



Predicting Pilot Training Performance Using Psychometry and Flight Performance Data

Jean-Francois Delisle, Manager AI Innovation, **CAE**

Karen Moore, CPsychol, CSci, AFBPsS, EuroPsy, MRAeS, **Symbiotics**



CAE and Symbiotics collaborated on a study on **pilot performance and pilot aptitude testing** using **AI methodology** to help the **pilot selection process** and enable a **better individualized adaptive training support** for our cadet during their initial training program.



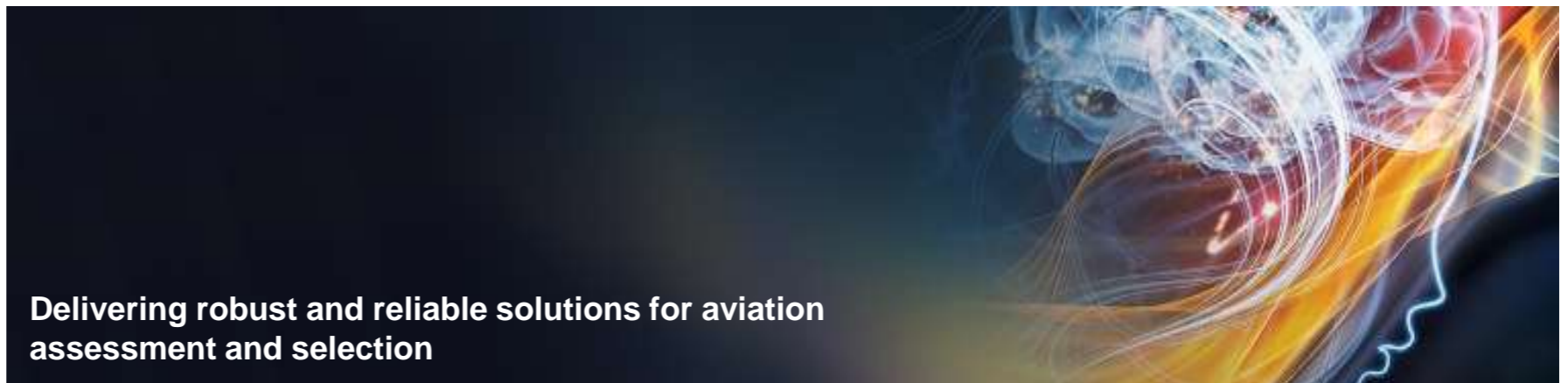
Making air travel safer



Prepare the defense forces for their missions



Helping medical staff save lives



Delivering robust and reliable solutions for aviation assessment and selection

Objectives of the study



Business Objectives

- Improve **predictability of success** rate on selection test & screening process
- **Accelerate cadet sourcing process** by refining the screening models
- **Individualized** support service based on aptitude and performance results
- Provide **adaptive learning** by suggesting Additional learning content
- Facilitate cadet **financing** through predictive license success
- Provide efficient **pilot pairing & assignment**



Study Objectives

- Create a dataset by combining Symbiotics and CAE **anonymize** data to comply with **GDPR**
- Explore if **Artificial Intelligence** can potentially help to determine which cadet groups (segments) will have the highest potential for success and support business objectives
- Identify the best **Learner Profile segmentation methodology** based on pilot aptitude testing and training performance
- Predict and explain a **cadet's probability of success**
- Provide a methodology to **standardize flight performance results** in a comparable way between courses

Flight Training Data Collection



CAE trains 1000 cadets / year

CAE trains approximately 1000 cadets annually across the globe (excluding CAE franchises)



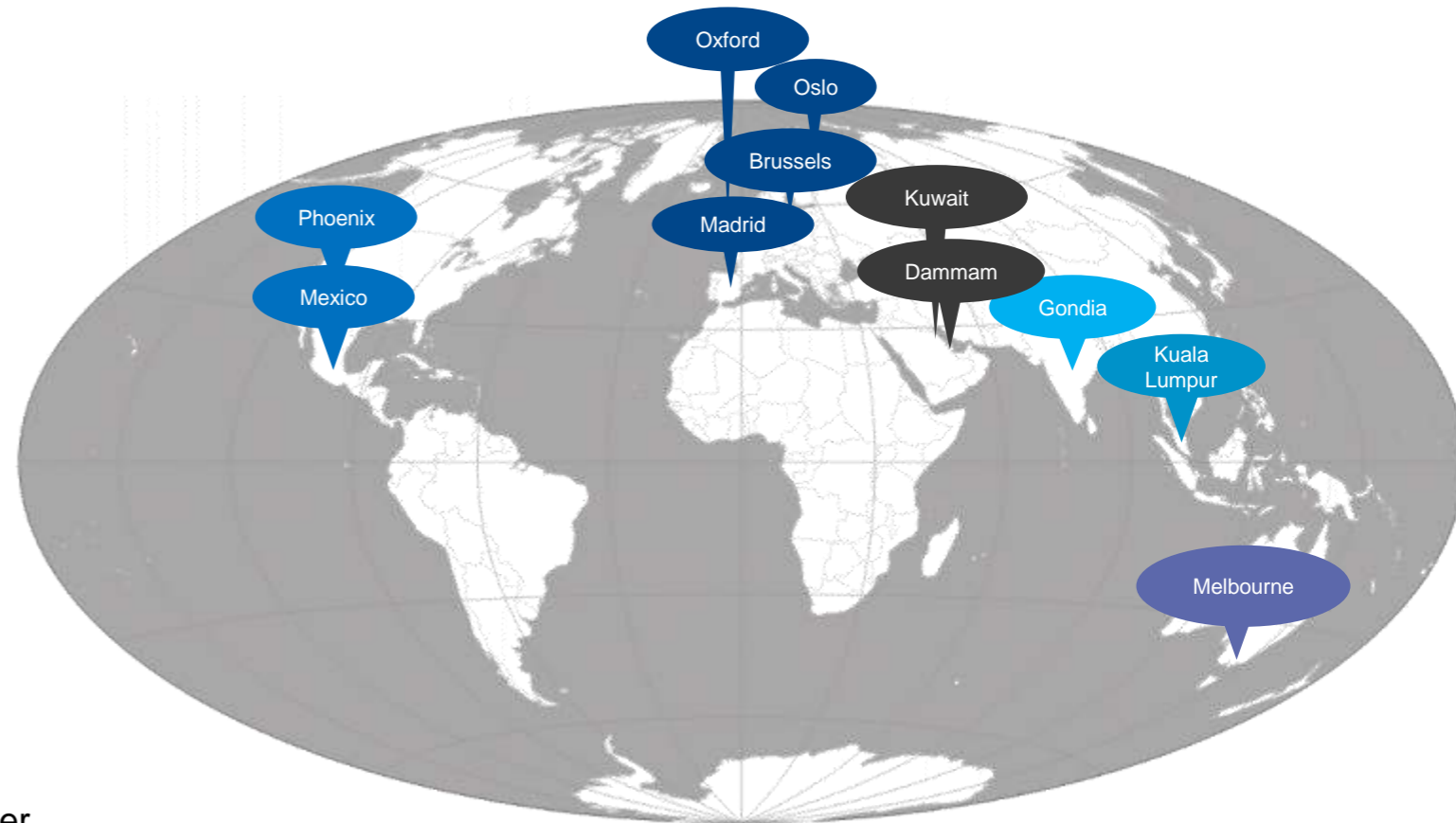
Phoenix and Oxford centers train the most cadets

The two training centers welcome about 50% of all cadets



The industry needs new pilots

Global requirement for 255,000 new airline pilots over the next 10 years
Need to develop 180,000 first officers into new airlines captains



- EASA / Europe
- FAA
- EASA / Asia
- CASA
- Directorate General of Civil Aviation
- CAE Authorized Training Center



Psychometric Data Collection



5500 aviation test completed monthly

Supporting assessments in over 130 locations, 80 countries, 5 continents



300 variations of assessments

Building on our current components tailored for specific client requirements and roles – e.g., Cadets, Pilots, Instructors.



Expertise, real world experience and benchmark data

Aviation specific norm groups coupled with over 2 decades of relevant industry expertise including delivering candidate interviews and group exercises.



Psychometric data collection



Data collected as part of cadet application process

- Delivered through ADAPT platform
- Selection of tests varies by school/client
- Common elements APQ (personality questionnaire) and Cognitive Reasoning
- APQ can be administered remotely, no right/wrong answers, untimed
- Cognitive Reasoning is timed test with right/wrong answers and requires invigilation
- These tests are often supported by tests of Maths, Physics and English Language
- Psychomotor skills tested through FAST, Flight Test and Ball Game



Output

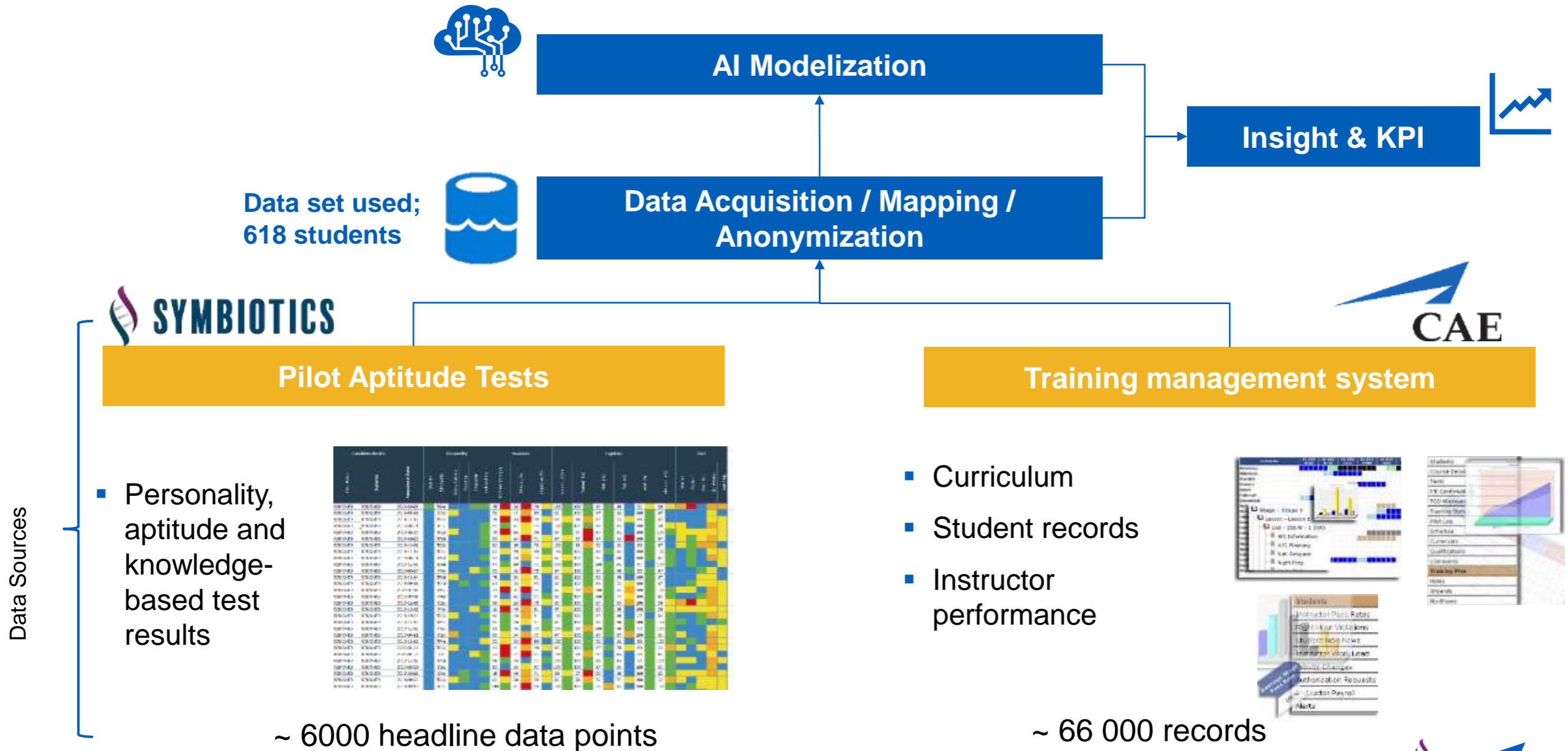
- Integrated report covering all test results
- Individual's scores are normed against aviation applicant norm group
- Traits reported as sten scores
- Scales overlaid with Red-Green 5 band scale to indicate levels of cadet readiness
- Cognitive Reasoning presented as percentile scores with Red-Green overlay
- Data held on ADAPT system for 24 months then anonymised and raw scores added to data pool

Our research question

Can we correlate Pilot Psychometric Test Results with their Performance during Flight Training?



Data Overview



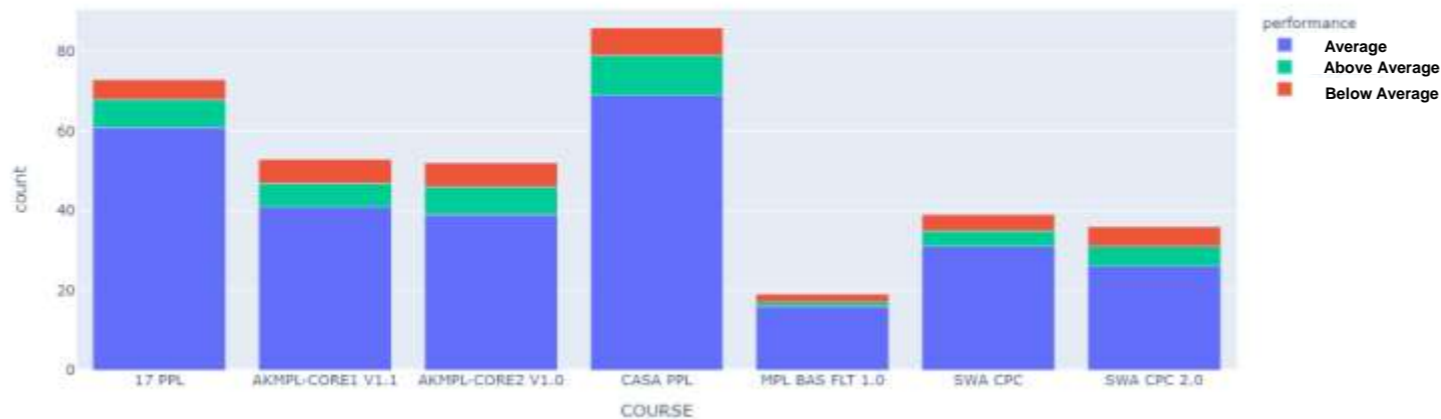
Pilot Performance Categorization

Mapping the performance KPIs to labels using statistical techniques.

- Performance has multiple dimensions
- Failing many line items in a single lesson is NOT EQUAL TO poor course performance
- Getting low grades in 10-line items in one lesson could be poor whereas in another lesson is not!

Using KPIs (lesson success rate, Probability of being a lesson topper, etc.), students are ordered and ranked.

Top 10% of the rank = **Above Average**, Bottom 10% rank = **Below Average**, Middle = **Average**.



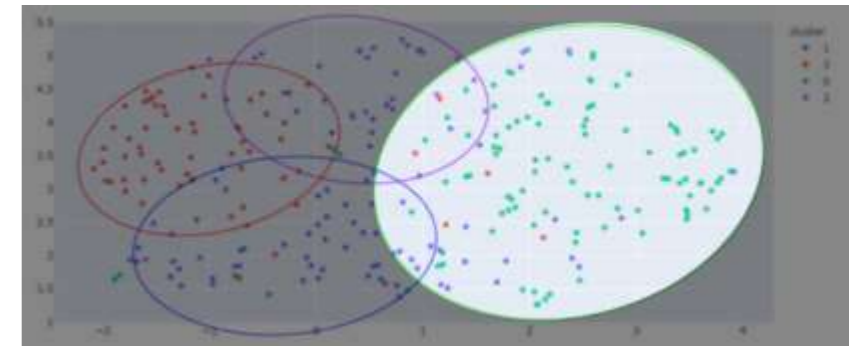
Overall performance distribution

Introducing Clustering using Symbiotics Psychometric Test Data



Components are values that correspond to a compression of 142 parameters in the aptitude test results. From the high dimension space, we are keeping data properties in order to evaluate distance between psychometric profile

Aptitude Cluster A

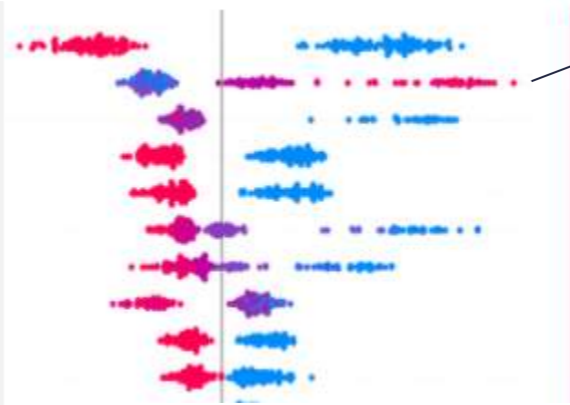


high stability score has a negative impact on cluster association

low stability scores have a positive impact on cluster association!

high Anger scores have a positive impact on cluster association

- Top 10 Contributing Aptitude Trait
- Stability
 - Anger
 - Character
 - Neuroticism
 - Derailers
 - Detached
 - Untrusting
 - Self Efficacy
 - Abnormal Traits
 - Development

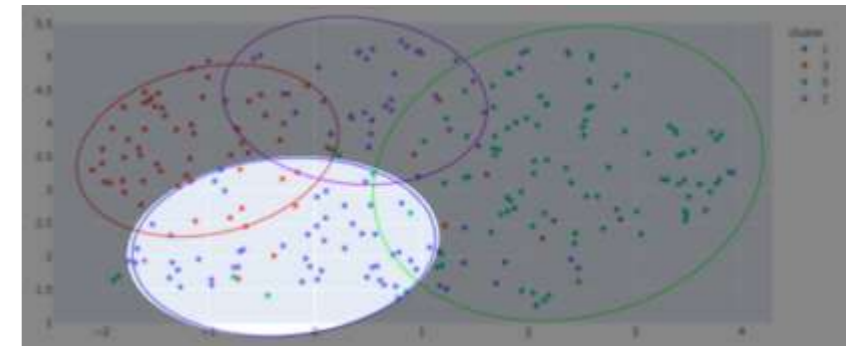


“SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations using the classic Shapley values from game theory and their related extensions”
<https://shap.readthedocs.io/en/latest/index.html>



Unstable, vulnerable, angry, introverted and non-trustworthy

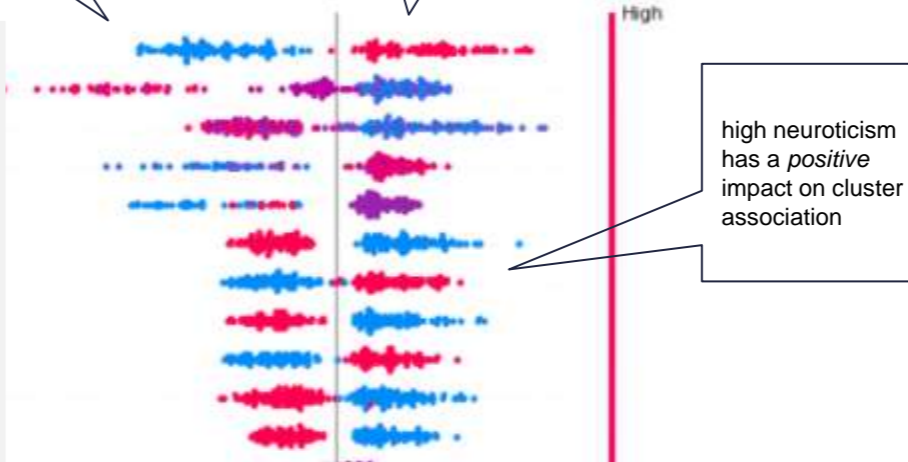
Aptitude Cluster B



low stability scores have a *negative* impact on cluster association

high stability scores has a *positive* impact on cluster association

- Top 10 Contributing Aptitude Trait
- Stability
 - Anger
 - Silhouette Match %
 - Perceptual Speed and Accuracy
 - Character Development
 - Neuroticism Profile
 - Abnormal Traits
 - Silhouette
 - Adherence

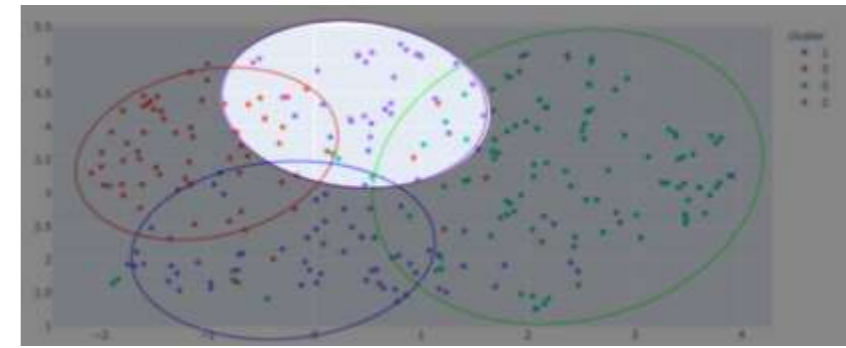


● attribute is high
● attribute is low

Not in cluster ← | → Part of cluster

Calm, stable, intelligent, introverted, low adherence, not trustworthy and not a stereotypical pilot

Aptitude Cluster C

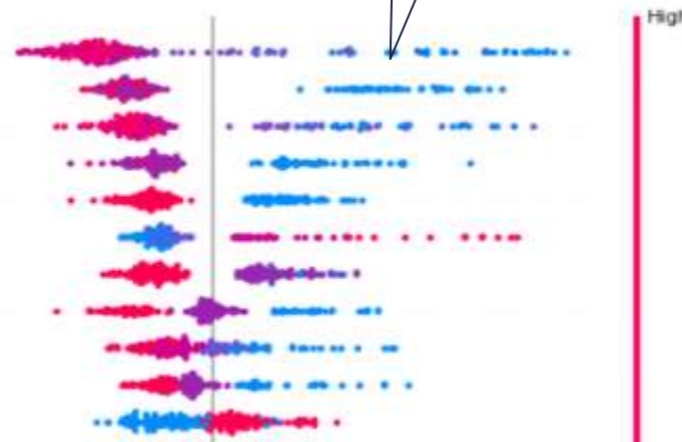


high cognitive scores have a *negative* impact on cluster association

low cognitive scores have a *positive* impact on cluster association

Top 10 Contributing Aptitude Trait

- Cognitive %
- Cognitive Sten
- Perceptual Speed and Accuracy
- FAST Cognitive overall
- FAST overall
- Psychomotor aptitude
- Working Memory
- FAST Cog 1
- Spatial Relationship Reasoning
- FAST Physical Stability

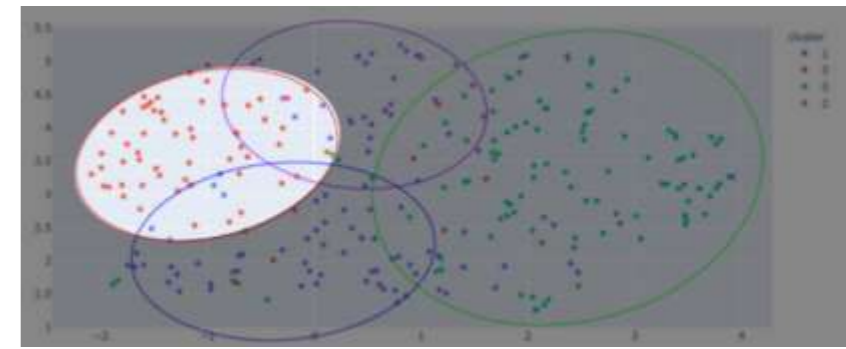


- attribute is high
- attribute is low

Not in cluster ← | → Part of cluster

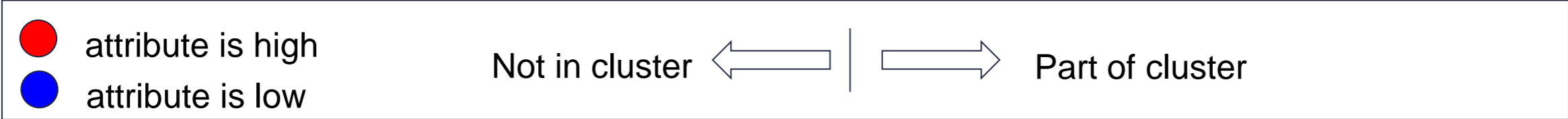
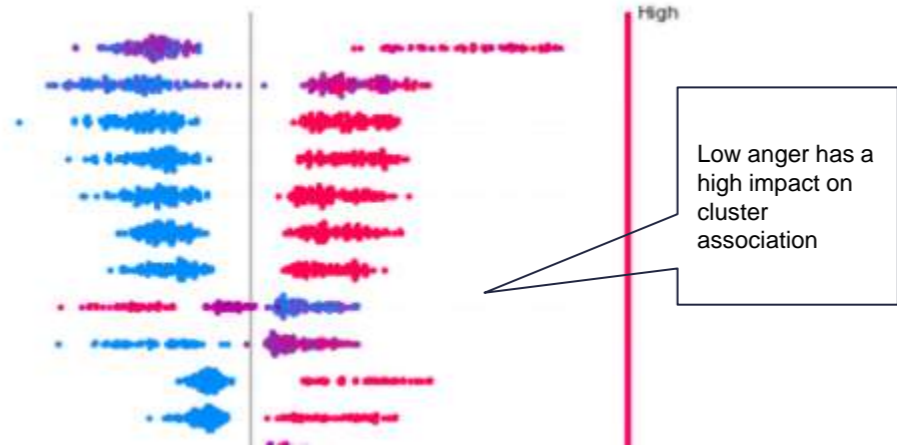
Stable, trustworthy, friendly and less cognitively capable compared to peers

Aptitude Cluster D



high scores in across the majority of these traits have a high impact on cluster association!

- Top 10 Contributing Aptitude Trait
- Self Efficacy
 - Silhouette Match %
 - Development Profile
 - Silhouette Stability
 - Adherence
 - Anger
 - Cognitive %
 - Overall Suitability
 - Company Mindedness

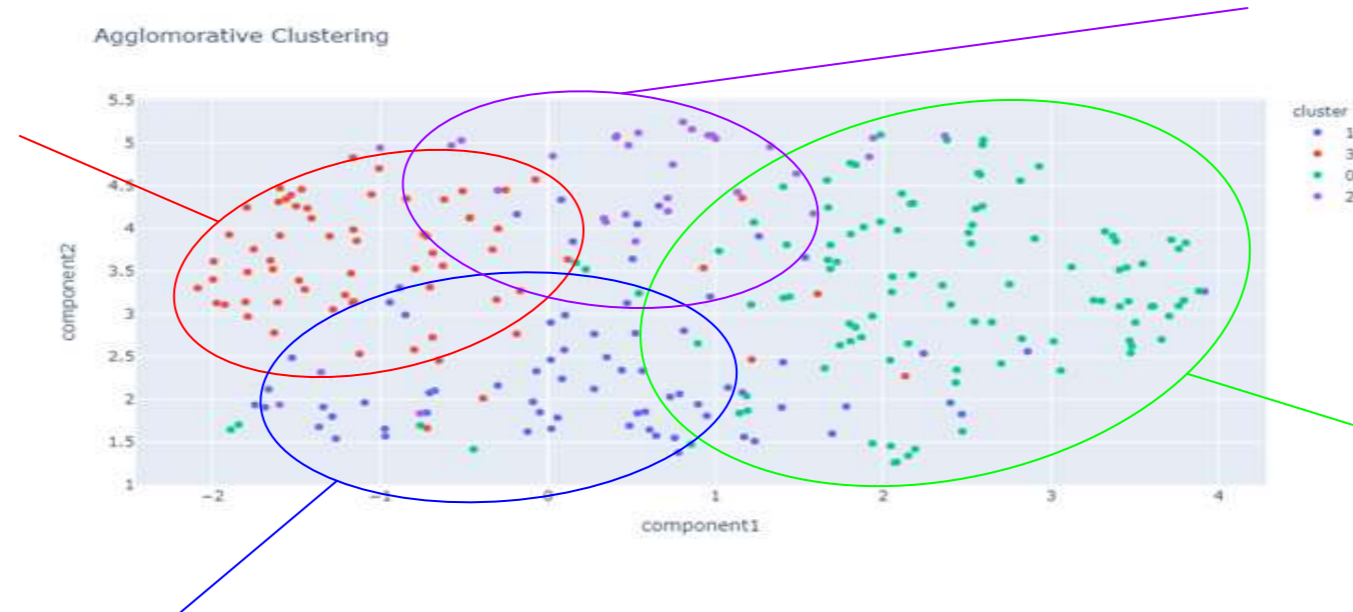


Self-aware, stereotypical pilot, stable, intelligent, friendly and scope to develop

Which Cluster Makes a Good Pilot?

Cluster D -
Self-aware, stereotypical pilot,
stable, intelligent, friendly and
scope to develop

Cluster C -
Stable, trustworthy, friendly and less cognitively
capable compared to peers



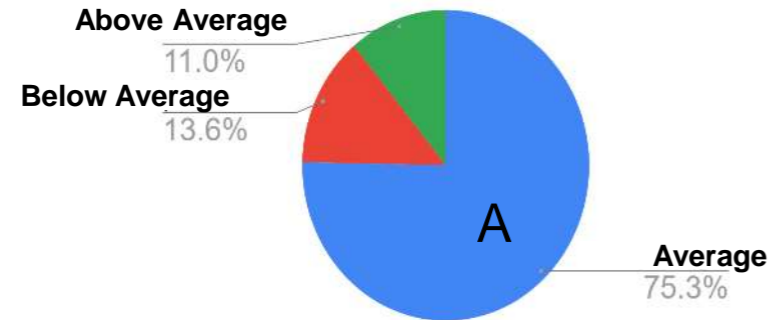
Cluster A –
Unstable, vulnerable, angry,
introverted, abnormal and non-
trustworthy

Cluster B –
Calm, stable, intelligent, introverted,
low adherence, not trustworthy and
not a stereotypical pilot,

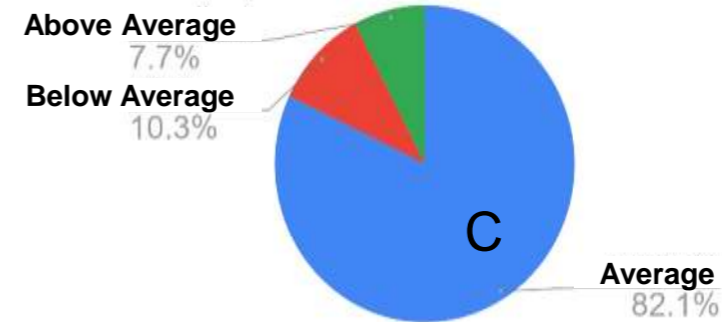
Exercise: Rank these clusters based on your own experiences! Who might perform best in Flight Training?

Comparing Performance against Aptitude Clusters

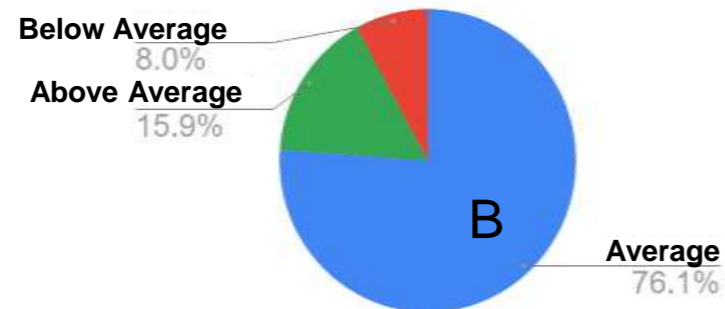
Cluster A - Unstable, vulnerable, angry, introverted, abnormal and non-trustworthy



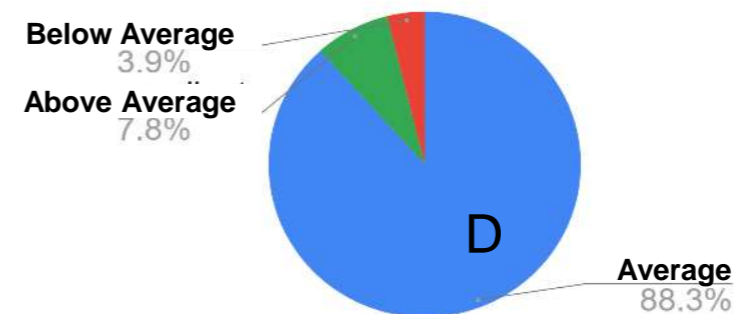
Cluster C - Stable, trustworthy, friendly and less cognitively capable compared to peers



Cluster B - Calm, stable, intelligent, introverted, low adherence, not trustworthy, not a stereotypical pilot.

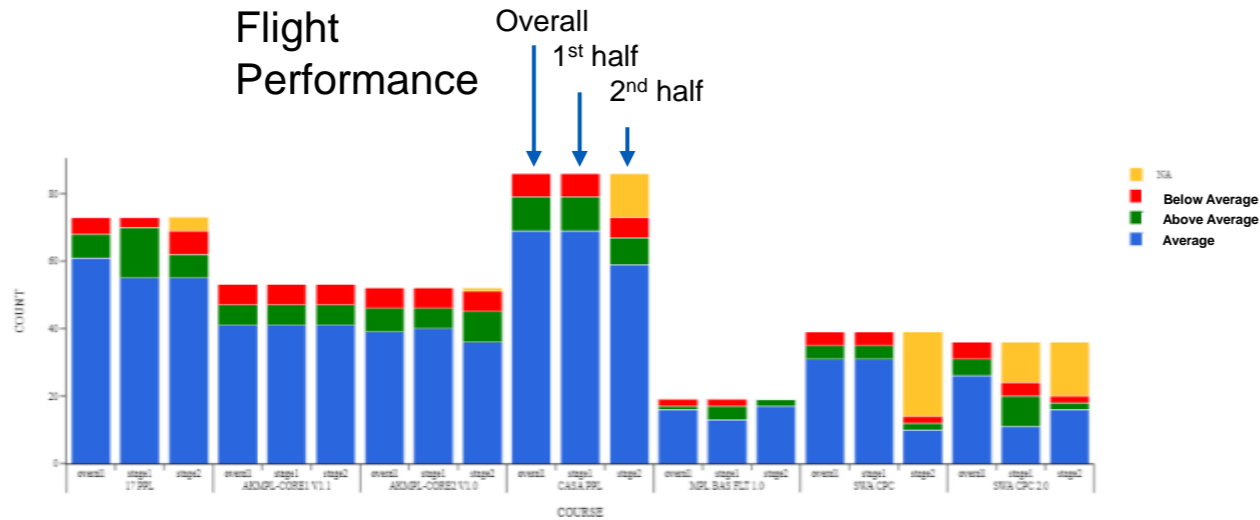


Cluster D - Self-aware, stereotypical pilot, stable, intelligent, friendly and scope to develop



Conclusions: Attitude problems lead to below average performance more than cognitive challenges. Intelligence has a role in excellence. Pilot stereotype doesn't mean excellence!

How about Temporal Performance?



Overall + Temporal performance distribution

This chart illustrates more variability in performance over time in clusters A and B. Whereas in clusters C and D their scores did not change 75% of the time.

Symb cluster	improvement	student count	perc count
A	0	84	61.3%
	-1	29	21.2%
	1	20	14.6%
	-2	2	1.5%
	2	2	1.5%
B	0	35	59.3%
	-1	16	27.1%
	1	5	8.5%
	2	2	3.4%
	-2	1	1.7%
C	0	22	75.9%
	-1	4	13.8%
	1	3	10.3%
D	0	47	75.8%
	-1	9	14.5%
	1	6	9.7%

1 => 1 step of improvement from stage 1 to stage 2.

Ex:

Below Average => Average (+1)

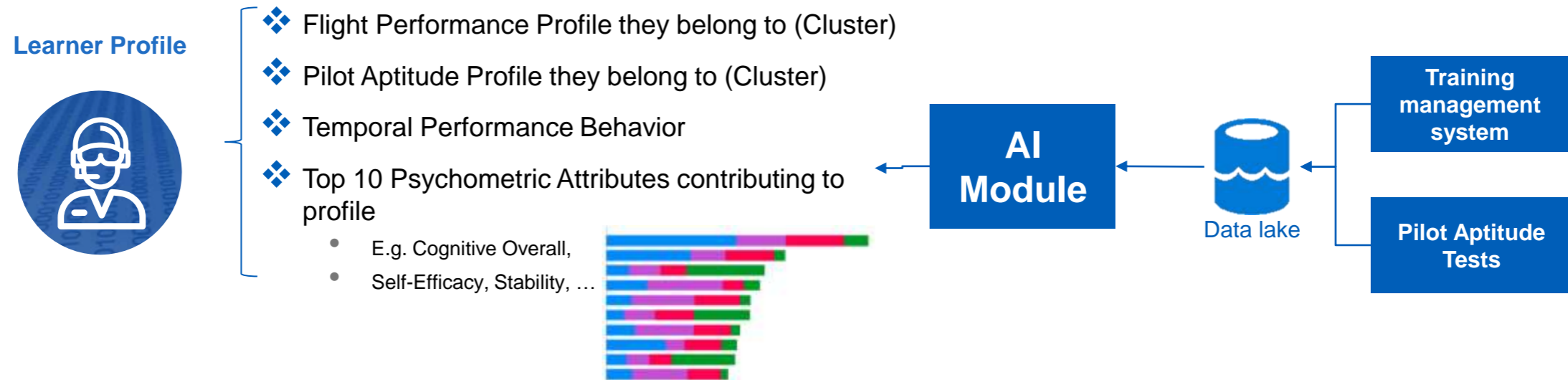
Above Average => Average (-1)

Below Average => Above Average (+2)

Cluster C & D have consistency in performance

How can we use this information as part of a Student Learner Profile?

- ✦ Predict flight performance or ideal career path
- ✦ Initial Aptitude test cluster assignment
- ✦ Key Traits of this Student contributing to their assignment to a cluster
- ✦ High level expected performance based on Aptitude Cluster assignment



Learner profile augmented with Pilot Aptitude can help us *identify Below Average* or average performers earlier and help them to *move to Above Average* profile with adaptive learning



Thank you!

