National University of Ireland, Galway

The New Science of Winning

CONSULTANT REPORT

APRIL 2018

Analytics in Sports:

J.E. Cairnes School of Business & Economics

Prepared by:

Tanay Chowdhury-17232842

Conor Harte - 17231352

Abhinay Arora-17233252

Pankaj Juneja -17231787

National University of Ireland, Galway

Prepared For:

LA Lakers

Jeanie Marie Buss,

President

Rob Pelinka,

General Manager, LA Lakers

Contents

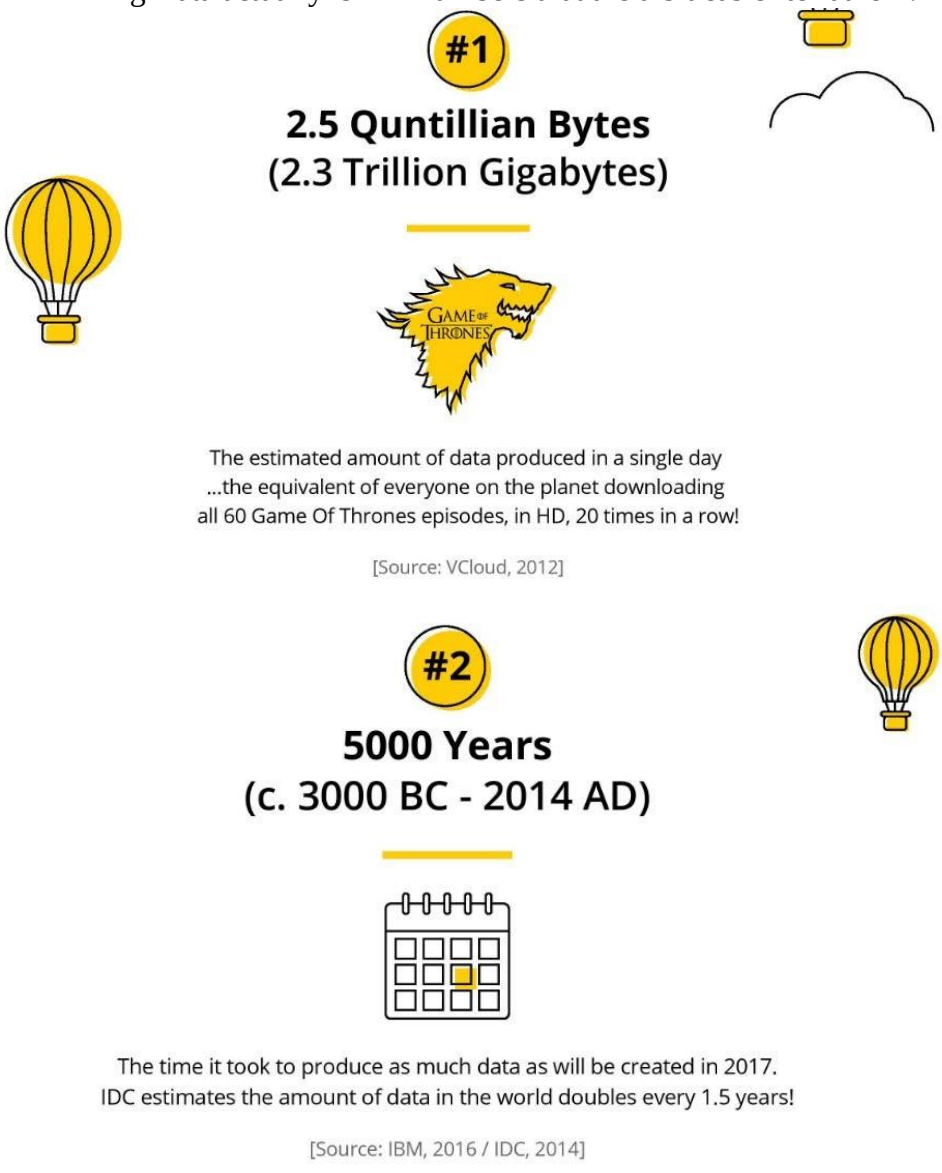
Executive Summary

LA Lakers, an American professional basketball team based in Los Angeles, attracted a gathering of authorities from the National University of Ireland's MSc Business Analytics program to provide recommendations concerning how big data & analytics can be used to up their preoccupation and the fan's understanding. The project scope involved analysis of the organisation and how big data & analytics can be used to improve the overall performance. An overview of the current sports market is provided, looking at the leaders & laggards of big data implementation. A big data strategy is provided, analysing LA Lakers in terms of athletic performance and business processes. An architectural strategy is provided which we believe will provide an easy to use platform for analysis. Applicable big data skills are researched and discussed followed by the importance and issues which may arise form big data governance and security.

Overview

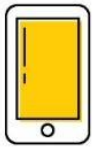
Big Data is Colossal. With nearly 2.3 trillion GBs of data created every single day and the data universe magnifying every two years, there's no denying that data is, and will continue to, shape our world. To your everyday user of tech, it may not seem like it, but Big Data trends are constantly transforming and evolving. Now a day, there are numerous articles or stories about how data is growing, irrespective of the source, that has the potential of changing the future. While the contextual information that are presented in them may be dissimilar, there's usually one thing that all of these Big Data stories have in collective: the numbers used to describe the immensity of Big Data are pretty hard to relate to.

What is a Zettabyte? How much data is actually generated every year? How large is the under-supply of data scientists needed to analyse all this big data?
So let's summarize what Big Data actually is – in numbers that are a bit easier to fathom.



#3

50 Billion
(6.7 Per Person on the Planet)



The number of smart devices that will be connected and producing data by 2020.

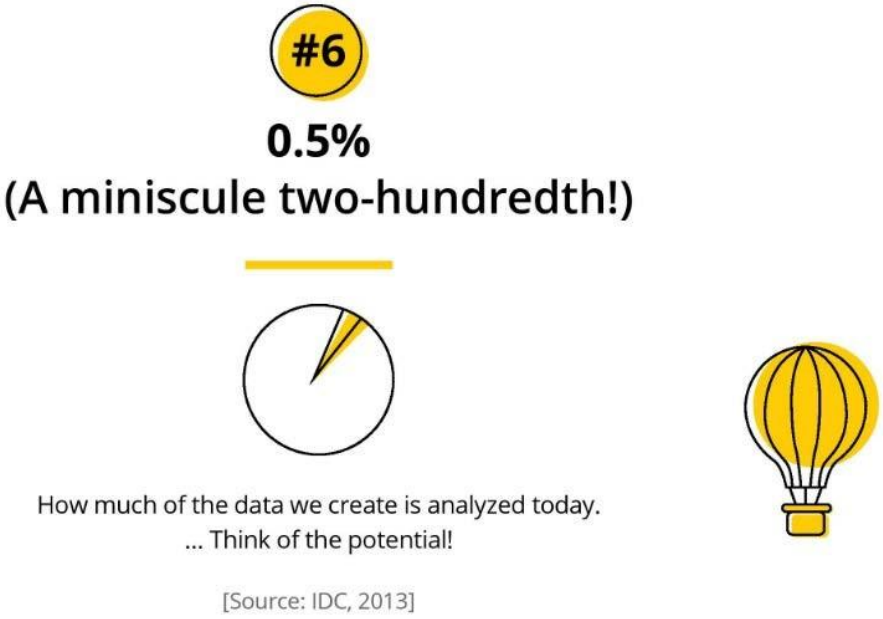
[Source: McKinsey, 2017]

#4

44 Zettabytes
(44 Trillion Gigabytes)

The estimated size of the "digital universe" - the amount of data created and copied each year - by 2020. Downloaded onto 128 GB iPad Airs, that would mean 6.6 stacks from the Earth to the Moon!

[Source: EMC, 2014]



How Big Data is Revolutionizing the games Industry

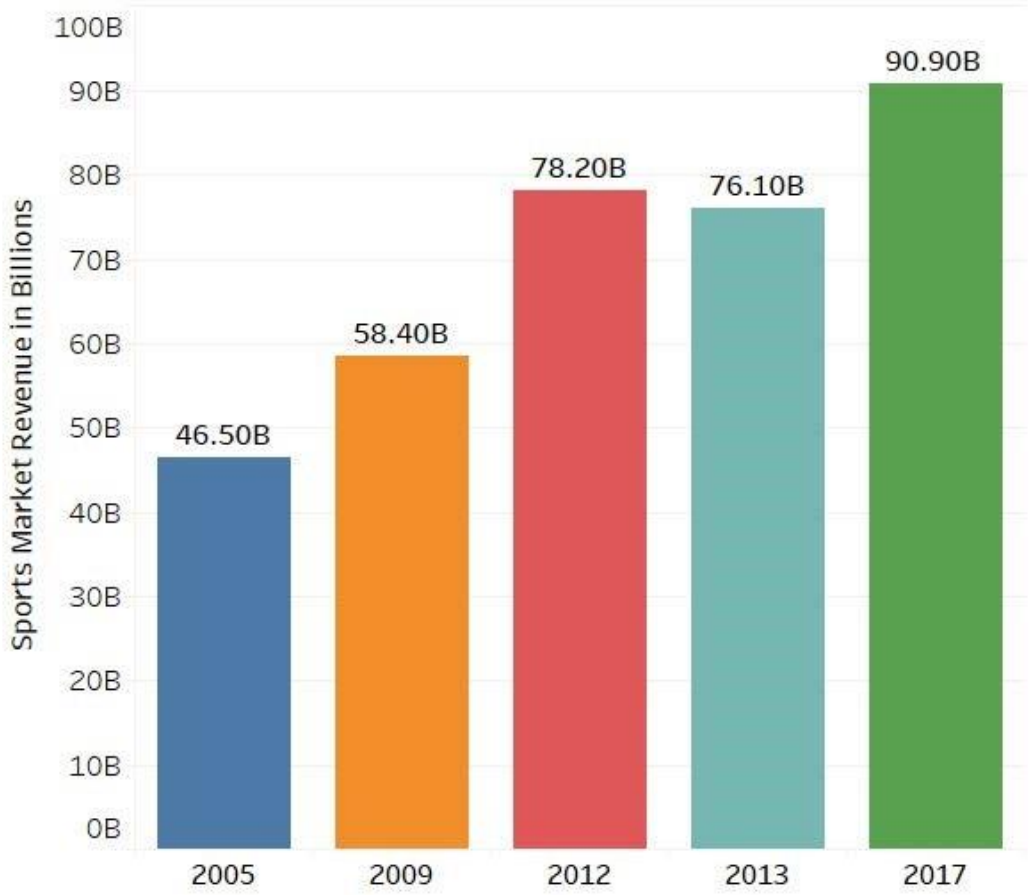
Sports related data is a huge big-data mint. Today, almost all major teams have data analytics experts on staff. Hollywood shone the light on data analytics in the famous Moneyball movie featuring the actor Brad Pitt. Teams often have to go through notes and hand them over to mathematicians who crunch figures that managers utilize to determine which player fits best – when, how, where and how much. All this data creates a comprehensive profile of a player to determine whether a player should be selected or traded or played on a given day.

Player analytics is a growing trend in professional sports and will have a wider jolt on the entire sports industry. Any team that does not apply them to its fullest extent is at a competitive disadvantage. But, it's not only about the team, data analytics can also be used to find what fans prefer.

Current State of sports analytics;

The current global sports market is huge with its total revenue projected around 90 billion dollars in 2017. Sports is one of the sectors in which data analytics has demonstrated great value and has a great potential with major professional sports teams putting them to use.

It is exciting to see how analytics is impacting the modern sports industry.



Leaders and Laggards in Sports Analytics:

Some of the world leaders in sports analytics are Stats and Zebra sports solutions:

Stats:

Stat is trusted by a huge number of fans and competitors around the world to dissect sport information from in excess of 100,000 games per year with unrivalled speed and exactness. Stats gives sports information, innovation, and substance to meet the developing needs of the Media, Broadcast, Leagues and Teams, Fantasy suppliers and players, and in addition major B2B and B2C brands.

Stats (product: SportVU) offers high tech Tracking Solution such as football tracking calculation that is produced utilizing a portion of the best catch, collection and investigative advances. Upheld by the honour winning STATS information science group, the best parts of these arrangements have been joined into a solitary, complete process that will give steady live and post-amusement rectified information.

The STATS SportVU following framework conveys execution insights by separating and preparing directions of players (X, Y) and the ball (X, Y, Z) through HD cameras and in addition advanced programming and factual calculations.

Insights delivered from the framework incorporate, however are not restricted to: Distance Travelled, Average Speed, Max Speed, Momentary Speed, Number of Sprints, Coverage Maps, Time of Possession, Player Possession, Ball Speed Ball Distance, Zone Coverage, Team Formation

In its main goal to make STATS SportVU the worldwide standard for exactness among optical-following frameworks, STATS charged a year free examination with the Technical University of Munich to approve

ZEBRA SPORTS SOLUTIONS

Zebra has intensely tested the norm in sports by introducing patented RFID innovation to deliver an outlook changing, player following framework in proficient games. Zebra Sports Solutions empower a trove of information for real-time insight. As its full range of capacities keeps on being utilized over various games properties, its effect will be critical. New player tracking insights will change how fans relate with sports through cross-channel engagement and everlastingly coaching and player work force assessments.

Athlete Performance: Zebra Sports is altering how scouts, mentors and coaches assess players. Their solution enables coaches to keep competitors at top execution by creating bespoke preparing plans in light of individual measurements and weariness edges. The exclusive programming further empowers NFL mentors to match up with official Game Day information and coordinate it to rapidly upgrade preparing.

*Instructing and Training***:** Zebra's solution enables a consistent, data-forward analysis method, allowing coaches to strategize in-game and in-practice using real-time data to adapt and be endlessly nimble in play-calling. Whether assessing route running patterns, evaluating grouping and formation, analysing the separation distances on pass coverage or assessing quarterback pressure, there are limitless, highly customizable applications.

*Fan Enhancement***:** This exclusive ongoing player tracking arrangement improves fan viewing through transmissions, portable engagement and in-stadium, empowering fans to see rarely seen continuous player information and convincing new execution experiences. This has made for a more immersive review encountering, expanding fan mindfulness, availability and the understanding fans have of players, groups and on-field methodology.

Laggards:

QSTC

In light of research which has shown that many businesses are still not getting value from their data science and analytics, it’s worth looking at some of the factors which differentiate those that are.

Observing athletic execution in swimming, particularly untamed water swimming as observed amid a marathon, was about outlandish. The shifting elements were just excessively steep. In any case, Queensland Sports Technology Cluster in Australia created miniaturized scale sensor information catch that gives point by point 3D investigation of a swimmer's development, including estimating the volume of water uprooted with each stroke. The model is still under development and is not perfect yet.

Big Data Strategy

When outlining the strategy for any sports organisation, two main areas exist. The initial and most obvious area is that of the sports team itself. The objective of the sports team is to perform to the best of their ability and in return win in the field of play. However, sports organisations need sources of revenue to fund their team. Therefore, we have decided to split the big data strategy into two sections, Athletic Performance and Business Processes, and look at how big data and analytics can add value in both these areas.

Athletic Performance

Player Analytics has had a massive impact on Professional Basketball in recent years. LA Lakers need to get up to speed and fully utilise analytics to take them to the next level. Analytics has provided major insights which have changed the way the game is played. It was recently revealed that the reward outweighed the risk when it came to shooting three pointers. This has resulted in drastic rise in the amount of three points teams are averaging per game, rising from 18.0 in the 2010-11 season to 23.8 in the 2015-16 season [1]. The power of analytics can also be used by each individual team to alter their style of play. SportsUV track the movement of each player during a match creating large data sets with detailed data on players performance. Through data mining LA Lakers can determine their strongest selection of players, and use this to alter the team selection to target opponent teams' weaknesses. Training sessions can be altered to suit the needs of individual players. Models can be created to predict injury or fatigue in players letting coaches know when players should be rested. Through the use of analytics, a data driven decision-making culture can be installed which ultimately eliminates biases and ensures coaches they are making the correct decisions.

Business Processes

While analytics is ensuring the correct decisions are being made on the court, it also a major player in the pursuit to maximise revenue. One main area is which analytics can play a big part is ticket sales. Analytics can be used to implement a variable ticket pricing strategy. Instead of having standard ticket pricing all year round, models can be produced to predict demand for certain games and prices can be matched accordingly which will help move towards a sold-out stadium for every game. Second hand ticket markets and hype gathering on social media can be analysed to predict the demand levels. One of the largest revenue streams for any basketball team is through season ticket holders. Prediction models can be created to identify season ticket holders who may not renew their ticket, by analysing how many games they have attended and how many season tickets they have bought previously, and then the season ticket holders classified as low likelihood of renewal can then be negotiated with.

Architectural Strategy:

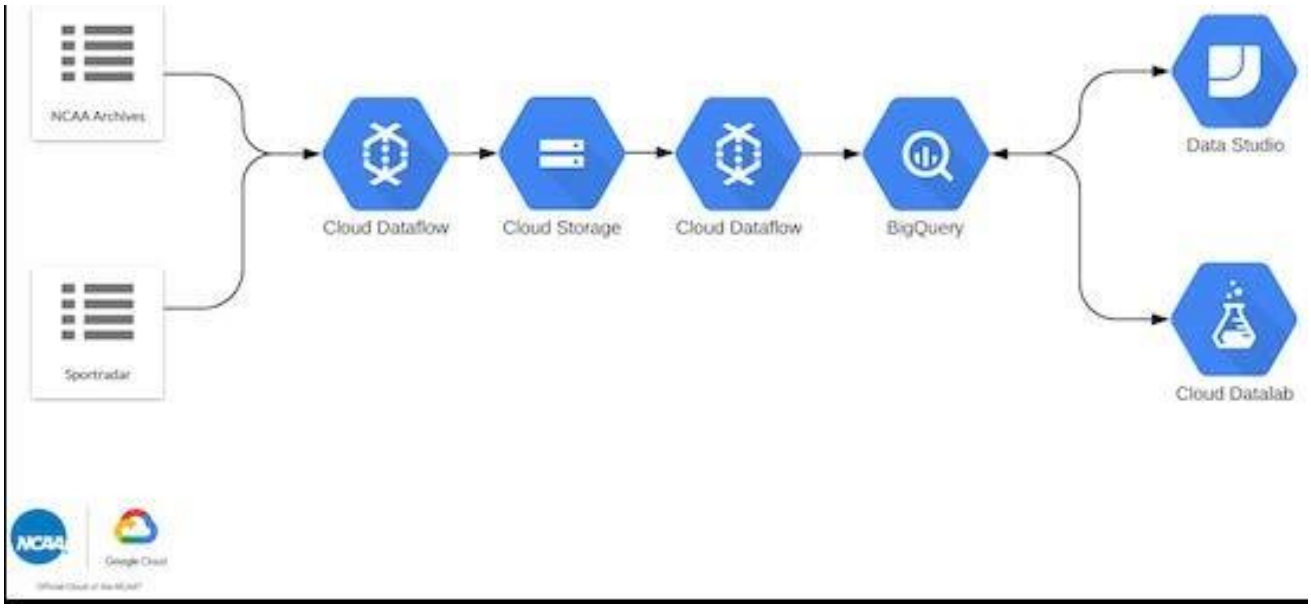


Figure 1: Architectural Framework

Extraction, transform, load(ETL)

Our primary data sources are Sports radar and NCAA Archives which contains data from play by play files and box score. Both source is official data source for getting all the details related to every game. Data is collected from these two-data source in various format like XML, JSON, CSV etc. The files are extracted through various file transfer protocols and integrated and stored in the system
To make the raw data actionable, the data in XML and JSON need to be transformed into a structure that can be loaded into Big Query for analysis. We are using Apache Beam graphs to do most of our ETL work. Cloud data flow provides the processing power of tens of thousands of files in a minute. The output of ETL jobs produces a public dataset on Big Query. This dataset enables data scientist to do their analysis work.

Analysis:

With the data being loaded in Big Query we will be good to go for analysis. Most of the analysis will be performed on the play by play data table. Big Query is an excellent tool for big and small data and allows anyone to express queries in standard SQL without worrying about scaling and index optimization. The user just need to write the query and run it and Big Query handles the rest.

Visualization:

We recommend two tools for visualisation as both the tools have different speciality. Google data studio is an excellent tool for data visualization. Data studio enables development of dashboards targeting data sources that includes google sheets. MySQL, Google Analytics, Big Query etc. Cloud Data lab combined with seaborne or Matplotib provides seamless connectivity to Big Query through pandas integration.

Real Time analysis:

Now based on the above task we need to prepare a predictive model which can generate useful insights during match by analysing the real time data.

Big Data Skills:

Some of the Big Data technologies that LA Lakers could use are-

Predictive Analysis:

SAS advanced analytics suite of software is the market leader in predictive analysis. Using fit bits and other types of sensors the LA Lakers can get statistics on heart rate, sleeping patterns of its players etc. LA Lakers can predict the chances of injury to a player. They can use predictive models for the same. Once they find that a player is highly fatigued, they should give him days off from practice or even games to revitalize himself, this will prolong the careers of the players as well as enable LA Lakers to field players who are at their hundred percent. MOCAP Analytics:

NBA is using the video technology that tracks the movement of players throughout the game , Six cameras track players' every move by way of a specific set of data points including, but not limited to players' names, numbers, and the ball. The league records these statistics and stores them in three data centres. This data is open to all the NBA teams; it's up to the teams to better analyse this data and come up with insights that help them achieve their Objectives.

MOCAP Analytics is a big data analytics firm which specializes in bridging gap between the coordinate data and basketball jargon. Using insights from the coordinate data the firm will advise the LA Lakers on how to improve the its defence. Also, it's heat map will record the player's frontcourt touches during the period of games, this will aid the coaching staff in positioning the player's better.

Bayesian Network Modelling:

Whereas other types of analytics focus on the individual performances of the players, Bayesian Network modelling focuses on individual performances of the player as well as the interaction between the player. There is saying in sports that 'whole is greater than the sum of its parts'. Using Bayesian Network Modelling we can find out which group of players play well together, or which group of players are good for playing certain style of basketball for e.g. which group of players are best for playing small ball.

Big Data Governance and Security:

Big Data Governance and Security is becoming a major issue as the industry grows. LA Lakers will be storing detailed information regarding each player and may gain insights which could have a large effect on the player. They also must be able to keep their analysis secure against the onlook of rival teams. A data leak could provide oppositions with a competitive advantage in the run up to major games.

LA Lakers will be holding detailed and specific information on many aspects of their players life. Biometric data like temperature, speed, heart rate, sleeping patterns, calorie intake and distance covered are all recorded. This data can be secured at a central location or a third-party vendor could be used to secure the data. It is of vital importance that the Lakers have consent of their players before storing the data. If analysis has predicted an injury for a player, this could have major impacts if released. Insurance policies would rise drastically while any sponsorship deals may be cut. Therefore, the consent of players should be obtained and all security practices must be followed to maintain the level of big data governance and security needed.

The club must also be aware of potential attacks they may face from competitors. The insights gained from the analysis are of vital importance To the Lakers but would also provide an advantage to their opponents. The data may reveal weaknesses in the Lakers play while also altering opposition to tactics the Lakers may apply. If the opposition team is aware of the insights, then these insights will no longer provide any competitive advantage. Therefore, security is of paramount importance to the club. LA Lakers should make sure the firms working on this sensitive data should have policies in place as to who has access to data. The data should be encrypted when stored in database, so even if somebody hacks the data, he/she still is unable to get meaning out of the data.

Big Data & Analytics is used to provide insights which will support the growth of the organisation. A breach of data relating to governance or security could have a major effect on the company. This could result in the loss games or even key players. Therefore, we recommend the Lakers invest adequate time and money needed to ensure security practices are up to date and fit for purpose.

References:

[1] Mashal, Jason (2017) **'The Three Point Revolution: An Economic Approach to Analytics'** Available at: https://edgeworthеconomics.com/experience-and-news/edgewords-blogs/edgewords/ article:03-29-2016-12-00am-the-three-point-revolution-an-economic-approach-to-analytics/

[2] Zebra Sports Solutions (2018) Available at: https://www.zebra.com/us/en/solutions/location-solutions/zebra-sport-solution.html

[3] Stats (2018) Available at: https://www.stats.com/about/