

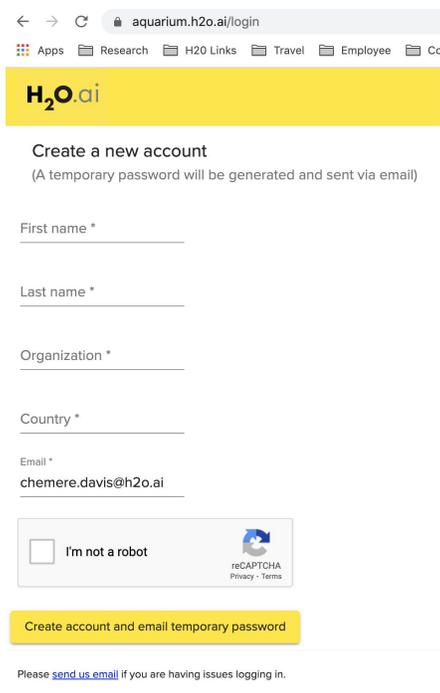
Module 1, Session 1: Getting the Right Dataset For Modeling: Data Munging & Data Curation

Supplemental Quiz Instructions

To ensure success with the quiz for Session 1, it would be helpful for the learner to use several hands-on labs through H2O.ai's Aquarium site. The following steps will walk you through how to launch the labs and exercises you will need for the assessment.

Create An Account on Aquarium

1. Navigate to the following site: <https://aquarium.h2o.ai/login>
2. Create a New Account



The screenshot shows the H2O.ai Aquarium login page. The browser address bar displays "aquarium.h2o.ai/login". The page features a yellow header with the H2O.ai logo. Below the header, the text "Create a new account" is displayed, followed by a note: "(A temporary password will be generated and sent via email)". The form includes several input fields: "First name *", "Last name *", "Organization *", "Country *", and "Email *". The email field is pre-filled with "chemere.davis@h2o.ai". Below the email field is a reCAPTCHA widget with the text "I'm not a robot" and a "reCAPTCHA Privacy - Terms" link. A yellow button labeled "Create account and email temporary password" is positioned below the form. At the bottom of the page, there is a link: "Please [send us email](#) if you are having issues logging in."

Launching the H2O-3 and Sparkling Water Lab

1. Once you've created your account on Aquarium and/or logged into the site, click on **Browse Labs** and navigate to **Lab 5: H2O-3 and Sparkling Water Training**

The screenshot shows a web browser at the URL `aquarium.h2o.ai/lab`. The page features a yellow header with the **H₂O.ai** logo. A navigation menu includes links for Apps, Research, H2O Links, Travel, Employee, Conferences, Github, Training, Python, and Meetings. On the left, a sidebar contains [Dashboard](#) and [Browse Labs](#). The main content area displays a list of labs, each with a lab ID, duration, and title. Lab 5, titled "H2O-3 and Sparkling Water Training", is circled in yellow. A yellow arrow points from the "Browse Labs" link to Lab 5.

Lab ID	Duration	Title
Lab 1	120 minutes	Driverless AI Test Drive (1.9.0 Experimental) (1 GPU)
Lab 2	120 minutes	H2O-3 and Sparkling Water Test Drive
Lab 3	120 minutes	Driverless AI Test Drive (1.8.7.1 LTS)
Lab 4	120 minutes	Driverless AI Training (1.9.0)
Lab 5	120 minutes	H2O-3 and Sparkling Water Training
Lab 8	120 minutes	Driverless AI Training (1.8.5.1)
Lab 10	120 minutes	Driverless AI LTS Training (1.8.4.1 LTS)

2. Click "Start Lab".

This step may take a few minutes for the AWS instance to be created and ready to access.

H2O-3 and Sparkling Water Training

Lab ID: 5
Lab duration: 120 minutes

Instructions:

1. Click the "Start Lab" button below.
2. *Note that the lab automatically ends after the number of minutes shown above!*
3. After the lab is fully started, URL links will appear below.
 - Jupyter: Click on this link and enter h2o as the token.
 - RStudio: Click on this link and enter h2o for the username and h2o for the password.
 - H2O Flow: This link is only available after starting H2O in Jupyter or R.
 - Spark: This link is only available after starting Sparkling Water in Jupyter or R.
4. Navigate to the h2o-3_hands_on (Jupyter or RStudio) or sparkling_water_hands_on (Jupyter only) directory.
5. Choose a lab to run.

Alternate ports

Note that standard ports have been internally mapped. This means:

- Jupyter Notebook: /jupyter/ instead of 8888
- RStudio: /rstudio/ instead of 8787
- Flow: /h2o/ instead of 54321
- Spark: /spark/ instead of 4040

3. Once the lab has been created, use the **Jupyter URL** link to launch the Jupyter environment
4. From the Jupyter home page, navigate to **/h2o-3_hands_on/lessons/Lesson 1: Introduction to H2O-3.ipynb**



Files Running Clusters Nbextensions

Select items to perform actions on them.

0 ▾ / h2o-3_hands_on / lessons

<input type="checkbox"/>	..
<input type="checkbox"/>	img
<input type="checkbox"/>	Lesson 1: Introduction to H2O-3.ipynb
<input type="checkbox"/>	Lesson 2: Gradient Boosting Models.ipynb
<input type="checkbox"/>	Lesson 3: Model Optimization and Grid Search.ipynb
<input type="checkbox"/>	Lesson 4: Feature Engineering Deep Dive.ipynb

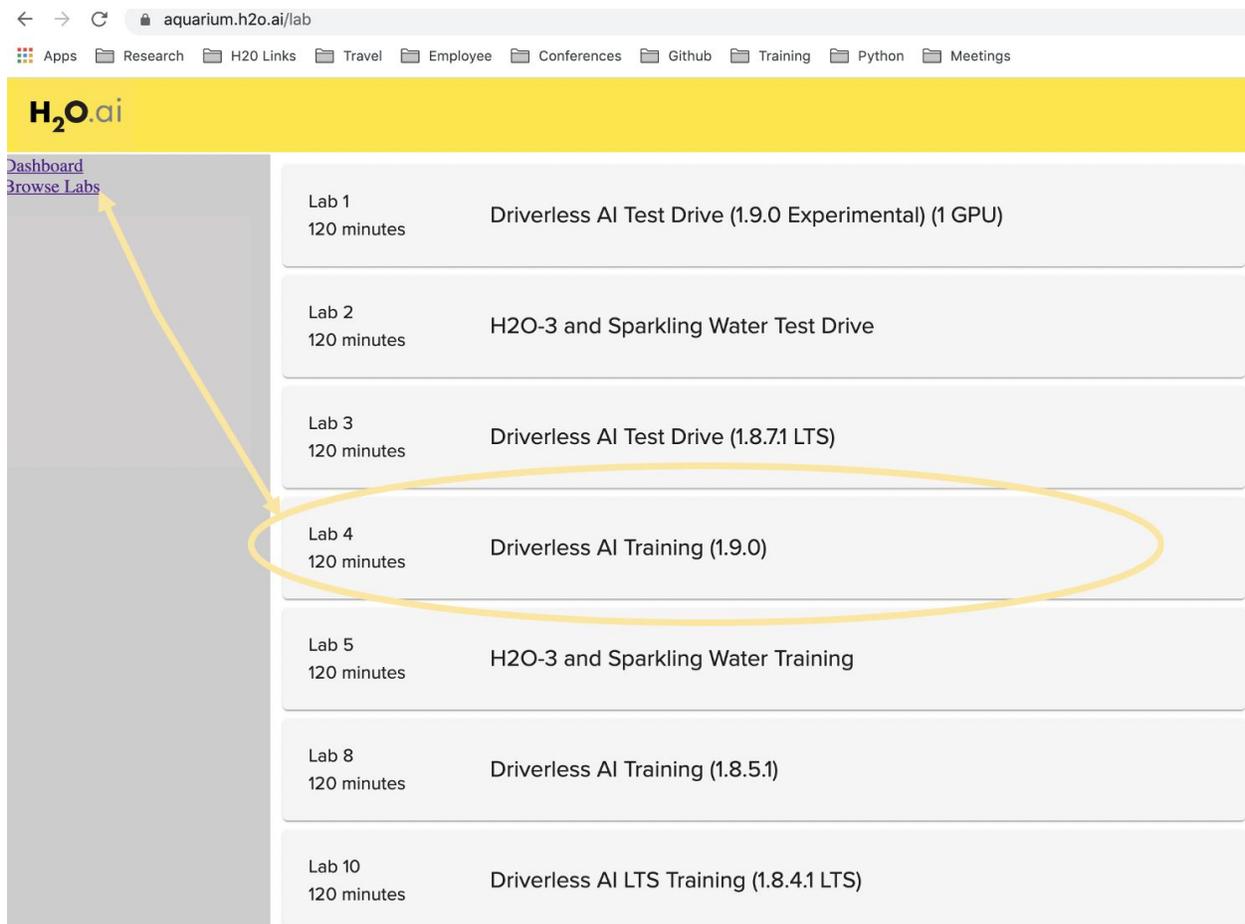
Exercise

From **Lesson 1: Introduction to H2O-3.ipynb**, follow steps 1-3. Note: You should also launch the Flow UI which needs to occur after the H2O Cluster is initialized in Step 1. You can either use the Flow link on the Aquarium page or follow the instructions in the notebook.

H2O-3 User Documentation: <http://docs.h2o.ai/h2o/latest-stable/h2o-docs/index.html>

Launching the Driverless AI Lab

1. From Aquarium, click on **Browse Labs** and navigate to **Lab 4: Driverless AI Training (1.9.0)**



The screenshot shows the H2O.ai Aquarium interface. The browser address bar displays 'aquarium.h2o.ai/lab'. The top navigation bar includes links for Apps, Research, H2O Links, Travel, Employee, Conferences, Github, Training, Python, and Meetings. The main content area features a list of labs, each with a duration of 120 minutes. Lab 4, 'Driverless AI Training (1.9.0)', is highlighted with a yellow oval. A yellow arrow points from the 'Browse Labs' link in the sidebar to Lab 4.

Lab ID	Lab Name
Lab 1	Driverless AI Test Drive (1.9.0 Experimental) (1 GPU)
Lab 2	H2O-3 and Sparkling Water Test Drive
Lab 3	Driverless AI Test Drive (1.8.7.1 LTS)
Lab 4	Driverless AI Training (1.9.0)
Lab 5	H2O-3 and Sparkling Water Training
Lab 8	Driverless AI Training (1.8.5.1)
Lab 10	Driverless AI LTS Training (1.8.4.1 LTS)

2. Click “Start Lab”.
This step may take a few minutes for the AWS instance to be created and ready to access.
3. Once the lab has been created, use the **Driverless AI URL** link to launch the Driverless AI platform
4. When the lab is loaded, accept the terms of use and enter the following credentials
Username: training
Password: training
5. View the **Driverless AI Data Recipes Demo video** linked in Module 1.

Exercise

- A. From the Datasets page, locate the CreditCard.csv dataset
- B. Click on [Click For Actions] and select the **Details**
- C. Review the details of the dataset
- D. Click on the **Modify By Recipe** button at the top of the Dataset Details page
 - a. Review the options available
- E. Navigate to the Live Code recipes in the H2O.ai Github repository using the below link.
- F. Click on the “sample_X.py” recipe. Copy the code and navigate back to Driverless AI.
- G. From **Modify by Recipe**, select Live Code and paste the code into the window.
 - a. Change the *fraction* variable to 0.2
 - b. For the *new_dataset_name* variable, rename the dataset to **creditcard_sample**
 - c. Click Save
- H. Review the dataset details of the new dataset.

Driverless AI Data Recipes Github Repository:

<https://github.com/h2oai/driverlessai-recipes/tree/master/data>

Driverless AI Data Recipes (Live Code) Github:

<https://github.com/h2oai/driverlessai-recipes/tree/master/data/livecode>

Driverless AI Documentation:

http://docs.h2o.ai/driverless-ai/latest-stable/docs/userguide/release_notes.html#version-1-9-0-july-27-2020