

Spring 2015

Saberlytics Software Suite

Dylan Hallagan

St. John Fisher College, dylanhallagain@msn.com

Cody Thomas

St. John Fisher College, cody.thomas2011@hotmail.com

How has open access to Fisher Digital Publications benefited you?

Follow this and additional works at: http://fisherpub.sjfc.edu/cs_undergrad

 Part of the [Software Engineering Commons](#)

Recommended Citation

Hallagan, Dylan and Thomas, Cody, "Saberlytics Software Suite" (2015). *Computer Science Undergraduate*. Paper 4.

Please note that the Recommended Citation provides general citation information and may not be appropriate for your discipline. To receive help in creating a citation based on your discipline, please visit <http://libguides.sjfc.edu/citations>.

This document is posted at http://fisherpub.sjfc.edu/cs_undergrad/4 and is brought to you for free and open access by Fisher Digital Publications at St. John Fisher College. For more information, please contact fisherpub@sjfc.edu.

Saberlytics Software Suite

Document Type

Undergraduate Project

Department

Computer Science

First Supervisor

Nicholas Bucciarelli

Subject Categories

Computer Sciences | Software Engineering

Comments

Final presentation for CSCI480: Senior Project.

SENIOR PROJECT PRESENTATION 2015
MATHEMATICS/COMPUTER SCIENCE DEPARTMENT



SABERLYTICS SOFTWARE SUITE
DYLAN HALLAGAN AND CODY THOMAS

PROJECT OVERVIEW

- **Project Objectives & Description**
- **System Technical Overview**
- **Engineering Process Methodology**
- **Engineering Test Methodology**
- **Engineering Key Work-Products**
 - Requirements Analysis
 - Architecture Development/Models
 - Test Strategy/Test Cases/Expected Results/Corrective Actions
- **Project Key Metrics**
- **Project Schedule/Key Milestones**
- **System Preview**
- **System Demonstration**
- **Academic Key Knowledge Acquired/Applied**
- **Strategic Value**
- **Questions**
- **Lessons Learned**

PROJECT OBJECTIVES & DESCRIPTION

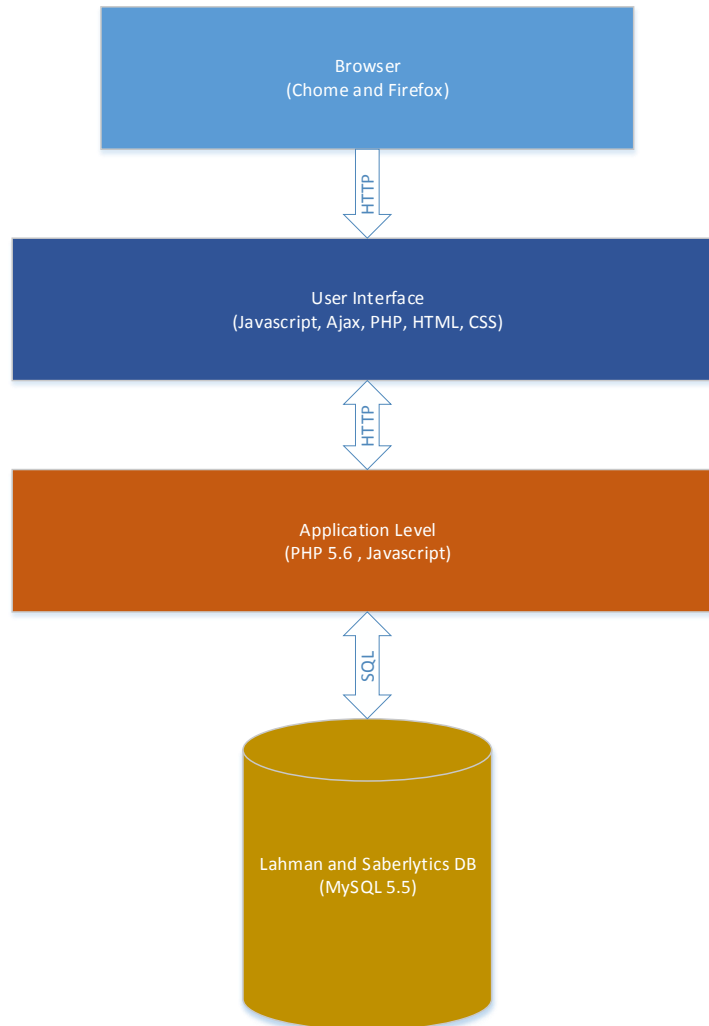
- **The primary objective of the Saberlytics Software Suite**
 - To provide GMs, coaches, scouts, and data analysts a tool to house separate research in one unified software
 - Readily access the research

SYSTEM TECHNICAL OVERVIEW

- **Hardware: Linux based server**
- **Soft Technologies:**
 - MySQL Database
 - PHP 5
 - JavaScript
 - HTML
 - Bootstrap CSS 3

SYSTEM TECHNICAL OVERVIEW

- **SYSTEM LEVEL DIAGRAM**



ENGINEERING PROCESS METHODOLOGY

- **AGILE: Used as a project development methodology**
 - Iterative design and build process
 - Allowed revisions and builds to be conducted in parallel
 - Ease of enhancement, refinement, and reduced errors

DEVELOPMENT PHASES - ITERATIVE

- **Phase 1:** Functionality Specifications
- **Phase 2:** Build a prototype for all tables and working data without a login system.
- **Phase 3:** Integrate Prototype 1 with user functionality
 - Login system
 - Project management
 - File uploads
 - Entering new reports
 - Player tear sheets
 - Realigned functionalities based on change in requirements

ENGINEERING TEST METHODOLOGY

- **Test Methodology**

- Software Engineering Test – Unit Testing
 - Correct functionalities (Based on requirements)
- Quality Engineering Test – Systems Acceptance Test
 - Compare SQLs results with industry leading resources
 - BaseballReference.com
 - Fangraphs.com

- **Value of Testing**

- Validates the capabilities of the application to work as stated

ENGINEERING KEY WORK-PRODUCTS

- **Requirements Document/Traceability**
- **Test Strategy (plans and executed results)**

- **Value of Work Products**
 - Iterative, Evolving
 - Based upon Final Version of Application

ENGINEERING KEY WORK-PRODUCTS

- Requirements Document/Traceability
























REQUIREMENTS TRACEABILITY MATRIX					
Project Name:	Statistical Baseball Analytic Software				
National Center:	St. John Fisher College				
Project Manager Name:	Dylan Hallagan & Cody Thomas				
Project Description:	Provide a software suite to analyse baseball players using R for custom projections, a player profile for scouts in the organization, and on demand reports for coaches during games				
ID	Req. ID	Business Requirement	Functional Requirement	Technical / System Requirement	System Design
001	001/002	We need a system that will provide a software suite to analyze baseball player statistics	The system shall take data input from the RetroSheet Database		
002	003/004/005		The system suite shall contain reporting services for custom projections based on the statistical analysis and data modeling.	The reporting services shall allow the user to create custom statistical selections to develop a model	The system shall allow the user to make selections through an SQL query builder
003	006				The system shall export the query to R for data modeling
004	007/008		The system suite shall contain a player profile	The system shall allow for scouts to add input to an athlete's player profile through a comment interface	
005	009/010			The system shall incorporate a queued comment section in the player profile containing scout name, scout picture, date added, tags, and the comment.	The comment section shall allow filters based on scout, date, and tags
006	011			The system shall display the athletes picture and basic statistics in the player profile	
007	012/013		The system shall contain GameDay Reporting functionality	The system shall be able to rapidly produce advanced opponent and player statistics for coaches	
008	014			The system shall display recorded Heat Map, Pitch FX, and matchup w/ Pitch FX images in the GameDay report	
009	015			The GameDay report shall provide an opponent pitch breakdown for analysis	
010	016			The GameDay report shall queue recorded analyst comments	
011	017/018/019		The system shall incorporate a user login database	The system shall provide specific login interfaces based on role	The system shall allow for the roles of admin, coach, scout, and analyst

PROJECT KEY METRICS

- **Number Open Tasks/Stories per Sprint**
- **Number of Changes in Requirements in Total**
- **Number of Changes to Source Lines of Code (SLOC)**

PROJECT SCHEDULE/KEY MILESTONES

- Standard Senior Project Class Schedule for All Teams

ID		% Work Complete	% Complete	Task Mode	Task Name	Duration	Start	Finish
1		0%	24%		Senior Project 2015	74 days	Mon 1/19/15	Thu 4/30/15
2		100%	100%		Executive Summary	10 days	Mon 1/19/15	Fri 1/30/15
3		0%	0%		Requirements	6 days	Mon 2/2/15	Mon 2/9/15
4		0%	0%		Key Metrics	2 days	Tue 2/10/15	Wed 2/11/15
5		0%	0%		Architecture	5 days	Thu 2/12/15	Wed 2/18/15
6		0%	0%		Test Strategy	3 days	Thu 2/19/15	Mon 2/23/15
7		0%	0%		Prototype Completed	15 days	Tue 2/24/15	Mon 3/16/15
8		0%	0%		System Reviews_1st Iterative	10 days	Tue 3/17/15	Mon 3/30/15
9		0%	0%		Test Strategy/Exection	2 days	Tue 3/17/15	Wed 3/18/15
10		0%	0%		Corrective Actions	3 days	Thu 3/19/15	Mon 3/23/15
11		0%	0%		Regression Test	2 days	Mon 3/23/15	Tue 3/24/15
12		0%	0%		System Reviews_2nd Iterative	5 days	Mon 3/23/15	Fri 3/27/15
13		0%	0%		Test Strategy/Exection	1 day	Wed 3/25/15	Wed 3/25/15
14		0%	0%		Corrective Actions	1 day	Thu 3/26/15	Thu 3/26/15
15		0%	0%		Regression Test	1 day	Fri 3/27/15	Fri 3/27/15
16		0%	0%		Review with Department	8 days	Mon 4/6/15	Wed 4/15/15
17		0%	0%		1st Review	1 day	Mon 4/6/15	Mon 4/6/15
18		0%	0%		Changes, if needed	2 days	Tue 4/7/15	Wed 4/8/15
19		0%	0%		2nd Final Review	1 day	Mon 4/13/15	Mon 4/13/15
20		0%	0%		Baseline Final Presentation	1 day	Wed 4/15/15	Wed 4/15/15
21		0%	0%		Submit to Department Chairperson	1 day	Wed 4/15/15	Wed 4/15/15

SYSTEM PREVIEW

- Dashboard Page

saberlytics.com/core.php?module=watchlist

Category	Total Work Orders
Category 1	3
Category 2	2

SUN	MON	TUE	WED	THU	FRI	SAT
29	30	31	1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17 Planned Performance Analysis	18
19 Planned Performance Analysis	20 Planned Performance Analysis	21	22	23	24	25

SYSTEM PREVIEW


- Team Page

Saberlytics Teams - Players - Leaderboards - Dylan Hallagan

Dashboard Watchlists Assignments - Schedule Reports Settings -

Dashboard / Team

Team



New York Yankees

The New York Yankees are an American professional baseball franchise based in the Bronx borough of New York City that competes in Major League Baseball's American League East division.

Total Payroll: \$231,978,886.00 Average Age: 31 Manager: Joe Girardi

Batting

Standard	Batted Ball	win Probability	Pitch Type	Pitch Value	Plate Discipline	Value	Pitch/x													
Season	Team	G	AB	PA	H	1B	2B	3B	HR	R	RBI	BB	IBB	SO	HBP	SF	SH	SB	CS	AVG
2013	NYA	162	5449	6044	1321	594	247	24	144	650	614	466	33	1214	57	36	3	115	31	0.2424
2012	NYA	162	5524	6231	1462	421	280	13	245	804	774	565	31	1176	62	49	2	93	27	0.2647
2011	NYA	162	5518	6306	1452	453	267	33	222	867	836	627	34	1138	74	51	8	147	46	0.2631
2010	NYA	162	5567	6379	1485	543	275	32	201	859	823	662	36	1136	73	44	8	103	30	0.2668
2009	NYA	162	5660	6447	1604	505	325	21	244	915	881	663	35	1014	54	39	8	111	28	0.2834
2008	NYA	162	5572	6257	1512	643	289	20	180	789	758	535	36	1015	80	39	11	118	39	0.2714

© Dylan and Cody 2015 Page rendered in 0.0002 seconds - Saberlytics Software

SYSTEM PREVIEW

- Player Page

Player

Derek Jeter
Positions: First Baseman, Third Baseman and Leftfielder
Bats: R Throws: R
Height: 75 Weight: 195lb.

Comparable Players

Similarity Scores Based Comparables	
Shannon Stewart	823.5
Paul Molitor	803.8
Robin Yount	802.9
Pete Rose	764.9
Mark Kotsay	760.3
Don Money	741.4
Rusty Staub	737.4
Dan Ford	733.1
Thurman Munson	728.9
Bob Watson	714.7

Aging Curve

War Projection for Derek Jeter
Source: Fangraphs 2015

Equation: $y = -0.04x^2 + 2.11x - 24.98$

SYSTEM PREVIEW

- Project Management

Sabertytics Search... Teams Players Leaderboards Dylan Hallagan

Dashboard Watchlists Assignments Schedule Reports Settings

Dashboard / Projects

Projects and Assignments

+ Add New Project Total Records : 3 Work Order

Search

No	#ID ↓	Player ↓	Assignment ↓	Status ↓	Start ↓	Deadline ↓	Actions
1	00011	Gian Carlo Stanton	Performance Analysis	Planned	April 17, 2015	Deadline Reached	Edit
2	00018	Derek Jeter	Prospect Analysis	Planned		Deadline Reached	Edit
3	00019	Manny Machado	Performance Analysis	Planned		Deadline Reached	Edit

<< First < Prev 1 Next > Last >>

© Dylan and Cody 2015 Page rendered in 0.0002 seconds - Sabertytics Software

SYSTEM DEMONSTRATION



ACADEMIC KEY KNOWLEDGE ACQUIRED/APPLIED

- **CSCI 231: Data Structure and Algorithms**
- **CSCI 353: Database Structures**
 - Database Architecture
 - Advanced SQL querying
- **CSCI 375: Programming the WWW**
 - PHP
 - Custom framework developed in PHP
 - Javascript/JQuery

STRATEGIC VALUE

- **Value Proposition**
 - Automated Valuation
 - Internal metrics
 - All franchise studies and work in one spot
 - Gameday Report Generation

LESSONS LEARNED

- **SQL Database Structuring**
 - With large databases time to render is a problem
 - Techniques for optimization
 - Crack a database into like queries/datasets
 - Example: Computed Player Statistics
- **Advanced PHP**
 - Login/Authorization
 - Session Handling
 - Background Fetchs
 - Dynamic pages rendering using AJAX Javascript/JQuery

QUESTIONS

