

FINAL PRACTICAL

UpSkill

Data

Analysis



Ready by **12th June 2020**

Covers **60% of the general score**

Put all your work in a zipped file

The data is originally from:

<https://github.com/ocatak/apache-http-logs>

Try to

- to analyse the log data.
- to detect vulnerability scans.
- to detect XSS and SQLI attacks.
- to examine access log files for detections.

Ctrl [Click me to get the data](#)

Expectation

1. **Analysis report** The report should contain: Executive summary, Data description, data analysis, results, discussion of results, conclusions and references in the cases of citations.
2. **Data analytics dashboard.** Build and share the link to your dashboard in your report if you built an online dashboard eg google data studio, chatio. OR share your dashboard file if you built the dashboard locally on shiny apps, powerbi etc.
3. **Machine learning or deep learning model.** You can perform lots of machine learning on the data supervised, unsupervised or semi-supervised. We expect your scripts or notebooks. Put the performance evaluation of your models among the results in the analysis report.
4. **Documented code** for each piece of code you present for this evaluation we expect proper documentation. I recommend using <https://docuaurus.io/> for documentation but you can use anything else for documentation.
5. **Production.** This is not an expectation but if you would like to move further you can push your model to production using frameworks like django, flask etc