

Duckies vs. Fishes

Bernhard Rippich

Technical University of Applied Sciences Regensburg
Regensburg, Germany
bernhard.rippich@st.oth-regensburg.de

Ivan Radeljak

Technical University of Applied Sciences Regensburg
Regensburg, Germany
ivan.radeljak@st.oth-regensburg.de

Abstract—With respect to the current reproducibility crisis of the fact that more than 70% of conducted experiments cannot be reproduced by other scientists [1]. This paper conducts a fully reproducible experiment for other scientists to reproduce. The important criteria of reproduction are being followed throughout the experiment as well as discussed later in this paper.

Index Terms—duckies, fish, reproducibility, docker, r, python, maximization

I. INTRODUCTION

The following paper will give a brief overview of the project and discuss in more detail the decision made for the reproduction package.

The program is based on an artificial business scenario from Head First Data Analysis by Michael Milton [4, p. 75-109]. In the given scenario, a company wants to know how many rubber ducks and fish can be produced and sold in the following three months to their current production year.

A data set is already given, which holds the information about all sales of ducks and fish in the last years from 2006 to 2008. Based on the given information the following three months in 2009 must then be predicted to get the right mixture of ducks and fish and hence to maximize sales.

The following figure 1 shows the sales of ducks and fish in the years of 2006, 2007 and 2008 per month as well as the total amount of sales for each month. In the figure, one can see a fluctuating behaviour of the graph regarding the amount of fish or ducks depending on the month. Both products are alternating between each other in regards to the amount. For instance, the graph of fish shows in the middle of 2006 to the end of autumn a comparably small amount of sold units, followed by a strong increase of the sales and vice versa for the graph of ducks.

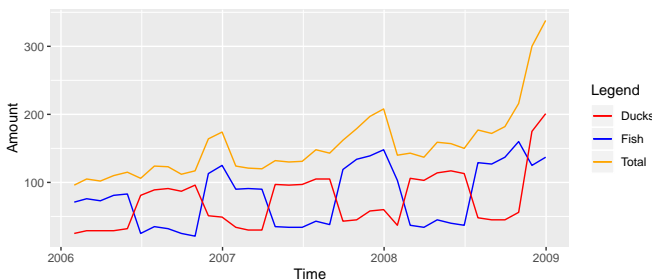


Fig. 1. History Sales Data

II. PREDICTION

Figure 2 shows the prediction of sales in the following three months to the given data set. The constraint for the prediction is the temporal dependency of ducks and fish to the respective month. The Prophet algorithm takes this dependency into account. Prophet is a forecasting algorithm implemented in Python and provides tuning functionality for data scientists and analysts [6].

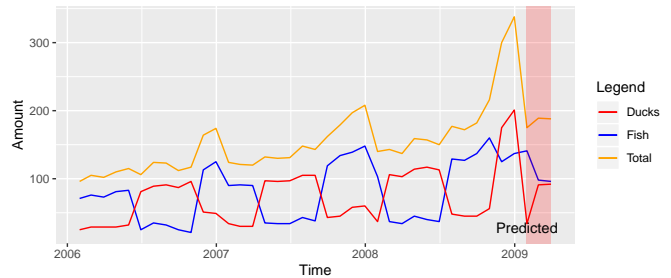


Fig. 2. Predicted History Sales Data

III. OPTIMIZATION PROBLEM

The problem with the calculation of the predicted profit is predicated on the constraint of limited resources and amount of factory machines, which ultimately makes it difficult to produce the right amount of fish and ducks. This problem is formulated as an optimization problem and is referred to as the maximization problem. The goal is to maximize profits with the available limited resource of pellets and the limit of factory machines. To achieve this goal it is crucial to calculate the best possible mixture of ducks and fish based on the predicted amount of sales. For this problem the Python package PuLP is used. PuLP is a linear problem modeler, which has the functionality to solve multiple optimization problems [5].

IV. RESULTS

The predicted data ensures a minimum amount of manufactured units of ducks and fish. Together with the constraints of the optimization problem the best possible profit is calculated. For the first three months of 2009 the estimated sales of fish and ducks is 141 and 34 in the first month, 98 and 91 in the second month and 96 and 92 in the third month. This results in a profit of 2179 \$ in the first, 2277 \$ in the second and 2284 \$ in the third month.

V. REPRODUCTION

The objective of reproductive results is to achieve identical results independent of external factors. The common factors are time, licensing, the operating system and the availability of the used dependencies. Apart from the technical aspects, the reproduction package should be structured in a way, that researchers have an easy understanding of the functionality. The reproduction package should offer an automated method for reproducing the results.

VI. DOCKER

Docker separates the infrastructure from the application. This concept allows to run the application on different host-systems without taking risk of incompatibility and therefore unreliable functionality [2]. Every application has to be packaged with its required dependencies into an image, which can be stored externally. The host system can access the storage, download the packaged application and run it without the need of administrative configurations. This package includes a docker recipe for generating this document and underlying data in a isolated environment for reproducible results.

VII. DISPATCHER

A dispatcher is an automated script, which takes care of all required commands for creating reproducible results. The benefits of using a dispatcher are the simple handling of generating of results and the consistency of the generated results. In this package, the dispatcher handles the building of the container, the generation process and the export of the results to the host system.

VIII. EVALUATION OF REPRODUCIBILITY

There are multiple areas, which have to be evaluated for the successful reproduction of a package.

A. Bitwise Identity

The bitwise identity is an approach to compare the results after multiple runs on the level of bits. This ensures that multiple runs of the package will not affect the results. The criteria of bitwise identity for this package is not fulfilled completely. The generated paper does not fulfill the criteria, since the pdf format stores additional metadata. If the metadata had been trimmed of, the documents would have been identical to the bits.

B. Dependencies

Depending on external packages poses the risk of future breaking changes. This can lead to unexpected behaviour or unvalidated results. To circumvent this exposure of risk, the used packages are locked. This disallows the system to update the packages automatically.

C. Runtime

The execution of an application involves many dependencies. These dependencies will not affect the reproducibility, except for the generation or usage of random values. To handle this problem, a fixed seed will ensure, that the generation of each random value is predictable.

IX. CHARACTERIZING THE EXPERIMENT

According to Freire et al. there are three aspects to cover in a reproducible environment [3]. The criteria are depth, Portability as well as Coverage and are covered in the following:

- 1) *Depth*: This paper comes with the full experiment given to the student at the beginning. A set of figures and a manuscript which are both generated (a) and a spreadsheet file used to generate the figures (b). During this experiment there is no data to measure, hence there is no necessity of providing such a data set (c). The system configuration and its initialization is covered by the usage of Docker, the scripts used in this project are deployed as the mentioned dispatcher and the Python scripts (d). The software system of this project can be seen as a white box (e). It provides source, configuration files and a build environment built inside of Docker.
- 2) *Portability*: The results of this experiment are reproducible on the original environment this experiment was conducted. It is also possible to replay the experiment on the original environment (a). Additionally, the experiment can be reproduced on similar environments with the same operating system being Windows (b). Furthermore, the experiment can be reproduced on a different OS, which is a Mac OS system (c).
- 3) *Coverage*: The experiment can be reproduced to the full extent. This includes the data set which is generated based on the data the task *Duckies and Fishes* provided as well as this paper along with the figures that are included. Thus, it makes the experiment fully reproducible (b).

X. CONCLUSION

This experiment gives any researcher the ability to easily reproduce the findings of this experiment. To meet the requirements of a fully reproducible experiment the three criteria of Freire et al. [3] can be met to the full extent.

REFERENCES

- [1] Monya Baker. "1,500 scientists lift the lid on reproducibility". In: *Nature* 533.7604 (May 2016), pp. 452–454. ISSN: 1476-4687. DOI: 10.1038/533452a. URL: <https://doi.org/10.1038/533452a>.
- [2] *Docker*. Date accessed: 02/19/2022. URL: <https://docs.docker.com/get-started/>.
- [3] Juliana Freire, Philippe Bonnet, and Dennis Shasha. "Computational reproducibility: State-of-the-art, challenges, and database research opportunities". In: June 2012, pp. 593–596. DOI: 10.1145/2213836.2213908.
- [4] Michael Milton. *Head First Data Analysis: A Learner's Guide to Big Numbers, Statistics, and Good Decisions*. "O'Reilly Media, Inc.", 2009.
- [5] *Optimization with pulp*. URL: <https://coin-or.github.io/pulp/main/includeme.html>.
- [6] *Prophet — Forecasting at scale*. Date accessed: 02/19/2022. URL: <https://facebook.github.io/prophet/>.