

Code Challenge

Lab Software Developer – Invention Challenge

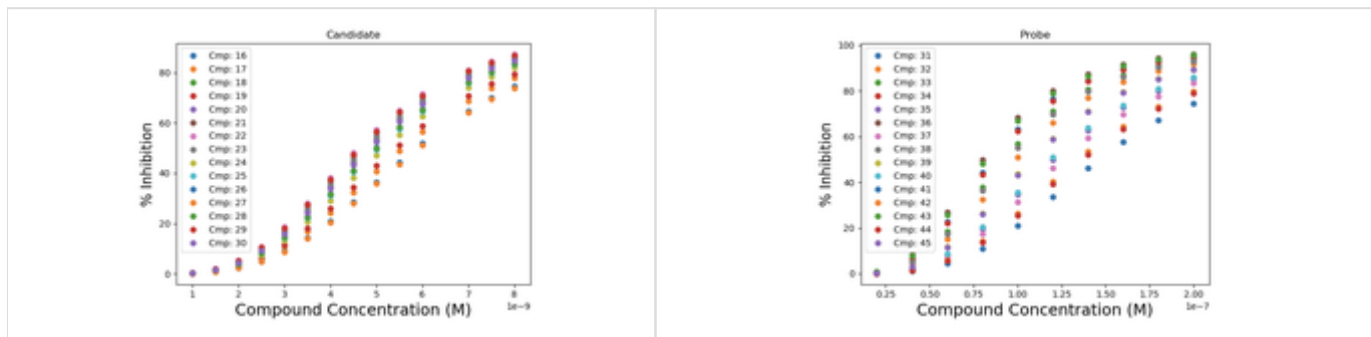
Working as a Lab Software Developer at Exscientia will require that you are able to think on your feet and are able to both follow a design brief but also able to propose innovative solutions of your own which the scientists might not have considered. This task is designed to give you to the chance to prove your skills as an innovator: build whatever application you want which does something interesting with the example dataset provided.

The only rule is that your application must be written in Python and can be either a: Desktop, Command Line, or Web-based application. If you are writing a web-based application, then of course you can utilise languages like JavaScript but there must still be a significant Python component to your project. Your program could for example have a command line interface that allows for the data to be filtered and for plots to be generated or your could have an interactive web-based tool which allows users to interact with the data - you are free to do whatever you want to demonstrate your coding and inventive skills.

As a rough guide we would suggest you spend between 2 - 4 hours on the challenge.

The example dataset (assay_results.csv) you have been given is for a typical type of experiment which is performed at Exscientia looking at how small molecules inhibit the activity of specific enzymes at increasing concentrations. We refer to these types of experiment as dose response experiments which for a typical inhibitor molecule will exhibit increased inhibition at higher concentrations. Note that a common way of comparing inhibition is to refer to the concentration of small molecule required to reduce the activity of an enzyme by 50%. We refer to this number as the IC_{50} .

To help you understand the provided dataset, the figure below plots some of the data. Each colour represents a single small molecule and for each small molecule multiple points have been plotted for the inhibition by the small molecule at increasing concentrations.



Dataset Files:

assay_results.csv -> % inhibition of the enzyme under study for each small molecule at increasing concentrations

compound_ic50.csv -> IC_{50} for each small molecule for the enzyme under study. Not necessarily from the results in assay_results.csv but from other experiments not included in this example

compound_labels.csv -> Depending upon the small molecule type and IC_{50} we can give labels to the small molecules which help to classify them and to indicate their level of activity. This file includes these labels for each small molecule

Note that this dataset has not come from any real experiments but from simulating various different small molecule activity profiles.

Please email back your complete program with instructions on how to run it and some documentation of what it does.