Data Analytics

Sarah Green - AFCAA

CLEARED
For Open Publication

Nov 02, 2022

Department of Defense
OFFICE OF PREPUBLICATION AND SECURITY REVIEW

VAULT Overview

- VAULT is an Air Force Cloud based platform
  - Access to modern data analytics tools – all the capability without issues that come with desktop versions
  - High speed compute in cloud with “clusters” of computers- can do in seconds what would take hours on your desktop with larger datasets and models
  - On NIPR & SIPR (subset of tools) – JWICS is on the Roadmap
- AFCAA “Tenant space” in VAULT
  - Dedicated storage and compute that only AFCAA can access
  - Opportunity for automation, collaboration and centralization

Realized Benefits

Legacy

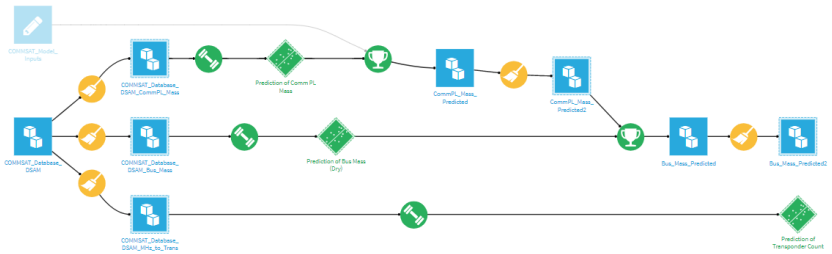
- Traceability is dependent on documentation & process used by analyst
- Only saved versions are kept- can lose trace to data in certain versions of models if not properly handled
- Often have issues with compatibility of desktop versions
- Mostly manual steps – not easily repeatable and often not well documented
- Extremely difficult to get desktop tools approved on high side
- Performance limited to desktop compute
- Models are tedious to update and are often several years outdated
- Org-wide changes like inflation updates need to be individually updated in each model manually
- Power of data science in cost community often limited due to very few programmers in the field
- Models built with manual steps tend to be very error prone

Cloud-based

- Complete step by step traceability to original, raw data
- Insight into who made change & when – and can revert to a previous version in modeling tools
- No compatibility issues with different versions of desktop software once in the cloud
- Automated steps from raw data to final product so that it's repeatable on new data that's received
- Can replicate environment on the high side (SIPR now, JWICS 2022)
- Performance will scale based on compute available in "cluster"-equivalent of groups of machines
- Can update models much more easily- often in days vs. months
- Ability to centralize updates so that analysts can pull them in to models quickly and easily
- Data science tools in a no code/low code environment
- Opportunity for significant reduction of modeling errors

Live Demo of AFCAA Data Analytics

Dataiku



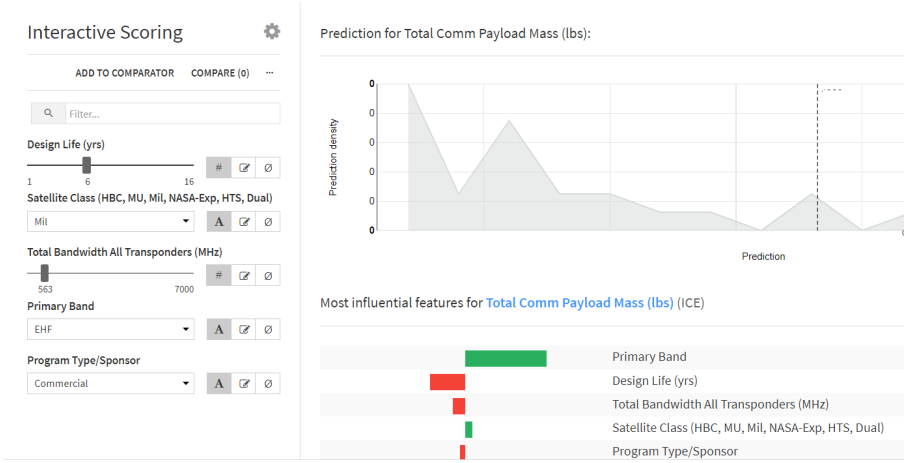
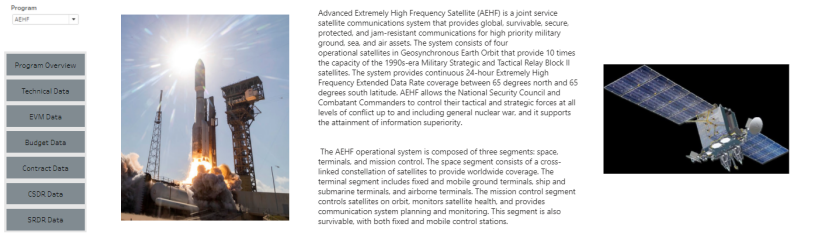


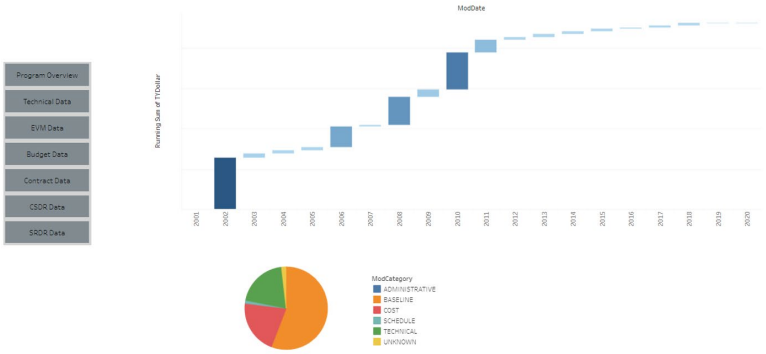
Tableau Server

Program Overview Dashboard



Timeline

Contract Data Dashboard



Progress to Date & Way Forward

- Began AFCAA VAULT Pilots in *Aug 2021*
- Platform Requirements Analysis (VAULT vs. Advana) was approved *Jan 2022*
- What comprises the Data Operating Model for Cost?
  - Cost Analysis VAULT Architecture + CADE + AFTOC
- Develop SAF/FMC POA&M/Roadmap for Data Operating Model *1QFY23*