**UNIVERSITY OF SUNDERLAND**

**FACULITY OF COMPUTER SCIENCE**

CETM25 - Data Visualisation  
Assignment Report

Cheng Man Kit

Student ID: 179195943

words count: 2718

Table of Contents

[Introduction 3](#_Toc19804098)

[Background 3](#_Toc19804099)

[Design Stage 3](#_Toc19804100)

[Data source and theme selection and specification 3](#_Toc19804101)

[Users’ requirement analysis 3](#_Toc19804102)

[Data storytelling process design 3](#_Toc19804103)

[Product functional & non-functional requirements specifications 4](#_Toc19804104)

[Product software architecture design 4](#_Toc19804105)

[Product use case specifications 5](#_Toc19804106)

[Development Stage 5](#_Toc19804107)

[The selection of visualization software and hardware methodologies 5](#_Toc19804108)

[Implementation methodology 5](#_Toc19804109)

[System testing method 6](#_Toc19804110)

[User evaluation plan & method 6](#_Toc19804111)

[Project Management 9](#_Toc19804112)

[Time Management 9](#_Toc19804113)

[Risk Assessment 9](#_Toc19804114)

[Quality Control 9](#_Toc19804115)

[Customer relationship management 9](#_Toc19804116)

[Future marketing strategy 10](#_Toc19804117)

[Conclusion 10](#_Toc19804118)

[Reference 11](#_Toc19804119)

# Introduction

In this report, the design and development cycle of the product and the management of the R&D project were descripted. The discuss and evaluation are following with three parts, Design stage, Development stage and Project management. In Design stage, both business and users’ requirements, data storytelling, functional & non-functional requirements of the product, software architecture and use case specifications were covered. For Development stage, the selection of visualization software and hardware, project implementation methodology, system testing method and user evaluation plan were discussed. Third, Project management have come over about the time management, risk assessment, quality control, user relationship and product marketing strategy.

## Background

The data visualization product was a part of the research and development project which requested by one of the top 10 online sport betting company. The target of R&D project was using Big Data Analysis and Machine Learning to build a new online betting web platform. As a data visualizer involved in this R&D project, the data visualization product requested to complete in the first phase with the scope of The FIFA World Cup and the data sets were come from Kaggle (https://www.kaggle.com/abecklas/fifa-world-cup).

# Design Stage

## Data source and theme selection and specification

According to the big data analysis become hotter and hotter in these few years, the online sport betting company requested to build a new online betting web platform with adopting the most updated technology of big data analysis and machine learning. Attracting more user and willing to place the bet on the new web platform was expecting by the online sport betting company. More statistical information of sports tournament and league will be provided and the results prediction which providing by machine learning and data analysis. The data source will be gathering from the open data of the official web site depending on the different type of sports and the variety of sports tournaments, leagues, matches.

## Users’ requirement analysis

One of the most significant factors of the success on a project was user requirements analysis. The poorly specified user requirements will lead to the project failure. The understanding of both user and organizational requirements are the key of the successful of product. Users’ requirements were gathering by surveys, interviews, focus groups, evaluation of an existing product.

The dashboard screen was required for showing the statistical information of tournament and league. Results prediction base on the statistical information and providing by machine learning and data analysis were needed.

## Data storytelling process design

There are three steps of the visual data storytelling process, included in finding insights (explore data), transforming these insights into a story (make a story), and giving this story to an audience (tell a story) (Lee et al., 2015).

In this R&D project, the story was start for the open data from the sports tournaments, leagues, matches, plotting charts after data analyzed, showing the statistical information, presenting by a dashboard in single page, and drill down by year, by season, by team, providing more details and sliced statistical information to the audience.

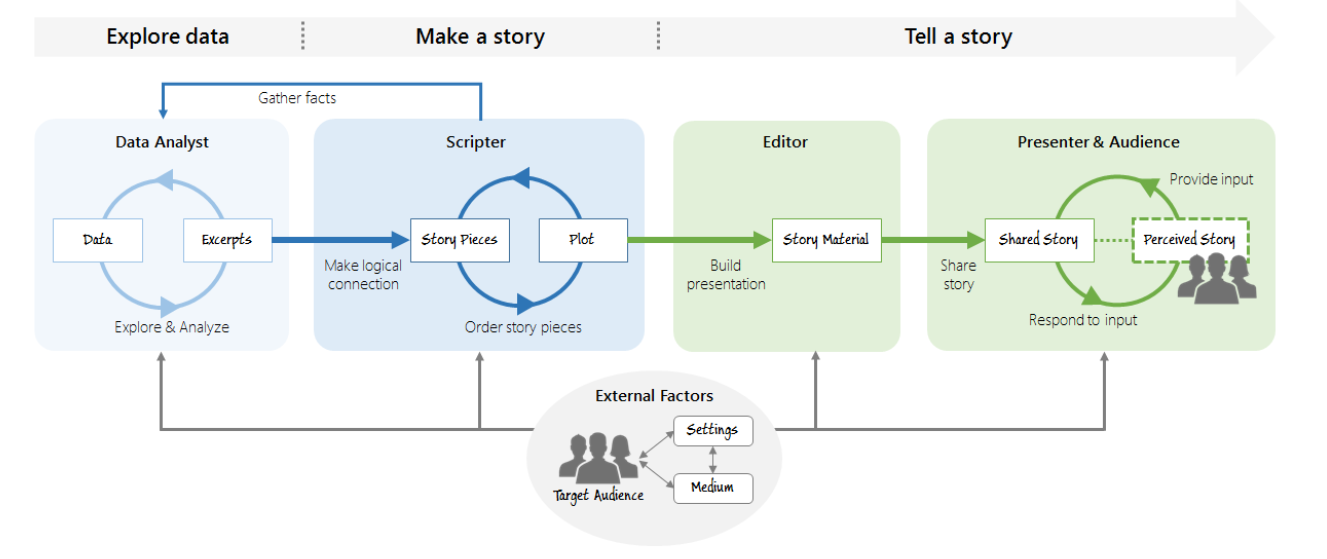


Fig 1. Storytelling process: transforming data into visually shared stories.  
adopted from Lee et al., 2015

## Product functional & non-functional requirements specifications

Functional requirements define what function and features of the product must-have. Nonfunctional requirements define the general properties of a system. The quality constraints of the system like portability, security, reliability, scalability, performance, and flexibility.

For the new web platform for this R&D project, the functional requirements are displaying the statistical information of the tournaments, leagues, matches with data analyzed with different levels and views. Second, providing the predicted results by machine learning algorithms.

And for the nonfunctional requirements, portability is a must, most of the peoples are living on the mobile phone now, the web platform should be suitable in the small screen. Security, reliability, and performance were needed also.

## Product software architecture design

Software architecture is the process of transforming software attributes such as flexibility, scalability, reusability, and security into a structured solution that satisfies the technical and business requirements (Vishnyakov and Orlov, 2015). The top 5 software architecture patterns are Layered (n-tier) architecture, Event-driven architecture, Microkernel architecture, Microservices architecture, and Space-based architecture. The Layered (n-tier) architecture was traditional and reasonably the most common because it is usually built encompassing the database, and many applications in business generally storing information in tables. But Microservices architecture should be more suitable in this R&D project. With the cloud base hosting, every single thing on the web platform is all delivered in separate batches by separate services. This approach is alike to the event-driven and microkernel approaches, but it’s used principally when the different tasks can be separated easily.

## Product use case specifications

The use case which has only one goal, with single start point and single end point, but with mulitle paths between the start and end point that achieve the goal. And the use case specification is a document used to obtain the particular details of a use case. Use case specifications giving a method to obtain the functional requirements of a system with combining all existing of different situations. They add more details base on a use case diagram. They are a useful tool in talking with project stakeholders, system users, business analysts, and developers. These specifications explain those requirements in another tone that all users of the project can understand, formulating a common word for the impacted people. This document will use in cooperate with project team members, including customers (the project owner), design team members, and programming developers.

# Development Stage

## The selection of visualization software and hardware methodologies

R programming language was selected for this R&D project. R is a free programming language, with statistical computing and graphics supported. And it also is a free software environment for development. With the packages - Shiny and Shiny dashboard, it provided an easy way to build outstanding dashboards and interactive web apps. Dashboard is one of data visualization tool and have recently gained a lot of importance. The status of business analytics metrics, key performance indicators (KPIs) and important data points were display on single screen for an organization, department, team or process. Information was presents in a way that is easy to read with the Dashboard. Tableau was in the considering list, which was claimed easy to use, with fast analytics, accepted different types of data source and have smart dashboard. Tableau was a business software but have a free public version with limited function. With the free public version, the dashboard can be exported as HTML web page for easy access. But according to the user requirements, the interactive web apps not only presenting the sports data with the visualization dashboard. The data analytics and prediction of results will be provided on the web platform. There was the limitation of Tableau, but it was fully supported by R programming language with required packages. Decision Trees, Linear Regression, Logistic Regression, Naïve Bayes, etc. all can be program by R with suitable packages. And those results are easy to publish to the web platform by Shiny and Shiny Dashboard.

For the hardware side, because this R&D project was going to build a web platform. It was suggested to use the existing cloud hosting services, like Amazon Web Services (AWS), Google Cloud Platform (GCP) or Alibaba Cloud. The cloud hosting services are become more and more secure and reliable, and the content delivery network (CDN) can provide the web content from the nodes that are nearest to the users across the globe. This is more suitable than build a web server and hosting the web platform in a single location.

## Implementation methodology

There are many of project management methodologies can pick in today, each of them with the rules, principles, processes, and practices.to maximize the use of resources and time, the selecting on a project management methodology is the way to success.

The key of Agile methodology is doing work in small iterative cycles with customer feedback at the end of each cycle. And this can increase the possibilities of developing a new product or service that people will acquire, apply and like (Denning, 2016). But the time performance in projects were not able to increase by agile project management systems. It was less effective than all projects on average in terms of time performance. even though the consequences of cost and time achievement are not effective, the ability of create more excellent product is clear in the research of agile project management (Suetin et al., 2016).

Beside of Alige, the traditional project management methodology is Waterfall, which is a downward-flowing stage model. In between the stages of the system development lifecycle (SDLC), including specifications, development, testing, and implementation the feedback was limited (Lakshman Mahadevan, Kettinger and Meservy, 2015).

In this R&D project, Agile methodology was more suitable to meet the flexibility and uncertainty.

## System testing method

To test the comprehensive and integration of the software product, system testing was performed. Functional testing is a sort of software testing in which the system is tested against the functional requirements and specifications. Functional testing guarantees that the requirements or specifications are accurately satisfied by the application. This sort of testing is only concerned about the result of processing. It concentrates on the simulation of real system usage but does not produce any system structure appropriations.

It is essentially described as a sort of testing which confirms that each function of the software application works with the requirement and specification. The source code is not concerned with this test. Each functionality is tested by giving proper test input, assuming the output and examining the actual output with the expected output.

Non-functional testing is a sort of software testing that is conducted to check the non-functional requirements of the application. It checks the behavior of the system is to meet the requirement or not. It tests all the features which are not tested in functional testing.

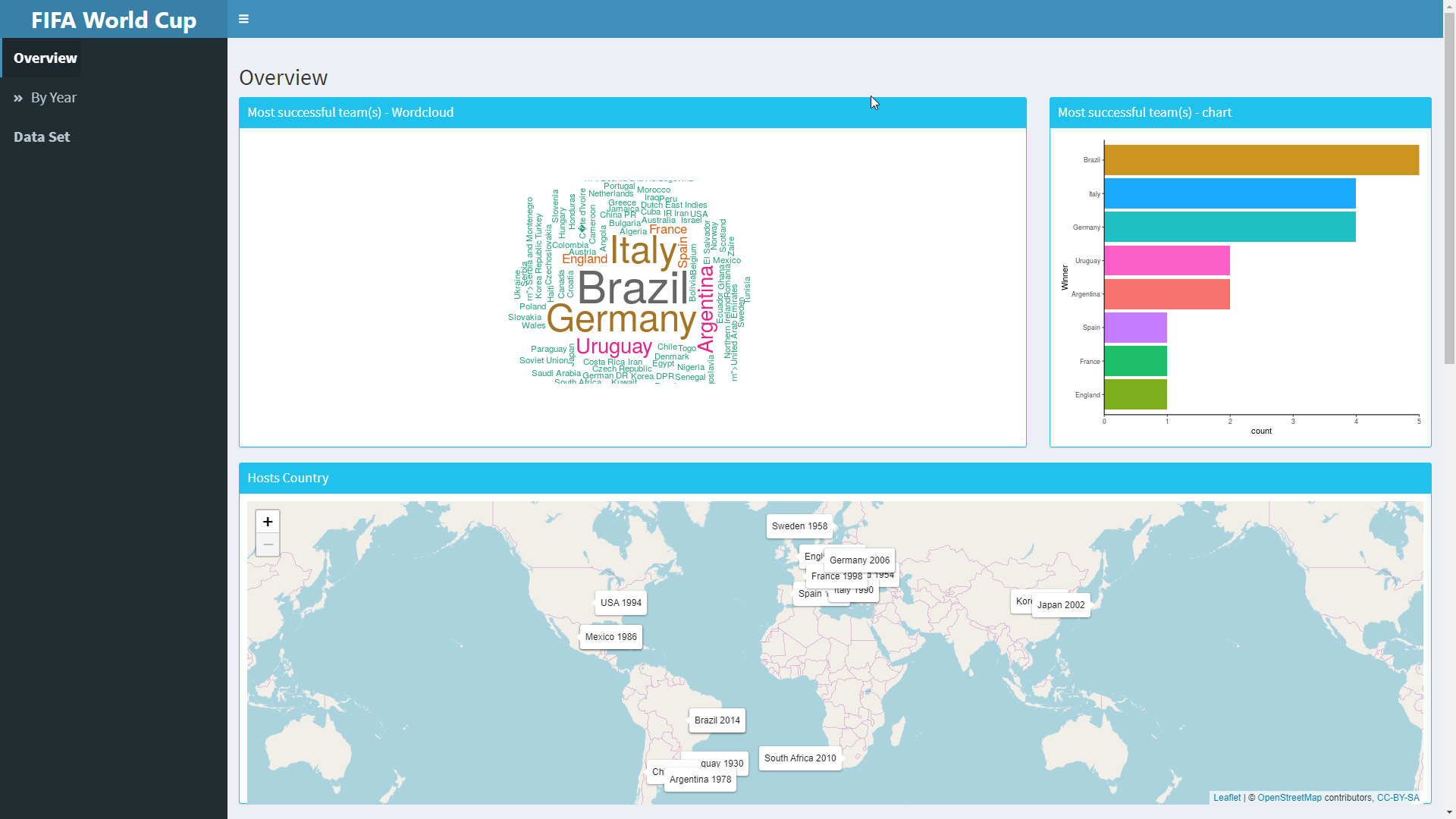
It is planned to test the readiness of a system as per nonfunctional parameters which are not included by functional testing. Non-functional testing is as important as functional testing.

## User evaluation plan & method

The user evaluation plan is the guides of every step to the process of evaluation. It helps to decide the information that the stakeholders need. And identify the best methods and approaches for getting the needed information. There were many different evaluation methods, Monitoring and feedback system, Goal attainment report, Interviews with key participants and Community-level indicators of impact.

In this R&D project, the goal attainment report was chosen for the evaluation method as it can clearly show the completed stage and track point.

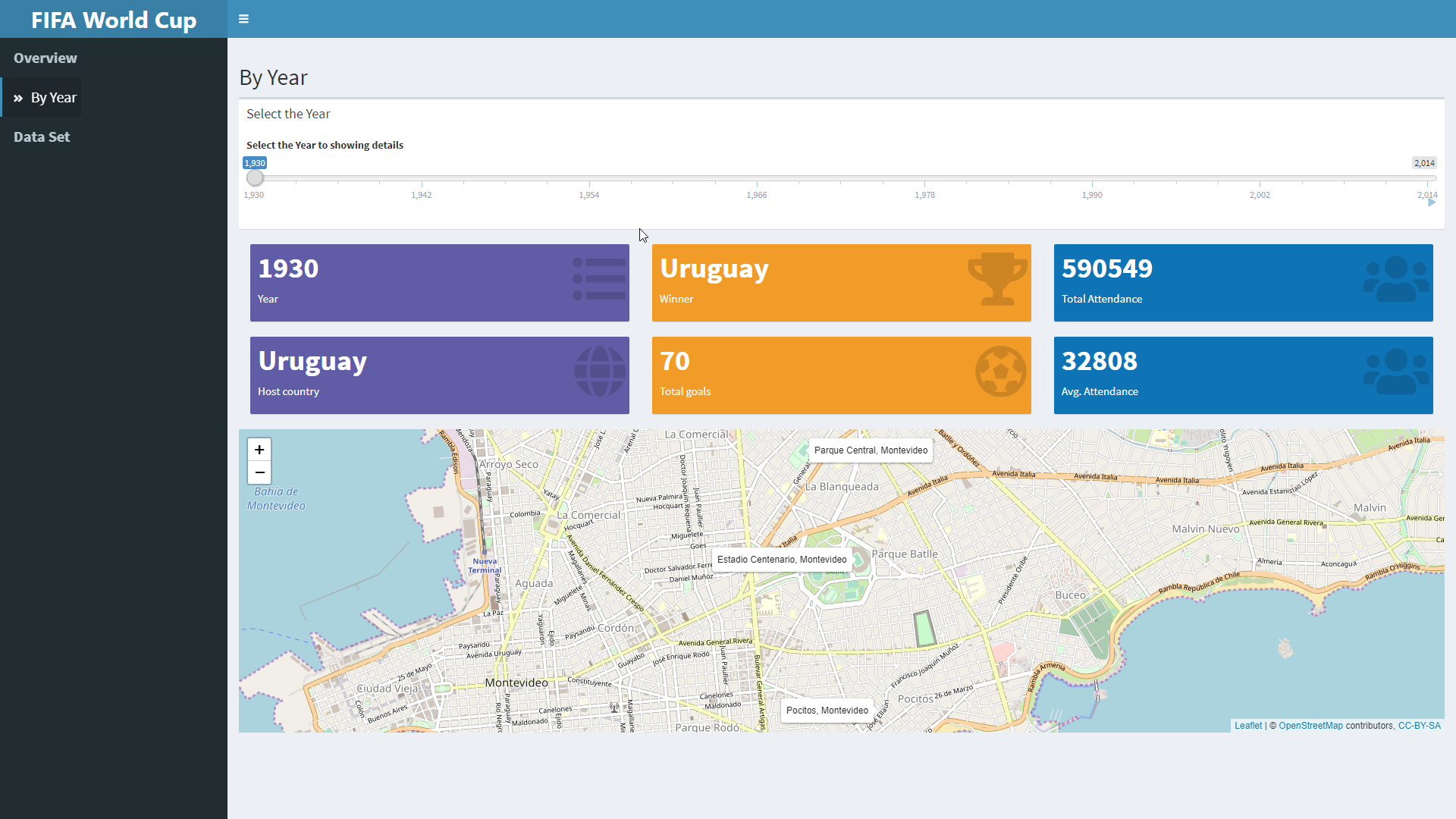
A prototype was built and hosted at the free platform www.shinyapps.io, which is easy to deploy the Shiny applications, no need to own a server, no hardware, installation, or annual purchase contract required.



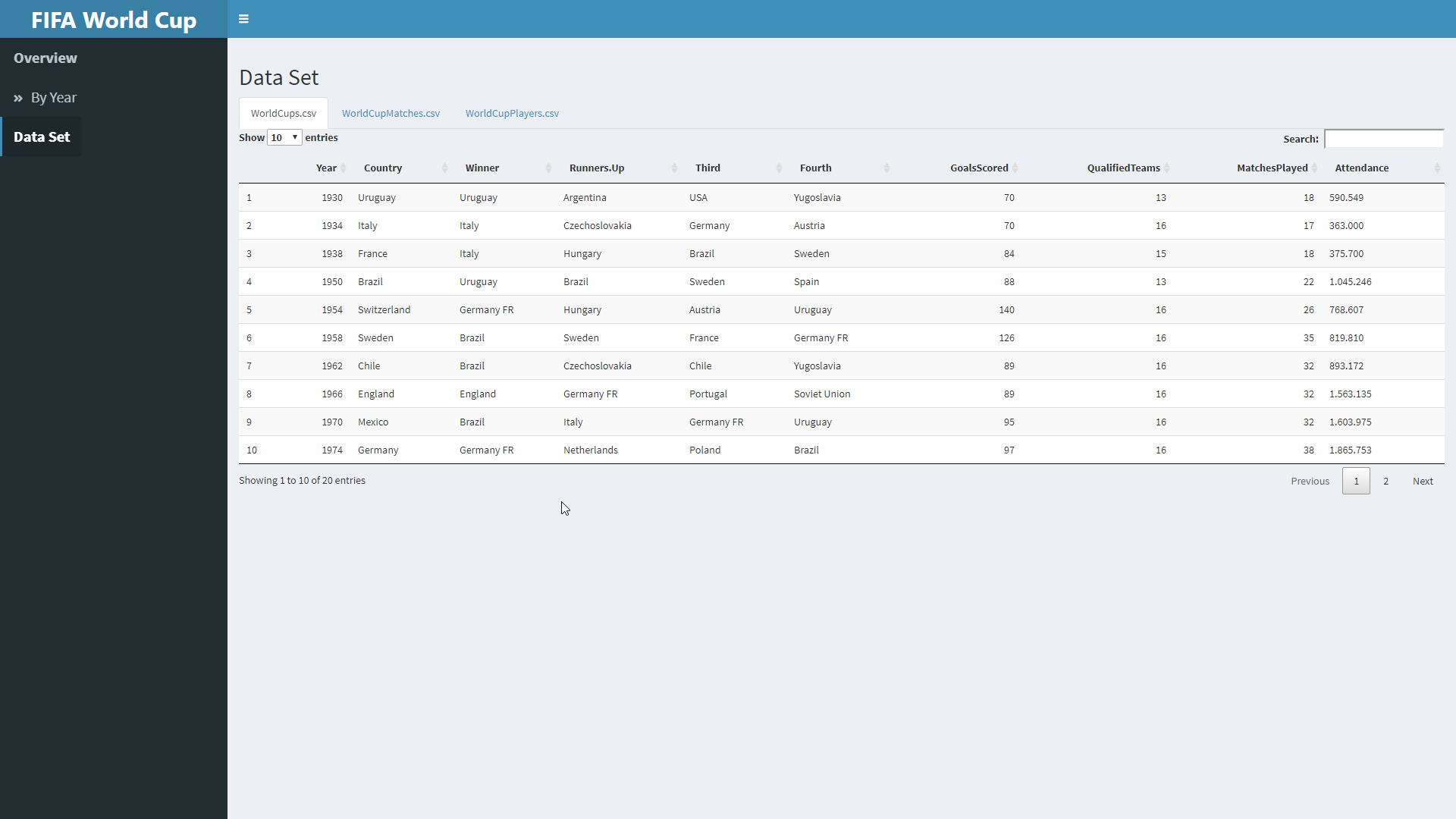
Pic 1. Screen capture of the Dashboard prototype – Overview 1



Pic 2. Screen capture of the Dashboard prototype – Overview 2



Pic 3. Screen capture of the Dashboard prototype – By Year



Pic 4. Screen capture of the Dashboard prototype – Data Set

# Project Management

## Time Management

There are many differences between traditional projects and on agile projects in time management. In traditional project management like Waterfall, at the project start, Project managers determine and try to predict the schedules. Development teams do not start the product development until the requirements gathering and design phases are complete. For the critical requirements and optional requirements, there is no schedule difference. But in the agile projects, the highest priority and highest value requirements were complete on first by scrum teams work in sprints. and during the project, scrum teams also evaluate and re-evaluate those work they can complete in a provided time slot.

## Risk Assessment

All the data used in this R&D project came from open data. There is no concern on data privacy. At the meant time the prototype did not collect and store any user data or personal information. Even on the later phase, the web platform for theis R&D project will not to collect and store the user data or personal information. And the web platform will be hosted by the cloud services that provide the security web access under HTTPS protocol. According to the we platform will function in global including EU area, the requirements from General Data Protection Regulation (GDPR) will be complies and conduct a Data Protection Impact Assessment (DPIA) when needed.

## Quality Control

Quality Control is a scheme that concentrates on accomplishing the quality requested and points to recognize and fix bugs. it is a procedure to check with the quality-validation and measure the executing program. Confirmed the work on the product was followed the standards and ensured that all the works done as per the requirements.

With Agile approaches, testing is performed every day as part of each sprint and is held in each requirement’s definition of done. it is easy to find the problems when test with a smaller amount of work. And easier to apply the fixes when the work just created.

Testing is convinced on agile projects because it is performed in every sprint.

## Customer relationship management

User involvement was proved that brings the system development to success(Bano and Zowghi, 2015). Agile project management proposes the value of face-to-face communication is the best way to gather information without wasting time. Documents are purposely simple and only include necessary information. working progress was communicated with showing working software regularly in the sprint review by project teams. Meetings on agile projects are as fast as possible and include only people who will benefit from the meeting. The creation of agile meetings is to enhance, not reduce, productivity.

documentation created in the simplest way as possible, visually communicating artifacts such as the vision statement, the definition of done, the impediments log, and important architectural decisions will be the best work. Pictures are worth than a thousand words.

## Future marketing strategy

A successful marketing strategy will gain new customers, boost sales amount, inspire the existing customers to spend more. In this R&D project, it would be translated to attracting more user who interests in sports games but never placed the bet before, encourage current users to place more bet or increase the bet amount with the results predication provided.

The essential elements of a product marketing strategy are Product, Audience, Messaging, People, Promotion and Analysis. The value which the product can provide is more mattering than its capabilities. The created of the product should be with the customer spirit and follows with their requirements. Beside of the product needs to be created with the customer spirit, the marketing strategy should also be created with the audience spirit. Connect the product and the audience using narrative product messaging. The people are naming to the whole R&D team, everyone in the team should on the same page and knows their part. Even the product was great, without any promotion, no one knows that was great. The product marketing strategy needed to adjust after analysis the outcome of above items.

As now only the first phase of the R&D project, the product marketing strategy will develop later, before the go-live launch.

# Conclusion

The prototype of the dashboard was created for the demonstration, which was developed with the agile project management methodologies. In the development stage, the R programming with Shiny and Shiny dashboard packages was proved the easy to use and interactive capabilities. this was matched with the agile design, the product can reflect rapidly with the change of user requirements. With the application hosting services provided by shinyapps.io, the development of the prototype can be focused on the functional requirements. Those nonfunctional requirements, like portability, security, scalability, performance, all relied on the application hosting service provider.

And the prototype was presenting the demonstration in phase one of the R&D project. More data analytics and results prediction were not performed in this stage. With the microservices architecture of software architecture design and the cloud base application platform, the development in the second phase and the laters will be easy to integrate.

The size of FIFA World cup data sets were small and not have many items and views can display. But that will be more with other tournaments or leagues.

# Reference

Baig, J.J.A., Shah, A. and Sajjad, F. (2017). Evaluation of agile methods for quality assurance and quality control in ERP implementation. *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)*.

Balalaie, A., Heydarnoori, A. and Jamshidi, P. (2016). Microservices Architecture Enables DevOps: Migration to a Cloud-Native Architecture. *IEEE Software*, 33(3), pp.42–52.

Bano, M. and Zowghi, D. (2015). A systematic review on the relationship between user involvement and system success. *Information and Software Technology*, 58, pp.148–169.

Brhel, M., Meth, H., Maedche, A. and Werder, K. (2015). Exploring principles of user-centered agile software development: A literature review. *Information and Software Technology*, 61, pp.163–181.

Clarke, R. (2015). Big data, big risks. *Information Systems Journal*, 26(1), pp.77–90.

Denning, S. (2016). Agile’s ten implementation challenges. *Strategy & Leadership*, 44(5), pp.15–20.

Galvan, S., Mora, M., O’Connor, R.V., Acosta, F. and Alvarez, F. (2015). A Compliance Analysis of Agile Methodologies with the ISO/IEC 29110 Project Management Process. *Procedia Computer Science*, 64, pp.188–195.

Guzmán, L., Oriol, M., Rodríguez, P., Franch, X., Je-Dlitschka, A. and Oivo, M. (n.d.). *How Can Quality Awareness Support Rapid Software Development?*

Ko, D.-G. and Kirsch, L.J. (2017). The hybrid IT project manager: One foot each in the IT and business domains. *International Journal of Project Management*, 35(3), pp.307–319.

KRISHNAN, M.S. (2015). Software Development Risk Aspects and Success Frequency on Spiral and Agile Model. *International Journal of Innovative Research in Computer and Communication Engineering*, 03(01), pp.301–310.

Lakshman Mahadevan, Kettinger, W.J. and Meservy, T.O. (2015). *Running on Hybrid: Control Changes when Introducing an Agile Methodology in a Traditional “Waterfall” System Development Environment*.

Lee, B., Riche, N.H., Isenberg, P. and Carpendale, S. (2015). More Than Telling a Story: Transforming Data into Visually Shared Stories. *IEEE Computer Graphics and Applications*, 35(5), pp.84–90.

Schön, E.-M., Thomaschewski, J. and Escalona, M.J. (2017). Agile Requirements Engineering: A systematic literature review. *Computer Standards & Interfaces*, 49, pp.79–91.

Serrador, P. and Pinto, J.K. (2015). Does Agile work? — A quantitative analysis of agile project success. *International Journal of Project Management*, 33(5), pp.1040–1051.

Sharp, H. and Hall, T. eds., (2016). *Agile Processes, in Software Engineering, and Extreme Programming*. *Lecture Notes in Business Information Processing*. Cham: Springer International Publishing.

Suetin, S., Vikhodtseva, E., Nikitin, S., Lyalin, A. and Brikoshina, I. (2016). Results of agile project management implementation in software engineering companies. *ITM Web of Conferences*, 6, p.03016.

Tiwari, S. and Gupta, A. (2015). A systematic literature review of use case specifications research. *Information and Software Technology*, 67, pp.128–158.

Vishnyakov, A. and Orlov, S. (2015). Software Architecture and Detailed Design Evaluation. *Procedia Computer Science*, 43, pp.41–52.

Wang, X., Wu, Z. and Zhao, M. (2008). The Relationship between Developers and Customers in Agile Methodology. *2008 International Conference on Computer Science and Information Technology*.

‌