BI PROJECT DOCUMENTATION

Link to original template

Student ID:
136446

Full name:
Mirav Hirji Bhojani

Industry:
Sports

Sector:
Cricket

Subsector:
International

Abstract

<The Abstract must be a concise yet comprehensive reflection of what is in your document. It must be self-contained, without abbreviations, footnotes, or references. It should be a microcosm of the full document. It must be between 150–250 words. It must be written as one paragraph, and it should not contain displayed mathematical equations or tabular material. Please ensure that it is grammatically correct.>

Section 1: Independent and Dependent Variables

1. **What is the title of your IS Project 2?**

Please refer to the following explanation of variables in research: https://youtu.be/MnfRdTCUIsc

Title of the project	Player Performance Prediction using Machine Learning Algorithm Case Base Cricket in Kenya.
Dependent variable (target/outcome) in the project *(Specify **only one** dependent variable)*	Selection of the starting eleven of players as it is the dependent variable since it depends on the players performance for a player to get selected they need to be classified as good or average.
Independent variable(s) (predictors) in the project *(this is in the form of the actual system that will be developed, e.g. an inventory system (independent variable) will affect stock taking (dependent variable))*	The independent variable will be the performance prediction model.
Hypothesis *(Provide a proposed*	Cricket in Kenya is declining due to the outdated technology and expertise skills they have since they are not able to choose the right combination of players for a match leading to them lacking required skills and hence

statement of how the independent variables (modules of the system) affect the dependent variable)	losing many matches.
Research Question *(Formulate a "How?" question using the hypothesis. Quote: "you will never get the right answer without asking the right question")*	How is the starting eleven currently chosen for matches and what are the challenges associated with this process.

Section 2: Problem Statement

2. **Answer the following 5 questions to <u>clearly</u> (by being brief and straight to the point) specify the Problem Statement based on your IS Project 2**

What is the problem?	Poor starting eleven team combination for a cricket match hence leading to losses in cricket matches as the team does not have the right players that fit in specific situations.
Where is the problem being experienced?	In Kenya
Since when has it been a problem?	Since 2003
Why is the problem still being experienced?	Due to lack of expertise from coaches and outdated technology being used for analysis
Who is affected by the problem?	Majorly Kenya, a cricket playing nation is being affected the most as the decline has also led to a decline in fame and money that was brought into the country by the sport. However, the players are also affected since the scope of cricket is slowly worsening and thus players

	cannot only rely on it to feed themselves and their family.

Based on the above answers, what type of AI-driven problem is it? *Options: classification problem, regression problem, association problem, or clustering problem.* ***Note:*** *what we have focused on in the BI Option is classification and regression problems.*	Classification Problem

Section 3: Proposed Solution (Data-Driven Solution in the form of a Data Product)

3. **Describe the proposed data-driven solution by answering the following questions**

Which algorithms (list more than 1) can be used to train the model based on whether it is solving a classification, regression, association, or clustering problem? *E.g., linear regression, logistic regression, Linear Discriminant Analysis (LDA), decision tree, Naive Bayes, k-Nearest Neighbours (KNN), Learning Vector Quantization (LVQ), Support Vector Machines (SVM), etc.*	Logistic regression, Decision trees and KNN
What reports does the business require from the data product? *For business-facing analytics*	The prediction of the performance of individual players so as to help Cricket Kenya to choose the best possible starting eleven needed to win the match.
What reports do the customers of the business require from the data product? *For customer-facing analytics*	The players will require this information to get motivated before they play the match so that they know that they will anyways perform well and thus no need to be nervous cause nervousness leads to under performance of players

What will the model be used to predict (identify only one variable and be specific. Specify the possible classes if it is a classification problem.)? *The target/outcome/dependent variable* *It should be related to the project's dependent variable.*	The model will be used to predict the performance of players for upcoming matches. The classification will be Good, Average or Poor. For the team selection the first with good classification will be selected and if they are exhausted and the Playing eleven is still not complete then those with average classification will be selected.
What variables does the model require to make the prediction (list multiple variables and be specific)? *The predictors/independent variables*	Previous score, average, venue, strike rate, number of wickets in previous match and opponent team

Where are you going to get a sample dataset? *Provide the link to a specific dataset that has similar variables. Possible sources of datasets include:* 1. *University of California (UC) Irvine: https://archive.ics.uci.edu/ml/datasets.php or https://archive-beta.ics.uci.edu/* 2. *Kaggle: https://www.kaggle.com/datasets* 3. *Data Science Dojo: https://datasciencedojo.com/blog/datasets-data-science-skills/ or https://code.datasciencedojo.com/datasciencedojo/datasets* 4. *Kenya Open Data: https://kenya.opendataforafrica.org/data/#menu=topic* 5. *openAFRICA: https://africaopendata.org/* 6. *Datahub.io: https://datahub.io/collections* 7. *Data.world: https://data.world/* 8. *Google Data Search: https://datasetsearch.research.google.com/* 9. *Google Public Data Explorer: https://www.google.com/publicdata/directory* 10. *Data.gov: https://www.data.gov/* 11. *Global Health Observatory Data Repository: https://apps.who.int/gho/data/node.home* 12. *UNICEF Data: https://data.unicef.org/* 13. *Earth Data: https://earthdata.nasa.gov/* 14. *CERN Open Data Portal: http://opendata.cern.ch/*	Cricket ODI - Players Performance | Kaggle

15. *FBI Crime Data Portal: https://crime-data-explorer.fr.cloud.gov/* 16. *New York City Taxi Trip Data: https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page* 17. *Our World in Data: https://ourworldindata.org/* 18. *World Bank Open Data: https://data.worldbank.org/* 19. *World Poverty: https://worldpoverty.io/map* 20. *Worldometers: https://www.worldometers.info/* 21. *European Data: https://data.europa.eu/en* 22. *Livestock Data for Decisions (LD4D): https://www.livestockdata.org/type/datasets* *etc.*	
What can't the user (both the employees in the business and the customers of the business) do now that they will be able to do in future using the data products (business-facing analytics dashboard, customer-facing analytics dashboard, and the prediction model)?	Currently the scoring is done on an external app and the and the starting eleven selection process is based on guts and feelings and not any other analytics or technology and thus in future once the system is ready the process of selecting the starting eleven will be eased up and be more reliable than it currently is hence resulting to favourable outcome for the match for Kenya and helping the country grow in the sports and allow improve living standards of the players and allow them to do well in their desired field of interest which is Cricket.

Section 4: Industry, Sector, and Subsector

4. **Can the problem statement involve a business? If yes, in which industry, sector, and subsector would the business operate in according to the Industry Classification Benchmark?**

Industry	Sector	Sub-Sector
Sports	Cricket	International

Section 5: Strategic Objectives based on Industry Standards/Expectations

5. **Using the industry's standards/exprBusiness Processes Perspective, the Customer Perspective, and lastly, the Financial Perspective.**

Kaplan & Norton's Balanced Scorecard Perspective	Strategic Objective
Innovation and Learning Perspective	SO1: To widen the scope of the system to work for other formats of cricket.
Internal Business Processes Perspective	SO2: To enable selecting the perfect combination of starting eleven for a match.
Customer Perspective	SO3: To motivate players to perform well in order to be selected in the starting eleven.
Financial Perspective	SO4: To attract more sponsors by the end of 2024 in order to help finance well for cricket tournaments and host more home matches.

Section 6: Leading and Lagging Indicators

6. **Identify leading and lagging indicators that can be associated with each strategic objective. Give each KPI a unique identifier.**

Strategic Objective	Leading Indicator	Lagging Indicator
SO1	KPI1a: Increase the number of formats that can be supported by the system	KPI1b: Total number of formats supported by the system
SO2	KPI2a: Availability of players previous performance data	KPI2b: Assessment of players previous performance
SO3	KPI3a: Performance in team competitions and internal matches	KPI3b: Player feedback on motivation
SO4	KPI4a: Number of sponsors showing interest in investing finances to sponsor the national team	KPI4b: Total sponsorship revenue generated from the sponsorship money

Section 7: Presentation Category for each Indicator

7. **Classify each KPI as raw numbers, progress, or change, with a preference for the KPIs presented in the form of a change indicator**

Leading Indicator	**Presentation Category** (raw numbers, progress, or change)	Lagging Indicator	**Presentation Category** (raw numbers, progress, or change)
KPI1a	Progress	KPI1b	Raw number
KPI2a	Progress	KPI2b	Change
KPI3a	Raw number	KPI3b	Raw number
KPI4a	Raw number	KPI4b	Raw number

Section 8: Presentation Attribute for each Indicator

8. **Describe the 5 attributes (measure, target, source, responsible user, and frequency) <u>for each KPI</u>**

For the source of the data: *In addition to specifying the actual source (the OLTP relational database), also specify the ethical considerations that need to be made as the data is being extracted from the source. Refer to the "The Kenya Data Protection Act No. 24 of 2019".*

KPI1a: Leading Indicator

Presentation Attribute	Description
Measure	Increase the number of formats that can be supported by the system
Target	Increase the number of supported formats from 1 to 4
Source	The source is the IT system database with an extraction ethical consideration of data minimization that is to only extract data that is required and not anything that is not required
Responsible User	Software Engineer
Frequency	Annually

KPI1b: Lagging Indicator

Presentation Attribute	Description
Measure	Total number of formats supported by the system
Target	System to support all the formats of the game
Source	The IT scoring and prediction system with and ethical consideration of not unethically retaining the data extracted for

	longer than it is required
Responsible User	Software Engineer
Frequency	Annually

KPI2a: Leading Indicator

Presentation Attribute	Description
Measure	Availability of players previous performance data
Target	To have at least 100% of players' previous performance data available
Source	Official team records and the ethical consideration while extracting this data should be protected such that no third party can get access to this data
Responsible User	Selection Committee
Frequency	After every match

KPI2b: Lagging Indicator

Presentation Attribute	Description
Measure	Assessment of players previous performance
Target	To assess all the players well according to their performance and skills in previous matches
Source	The source is Player Statistics with a consideration that the data should only be used for the purpose it was collected for and not any other reason

Responsible User	Selection Committee
Frequency	After every match

KPI3a: Leading Indicator

Presentation Attribute	Description
Measure	Performance in team competitions and internal matches
Target	To achieve a top-three finish in a league or tournament
Source	Match statistics and the ethical consideration should be limited to the purpose of collection only
Responsible User	Players
Frequency	After every match

KPI3b: Lagging Indicator

Presentation Attribute	Description
Measure	Player feedback on motivation
Target	All players expressing satisfaction with their motivation levels
Source	Players and the ethical consideration should be limited to the purpose of collection only
Responsible User	Selection Committee
Frequency	After every match

KPI4a: Leading Indicator

Presentation Attribute	Description
Measure	Number of sponsors showing interest in investing finances to sponsor the national team
Target	To have at least 10 sponsors showing interest in sponsoring the national team in 2023-2024 cycle
Source	Sponsorship details provided by the financial department extraction consideration of Data subject rights that is to respect the privacy of the data of the sponsors
Responsible User	Cricket Kenya
Frequency	Annually

KPI4b:Lagging Indicator

Presentation Attribute	Description
Measure	Total sponsorship revenue generated from the sponsorship money
Target	To achieve 10 million shillings in sponsorship revenue for the 2023-2024 cycle
Source	Financial Records of the year provided by the Finance department. Data to be extracted following the lawful basis of using the data that should have a meaningful reason to use the data.
Responsible User	Cricket Kenya
Frequency	Annually

Section 9: Measures and Dimensions

9. Identify the measures (categorical data) and the dimensions (quantitative data) for the applicable KPIs

Categorical Data (Measures): Nominal, ordinal, or time-series data *(State the actual category and whether it is nominal, ordinal, or time-series data)*	**Associated KPIs**
General match statistics (Nominal Data)	KPI3a, KPI3b
Specific player performance ranking in a match (Ordinal Data)	KPI3a
Revenue generated from sponsorship money in a specific year (Time-series Data)	KPI4b

Quantitative Data (Dimensions): Interval scale or ratio scale data	**Associated KPIs**
Number of cricket format supported by the system (Ratio Scale)	KPI1a, KPI1b
Previous player performance data (Interval Scale)	KPI2a, KPI2b, KPI3a
Number of sponsors interested (Ratio Scale)	KPI4a, KPI4b

Section 10: Type of Message

10. Describe the type of message (one message per KPI) that each KPI intends to communicate

KPI	**Proposed Message Type** *(comparison, distribution, composition, or relationship; refer to this article: https://extremepresentation.typepad.com/blog/2006/09/choosing_a_good.html)*

KPI1a	Comparison specifically to one variable or non-cyclical data
KPI1b	Distribution specifically to single variable with few items
KPI2a	Composition changing over time with many periods
KPI2b	Comparison changing over time with many periods
KPI3a	Comparison changing over time with many periods
KPI3b	Relationship with two or three variables
KPI4a	Distribution with different number of variables
KPI4b	Composition with static variable

Section 11: Visualisation Type

11. Identify visualisations that can be used for each KPI

KPI	**Proposed Visualisation Type** *(at least 2 types for each KPI)*
KPI1a	Bar chart or Line chart
KPI1b	Column histogram or Line histogram
KPI2a	Stacked 100% area chart or Stacked area chart
KPI2b	Multiple bar chart or Column Chart
KPI3a	Column Chart or Line Chart

KPI3b	Scatter Chart or Bubble Chart
KPI4a	Column Histogram or Scatter Chart
KPI4b	Waterfall Chart or Pie Chart

Section 12: Business-Facing Analytics Dashboard Layout

12. Use a mockup to present the proposed position of each visualisation on a dashboard for the employees of the business

Proposed (not mandatory) applications that can be used to design the mockup of the dashboard include: https://clickup.com/features/dashboards

Note: This is a mockup of the dashboard. A mockup is a visual design of the final version of a User Interface, not the final version.

Also Note: The dashboard must have all the 4 Kaplan & Norton's Balanced Scorecard Perspectives

Section 13: Customer-Facing Analytics Dashboard Layout

1. **Take a screenshot of the developed dashboard for the customers of the business**

Note: The dashboard must be functional (connected to a database)

The following site can be used to generate dummy data required by the dashboard: https://generatedata.com/generator

Section 14: DataOps Strategy

2. **Describe the DataOps strategy that will be applied to deliver the solution**

Describe how functional testing, unit testing, integration testing, and API reliability testing will be carried out on the data products	Functional testing mostly was done where data was pre processed that is during the data imputation, data cleaning and data transformation stage where all the data was checked to see if there are missing values and if they were present then appropriate measures were taken such as what to replace with the missing values that is mostly the mode or the mean value of the dataset. For unit testing specific test cases were taken into consideration, one being through code where the missing data was checked for and printed out the results to check if there is any missing data. Integration testing was done while loading the pre-processed data into the training model to assure that no data was lost in the process; this was evident when the data that was loaded in the training model was the same one that was assigned after the data was transformed. For API reliability testing there was nothing much since there was no API used the only API that I wanted to use was for deployment but since i did not deploy the model there was no API used thus no need to do API reliability testing.
Describe how the prediction models will be validated	Validations were done such that first of all the dataset was split into 80:20 ratio where 80% of the dataset was used to train the model and the remaining 20% was used to test whether the unseen data will return an accurate result. Moreover, K-fold cross-validation was also done to confirm that the result returned by the model is more robust and is not influenced by any factor. Once the model was trained hyper parameter tuning was also carried out to find the optimal hyper parameters to ensure high performance of the model.
Specify all the Minimum Viable Products that can be delivered to form the whole data product (to form 2 dashboards and 1 model)	Dashboard 1: Business Facing Analytics Dashboard. This dashboard is to help the business, in my case the cricketing body, to carry out statistics calculations that will benefit them when calculations will be used to make informed decisions. Dashboard 2: Customer facing Analytics Dashboard: This dashboard is to help

	the customers, in my case it is the players who will use this dashboard and carry out statistical calculations to improve their overall game and performance. Model: This is the developed model that will be used to classify the players according to their performance that is based on the runs they score and this then will be sued by both the players and the cricketing body to help improve performance of the players so that they do not remain in the same category and instead be advancing to the higher categories.
Specify when each Minimum Viable Product will be delivered (the exact date according to the IS Project 2 Gantt chart and according to the BI semester dates)	Dashboard 1: 19/07/2023 Dashboard 2: 27/11/2023 Model: 27/11/2023

Section 16: Link to Model Training Markdown

3. **Specify the link to the markdown hosted on https://github.com/ that displays how the model was trained**

Link: BIProject--136446/README.md at main · MiravBhojani/BIProject--136446 (github.com)

Section 17: BI Implementation Strategy

Many businesses fail by attempting to adopt new technologies too fast across the entire business without a proper plan. A proper plan should outline how the technology will be used to solve a clearly defined problem.

4. **Propose a BI implementation Strategy that the business should follow to make the project a success:**
    a. **Data Quality and Validation**
    This should be guided by the BI1 Concept 1 theory on DataOps and it should specify business policies (rules) that should be enforced for the following principles:
        i. ***Data governance:*** *Policies to ensure that data is accurate, consistent, reliable, and easily accessible to authorised users*

        ii. ***Data auditability:*** *Policies to provide the ability to track and review the changes made to data over time to guarantee its accuracy, completeness, and compliance with regulatory and business requirements*

        iii. ***Data lineage:*** *Policies to provide the ability to track the origin, movement, and transformation of data from its source to its destination throughout its lifecycle*

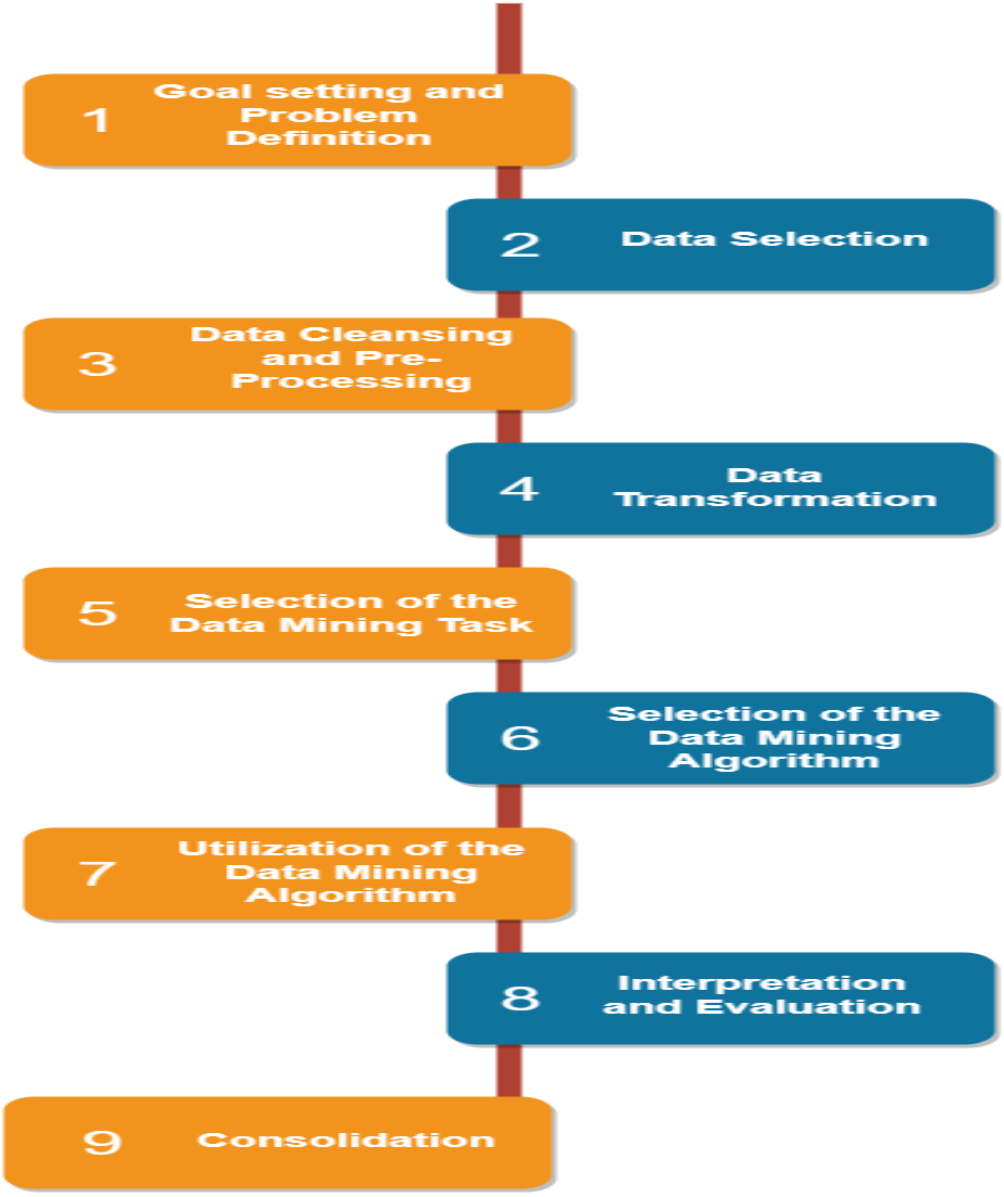
For data governance, ensuring standard data collection for features like runs, balls faced, and batting positions ensures accuracy and reliability. Data should only be accessible to the authorised personnel and not everyone this is to maintain integrity of data and not allowing any modification that can lead to storing inaccurate data which if decisions are made on this data then those decisions may turn out to be inaccurate as well which may harm the players and the cricketing body itself. For Data Auditability, creating change logs enables tracking and reviewing

edits to the dataset, maintaining its accuracy, completeness, and compliance, also having backups of the original data and then comparing it after changes are made to check what has been edited. Data lineage documentation will detail the journey of metrics from their sources through transformations, ensuring transparency and accountability in the analysis process. By implementing these policies within the project, a robust foundation for accurate, accessible, and traceable data, conforming to DataOps principles is laid off. This approach minimises risk effect and ensures a controlled integration of technology into the project, avoiding pitfalls associated with rapid, unstructured technology adoption.

b. Model Lifecycle (based on MLOps)

This should be guided by the BI1 Concept 1 Theory on Knowledge Discovery in Databases and the BI2 Concept 2 Labs we have gone through sequentially. Make use of a diagram.

The model starts by following the knowledge discovery process of databases where first the problem is defined and the goals of the project are set, for my case it is the batting prediction model scope and expectations. The next step is to select and collect data that is unbiased and in the most ethical way. For the batting prediction model I obtained the dataset from Kaggle which has readily available data that can be used for the prediction. The next step is to clean the data so what I did was first loaded the data then checked for any missing values and then transformed the data such that if there are any missing values they are then replaced with the mode or the mean values of the dataset. Once this was done the data was ready to undergo data mining thus had to select which task needs to carried out and for my case it is classification model and then to select the algorithms to use and after evaluation random forest turned out to be the best approach for the algorithm to be used after that it was just a matter of training the model and evaluating the result to check if it is accurate and with that the model was complete and ready for use by the cricketing body and the players through the consolidated dashboard.


1
Goal setting and Problem Definition
2
Data Selection
3
Data Cleansing and Pre-Processing
4
Data Transformation
5
Selection of the Data Mining Task
6
Selection of the Data Mining Algorithm
7
Utilization of the Data Mining Algorithm
8
Interpretation and Evaluation
9
Consolidation

c. Data Warehousing Approach

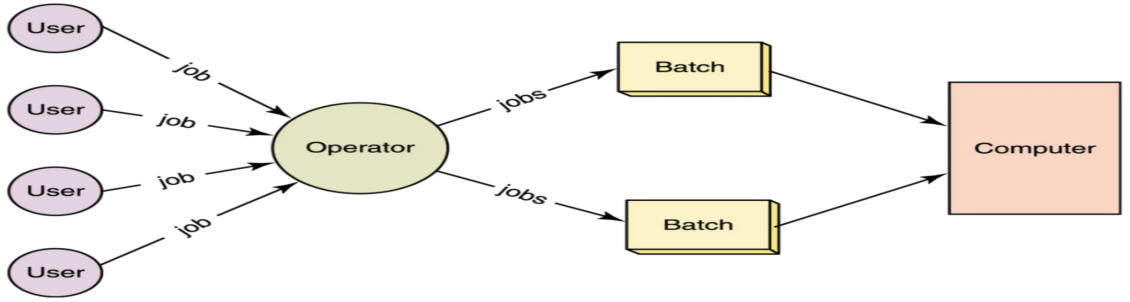
Discuss either the "Dimensional Modelling Approach by Kimball" or "Normalisation Modelling Approach by Inmon"

For the cricket player model Dimensional Modelling is better this is because this methodology's emphasis on organising data for analytical purposes suits cricket data analysis well. By utilising star schemas or snowflake schemas, this approach allows for a clear representation of batting performance metrics, player statistics, match details, and other pertinent factors. This facilitates efficient querying and reporting, enabling in-depth analysis of cricket batting trends, player comparisons, and performance insights. The dimensional model's structure provides a logical and comprehensive framework to explore the intricate relationships within cricket data, aiding in uncovering valuable insights for strategic decision-making and performance evaluation.

d. Data Pipelines

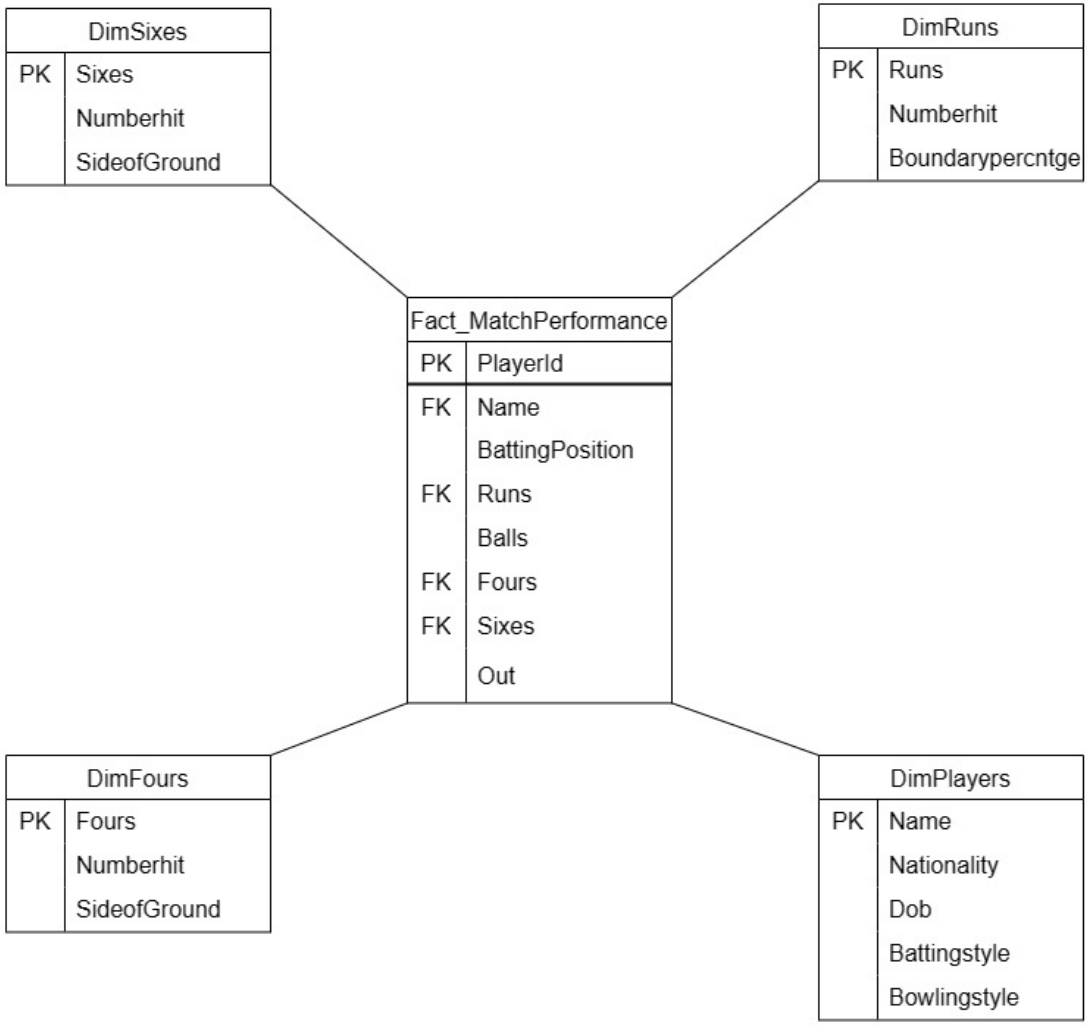
Discuss either "batch processing" or "stream ingestion", as well as the data sources and the ethical considerations to be made. Make use of a diagram and the content you are learning in the "BBT4202: Cloud Computing" course.

Batch processing appears to be more suited for your project. This method allows you to gather historical match data, player statistics, and other related information in discrete chunks or batches. It aligns well with cricket tournaments, enabling comprehensive trend analysis and long-term pattern identification in batting strategies. Moreover, it offers the ability to perform analysis on consolidated, historical datasets, which is beneficial for cricket data analysis involving player performances across different seasons.



e. The Data Warehouse Schema

Design a star schema or a snowflake schema for a Multidimensional OLAP architecture. This is informed by the measures and dimensions identified earlier in this project.



Section 18: Essay

5. **Business Intelligence is not solved using technology only. You need organisational buy-in. Write a 1,500-word essay to discuss a business-first, technology-second perspective in relation to the BI project.**

Business first perspective refers to the emphasis on prioritising the business aim and goals when developing business Intelligence solutions, while on the other hand technology first refers to when the business issues are resolved with the help of technology only. Technology is a vital part of solving modern business issues however, this is not always the case. Some issues need to be resolved with non-technical aspects where there is a need for human thought processes to work. This is because as much as there is the evolution of technology in machine learning and Artificial intelligence which is trying to act like the human brain and imitate it however, it is not the actual human brain. The human brain can never be exactly copied and that is why as much as there is so much reliance on Artificial intelligence and machine learning they are still not hundred percent accurate they still contain errors though there is no doubt that technologies have made life easier. A lot of cumbersome and repetitive work has been automated and that work has become more efficient but at what cost, at the cost of some employees losing their jobs which also has led to major unemployment rates all over the world especially in the developed countries where people who are incompatible with technology are said to be a step behind as they cannot cope up with technology. This is the same case with businesses. This is because those businesses that are still relying on traditional approaches are a step behind and losing their business to technology prone businesses. But is it the right approach?

For the project being carried out that is a player performance analytics model for cricket players where the cricketing body of the national team that is Cricket Kenya is the business facing side while the players playing for the national team that is for Kenya are the customers of the business and they both will be using the model for their benefit. Cricket Kenya will use it so that they can become a successful cricketing body all over the world and for the players who are the customers will be using this model so that they can make a successful career out of cricket and attain world recognition and wealth.

Business-first perspective is very important. For my case for Cricket Kenya their goal and aim is to become successful cricketing body and this will only happen when Kenya wins all the matches and major cricket leagues and for this to happen the players need to be performing well thus the main thing Cricket Kenya will be focussed on is the high performance of the players and this high performance is only possible with dedicated practise and overcoming the weaknesses that the players have and this does not have anything majorly to deal with technology thus for my case there is need to have business-first and technology second perspective since technology will only be used to classify the performance of the players but the performance will come from practise only.

Moreover, a user-centric design is crucial in this scenario, ensuring that the classification model generates insights that are intuitive and directly applicable to cricket management. By involving cricket experts in the model's development, the emphasis is on delivering decision-driven analytics. These analytics should directly aid in player evaluation, predicting performance based on various metrics, and recommending strategies for upcoming matches. This approach ensures that the classification model directly supports decision-making processes within the cricketing context, resonating with a business-first mindset.

Also the business-first perspective also extends to the model's scalability and adaptability. The technology chosen should seamlessly integrate with existing cricket databases, accommodating the vast and ever-growing volume of player data and match statistics. Scalability ensures that the model remains robust as the dataset expands, allowing it to handle the intricacies of player performance across different formats of the game. Additionally, the technology should be adaptable, capable of evolving along with the changes in cricketing strategies, rules, and player dynamics.

Moreover, the business-first perspective ensures that the cricket performance classification model is not developed in isolation but as an integral part of the cricket team's overall strategy. It prioritises the practical needs of the end-users and the unique intricacies of the sport, steering clear of a technology-centric approach that might overlook these critical aspects. This approach fosters a holistic integration of technology into the

cricketing ecosystem, creating a symbiotic relationship between the data-driven insights provided by the model and the strategic decisions made on the cricket field.

Furthermore, a business-first mindset emphasises user-centric design in the development of the classification model. The end-users, primarily cricket management professionals, should find the insights generated by the model intuitive and directly applicable to their decision-making processes. This requires a focus on delivering actionable analytics that assist in player evaluation, performance prediction, and strategic planning. The model should not only be a technological marvel but a practical tool that enhances the efficiency and efficacy of cricket management.

Also, from a technological standpoint, adopting a technology-second approach involves selecting tools that best align with the cricketing data's specific requirements. The focus is on integrating these technologies seamlessly with existing cricket performance databases, ensuring scalability to handle vast amounts of player data, match statistics, and historical records. Moreover, the emphasis lies in the technology's ability to automate data processing and analysis, allowing cricket experts to focus more on interpreting insights rather than the technical complexities of the model.

Moreover, the integration of these perspectives in the context of cricket performance classification ensures a harmonious balance. It emphasises the need to strike equilibrium between understanding the game's intricacies and leveraging technological advancements. Effective collaboration between cricketing experts and data scientists facilitates continuous feedback loops, allowing the model to evolve in tandem with the ever-evolving dynamics of the sport. Real-world case studies showcasing successful implementations of this integrated perspective reinforce its effectiveness in revolutionising cricket management strategies.

Aso, challenges might emerge in balancing technological advancements with the nuanced understanding of cricket. However, fostering a culture that values data-driven decision-making in cricket management serves as a catalyst for overcoming such obstacles. This culture encourages

continuous learning, enabling the organisation to adapt the classification model to the evolving needs of the sport. Ultimately, the integration of a business-first, technology-second perspective in the cricket performance classification model ensures not just short-term gains but also long-term value addition to the sport by enhancing team performance and strategic decision-making.

Furthermore, the technology-second perspective in cricket performance classification emphasises the significance of selecting appropriate technology that seamlessly integrates with the existing cricket management infrastructure. Rather than chasing the latest technological trends, it focuses on the compatibility, scalability, and sustainability of the chosen tools within the team's ecosystem. This ensures smoother adoption and integration of the classification model into the existing workflows, minimising disruptions and maximising the model's usability.

Additionally, adopting a technology-second perspective underscores the importance of continuous learning and upskilling within the cricket management team. It promotes the development of expertise and understanding of the technological tools used in the classification model, empowering the team to leverage these tools effectively for data-driven decision-making. This focus on capacity building ensures that the team maximises the potential of the technology to enhance cricketing strategies and player performance evaluation.

Also, ethical considerations play a significant role within the business-first perspective. The responsible use of technology in cricket performance analysis involves evaluating the ethical implications of data collection, analysis, and subsequent decision-making. This perspective ensures that technology is employed ethically, respecting player privacy and adhering to ethical standards within the cricketing community. It's crucial to harmonise technological advancements with ethical considerations, emphasising the importance of integrity and fairness in the sport.

Furthermore, considering the long-term sustainability of the classification model becomes imperative in a business-first approach. This encompasses aspects like ongoing maintenance, updates, and scalability of the technology. Ensuring that the model remains relevant and

effective in the years to come aligns with the business-first focus on sustained success for Cricket Kenya, underlining the need for technology that grows and evolves in tandem with the cricketing landscape.

Finally, A business-first perspective within Cricket Kenya's strategic objectives is pivotal as it intricately ties the model's development to the core goals of the cricketing body. Prioritising the high performance of players aligns directly with Cricket Kenya's ambition to elevate its status in international cricket. The focus primarily revolves around enhancing player performance through dedicated practice and targeted skill improvement. While technology plays a significant role in classifying and analysing player performance, the crux of achieving success lies in relentless practice and skill development, highlighting the necessity for a business-first approach over a technology-driven one.

In conclusion, the integration of a business-first, technology-second approach in developing the cricket performance classification model for Cricket Kenya embodies a strategic balance between leveraging technological advancements and prioritising the sport's objectives. Rooted in a user-centric design, this approach emphasises actionable insights tailored to cricket management needs while ensuring scalability, adaptability, and ethical use of technology. By fostering collaboration between cricket experts and data scientists and emphasising sustained success, ethical practice, and continuous evolution, this integrated perspective delivers a model that harmonises technological prowess with the sport's strategic vision, enhancing decision-making and performance for Cricket Kenya and its players.