



✓ **Congratulations! You passed!**

TO PASS 80% or higher

Keep Learning

GRADE
100%

Graded Quiz: Test your Project Understanding

LATEST SUBMISSION GRADE

100%

1. How do you start an H2O cluster using the Python API?

1 / 1 point

- ☐

```
1 import h2o
2 h2o.start()
```
- ☒

```
1 import h2o
2 h2o.init()
```
- ☐

```
1 library(h2o)
2 h2o.init()
```

✓ **Correct**
Correct!

2. How do you convert a pandas DataFrame **df** into an H2OFrame?

1 / 1 point

- ☒

```
1 import h2o
2 df = h2o.H2OFrame(df)
```
- ☐

```
1 import h2o
2 df = df.to_h2oframe()
```

✓ **Correct**
Correct! We did exactly this in Task 4 of this hands-on project.

3. What does the *enum* data type represent in an H2OFrame?

1 / 1 point

- ☐ Numeric data
- ☐ Ordinal data
- ☒ Categorical data

✓ **Correct**
Correct! Enum is a categorical data type. The name is derived from the name for categorical data in Java.

4. What aspects of the data science and machine learning pipeline does H2O AutoML seek to automate? (Select all that apply)

1 / 1 point

☒ Data Preprocessing

✓ **Correct**
Correct! H2O AutoML takes care of data imputation, one-hot encoding, and standardization as part of data preprocessing

☒ Model Generation

✓ **Correct**

Correct! H2O AutoML automates the process of training and tuning various families of models using random grid search. Random grid search combined with stacked ensembles is a powerful combination.

☒ Ensembles

✓ **Correct**

Correct! Ensembles often out-perform individual models. H2O AutoML trains two ensembles using stacking. Ensembles perform particularly well if the models they are based on:

1. are individually strong, and
2. make uncorrelated errors

Stacking uses a second-level metalearning algorithm to find the optimal combination of base learners.

☐ Feature Engineering

5. Once H2O AutoML trains the models, cross-validations must be done separately. True or False?

1 / 1 point

☐ True

☒ False

✓ **Correct**

Correct! H2O AutoML automatically trains and cross-validates models, where 5-fold cross-validation is performed by default. One can specify the *nfolds* parameter ≥ 2 for the number of folds for k-fold cross-validation of the models in the AutoML run.