**Welcome to VAULT 101, *Intro to Air Force Data Lab (VAULT)***

**15-16 February 2022**

VAULT 101 is a 2-day course designed to help analysts & engineers understand how to access Air Force Data Lab (VAULT) and utilize some of the applications within VAULT to perform data analysis in support of their mission.

The course will include instruction on:

* Python and R coding in Zeppelin and Hue
* Databricks and Tableau
* Data cleaning with Pandas
* Data analysis with Regression
* Delta Tables (advantages and basic operations)
* Data sources and Databricks integration
* Data types and calculated fields
* How to create a Dashboard

Course schedule overview

**15-16 Feb, Tues-Wed**: 0900 – 1530 (Mountain Time), with a lunch break planned for 1130-1300

***To get the most out of this course*, please accomplish the following prior to the course start date:**

***(Mandatory)* Ensure that you have access to AF Data Lab and the AFAC Tenant**

- Reference page 2 of this document & “*Getting Started*” slides for ‘step-by-step’ & visual guide

- Access to AFAC & AFOTEC Tenants

- Access to Databricks & Tableau

* AFAC Notebook Cluster on Databricks (should come with AFAC Tenant access)
* If you **don't have Tableau** in your VAULT applications, please go to: <https://public.tableau.com/> to create a **free** account! You'll be able to participate in the course exactly the same

***(Highly encouraged)* Take time to learn/refresh your coding skills, utilizing these training resources**

1. DataCamp
   * <https://app.datacamp.com/learn>
2. UDemy
   * <https://www.udemy.com/topic/python/>
   * <https://www.udemy.com/topic/r-programming-language/>
3. DigitalU
   * <https://digitalu.af.mil/app>
4. YouTube also has many good resources for learning about how to code in various languages
   * Focus on learning (or, refreshing) your R & Python programming skills
5. Additional courses & resources:
   * Automate the Boring Stuff with Python <https://automatetheboringstuff.com/>
   * Google Python Course <https://developers.google.com/edu/python>
   * MITx Computational Thinking using Python

<https://www.edx.org/xseries/mitx-computational-thinking-using-python>

*If you have any questions, please contact Capt Antony Brown at antony.brown@spaceforce.mil and*

*Lt Alec Sekelsky at alec.sekelsky.1@us.af.mil; they are the lead instructors & main POCs for this course.*

**To gain access to the AFOTEC and AFDV tenants, and the Databricks application…**

**First, establish an AF Data Lab (VAULT) account:**

* Get on AF Data Lab (<https://afdatalab.af.mil/>)
  + Select DOD EMAIL certificate to authenticate
* Create a new account

And then, once your new account is created…

**Gain access to the AFOTEC and AFDV tenants:**

* Go to “Help Desk & Requests” -> “Tenant & Application Requests” (menu across the top) -> select “Tenant Instances” (under Categories) -> pick ***AFDV***, then click ‘Order Now’, then ‘Checkout’
  + [Service Portal - Service Portal (servicenowservices.com)](https://afcdoprod.servicenowservices.com/sp) (“Help Desk & Requests”)
  + [Catalog Categories - Service Portal (servicenowservices.com)](https://afcdoprod.servicenowservices.com/sp?id=sc_category) (“Tenant & Application Requests”)
  + [Catalog Categories - Service Portal (servicenowservices.com)](https://afcdoprod.servicenowservices.com/sp?id=sc_category&sys_id=4e1492fa1b778c10870b419d9c4bcb37&catalog_id=-1) (“Tenant Instances”)
  + After completing above steps to request access to ***AFDV*** tenant, find (click ‘Show More’ at bottom of “Tenant Instances” page) & request access to the ***AFOTEC*** tenant

**Gain access to the Databricks application:**

* Go to “Help Desk & Requests” -> “Tenant & Application Requests” ->  select “Applications” (under Categories) -> pick Databricks, then click ‘Order Now’, then ‘Checkout’
* [Databricks - Service Portal (servicenowservices.com)](https://afcdoprod.servicenowservices.com/sp?id=sc_cat_item&sys_id=fb6a52b21bb78c10870b419d9c4bcbb7&sysparm_category=17a4d6ba1b778c10870b419d9c4bcbda) (“Applications”)
* Create a ticket with the help desk asking for AFOTEC Cluster permissions on both NIPR and SIPR.  This is necessary for you to use the cloud computing resources we have

**To request a SIPR VAULT account:**

* Go to the NIPR VAULT, Help Desk and Requests tile
* Open a Trouble Ticket requesting access to the SIPR VAULT environment
  + Also specify that you’d like to join the SIPR ***AFOTEC*** tenant
* Provide your SIPR Email and your DODID number (on your CAC).

Your account will be created in < 48 hours.

**Are you a 15A?  You get more free stuff:**

For 15As and 1515s you should also request access to the following tenant, this is owned by HAF/A9 and will continue to be available to you wherever you go:

* Go to [https://afdatalab.af.mil](https://afdatalab.af.mil/) using a CAC enabled computer, login with EMAIL cert (first time)
* Click on Help Desk and Requests block
* Click Tenant Applications
* Click Tenant Instances, Request Access to Existing Tenant
* Find ***AFAC***, click it
* Click Order Now

**Advantages of AF Data Lab**

* Store your data securely via the AFOTEC Tenant (NIPR is PII capable, SIPR)
* Use DoD datasets (growing all the time)
* Import public data from the web to help in your analysis work
* Free form ability to analyze
  + Import any package related to Python, R, Scala, Julia
* Share your analysis from data in real time
* Use other people's work, and improve on it
* Official solution