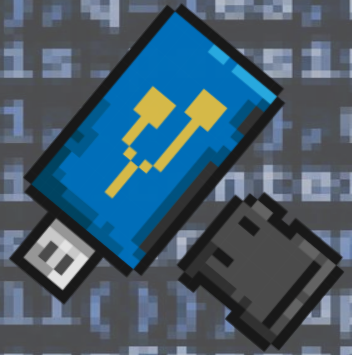
start [] = LEVEL 3 .exe



More data

You found the password for the ZIP file! Great! Now you can open it and look at the files inside (or you can ownload them from the contest platform).
The ZIP file contains two files:

- train_data.csv
- test_data.csv

The data scientist in you immediately recognizes that this is a dataset to train a machine learning model with. Like any good data scientist, before training a model you familiarize yourself with your data.

More data

While doing that you find out that the training data consists of these 9 columns:

Column Name	Data Type	Possible Values / Range
MODE	String values	auto, beam, burst, REDACTED
POWER	String values	high, low
AMPS	Floating-point values	[0,1]
VOLTS	Floating-point values	[0.02,8.7]
TEMP	Floating-point values	[-100,373.15]
UNIT	String values	K,C,?
DELTA	Floating point values	[-1,1]
GAMMA	Floating point values	[-2,2]
OUTPUT	Floating-point values	[-8.4,9.3]

notice that the test data contains the same columns, except for the OUTPUT column. So your plan is to train a model that can predict the missing OUTPUT values in the test set.

More data

Before you can start training you have to prepare your training data and fix some problems in the data set.
There are two problems with the training data:

- One column contains several NaN (Not a Number) values.
- One of the floating-point columns contains some outliers that do not fall into the range given in the last slide.

Task:
Find and remove these problematic rows from the training data.

Output:
What is the number of rows remaining in the training data after removing the problematic rows?