

# New Basketball Analysis

Caio Rocha Pereira      Pedro Lins de Araujo Lima  
Bruno Vitorino Cortez de Lira

November 17, 2017

## 1 Objetivos

Nosso objetivo é analisar e coletar dados de times e jogadores da NBA, buscando obter informações úteis e tentando prever as posições dos times, vistas as informações observadas e coletadas. Dessa forma, exercitaremos o conhecimento de diversas ferramentas e técnicas que foram vistas durante o andamento da disciplina.

## 2 Metodologia

O projeto consiste de 5 etapas: coleta, pré-processamento, análise, predições e conclusões. Cada etapa será detalhada a seguir.

### 2.1 Coleta

A coleta dos dados será realizada fazendo um scraping a partir do site: <http://stats.nba.com/>, nessa etapa usaremos ferramentas como BeautifulSoup para obtermos os dados que julgarmos interessantes de cada time a partir da lista que ele disponibiliza, armazenando esses dados em data frames de pandas para que o acesso desses dados nas etapas posteriores seja facilitado.

### 2.2 Pré-processamento

A partir dos data frames obtidos na seção anterior, realizaremos uma etapa de pré-processamento para lidar com dados inconsistentes ou que estejam faltando no dataset para que esses tipos de dados não prejudiquem o andamento do projeto nas etapas seguintes.

### 2.3 Análise

Tentaremos obter informações e relações entre os dados que não são visíveis apenas observando o dataset e as informações complementares. Essa análise consistirá de comparações, obtenção de médias, dependência de variáveis, e visualizações que nos levem a conclusões relevantes.

### 2.4 Predições

Tentaremos prever posições de times com base em seu elenco e desempenho, podendo estimar posição final na tabela, média de pontos por jogo e quantidade de vitórias, por exemplo. Também tentaremos obter predições sobre jogadores, tentando estimar fatores de seu desempenho e a relação desse com o desempenho do time.

### 2.5 Conclusões

Iremos mostrar o que foi aprendido durante cada fase do projeto, incluindo as partes mais simples e as dificuldades encontradas, assim como os resultados obtidos ao fim do processo.