

Visualization of Football Data

Rasmus Bo Adeltoft
Sebastian Seneca Haulund Hansen
Steffen Berg Klenow
Christian Bjørn Moeslund
Andreas Staurby Olesen
Henrik Sejer Pedersen
Supervisor: Marco Chiarandini

17th May 2016

1 Abstract

2 Preface

Contents

1	Abstract	2
2	Preface	2
3	Introduction	4
4	Theory	4
4.1	Something	4
4.1.1	Design Process	4
4.1.2	Design	6
4.1.3	Data Types and Data Sets	6
4.1.4	Idioms	6
4.1.5	Analysis and Complexity	6
4.1.6	Facets and View Manipulation	6
4.2	Exploratory Analysis	6
4.3	Tools and Technologies	6
4.3.1	R	6
4.3.2	D3.js	6
5	Results	6
6	Discussion	6
7	Conclusion	6
8	Usability	6
9	Bibliography	6
10	Appendix	6
11	Process Evaluation	6

3 Introduction

Data is collected at a rapidly increasing rate in all fields and it becomes necessary to present data in different ways in order for humans to make sense of it. One way to do this is through data visualization. Visualization can help human’s understanding of large data sets, as the data can be summarized very effectively, and patterns can quickly be recognized by humans. When making visualizations it is important to understand how the human cognitive system works, such that visualizations can be designed to make it easier for humans to understand the data. In order to do this, we will apply principles from the field of visualization to present football data. We will use tools such as R to process data and plot static visualizations, and use D3 to make interactive and dynamic visualizations.

Specifically, we will do this both by making visualizations that can help explore the questions that we present below, and by doing exploratory analysis such that new patterns can be discovered. The specific questions that we will be investigating are:

- How does a team evolve throughout a season in terms of goals, points, etc.?
- How does a team’s playing style (for example passes, possession and tackles) change throughout a match?
- How does a winning team differ from a losing team?

During the visualization process we will consider different visualization techniques and choose a suitable one based on principles and analysis tools given by Tamara Munzner in “Visualization Analysis and Design” to make sure that the data is presented in an accurate and easily understandable manner. This includes considerations regarding the human cognitive system.

4 Theory

4.1 Something

4.1.1 Design Process

This section describes the typical work flow of a data scientist. We will focus on the following four faces: Preparation, Analysis, Reflection and Dissemination.

The process of getting the data, understanding the data and produce results is an iterative process. The process is seen on Figure 1.

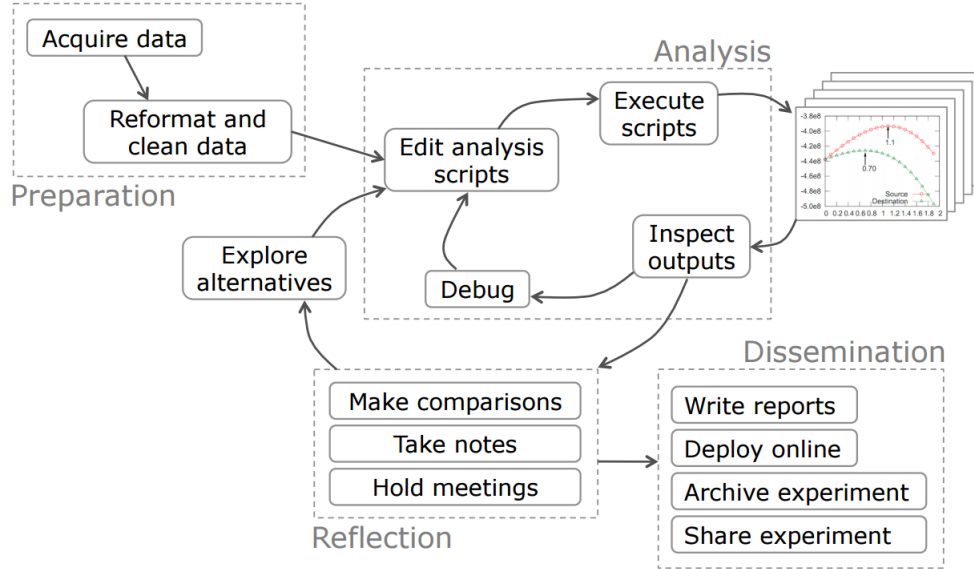


Figure 1: The model showing the iterative process[Guo12, Chapter 2]

- The first face is the preparation face. Here the you have to acquire the data, that could be from hard disks, servers, through an API ect. Where to store and how to organize the data files should also be considered, so it is easy to replace the right files if the data gets outdated. Then the data should be cleaned, meaning tuples with missing values, change the formatting, sorting the data ect.
- The second face is the analysis face. Here the data is analysed to get more information about the data. This is an iterative process, where the you create and run scripts, look at the output, maybe find some mistakes, debug these and run it again.
- The third face is the Reflection face. Here the output results is discussed, for example by making comparisons between outputs, and exploring alternatives.
- The fourth and last face is the dissemination face. Here the the results are reported and maybe published in a report. [Guo12, Chapter 2]

4.1.2 Design

4.1.3 Data Types and Data Sets

4.1.4 Idioms

4.1.5 Analysis and Complexity

4.1.6 Facets and View Manipulation

What is facets, why facet

4.2 Exploratory Analysis

4.3 Tools and Technologies

4.3.1 R

4.3.2 D3.js

5 Results

6 Discussion

7 Conclusion

8 Usability

9 Bibliography

References

[Guo12] Philip Jia Guo. *Data Science Workflow: Overview and Challenges*.
PhD thesis, Stanford University, 2012.

10 Appendix

11 Process Evaluation