**Macintosh HD:Users:rruzzo:Desktop:logo.pngCapstone Analytics**

Programmer’s Manual

Version 1.0

Table of Contents

[1 Introduction 2](#_Toc417668076)

[1.1 Purpose 2](#_Toc417668077)

[1.2 Scope of the Product 3](#_Toc417668078)

[1.3 Definitions, acronyms, and abbreviations 3](#_Toc417668079)

[1.4 References 3](#_Toc417668080)

[1.5 Overview 3](#_Toc417668081)

[2 General Description 3](#_Toc417668082)

[2.1 Product Perspective 3](#_Toc417668083)

[2.2 Product Functions 3](#_Toc417668084)

[2.3 User Characteristics 4](#_Toc417668085)

[2.4 General Constraints 4](#_Toc417668086)

[2.5 Assumptions and Dependencies 4](#_Toc417668087)

[3 Specific Requirements 4](#_Toc417668088)

[3.1 Database 4](#_Toc417668089)

[3.1.1 Master 4](#_Toc417668090)

[3.1.2 Salaries 4](#_Toc417668091)

[3.1.3 Teams 4](#_Toc417668092)

[3.1.4 Appearances 4](#_Toc417668093)

[3.1.5 Batting 4](#_Toc417668094)

[3.1.6 Fielding 5](#_Toc417668095)

[3.1.7 Pitching 5](#_Toc417668096)

[3.2 Security 5](#_Toc417668097)

[3.2.1 Username 5](#_Toc417668098)

[3.2.2 Password 5](#_Toc417668099)

[3.3 Application 5](#_Toc417668100)

[3.3.1 Linear Regression 5](#_Toc417668101)

[3.3.2 Standard Deviation 5](#_Toc417668102)

[3.3.3 Minimum Plate Appearances 5](#_Toc417668103)

[3.4 Application GUI 5](#_Toc417668104)

[3.4.1 Statistic Drop Down Menu 6](#_Toc417668105)

[3.4.2 Minimum Plate Appearances Drop Down Menu 6](#_Toc417668106)

[3.4.3 Data Results Return 6](#_Toc417668107)

[4 Appendices 6](#_Toc417668108)

[4.1 Appendices 6](#_Toc417668109)

[5 Index 6](#_Toc417668110)

[5.1 Index 6](#_Toc417668111)

# 1 Introduction

## 1.1 Purpose

Purpose will be ignored in accordance with the Grading Rubric.

## 1.2 Scope of the Product

The product has extracted annual statistics and salaries from publicly provided databases. With both performance based statistics and current salaries, we have done statistical analysis to determine the fair market value of a player.

Once the statistical analysis is complete, players that are underpaid based on performance would be identified as optimal candidates for acquisition. Conversely overpaid players could be avoided. Ultimately the user could use these tools to maximize their profits by ensuring they avoid signing bad contracts.

## 1.3 Definitions, acronyms, and abbreviations

LAMP

Regression

Mean

Standard Deviation

Sum

Intercept

Slope

## 1.4 References

<http://www.seanlahman.com/baseball-archive/statistics>

Used for statistics references.

## 1.5 Overview

Overview will be ignored in accordance with the Grading Rubric.

# 2 General Description

## 2.1 Product Perspective

Forbes evaluates Major League Baseball at approximately 36 billion dollars and the average team being worth 1.2 billion dollars. Even with these large values, payroll for players make up a large part of the expenses for baseball clubs. Depending on revenue and revenue sharing amongst the teams, MLB has attempted to minimize the advantage larger clubs have over small clubs. However, the 2014 payroll shows a gap of 190 million dollars between the highest and lowest payroll. This product would help smaller clubs evaluate players on an objective basis in order to identify players that should be acquired at a minimal cost.

## 2.2 Product Functions

The software will allow users to select up to four statistical categories to evaluate. The user will be able to weight the importance of each statistic being evaluated. This will allow the user to increase the importance of power (homeruns) or contact (hits) hitting styles as appropriate to find their ideal replacement player. After the search is complete, the program will return the players the most undervalued players in a list.

## 2.3 User Characteristics

The initial market for users of this program would be teams and agents involved in MLB. However, expansions should be made into the fantasy sports market as there is a much larger audience that could offset the cost of production. Further releases could focus on development of a mobile specific web application or smart phone application using the same design to evaluate players.

## 2.4 General Constraints

The current version 1.0 will be a web based only application. The user will require internet access in order to use the application. Other limitations are that the user will require their own web browser in order to access the site.

## 2.5 Assumptions and Dependencies

The assumptions will be that the user will require a compatible web browser. Web browsers supported include Chrome, Internet Explorer, Safari and Firefox. Versions updated as of the 15th of April 2015 should be compatible. View of the website will also depend on firewall settings and not being blocked by other safe viewing mechanisms.

# 3 Specific Requirements

## 3.1 Database

A data base will need to be stored on the web application server in order to allow the front end to access statistics of players to be evaluated.

### 3.1.1 Master

Master will account for one member fields in the database and will contain information to identify the baseball player.

### 3.1.2 Salaries

Fielding will account for multiple member fields in the database and will contain multiple variables that are standard to baseball.

### 3.1.3 Teams

Teams will be a member field in the database to account for the team a player is currently under contract with.

### 3.1.4 Appearances

Plate appearance will be a member field in the database to account for how many times the player has appeared for home plate.

### 3.1.5 Batting

Batting will account for multiple member fields in the database and will contain multiple variables that are standard to baseball.

### 3.1.6 Fielding

Fielding will account for multiple member fields in the database and will contain multiple variables that are standard to baseball.

### 3.1.7 Pitching

Pitching will account for multiple member fields in the database and will contain multiple variables that are standard to baseball.

## 3.2 Security

The web application will require security to ensure that subscription services are being used. Username and Password will be required. Those will be obtained via e-mail through subscription services.

### 3.2.1 Username

Usernames will be verified as a unique id prior to being issued by subscription services. Once username is established a field will be provided for login and correct entry will be required to gain access to web application.

### 3.2.2 Password

Passwords will be issued by subscription services. Passwords can be changed but must be verified as 8 characters at least one upper and lowercase letter, one number, and a special character (!@#$%^&\*()\_-+/\’:,?{}[]~). Password entry field should be next to or below Username.

## 3.3 Application

The web application will provide analysis of baseball players using statistics with salary to determine what the player’s market value is.

### 3.3.1 Linear Regression

Linear regression is an approach for modeling the relationship between a scalar dependent variable y (salary) and one or more explanatory variables or independent variable (this will be a baseball statistic) denoted X. The application will utilize linear regression to determine fair market value.

### 3.3.2 Standard Deviation

The standard deviation is a measure that is used to quantify the amount of variation or dispersion of a set of data values. The application will apply this variation to determine the variation in salary. This will also be based upon the choice of statistic.

### 3.3.3 Minimum Plate Appearances

This will be used as population to control to provide more meaningful data sets. It will serve to remove players who did not play a statistically significant amount of time.

## 3.4 Application GUI

The web application GUI will provide choice and feed back to the user to manipulate the data for their use.

### 3.4.1 Statistic Drop Down Menu

This will provide a drop down menu to choose from batting statistics in order to perform analysis on the database and provide feedback to the user.

### 3.4.2 Minimum Plate Appearances Drop Down Menu

This will provide a drop down menu to choose minimum plate appearances in order to filter players who did not have a significant amount of plate appearances. This will remove players from consideration to provide tailored results.

### 3.4.3 Data Results Return

Results will return players sorted by the difference between salary and expected salary. Players with the largest negative difference will be listed first. Data table will provide player name, statistic being used, salary, and expected salary.

# 4 Appendices

## 4.1 Appendices

Appendices will be ignored in accordance with the Grading Rubric.

# 5 Index

## 5.1 Index

Index will be ignored in accordance with the Grading Rubric.