Optimizing Sales of Ducks and Fish

PATRICK LOWE, University of Passau, Germany

The reproducibility of sales optimization, and how to future proof the tools used. The Excel Solver used in the Head First Data Analysis's Optimization chapter may not always be accessible, the challenge is to create an open-source alternative. While Google Sheets offers a similar solution, this project investigates the more widely accessible Python programming language and the package pulp. This tool was able to replicate results as Excels Solver while being more versatile.

Additional Key Words and Phrases: datasets, optimization, python, excel solver

1 INTRODUCTION

Excel is becoming outdated for big data analysis, python is one tool which can replace it. The book Head First Data Analysis's (HFDA) chapter 3 on optimization utilizes Excels Solver functionality to find the optimal number of rubber ducks and fish needed to maximise profits with limitations on resources. It then looks at a pitfall, not recognising seasonal changes in sales. However, Excel requires a subscription and is less accessible than open-source programming languages. This project looks at replicating the results from the book using Python and its available packages. There are 2 input files (1) for the data constraints and values of profit and (2) for the historical sales of both products. The python script utilises the Pulp package to analyse both files to replicate the results of the Excel Solver.

2 METHODOLOGY

Docker is used to create a replicable environment of HFDAs Chapter 3 on optimization since it has the ability to create an environment which can replicate the results in 1 step. Other alternatives such as Podman and VirtualBox however these have more set-up and space requirements than Docker. It is also a widely used, open-source, virtual container which increases its chance of being available long term.

Since Excel requires a license to be run, we have opted to use Python since it is open source, requires minimal set-up, and is human-readable. We use the python package pandas to read the XLS data file, math is used to for calculations, datetime is used for graph data, matplotlib is used for the graph, and the main package we need is PuLP. The PuLP package is a linear problem solver package, and has a BSD license allowing for redistribution.

In order to create a report on-demand, we have chosen LaTeX
* Solver for excel needed to be manually added through the "addons" section
* Selecting 2nd constraints is a bit arbitrary although chapter 3 has provided their answer
* The 2 excel workbooks come with 2 empty sheets each

3 RESULTS
HERE

4 FIGURES
HERE

5 CONCLUSION

6 REFERENCES
* Github for continuing work
    * Archive in Zenodo
    * Conda for environment dependencies
    * Docker for environment container

Author's address: Patrick Lowe, lowe03@ads.uni-passau.de, University of Passau, Innstr. 41, Passau, Germany, 94032.