# Business Understanding Document
# IPL Analytics

Abhinav Chawla IMT2013002
Aditya Naidu IMT2013003
Shivam Kumar IMT2013042
Siddartha Padhi IMT2013043

16 October 2016

## 1 Problem Area Description

Coaches and scouts since long have been using intuition and gut feeling to select players in the squad and to decide the strength and the weakness of a player. If we are able to analyse the sports data, coaches will have analytics computed by us supporting or contradicting this intuition which will help in better understanding of the qualities of a player. As a big cricketing nation, we have chosen to take cricket as the sport to analyse upon.

Our project broadly aims at analysing the previous IPL seasons to find out various weaknesses, strengths and abilities of various players. We aim to find insights about which players have a mutual understanding with each others, which player is strong and weak against which type of bowler, how does a player perform under pressure etc.

## 2 Current Existing Solution

Most of the high level teams have hired a data analytics team who help them in this task. We want to emulate them and if possible, try to find new correlations.

## 3 Business Objectives Definition

Our project aims at coming up with an analytics driven solution which assists the team management with player selection and devising game play strategies based on several factors such as individual and team performance of a player against the opponent team and/or players. The business objectives we hope to achieve with our solution are implementation of better player selection methods than just intuition/experience and amplification of team's success rate.

# 4  Business Success Criteria

We, as the end output, want to predict which squad should be chosen by a specific team for a particular match, decide the batting order and decide which bowler to bowl to exploit the opponents's weakness and provide the best results.

# 5  Situation Assessment

We will describe the dataset in this section. We are provided with 578 Excel files, one for each IPL match since 2008. Each excel file includes metadata as description of the matchfor first 20 rows which includes the teams contesting the match, date, venue, toss result, umpire name, man of the match, the winner of the match and other extra details. The rows after that include the actual data. Each row consists information of each delivery which includes bowler name, batsman on strike, batsman off strike and runs scored on that ball.

# 6  Resource Inventory

We will be using data set provided by "cricsheet.org" for a ball by ball detail of every IPL match conducted. We would be scraping data from "3rd party websites to gather basic information about playing styles of every player. We have a competent team of 4 members with the required skill set to carry out the analysis. We have the required computing power i.e. Personal Laptops to carry out the data mining tasks.

# 7  Requirements, Assumptions and Constraints

There are no restrictions in legal or economic restrictions. The dataset from "cricsheet.org" is publicly available and free to use. Our main assumption is the validity of data. We are assuming that the data is correct and up to date. The data is limited to only IPL matches. A players performance is judged only by the way he performs in these matches. Performance in other formats of cricket matches such as International T20, ODI and Test Cricket are not being considered.

# 8  Risks and contingencies

The data set is limited to only IPL matches. We have not considered a players performance in other formats of the game which might have an important role in analysing a players performance accurately.

# 9    Cost/Benefit Analysis

There is no economic factor associated with this project. The data is publicly available and free to use. The softwares (Tableau) being used have an academic liscence and are free to use for this project. Out insights can help coaches form the optimum squad to play against a team and increase their chances of winning the match.

# 10    Project Plan

| Phase | Time | Duration |
|---|---|---|
| Business Understanding | 1 week | 9/10/2016 - 16/10/2016 |
| Data Understanding | 1 week | 17/10/2016 - 24/10/2016 |
| Data Preparation | 1 week | 25/10/2016 - 1/11/2016 |
| Data Modeling | 2 weeks | 2/11/2016 - 16/11/2016 |
| Evaluation | 1 week | 17/11/2016 - 24/11/2016 |
| Deployment | 1 week | 25/11/2016 - 2/12/2016 |