answer to action, organisations go through four distinct phases. Each of these requires varying levels of flexibility and carry varying levels of cost. Each also require different skill sets and often imply different cost recovery models.

- **Experimentation:** Organisations often search for the “unknown unknowns”, trying to discover new opportunities or solutions through novel uses of information. At this point there is rarely a well-defined business problem to solve.
- **Exploration:** Once an interesting opportunity has been found, the organisation needs to find the answer to the identified opportunity. Exploration focuses on clarifying the “known unknowns” and is often driven by specific business requirements.
- **Prototyping:** Once a solution has been identified it needs to be validated for use. Prototyping is focused on generating and testing a candidate solution for business appropriateness.
- **Automation:** Finally, once the analytical solution has been identified, it needs to be migrated into an automated and robust process.

Appendix 2. “The Analytics Continuum”

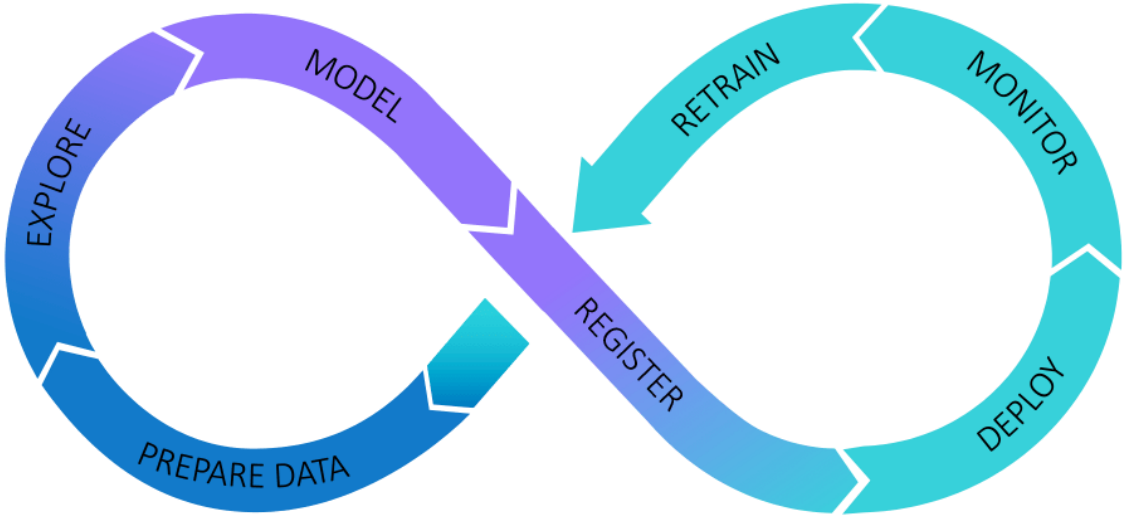
When working within analytics you are moving along a continuum of capabilities that cater for different needs.

Descriptive analytics - "What is currently happening or what has happened?" by querying data and summarizing key performance metrics, identifying segments of groups of entities etc. These types of queries provide valuable insights into the past and present to leaders/decision makers, the goal is to identify whether we are on track or not. If we have problems or obstacles then the goal is to ask additional questions.

Diagnostic analytics: "Why did it happen?" and is a more detailed, interactive type of analysis. Often utilizes different applications such as interactive visualization and data discovery tools. Identify outliers, determine correlation, trends, etc. Diagnostic analytics benefit from connection of data sources (and lots of other contextual data, which can make the reasoning about

Predictive analytics "What will happen?“ assesses patterns and likely outcomes using statistical or machine learning techniques. Two approaches - Driven by curiosity to find out whether observed trends will continue or driven by a ‘business’ need to assess a future outcome, for example, credit scoring to assess future risk of repayment.
Predictive analytics are considered more powerful because it provides the user with scope to take corrective action before the impact. Increasingly, and for more dynamic environments, predictive analytics can be adaptive, with a shorter time between model development and execution. In other words, as more data is gathered during the execution process, models can be retrained frequently.

Appendix 3: Adopt the "Analytics Lifecycle" and "ModelOps"

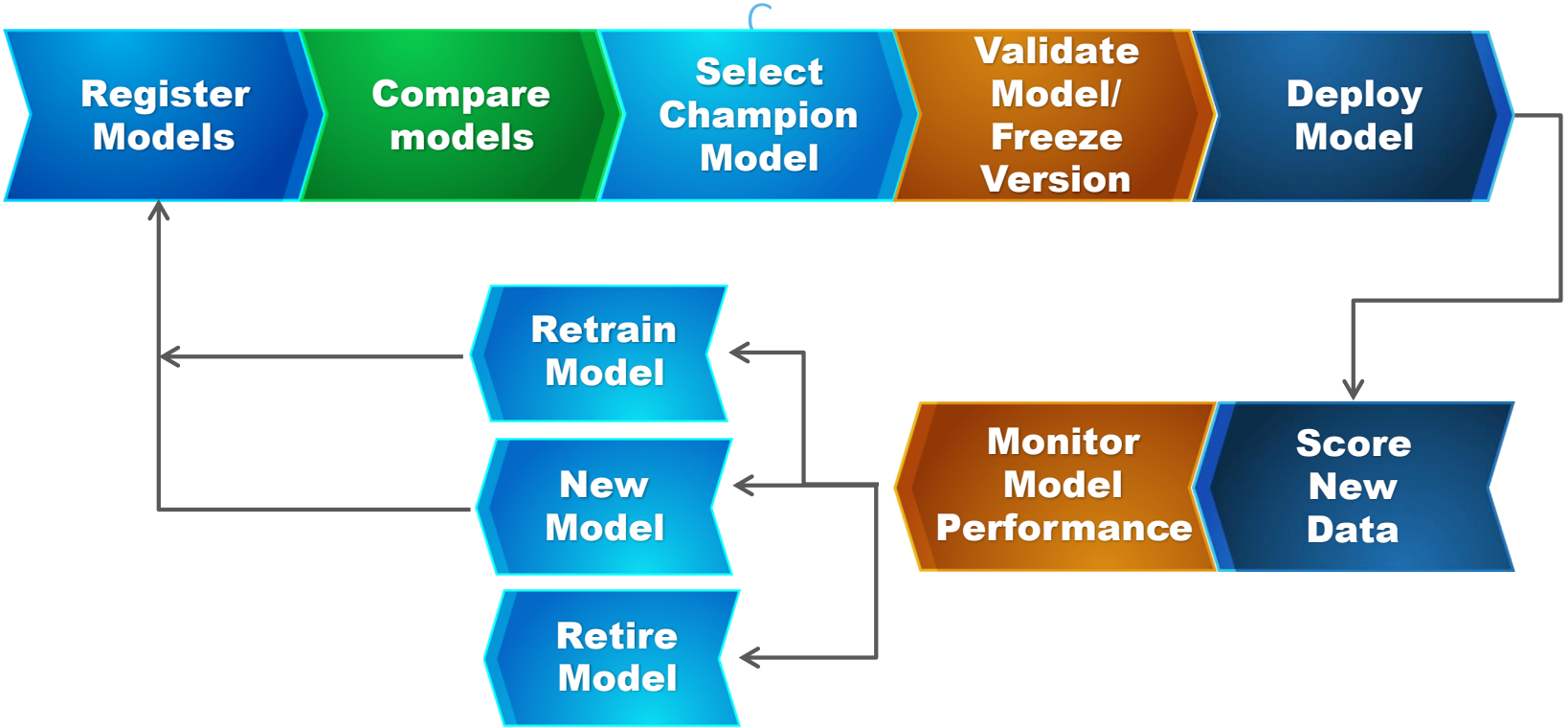


ModelOps is a holistic approach for rapidly and iteratively moving models through the analytics lifecycle. While ModelOps is based on the application development community's DevOps approach, where DevOps focuses on application development, ModelOps focuses on getting models from the lab, to validation, to testing, to deployment as quickly as possible, while ensuring quality results.

ModelOps – which encompasses culture, processes and technology – enables a smooth, efficient and continuous path to develop and deploy models.

Appendix 4: Model Comparison and "Champion / Challenger"

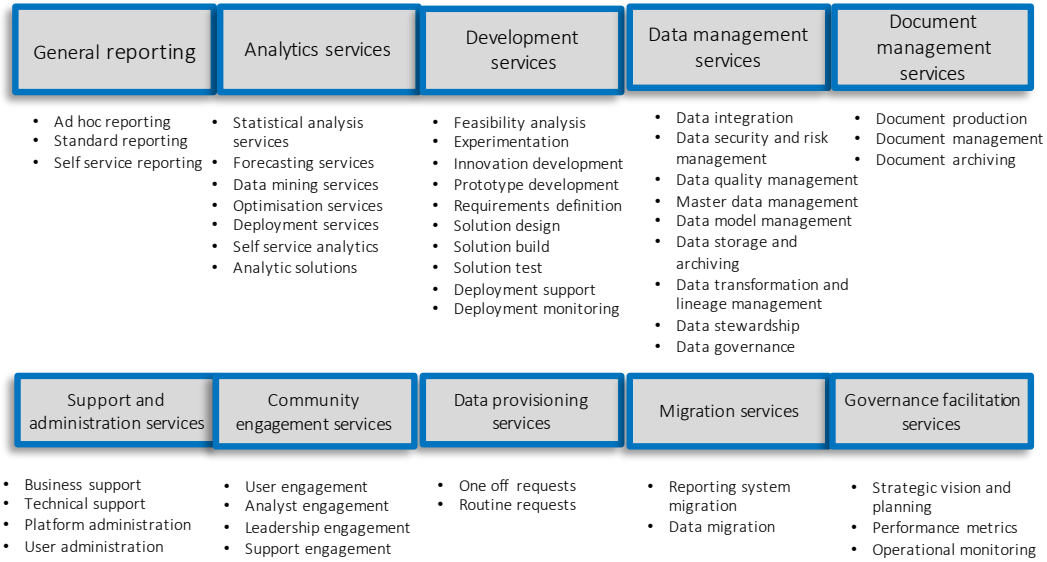
Prototype 3. Compliance Intelligence and Analysis



Appendix 5: Service Catalogue – Example Framework

Identify Service catalogue categories

As an example, the service catalogue may cater for the following categories of reporting and analytical services. Services are produced to serve the collective and specific needs of the internal and external stakeholder priorities. The following example is not Cluster specific but aims to show key categories and options



Appendix 6: Change Management to support new initiatives

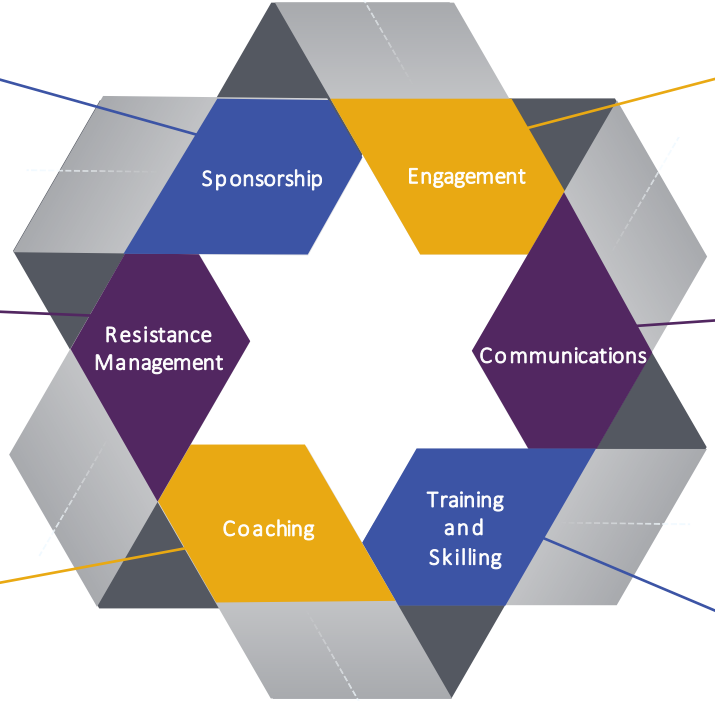
- This will be *multi-speed* – depending on clusters and branch, current and aspired position on the maturity continuum
- On-boarding – will be a key element of successful change
- Focus on transition to new platforms that operate in a different way to current status
- SAS will need to work alongside teams to make sure continuity of delivery of services
- Being aware that existing practices on new platform won't necessarily work – "it's what we've always done" attitude will have to change
- Change Management work so far covers not just NEEB's function of connecting the datasets, it covers the NEEB program as a whole.

NEEB Program Change Management

Approach: Applied a structured change management process and set of tools for leading the people side of change to achieve the desired business outcomes and benefits.

People: Embrace | Adopt | Use

Aim: Used the CDO and SRO to build a coalition of sponsorship with senior executives to ensure that change is communicated, supported and successfully implemented. Included:
- Executive briefing packs
- Executive communications
- SRO and CDO executive engagement



- User Forums, seminars, and showcases
- Data Management Sub Committee meetings
- NEEB roadshows
- Presentations and demonstrations in theatre and at team meetings
- NEEB program status reports at boards and Committees e.g. IC, EC and NEEB Program Board

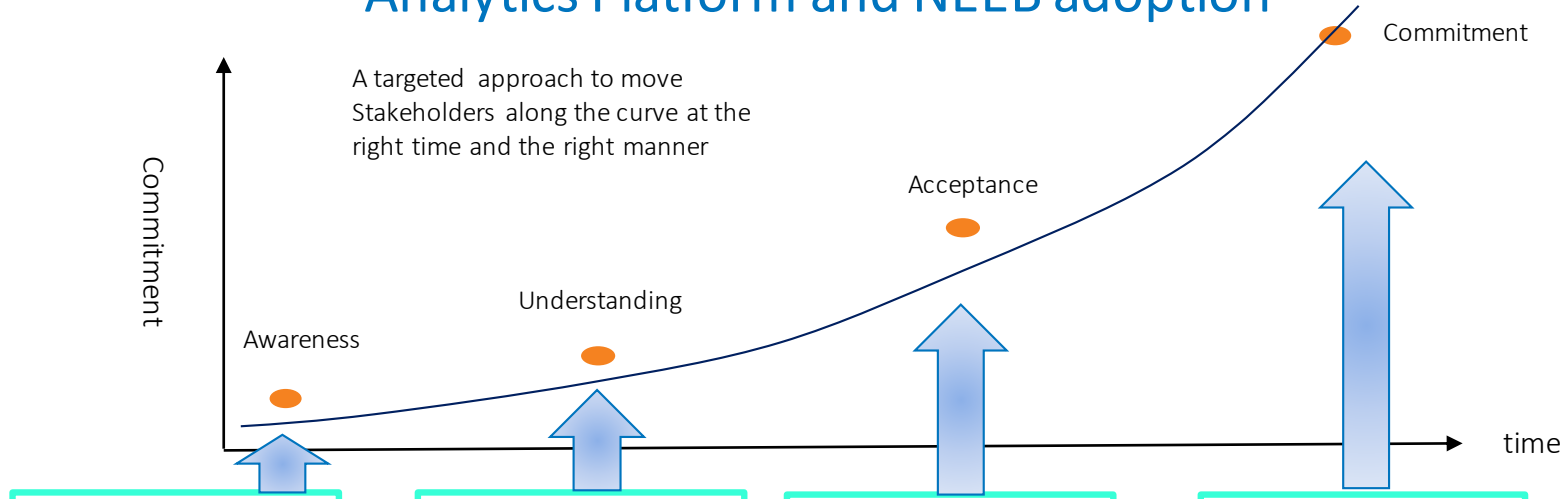
Working with stakeholders to:
- Listen to concerns and issues
- Removing barriers for participation
- Articulating the future vision (creates hope)
- Focusing on what need to change, why and what's in it for me?
- Demonstrating the benefits in a tangible way

Aligned with NEEB, Power BI, SAS etc. campaigns e.g.
- Data Stewards blog
- Intranet sites
- Case Studies
- Videos
- Corporate departmental communications e.g. tv slides, Staff Notices, CDO messages
- NEEB targeted communications

- Coaching showcase business to understand, adapt and manage changes
- Manage and work with key business areas and their managers
- Define and communicate the impact on business areas and stakeholders

- Data Discovery Seminar Series
- Power BI Training
- Power BI User Forums
- SAS Training
- SAS User Forums
- SAS Jump Starts
- Analytics in Action Day

Change Management Approach
Analytics Platform and NEEB adoption



Awareness
- Realises that a change in the process is imminent
- Articulates some of the changes the process involves
- May have some curiosity based on communications

Understanding
- Can articulate the differences between the previous and new process
- Asks questions to clarify concepts

Acceptance:
- Demonstrates positive support and endorsement
- Actively seeks information about their role in the new process
- Proactively inputs data and knowledge into system
- Champions process

Commitment:
- Incorporates the new processes as normal practice
- Demonstrates involvement with process and system
- Proactively sets objectives, initiates discussions (with senior staff)

Summary of NEEB Change Management Action Plan

✓ Done
⧖ In progress
➢ Future work

National Education Evidence Base

Sept to Oct 2019

Focus: Executive Engagement

Key Theme: What is NEEB?
(Awareness)

Progress:

- ✓ Intranet update
- ✓ 13/9 – update Data Management Sub Committee
- ⧖ TBA Nov - Email Communication from SRO to Exec colleagues
- ⧖ Executive face to face briefings
  - ✓ 18/9 – Corporate Strategy and HERI
  - ✓ 14/10 – Dep Secs and GMs
  - ✓ 18/10 – ECCC
  - ⧖ TBA Nov – Schools
  - ➢ TBA Skills, Secretary, Minister's Office
- ⧖ 22/11 - Implementation Committee
- ⧖ TBC - Executive Board

Oct to Nov 2019

Focus: Department-wide engagement

Key Theme: What is NEEB?
(Awareness)

Secondary theme:

- NEEB activities already undertaken in conjunction with Data Analytics Branch

Progress:

- ✓ Intranet update
- ⧖ Engagement with policy teams, data and middle managers (CX team)
  - ✓ 18/10 – Schools Policy Group Presentation
  - ✓ 25/10 – CDO email to stakeholders and program board re: PI4 Showcase
- ⧖ Departmental communications products approved for implementation. Includes:
  - 23/10 - Digital Sign
  - 21/10 - CDO Communication
  - 28/10 - Staff Notice
- ⧖ Develop communications and engagement plan and products for Nov-Dec theme

Nov to Dec 2019

Focus: General department-wide engagement, Middle Managers of Data areas

Key Theme: Imagine the potential of NEEB

(Awareness and Desire)

Secondary themes:

- The value of NEEB data for evidence-based policy
- More accessible data to support policy decisions

Progress:

- ⧖ Use schools NAPLAN CTG and Targets visualisations to show potential policy questions that can be answered by NEEB
  - 01/11 - Demonstrate to schools to get approval to use
  - ➢ 03/12 – Present at PBI Cloud launch showcase in Theatre

Feb to June 2020

Focus: Impacted stakeholders

Key Theme: What is in it for me?
(Desire)

Secondary theme:

- Role-based benefits tangible business deliverables showing progress

Progress:

- ➢ 7/2 - DAB and NEEB showcase in Theatre
- ➢ Change Management activities to be defined based around NEEB deliverables from each Program Increment which have a material impact on stakeholder/s e.g. job changes, business process changes
- ➢ Potential policy questions that can be answered by NEEB
  - ➢ Low socio-economic
  - ➢ Closing the gap

Appendix 7: Recommended Training Courses

Course	Early Childhood and Child Care Compliance Intelligence and Analysis	Schools & Youth Assessment & Analysis	Schools & Youth Payments and Collection	Higher Education, Research and International Market Analysis & Data	National Education Evidence Base (NEEB) [Initial users & early adopters)
Neural Networks: Essentials	✓				
Programming for SAS Viya		✓		✓	
SAS Viya and R Integration for Machine Learning	✓				
Self Service Data Preparation in SAS Viya		✓	✓	✓	✓
SAS Visual Analytics 1 for SAS Viya: Basics		✓	✓	✓	✓
SAS Visual Analytics 2 for SAS Viya: Advanced	✓				✓
Managing Models and Building Decision in SAS Viya	✓				
SAS Visual Statistics in SAS Viya: Interactive Model Building	✓				
SAS Visual Data Mining and Machine Learning in SAS Viya	✓	✓			

End of Document