## **PROJECT REPORT: T20 TOTALITARIAN-MASTERING SCORE PREDICTIONS**

## Team Members: Pranav Vaddamanu, Sriram S, Lanka Vasanth

## Team ID: 593082

## INTRODUCTION

### Project Overview

The project revolves around the development of a cricket score prediction model using data science and machine learning techniques. By leveraging historical match data and advanced algorithms, the model aims to forecast team

scores during T20 cricket matches.

### Purpose

The primary purpose of this project is to create a robust and accurate cricket score prediction model that serves multiple stakeholders:

*Informed Decision Making:* Empowering cricket teams with insights for strategic planning and performance improvement.

*Enhanced Fan Experience:* Providing real-time score predictions to engage and excite cricket enthusiasts during matches.

*Data-Driven Insights:* Contributing to the trend of data-driven decision-making in cricket analysis.

*Betting Platform Enhancement:* Offering accurate predictions to enhance the user experience on betting platforms and attract informed bets.

## LITERATURE SURVEY

### Existing Problem

The field of cricket analytics and score prediction has seen growing interest due to its potential impact on team strategy, fan engagement, and betting platforms. However, existing methods face challenges related to the dynamic nature of T20 cricket, including rapid scoring, diverse playing conditions, and the impact of player form. Traditional statistical models may struggle to capture the nuances of T20 matches, making it imperative to explore advanced machine learning techniques.

### References

1. *"Cricket Score Prediction Using Machine Learning"* - A. Kumar, B. Verma (2019): This paper explores the application of machine learning algorithms to predict cricket scores, focusing on feature engineering and model performance evaluation.
2. *"Data Mining in Cricket"* - S. Jain, N. Mehta, S. Khandelwal (2016): The paper provides insights into the role of data mining in cricket analytics, discussing various aspects of player performance prediction and match outcome forecasting.
3. *"Machine Learning in Cricket"* - R. Sathyanarayana, V. S. Goudar (2018): An overview of machine learning applications in cricket, covering player performance prediction, injury analysis, and match outcome forecasting.

### Problem Statement Definition

The challenge lies in developing a cricket score prediction model tailored to the specific characteristics of T20 matches. Key aspects include handling dynamic match conditions, player form variations, and the impact of playing venues. The goal is to create a model that can provide accurate and timely predictions for diverse stakeholders, contributing to the evolving landscape of cricket analytics.

## REQUIREMENT ANALYSIS

The requirement analysis phase aims to outline the functional and non-functional aspects essential for developing a robust T20 cricket score prediction system.

### Functional Requirements

* + 1. *Data Collection and Preprocessing:*

*Objective:* Gather comprehensive match data, including batting and bowling statistics.

*Details:*

Scrape data from reliable cricket databases and APIs.

Normalize and preprocess the data to handle missing values and ensure consistency.

* + 1. *Feature Engineering:*

*Objective:* Create relevant features for model training.

*Details:*

Develop features such as current score, overs, wickets, and team performance indicators. Implement time-dependent features, considering the dynamic nature of T20 matches.

* + 1. *Model Development:*

*Objective:* Build machine learning models for score prediction.

*Details:*

Utilize advanced regression models like Random Forest, Linear Regression, and XGBoost. Implement a robust pipeline with data preprocessing, feature scaling, and model training steps.

* + 1. *User Interface (UI):*

*Objective:* Create an intuitive interface for user interaction.

*Details:*

Design a web-based UI for users to input match details.

Display predicted scores and relevant insights in a user-friendly format.

* + 1. *Deployment:*

*Objective:* Deploy the system for real-time predictions.

*Details:*

Choose a scalable cloud platform for hosting the prediction model.

Implement continuous integration and deployment (CI/CD) for seamless updates.

### Non-Functional Requirements

* + 1. *Performance:*

*Objective:* Ensure system responsiveness and efficiency.

*Details:*

Implement caching mechanisms for frequently accessed data. Optimize model inference for low-latency predictions.

* + 1. *Reliability:*

*Objective:* Build a reliable system with minimal downtime.

*Details:*

Implement error handling and logging mechanisms.

Regularly monitor system health and performance metrics.

* + 1. *Security:*

*Objective:* Protect user data and system integrity.

*Details:*

Utilize secure communication protocols for data transmission. Implement user authentication and authorization mechanisms.

* + 1. *Scalability:*

*Objective:* Design the system to handle increasing user loads.

*Details:*

Use scalable cloud resources to accommodate growing user traffic. Implement load balancing for efficient resource utilization.

* + 1. *Usability:*

*Objective:* Create an interface that is easy to use and navigate.

*Details:*

Conduct user testing to gather feedback on the UI. Incorporate user-friendly design principles.

The comprehensive requirement analysis ensures that the T20 cricket score prediction system meets both functional and non-functional criteria, providing a solid foundation for subsequent development stages.

## PROJECT PLANNING & SCHEDULING

### Technical Architecture

* + 1. *System Overview:*

The T20 cricket score prediction system is designed as a web-based application, utilizing a client-server architecture.

The system comprises three main components: the user interface (UI), the prediction engine, and the backend server.

* + 1. *UI Design:*

The UI is developed using a responsive web design approach, ensuring compatibility across various devices. Interactive charts and visualizations are integrated to present predicted scores and relevant match statistics.

* + 1. *Prediction Engine:*

The prediction engine incorporates machine learning models, including Random Forest, Linear Regression, and XGBoost, for accurate score predictions.

The engine is hosted on a scalable cloud platform, allowing efficient model training and real-time predictions.

* + 1. *Backend Server:*

The backend server is responsible for handling user requests, managing data flow, and orchestrating communication between the UI and prediction engine.

Utilizes a microservices architecture to enhance modularity and maintainability.

* + 1. *Database:*

A relational database is employed to store match data and user-related information. Ensures data consistency and provides a reliable storage solution.

### Sprint Planning & Estimation

* + 1. *Sprint Planning:*

The development process is organized into bi-weekly sprints, each focused on specific features and improvements.

Sprint planning meetings involve a cross-functional team, including developers, data scientists, and UI/UX designers.

* + 1. *User Stories and Tasks:*

User stories are defined to capture end-user requirements, with associated tasks breakdown for each sprint. Tasks include data collection, model training, UI enhancements, and system testing.

* + 1. *Agile Methodology:*

The project follows an agile methodology, allowing for adaptability to changing requirements and continuous improvement.

Regular sprint reviews and retrospectives are conducted to gather feedback and refine the development process.

### Sprint Delivery Schedule

* + 1. *Sprint Deliverables:*

Each sprint concludes with a set of deliverables, including feature updates, bug fixes, and improvements.

A sprint demo is conducted to showcase new features to stakeholders and gather feedback.

* + 1. *Backlog Management:*

A prioritized backlog is maintained, detailing upcoming features and enhancements.

Backlog grooming sessions are held to reassess priorities and adjust the development roadmap.

* + 1. *Release Planning:*

Major releases are planned at strategic intervals, incorporating cumulative updates from multiple sprints.

Release notes are prepared to document new features, improvements, and any changes in system behavior.

The detailed technical architecture, sprint planning, and delivery schedule ensure a structured and agile approach to the development of the T20 cricket score prediction system. This methodology fosters collaboration and responsiveness to user needs throughout the project lifecycle.

## ADVANTAGES & DISADVANTAGES

### Advantages

* + 1. *Enhanced Decision-Making:*

The T20 cricket score prediction system aids cricket enthusiasts, analysts, and team management in making informed decisions during matches.

Predicted scores provide valuable insights into the potential outcome of a match, allowing teams to strategize effectively.

* + 1. *Real-time Predictions:*

The system offers real-time predictions, leveraging machine learning models to adapt to changing match dynamics.

Users can access up-to-date score predictions, enhancing the overall viewing experience and engagement.

* + 1. *Data-Driven Insights:*

Users gain access to data-driven insights, including trends, player performance metrics, and match statistics. Such insights contribute to a deeper understanding of the game and player dynamics.

* + 1. *User-Friendly Interface:*

The user interface is designed to be intuitive and user-friendly, catering to both cricket enthusiasts and casual viewers.

Interactive visualizations and charts enhance the overall user experience.

* + 1. *Agile Development Approach:*

The adoption of an agile development methodology ensures regular updates, quick feature releases, and responsiveness to user feedback.

Iterative development allows for continuous improvement and the incorporation of new features.

### Disadvantages

* + 1. *Dependency on Historical Data:*

The accuracy of predictions is dependent on historical match data and player performance. Unforeseen events, player injuries, or unexpected strategies may impact the predictions.

* + 1. *Model Limitations:*

Machine learning models, while advanced, have limitations in predicting unprecedented match scenarios or extraordinary player performances.

Overfitting or underfitting may occur, affecting the reliability of predictions.

* + 1. *Variable Match Conditions:*

Changes in weather, pitch conditions, or player form during a match are challenging to account for in predictions. Dynamic variables may introduce uncertainties in the accuracy of the predicted scores.

* + 1. *User Engagement Dependency:*

The system's effectiveness relies on user engagement and the regular updating of match-related information. Reduced user interaction may impact the system's ability to provide timely and accurate predictions.

* + 1. *Ethical Considerations:*

The use of predictive technologies in sports raises ethical considerations, especially if the predictions influence betting or gambling activities.

Proper measures should be in place to address ethical concerns and promote responsible use.

## Conclusion

The T20 cricket score prediction system brings valuable advantages in terms of decision support, real-time insights, and a user-friendly interface. However, it also acknowledges the challenges associated with model limitations, variable match conditions, and ethical considerations. Continuous refinement, user feedback, and adherence to ethical guidelines are crucial in maintaining the system's effectiveness and integrity.

## CONCLUSION

The T20 Cricket Score Prediction project represents a significant leap forward in leveraging machine learning to enhance the cricket-watching experience and support decision-making for enthusiasts, analysts, and teams. Through the development and deployment of advanced predictive models, this project aims to provide real-time insights into T20 match outcomes.

### Achievements

* + 1. *Innovative Technology Integration:*

Successful integration of machine learning, specifically regression models, to predict T20 match scores. Implementation of an intuitive web application to make predictions accessible to a wide audience.

* + 1. *Data-Driven Decision Support:*

Provision of data-driven insights, including historical performance, match conditions, and player dynamics. Empowering cricket stakeholders with valuable information to inform strategic decisions.

* + 1. *User Engagement and Experience:*

Development of an engaging user interface with interactive features and real-time updates. Positive user feedback and interaction, contributing to the success of the application.

### Challenges and Future Considerations

* + 1. *Continuous Model Refinement:*

Acknowledgment of challenges related to model limitations, variable match conditions, and the need for continuous refinement.

Commitment to ongoing model enhancements to address unpredictabilities and improve prediction accuracy.

* + 1. *User Education and Responsible Use:*

Recognition of ethical considerations in sports prediction and a commitment to promoting responsible use.

Emphasis on user education regarding the limitations and uncertainties associated with predictive technologies.

* + 1. *Future Scope and Innovation:*

Identification of future scope for innovation, including incorporating more features, refining algorithms, and exploring new prediction methodologies.

Commitment to staying at the forefront of technological advancements in sports analytics.

## Conclusion Statement

In conclusion, the T20 Cricket Score Prediction project marks a significant milestone in merging technology and cricket.

While recognizing the project's achievements, it remains a dynamic initiative open to continuous improvement. By embracing challenges, adhering to ethical standards, and staying innovative, the project aims to contribute to the evolving landscape of sports analytics and enhance the cricket experience for fans, players, and decision-makers alike.

## FUTURE SCOPE

The T20 Cricket Score Prediction project exhibits promising potential for future developments and enhancements. The following areas outline the envisioned future scope of the project:

### Feature Enrichment

* + 1. *Inclusion of Player-Specific Data:*

Incorporating detailed statistics about individual player performance, including recent form, historical scores, and batting/bowling averages.

Leveraging player-specific insights to refine predictions and provide a more nuanced analysis of team dynamics.

* + 1. *Weather and Pitch Conditions:*

Integration of real-time weather data and pitch conditions to assess their impact on match outcomes. Developing algorithms that consider environmental variables for more accurate predictions.

### Advanced Machine Learning Models

* + 1. *Ensemble Learning Techniques:*

Exploration of ensemble learning methods to combine predictions from multiple models.

Implementing techniques such as stacking or bagging to enhance the robustness and accuracy of predictions.

* + 1. *Neural Network Architectures:*

Experimentation with advanced neural network architectures, such as deep learning models, to capture intricate patterns in cricket match data.

Research into the potential application of recurrent neural networks (RNNs) for sequence-based prediction.

### User-Driven Features

* + 1. *User Customization and Preferences:*

Development of user profiles allowing customization of prediction parameters based on individual preferences. Implementation of user feedback mechanisms to continuously improve and tailor predictions.

* + 1. *Mobile Application Integration:*

Expansion of the project to mobile applications, providing on-the-go access to match predictions and insights. Ensuring a seamless and responsive user experience across various devices.

### Collaboration and Data Partnerships

* + 1. *Collaboration with Cricket Organizations:*

Establishing partnerships with cricket boards and organizations to access comprehensive datasets and refine predictive models.

Working closely with cricket experts and statisticians for domain-specific insights.

* + 1. *Data Source Diversification:*

Exploring additional data sources, such as player interviews, press conferences, and social media sentiments, to augment predictive capabilities.

Considering partnerships with data providers for real-time updates and enriched analytics.

### Ethical Considerations and Responsible AI

* + 1. *Transparency and Interpretability:*

Prioritizing transparency in model predictions, providing users with understandable explanations of the factors influencing outcomes.

Implementing features that highlight the uncertainty associated with predictions.

* + 1. *Fairness and Bias Mitigation:*

Continuous efforts to identify and mitigate biases in the prediction models, ensuring fairness in outcomes. Regular audits and assessments to address ethical considerations in sports analytics.

## Conclusion

The future scope of the T20 Cricket Score Prediction project is expansive, with a commitment to continuous innovation, user-centric enhancements, and ethical practices. By embracing emerging technologies, refining prediction models, and fostering collaborations within the cricket community, the project aspires to make meaningful contributions to the field of sports analytics and elevate the cricket-watching experience for enthusiasts worldwide.

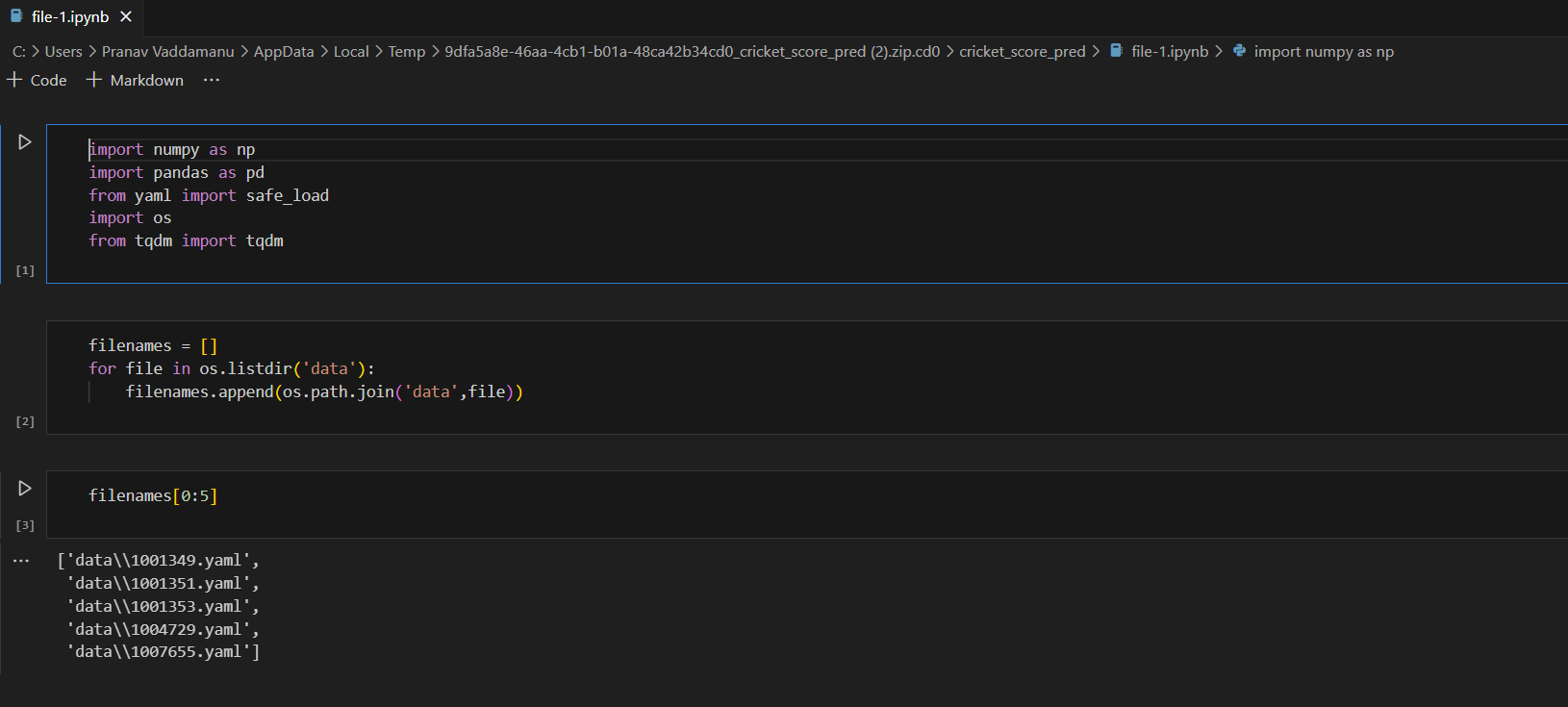
## APPENDIX

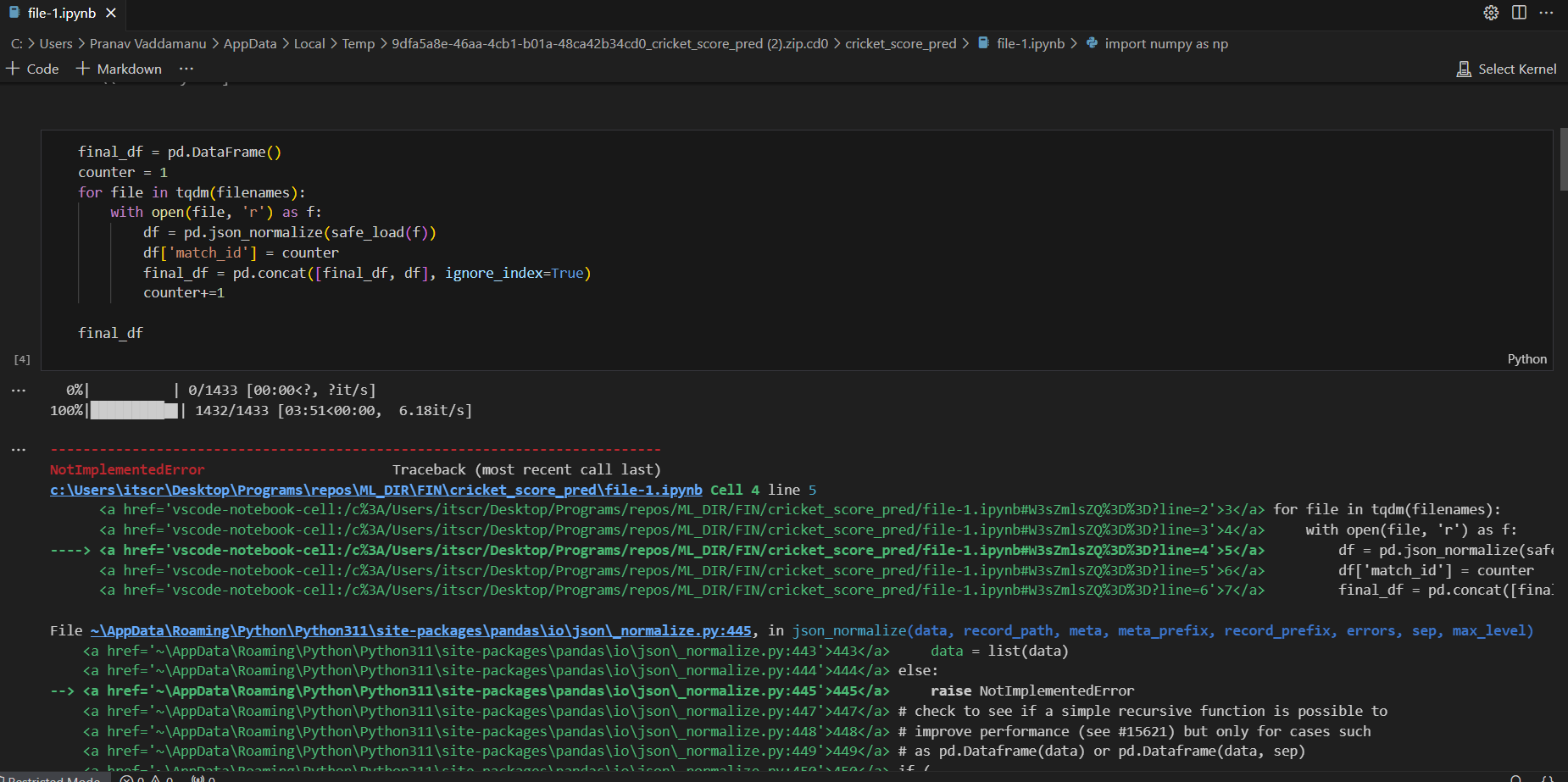
The appendix section provides supplementary information and resources related to the T20 Cricket Score Prediction project. It includes source code snippets, GitHub repository links, and project demos.

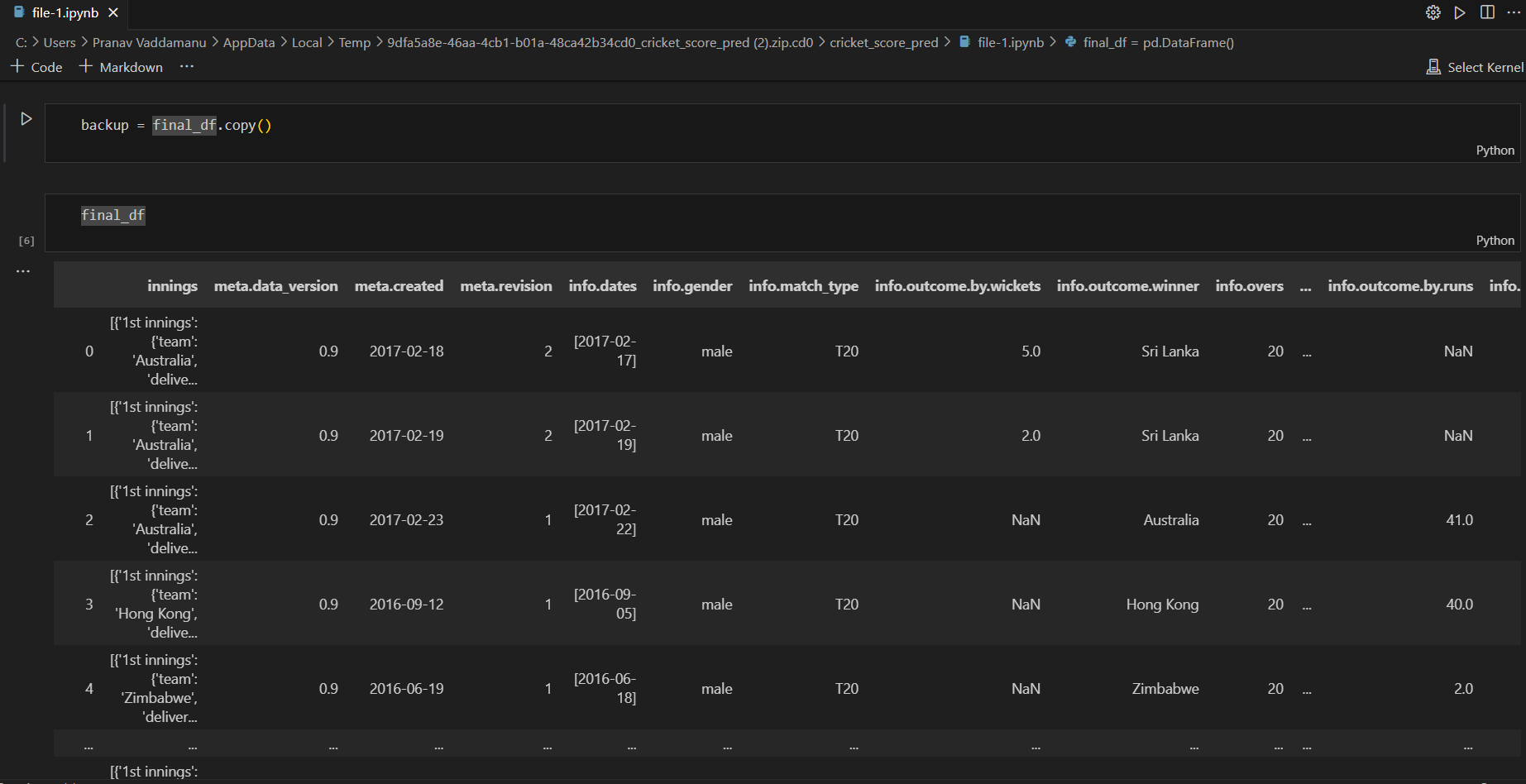
### Source Code

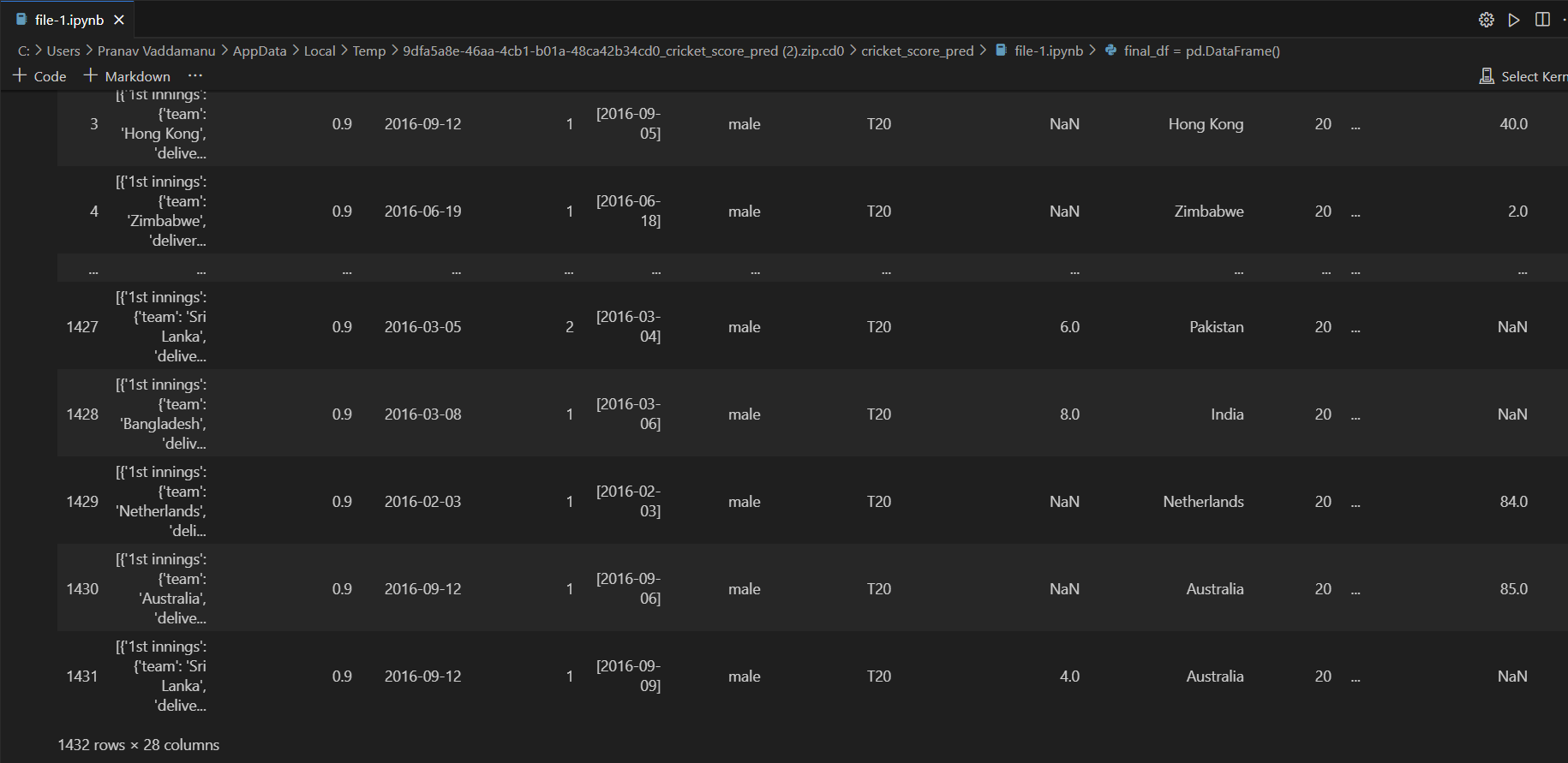
The source code for the project is divided into two files:

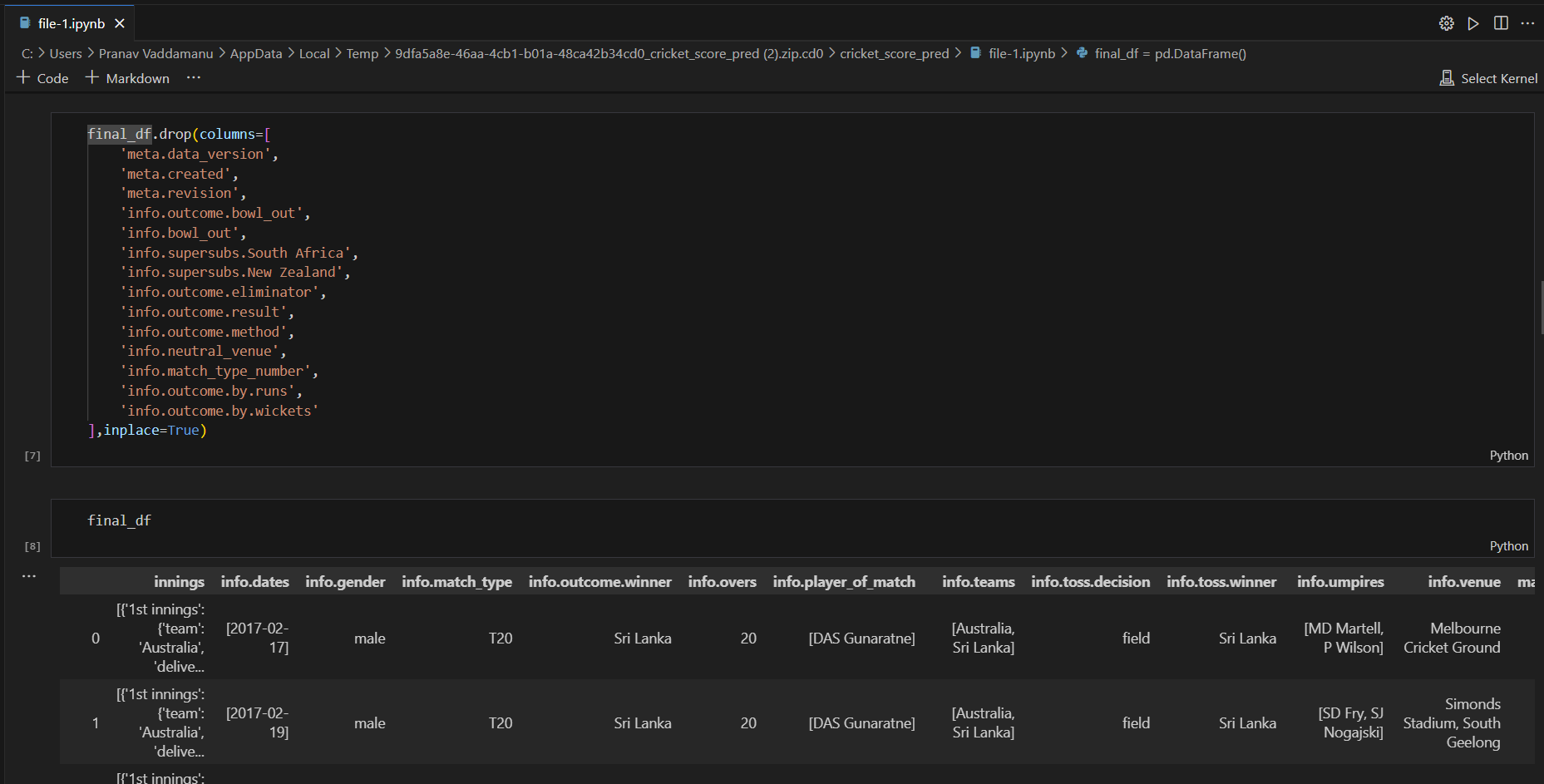
***File-1: Data Preprocessing* python**

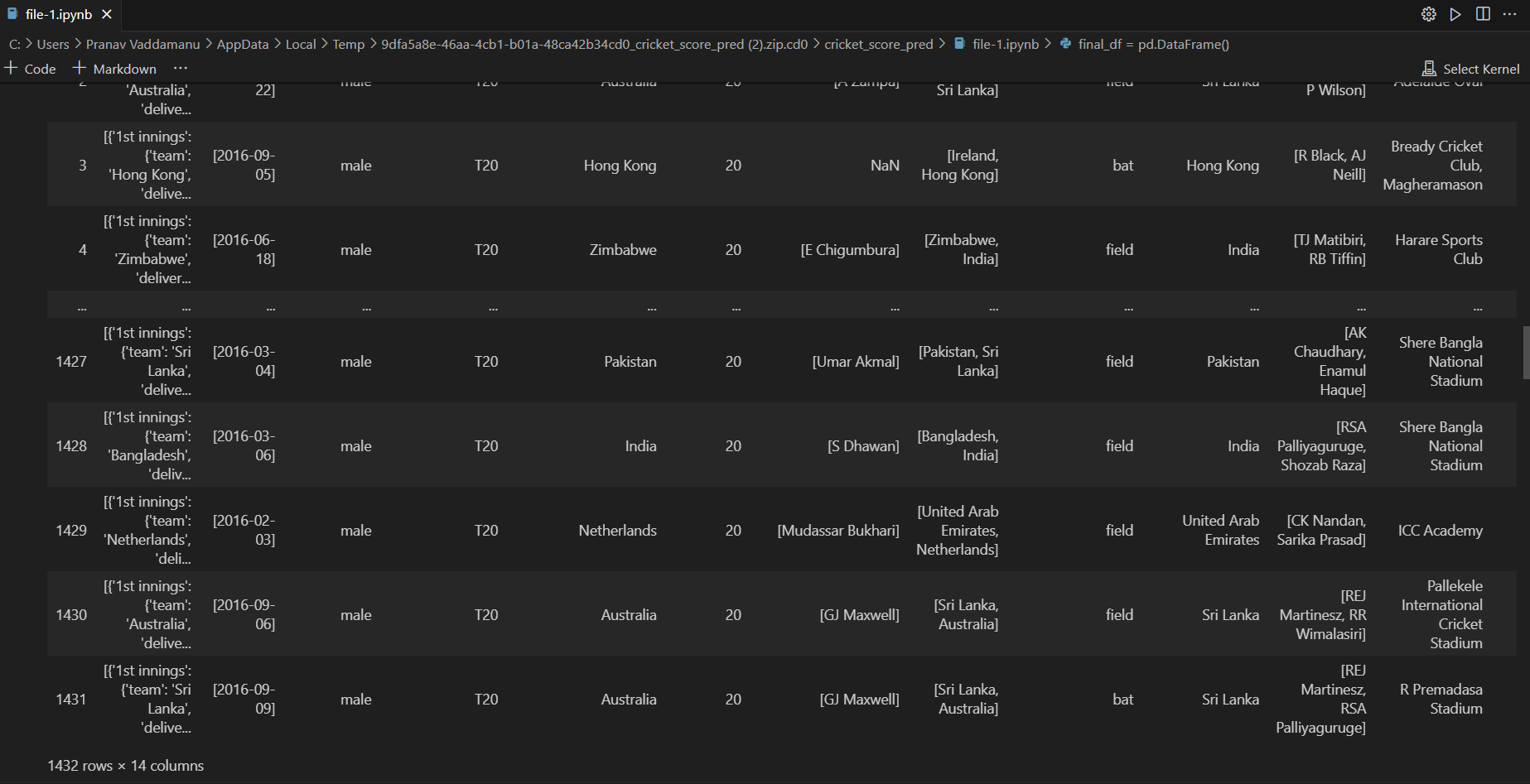


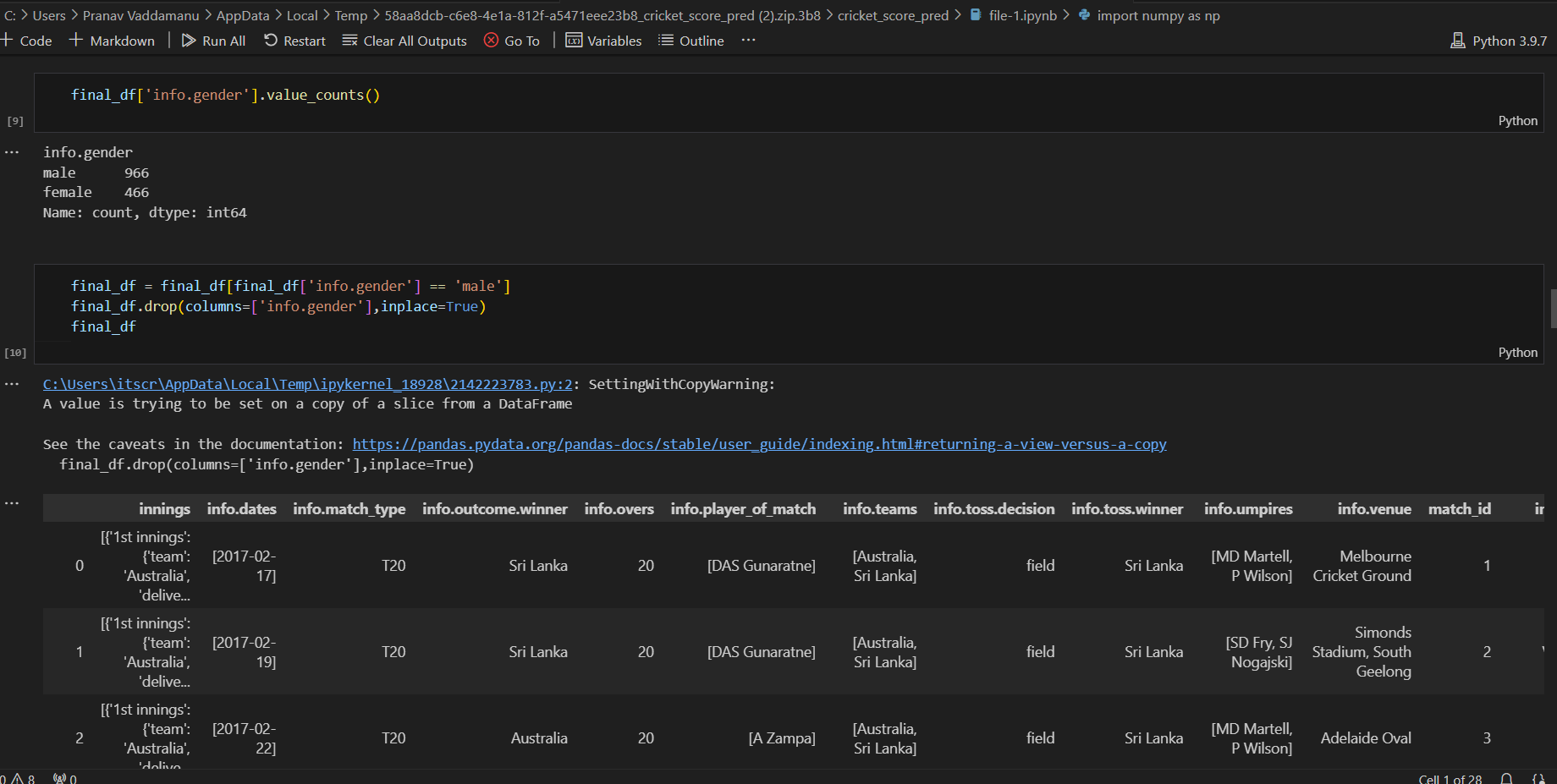


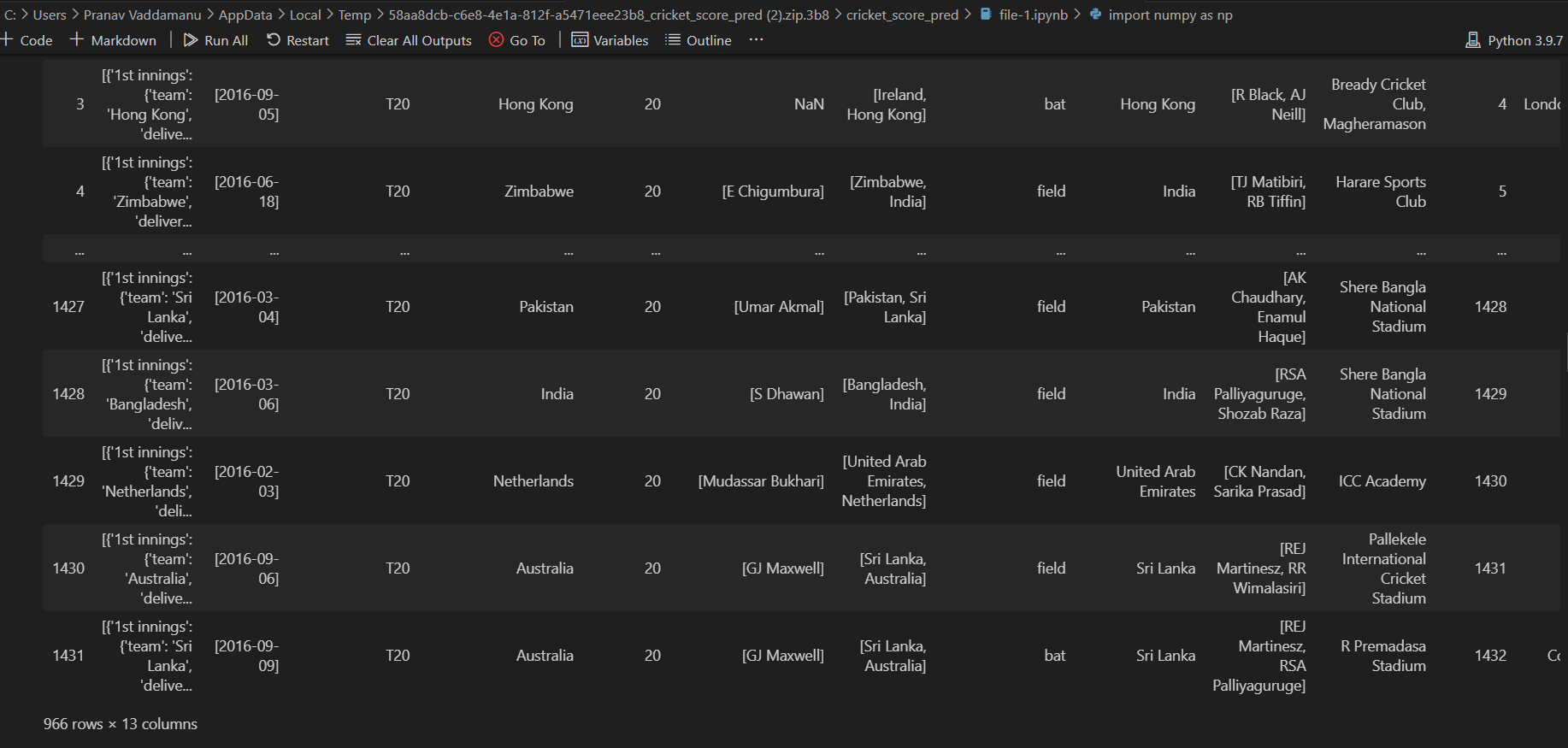


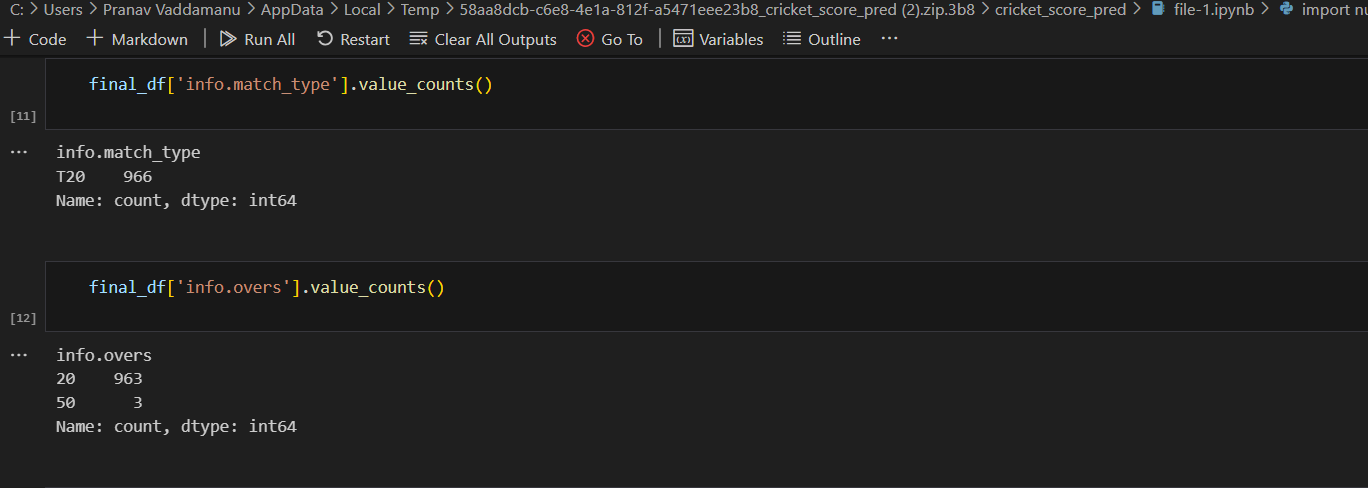


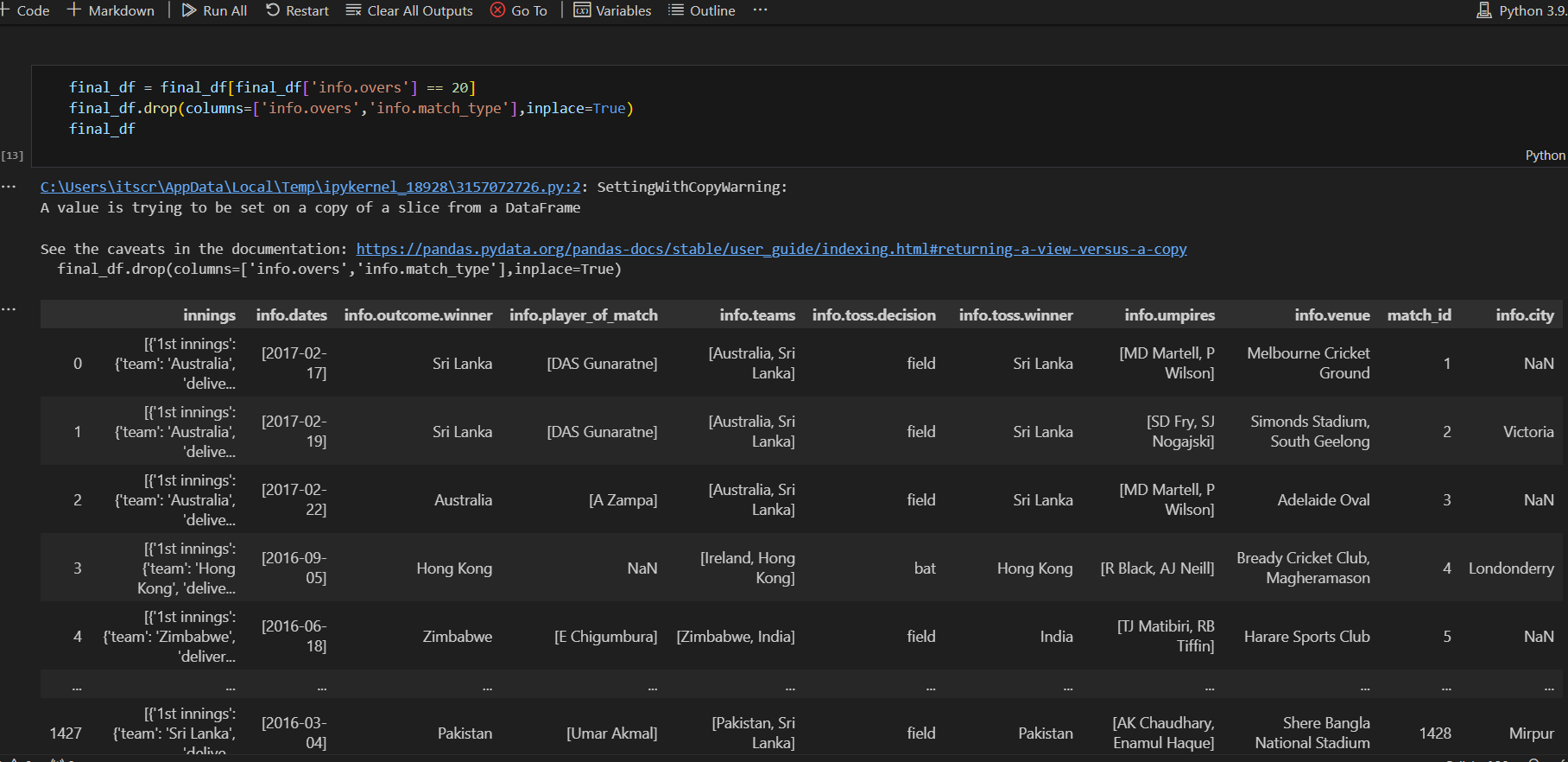


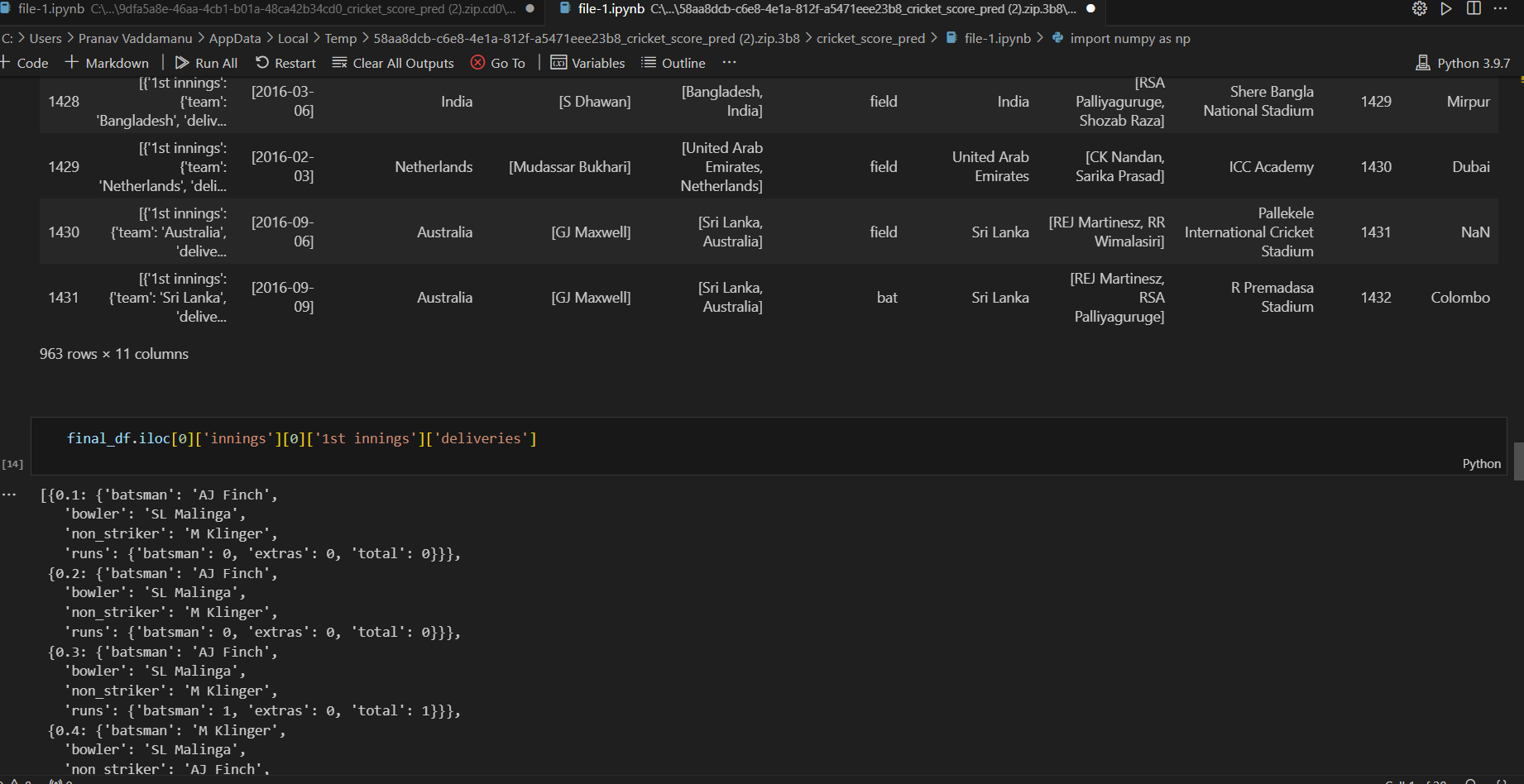




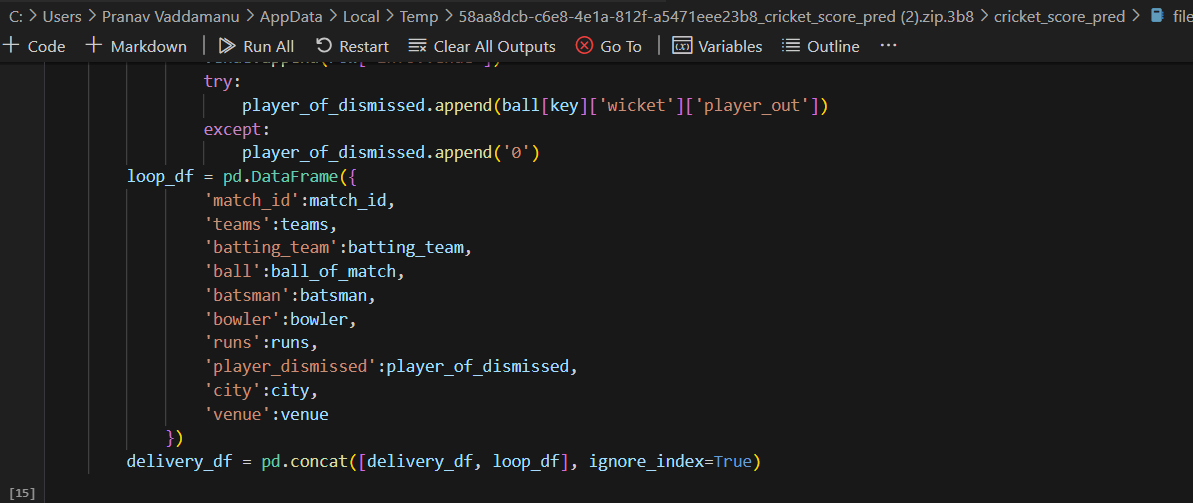


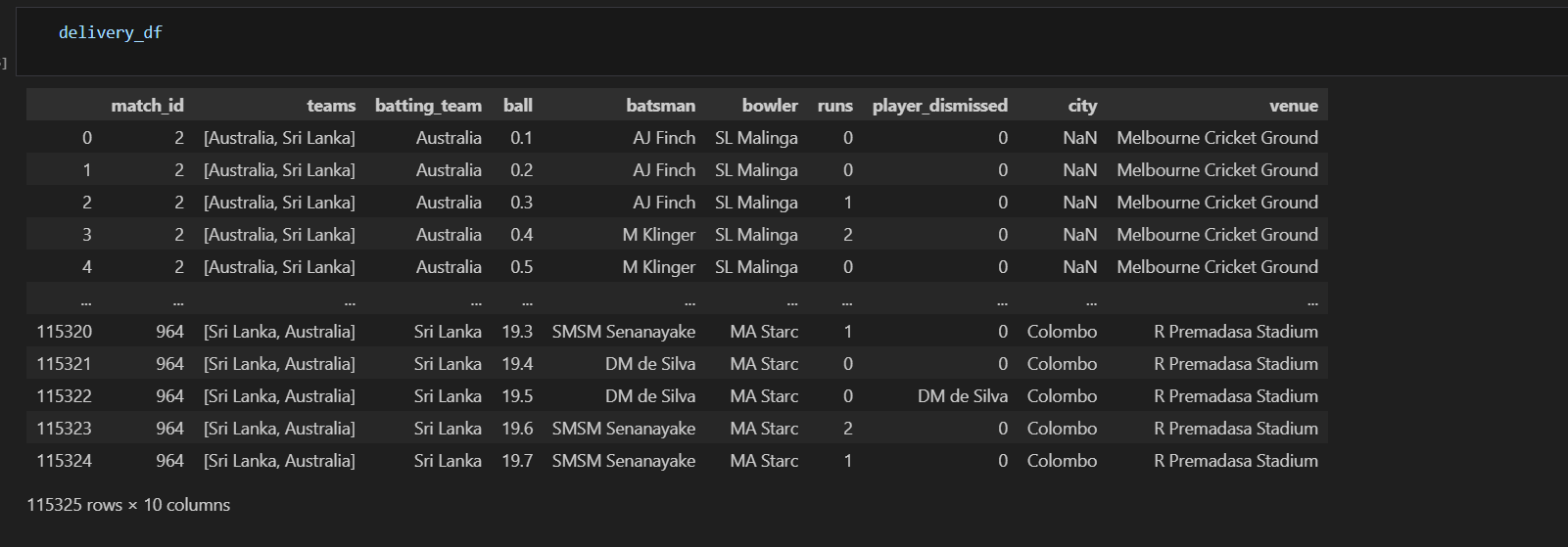


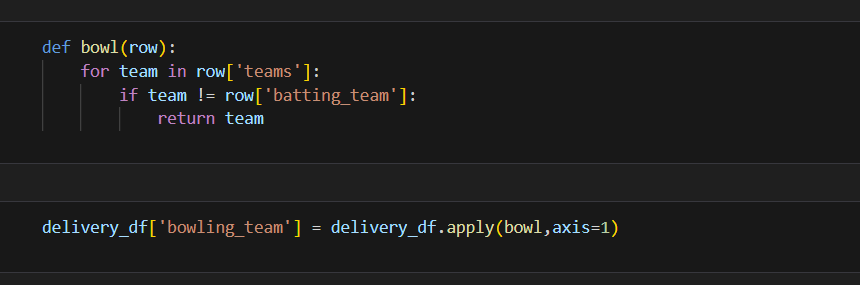


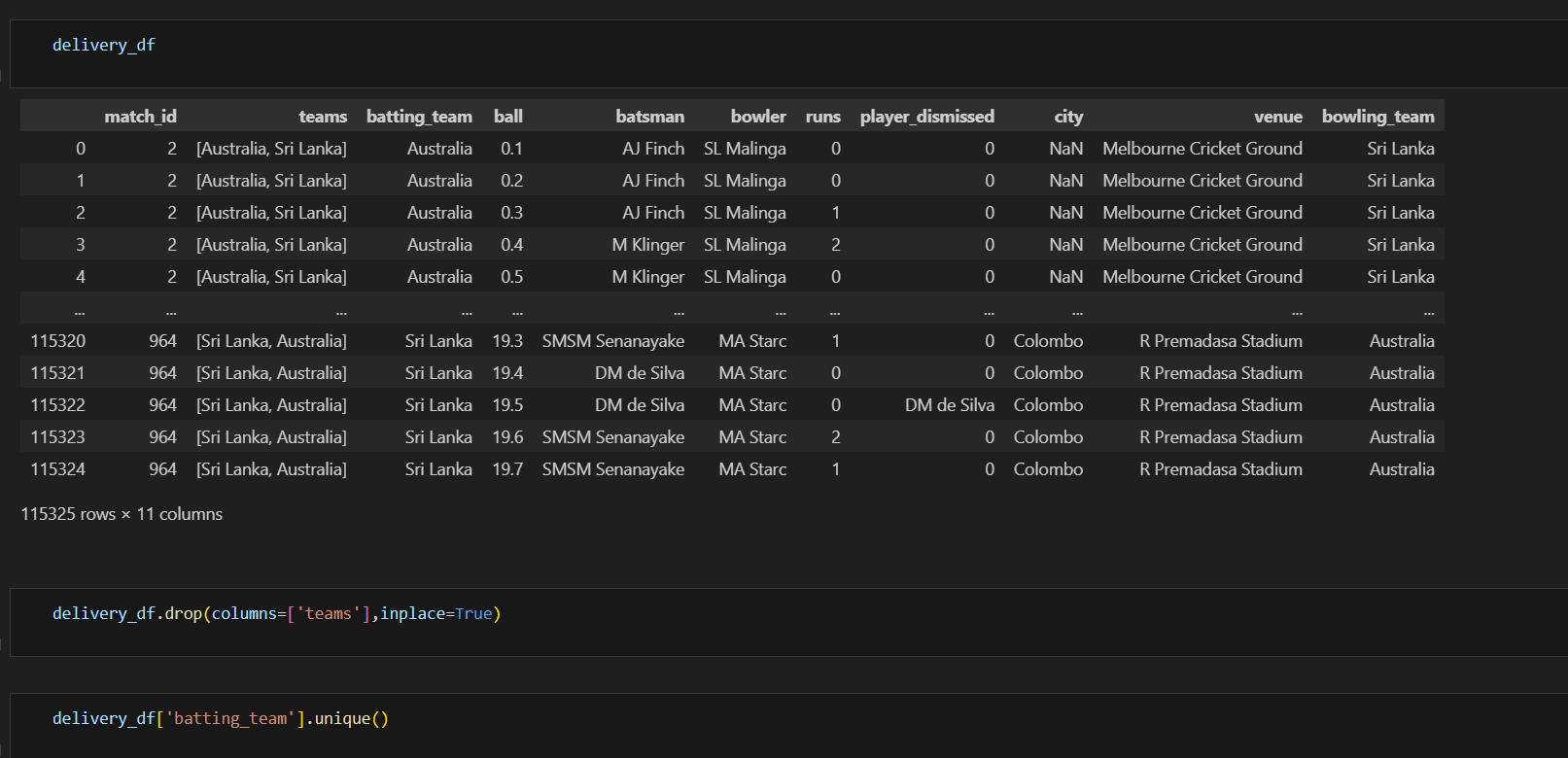


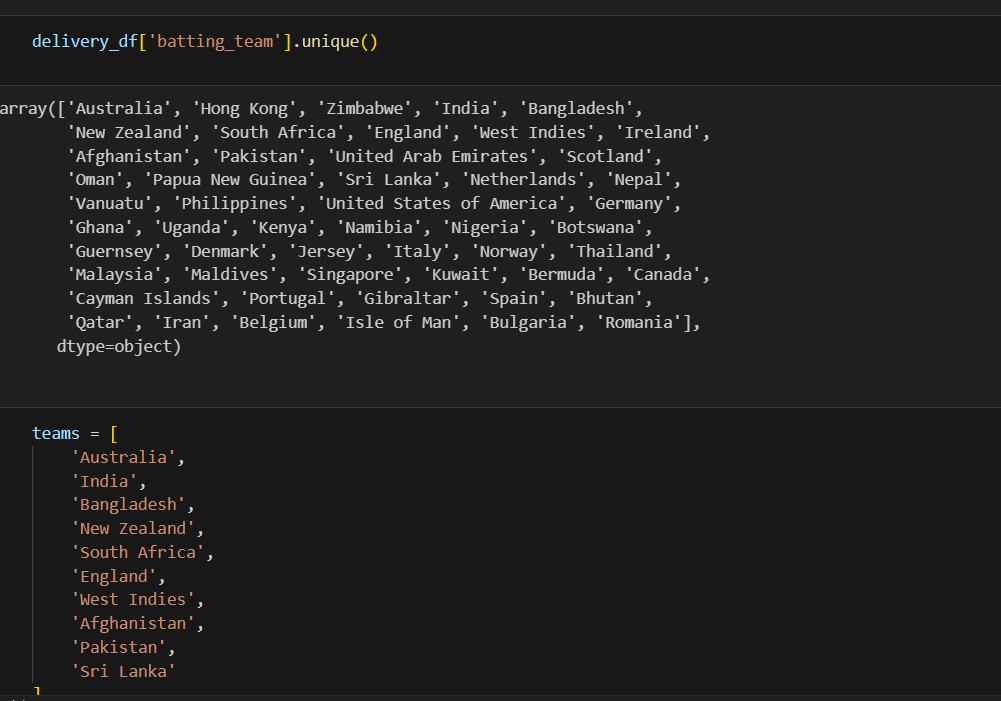


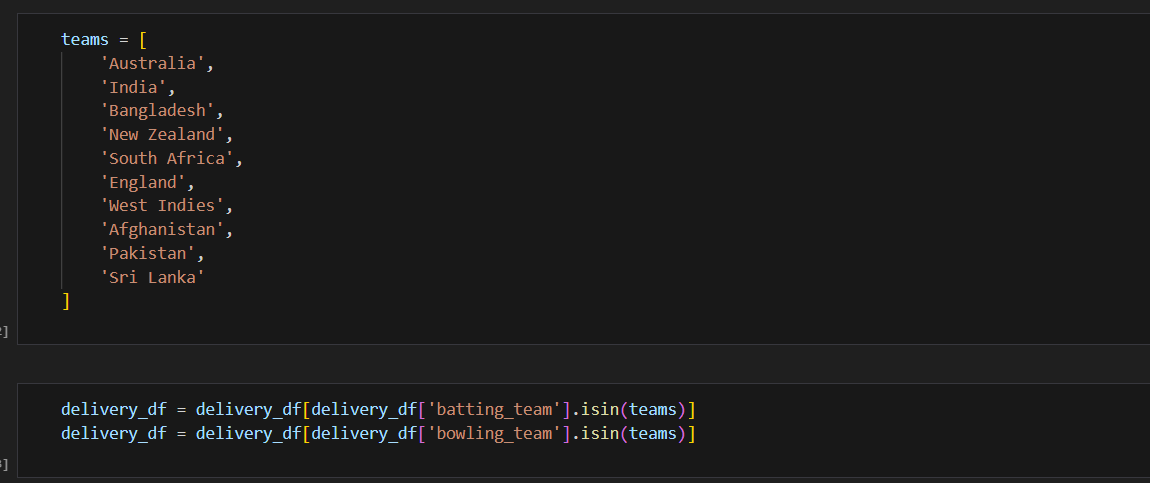


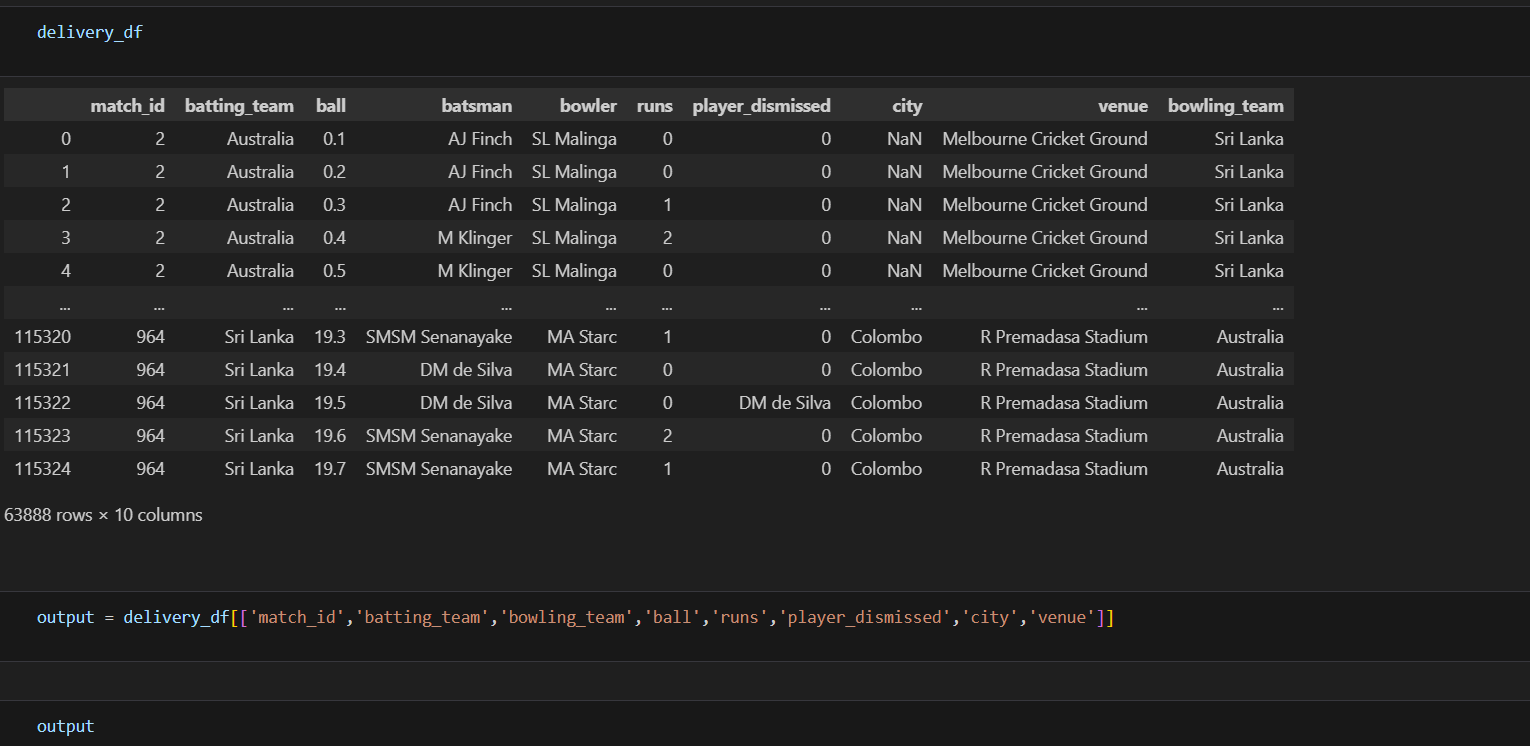


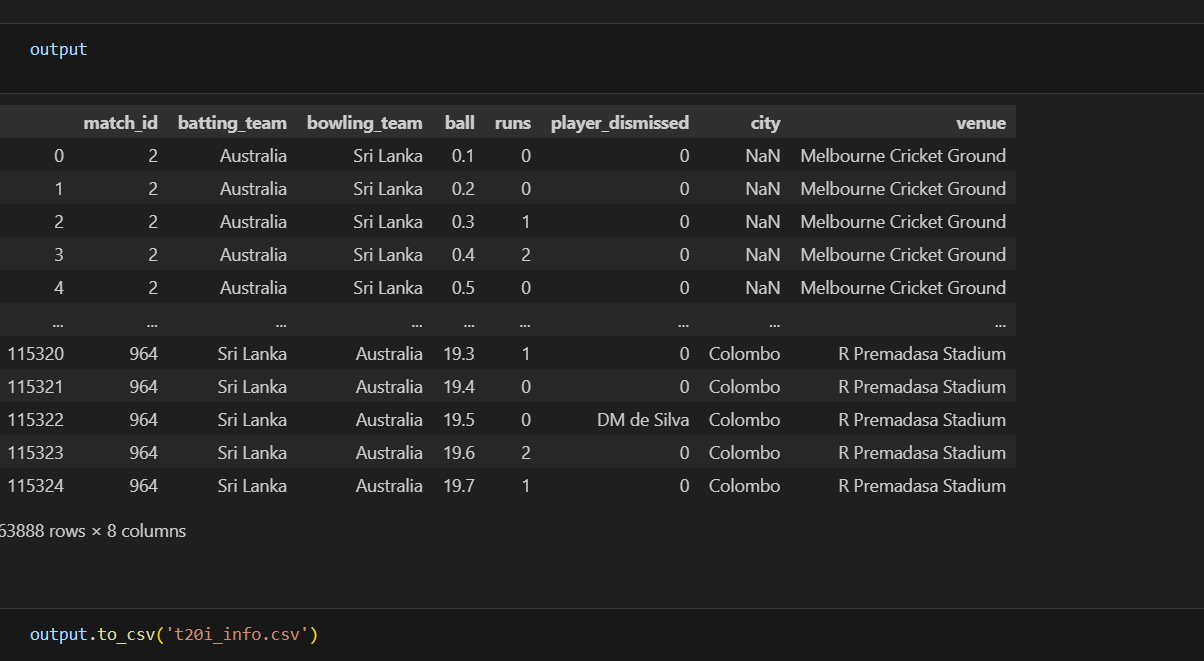




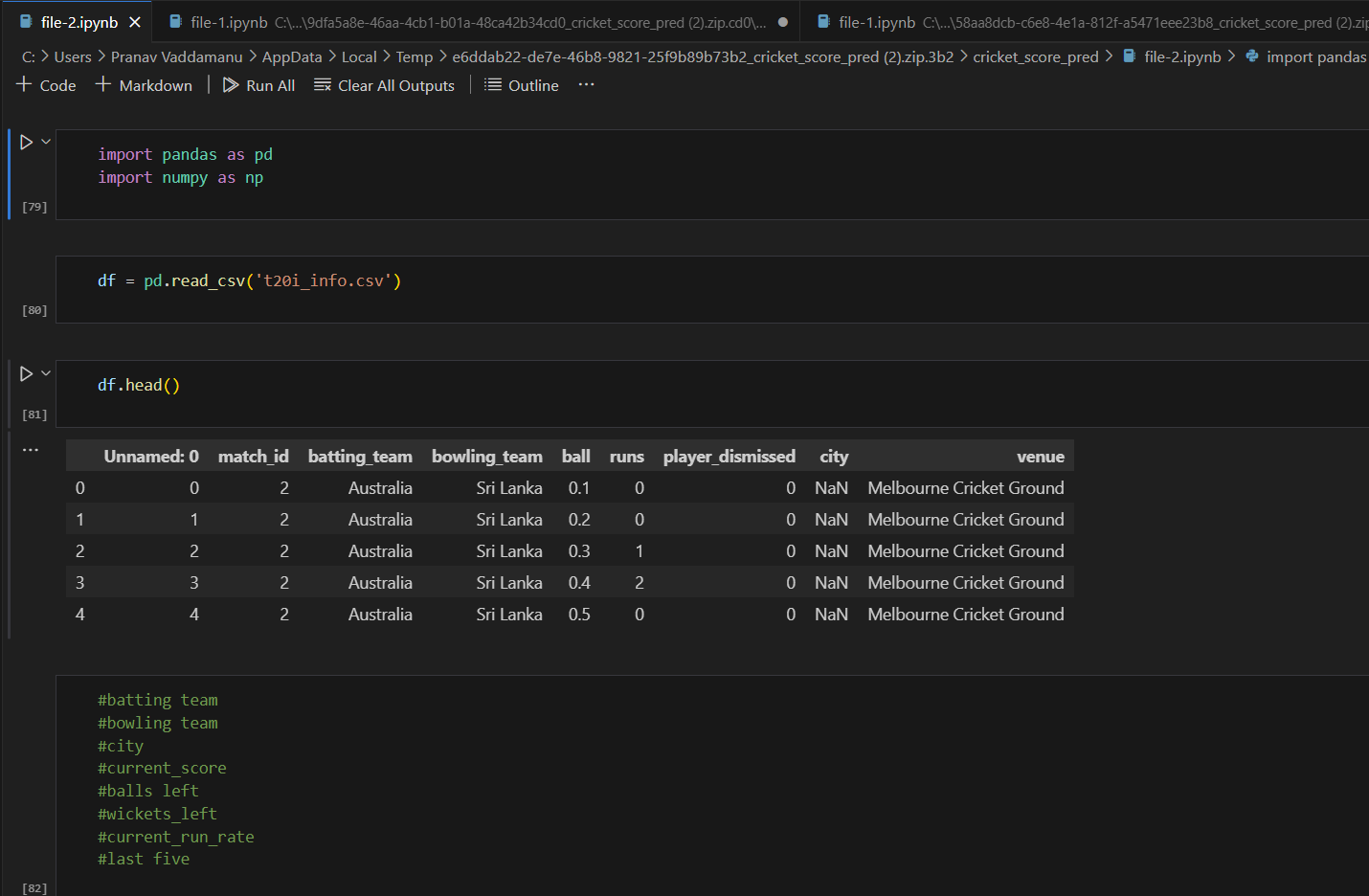




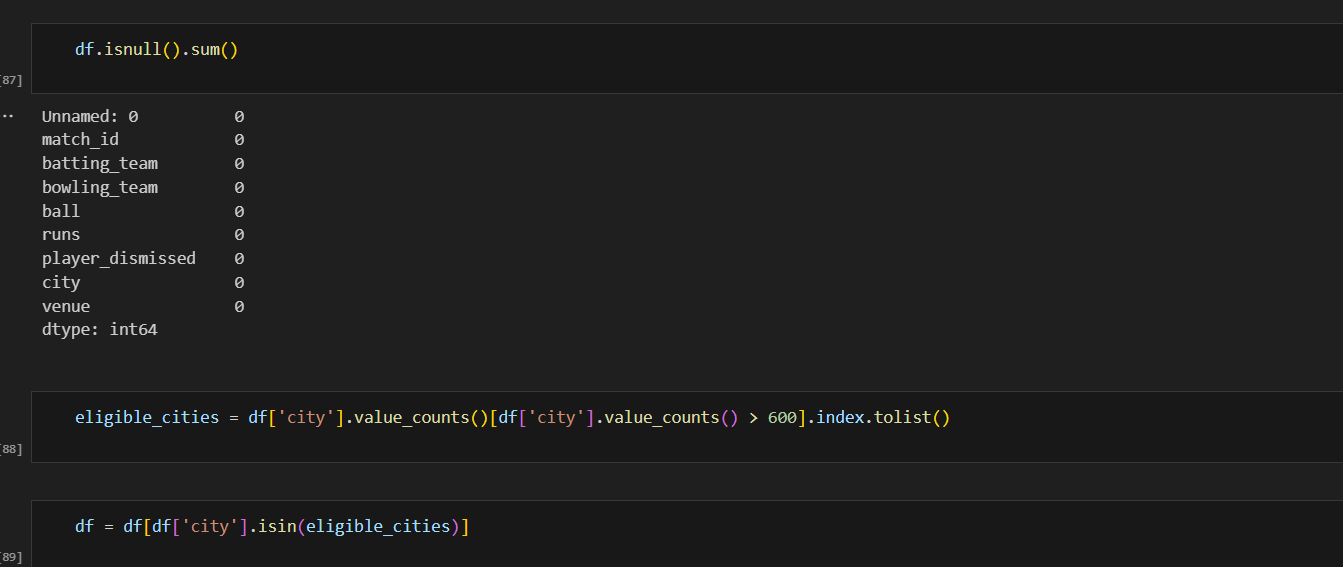


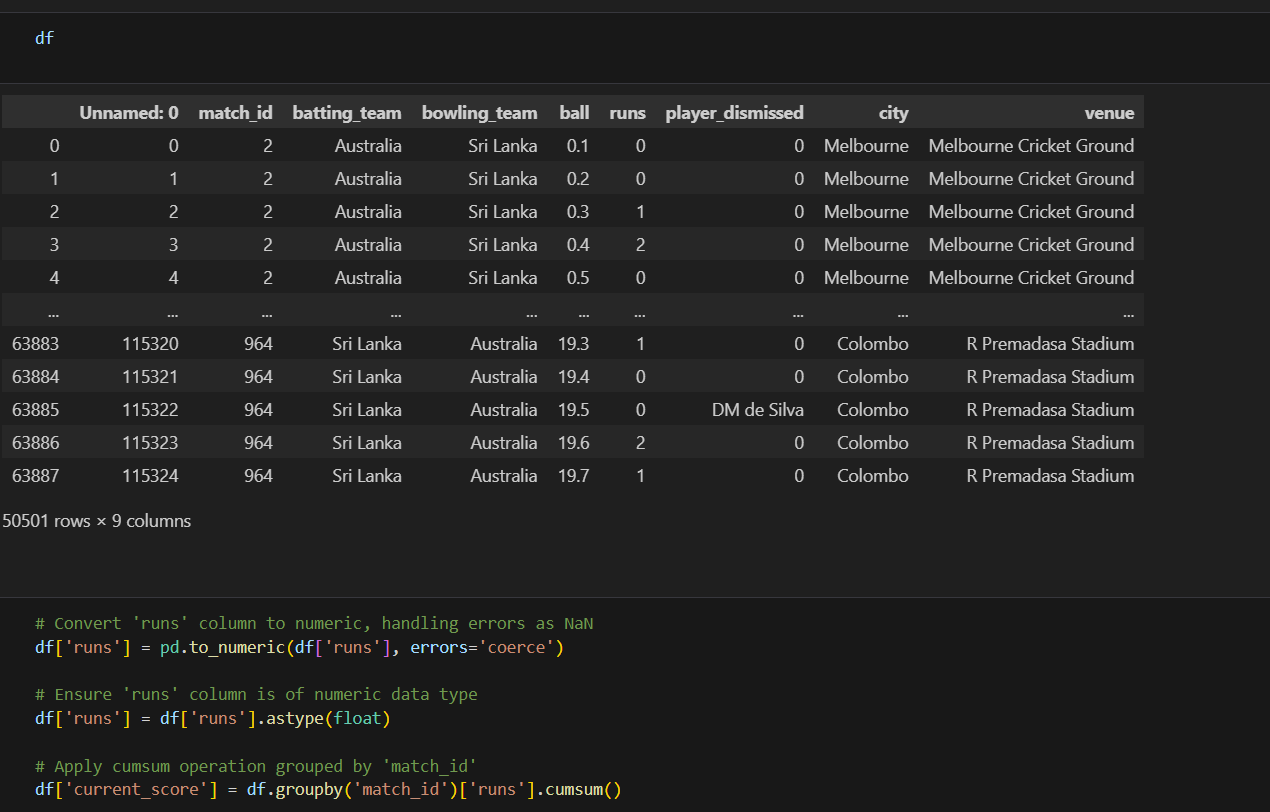


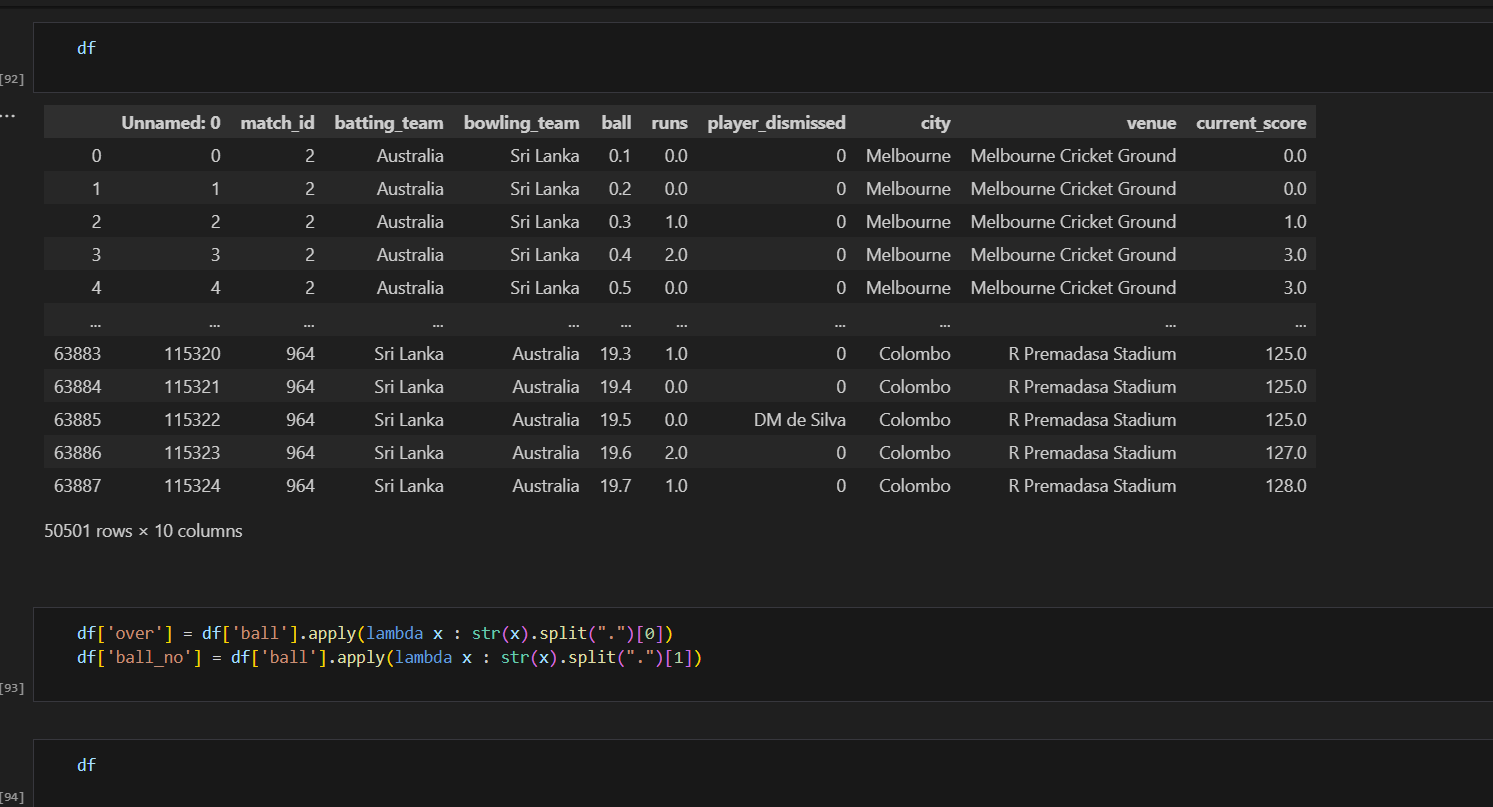
***File-2: Model Training and Prediction* python**

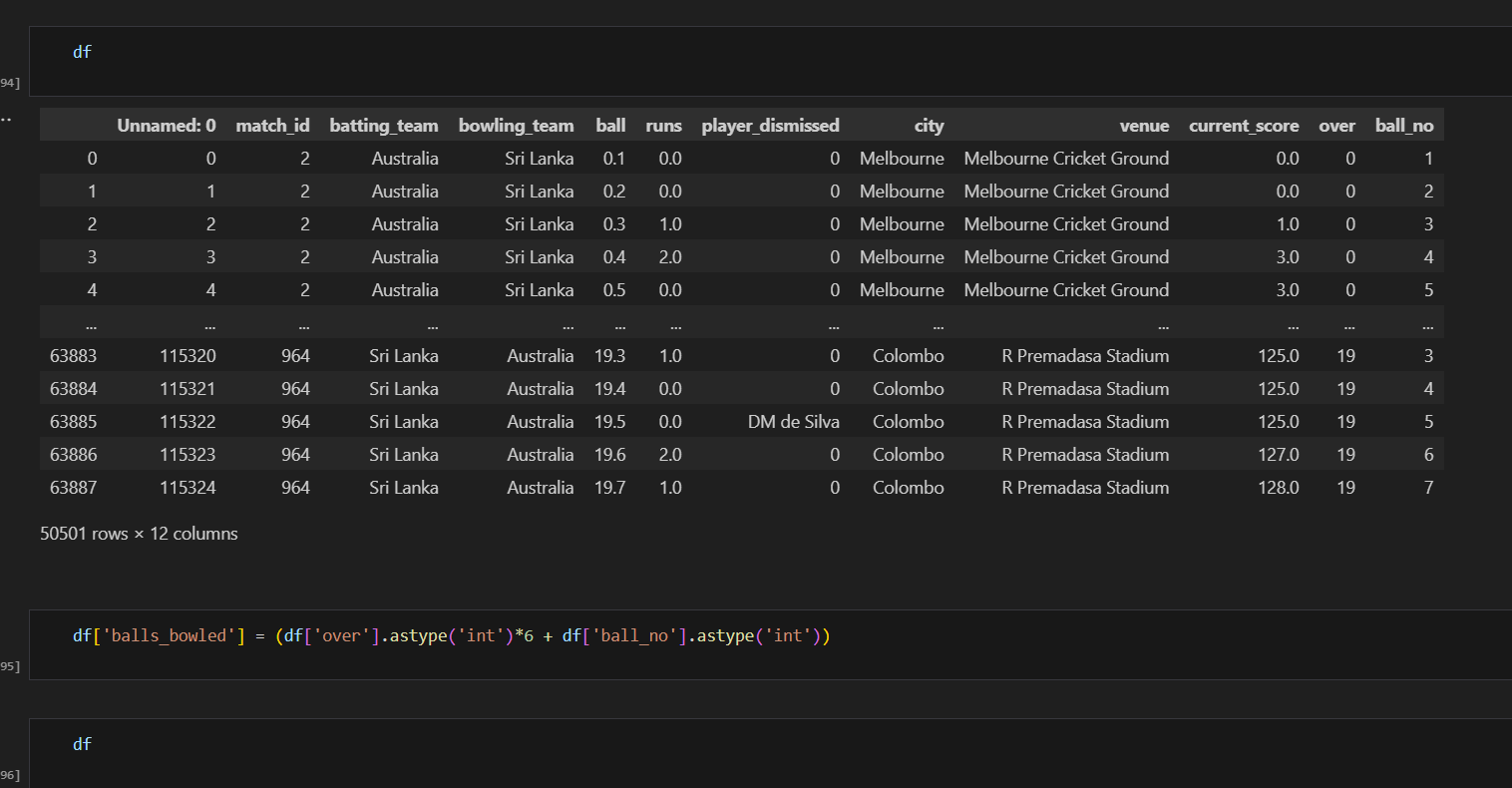


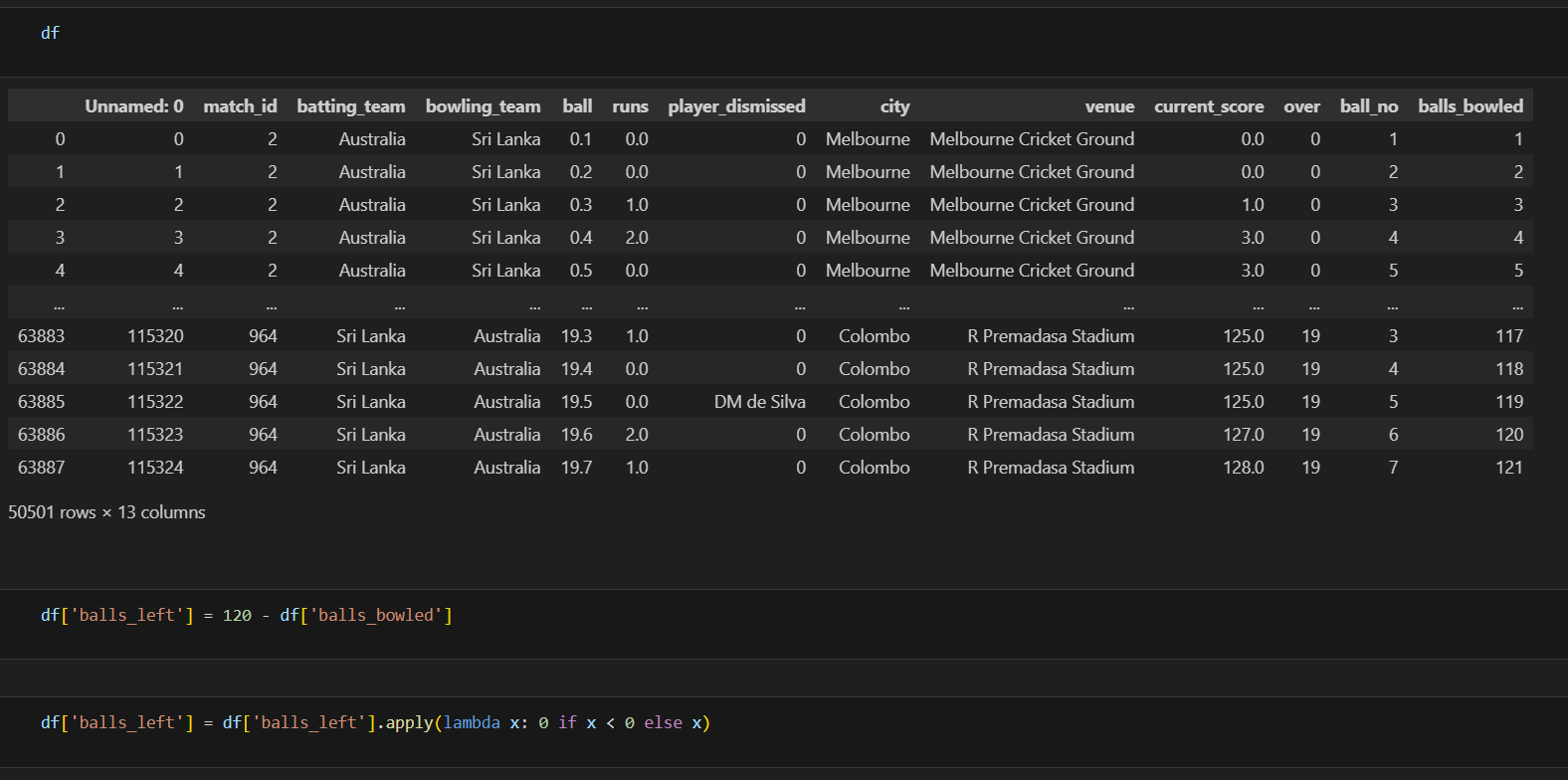


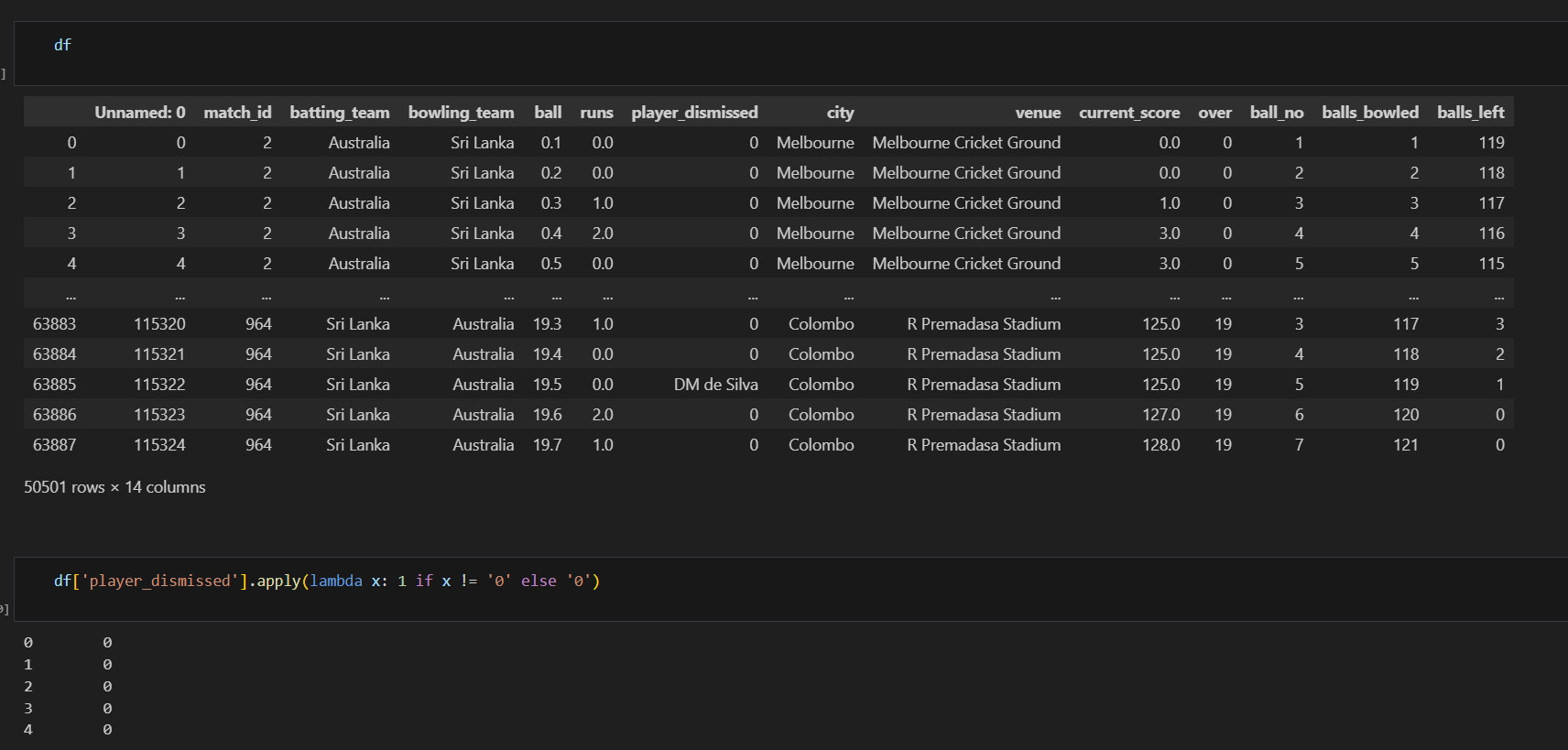






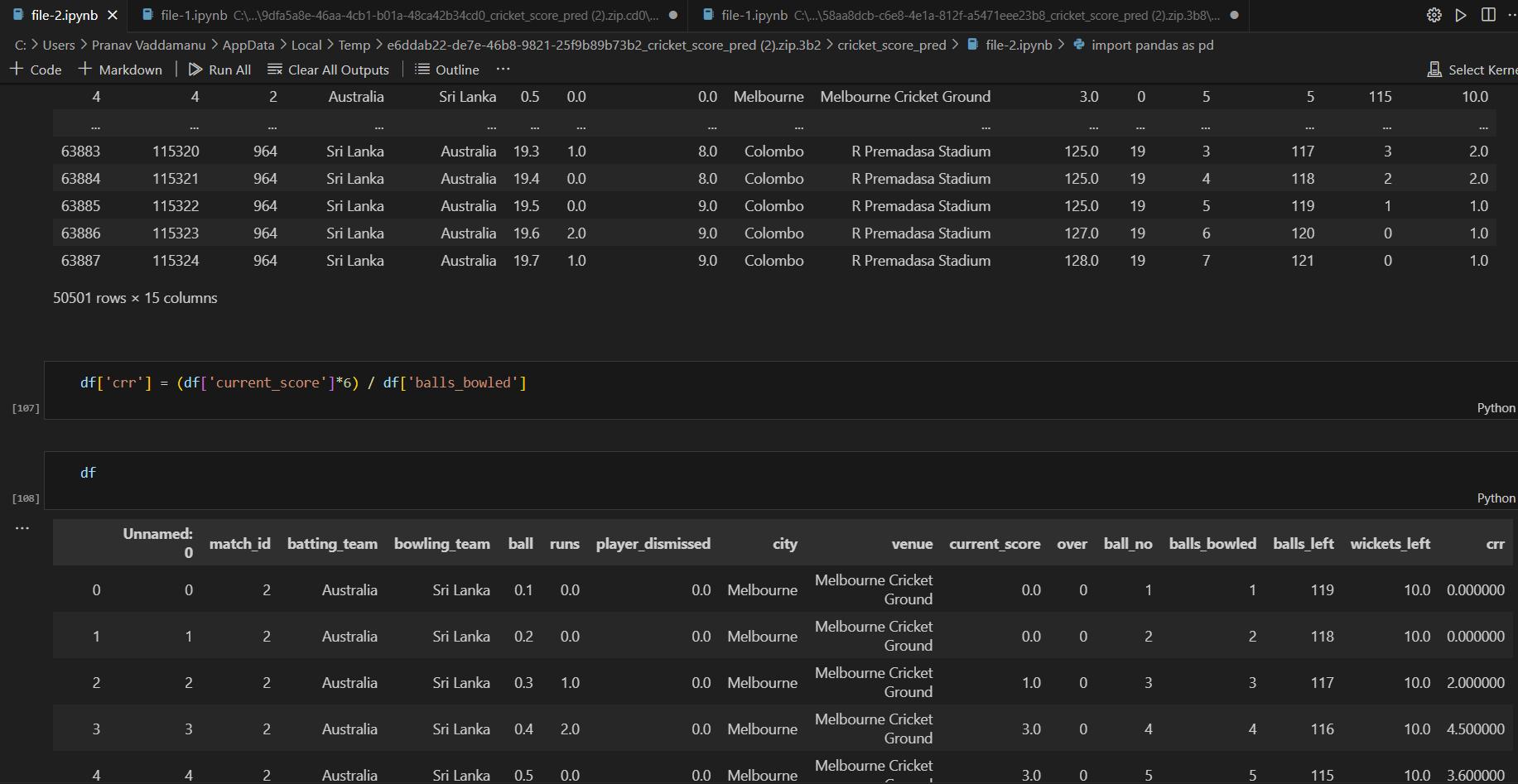


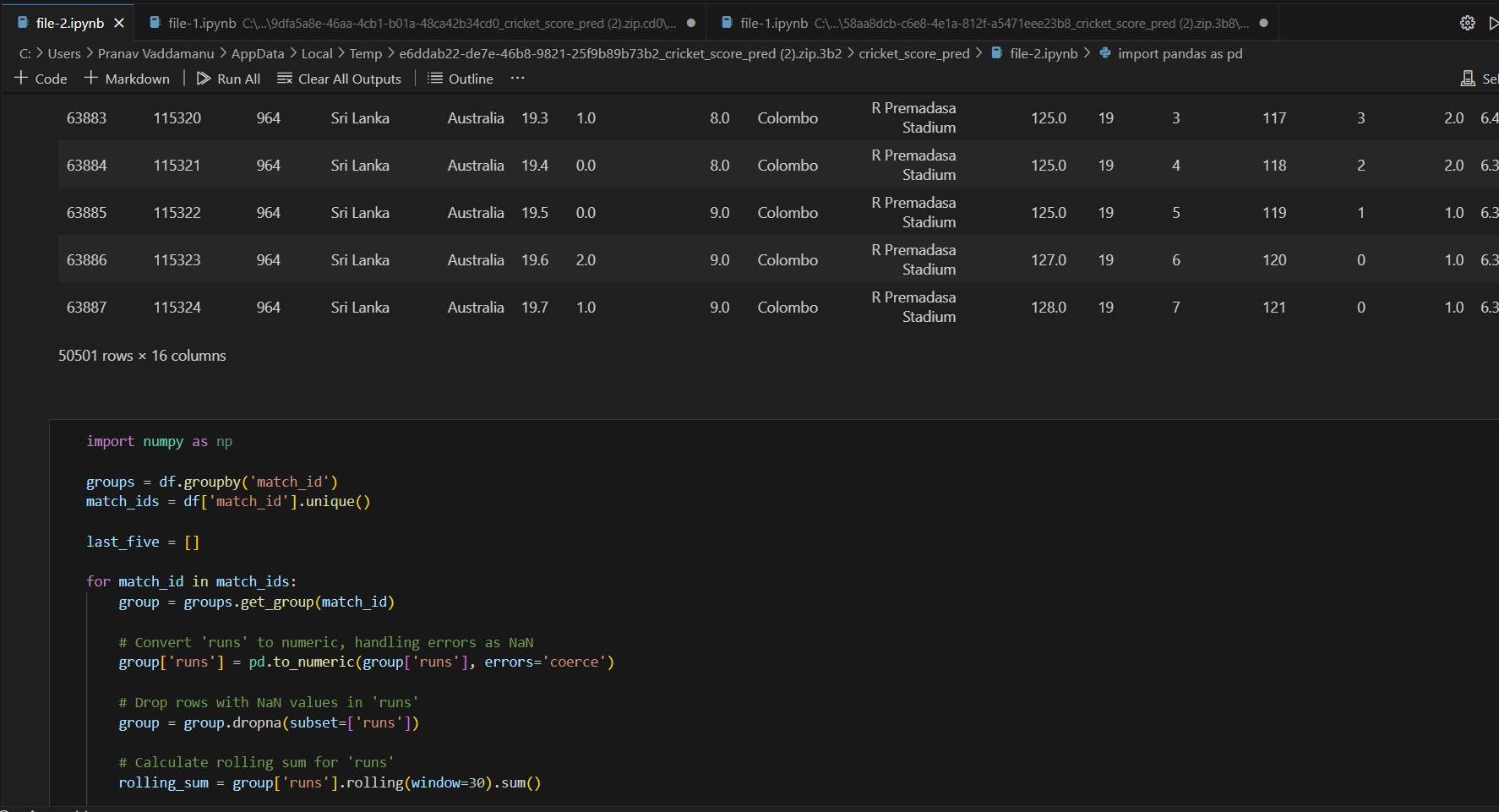


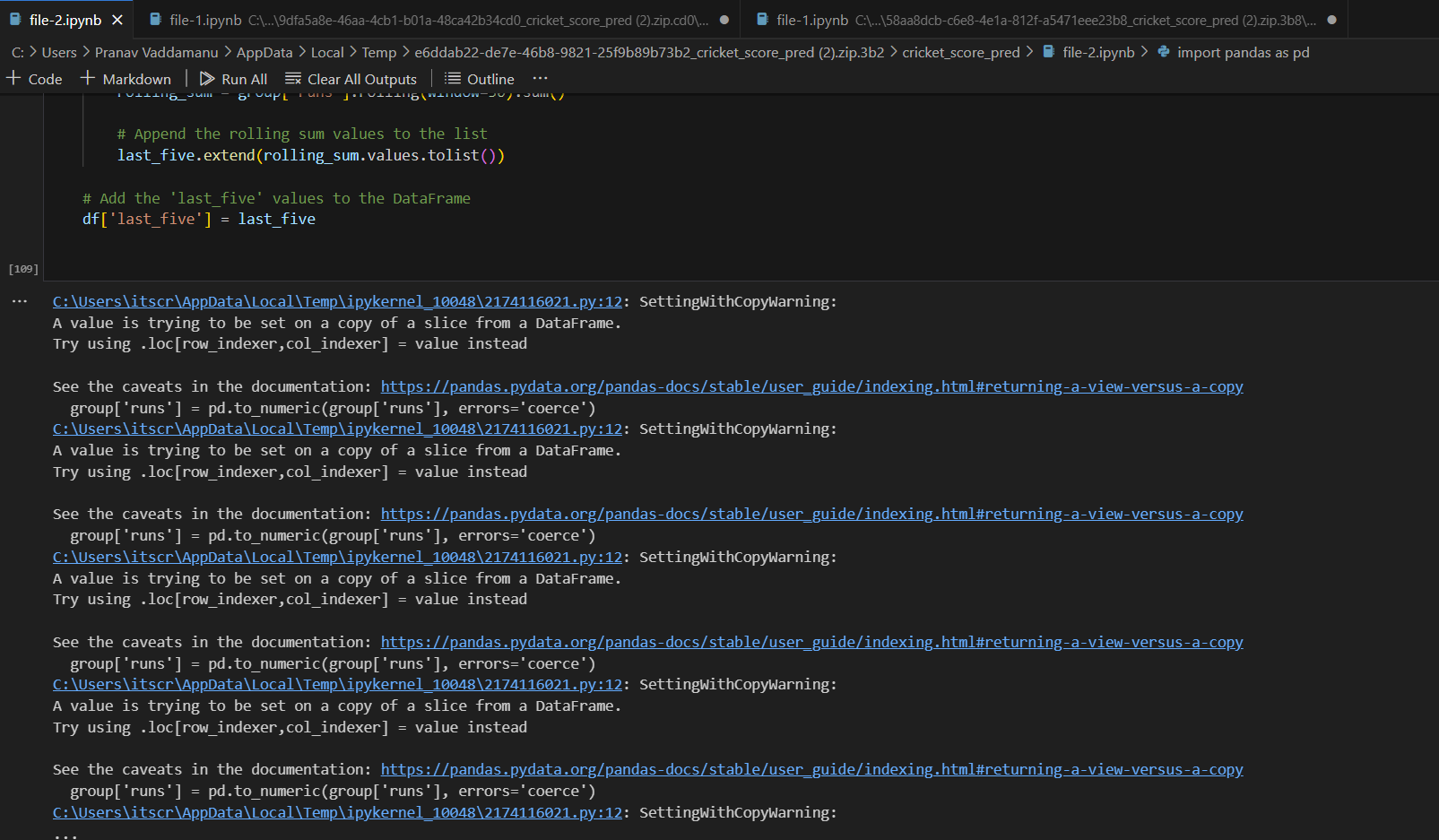


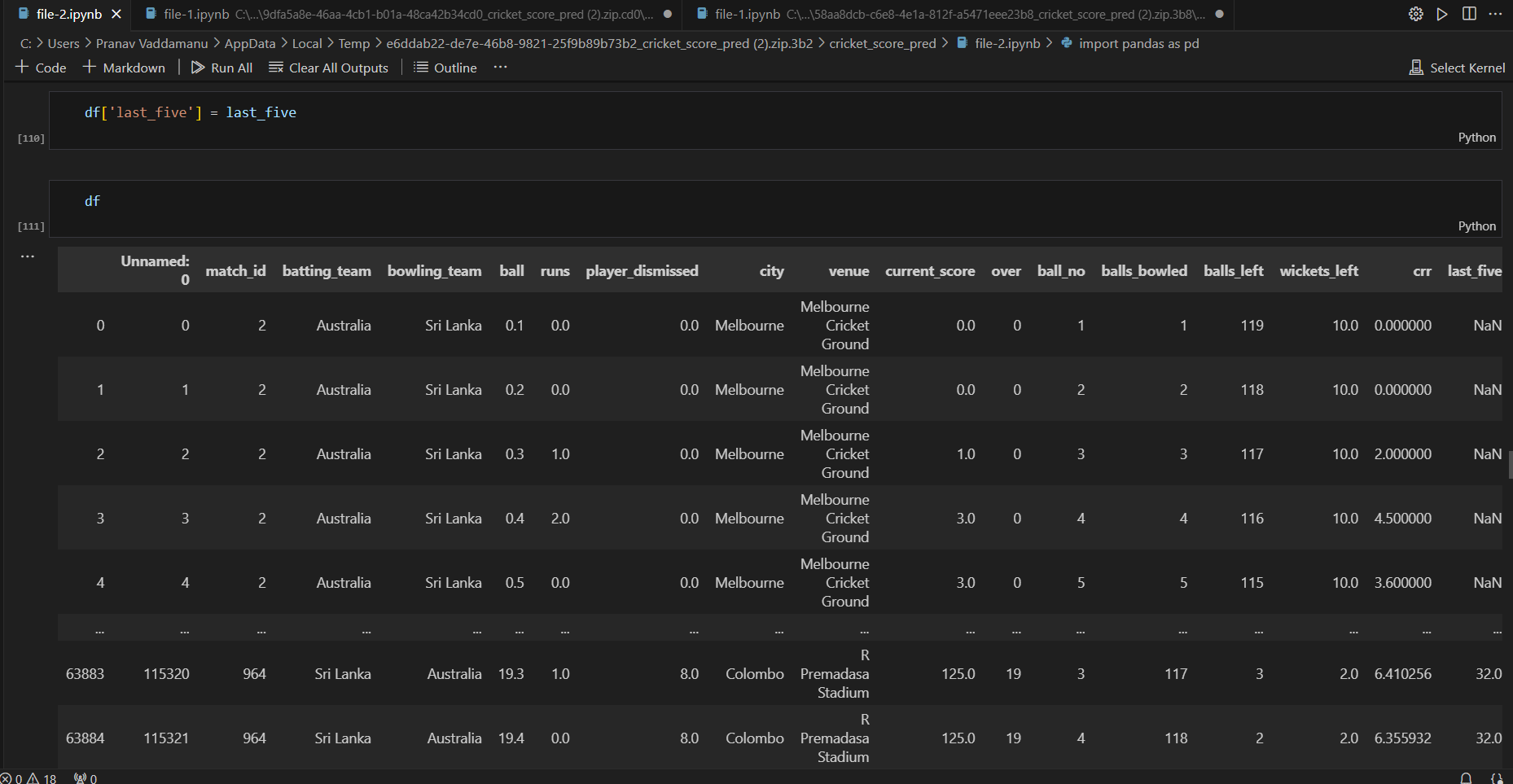


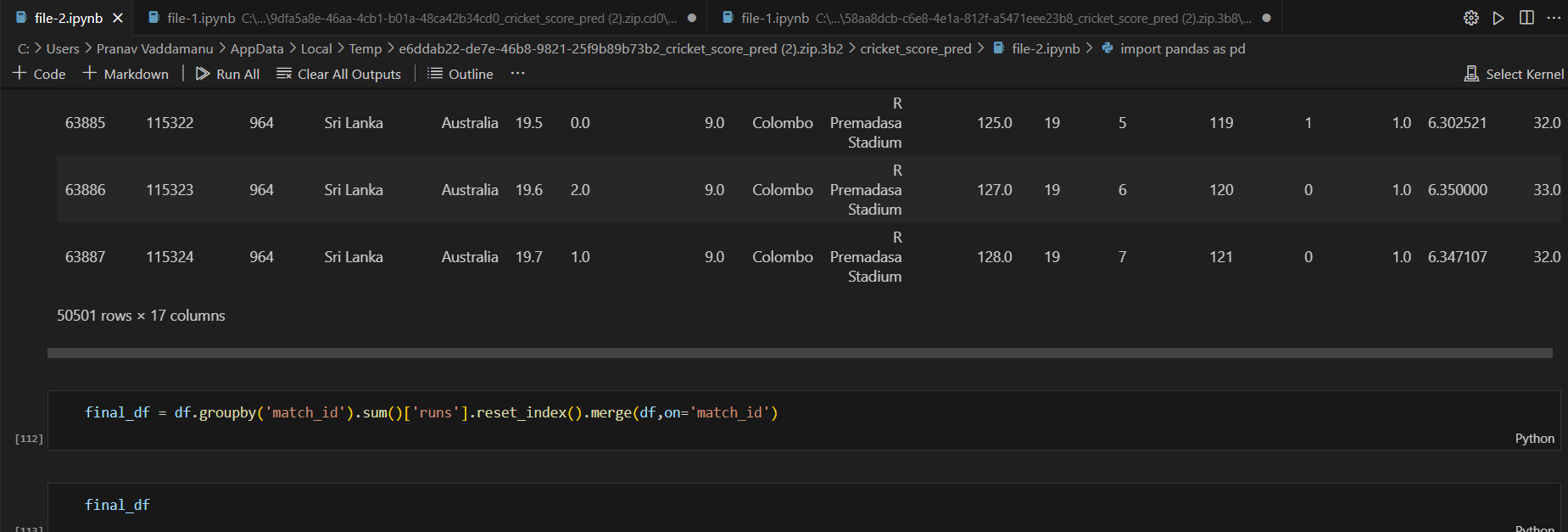


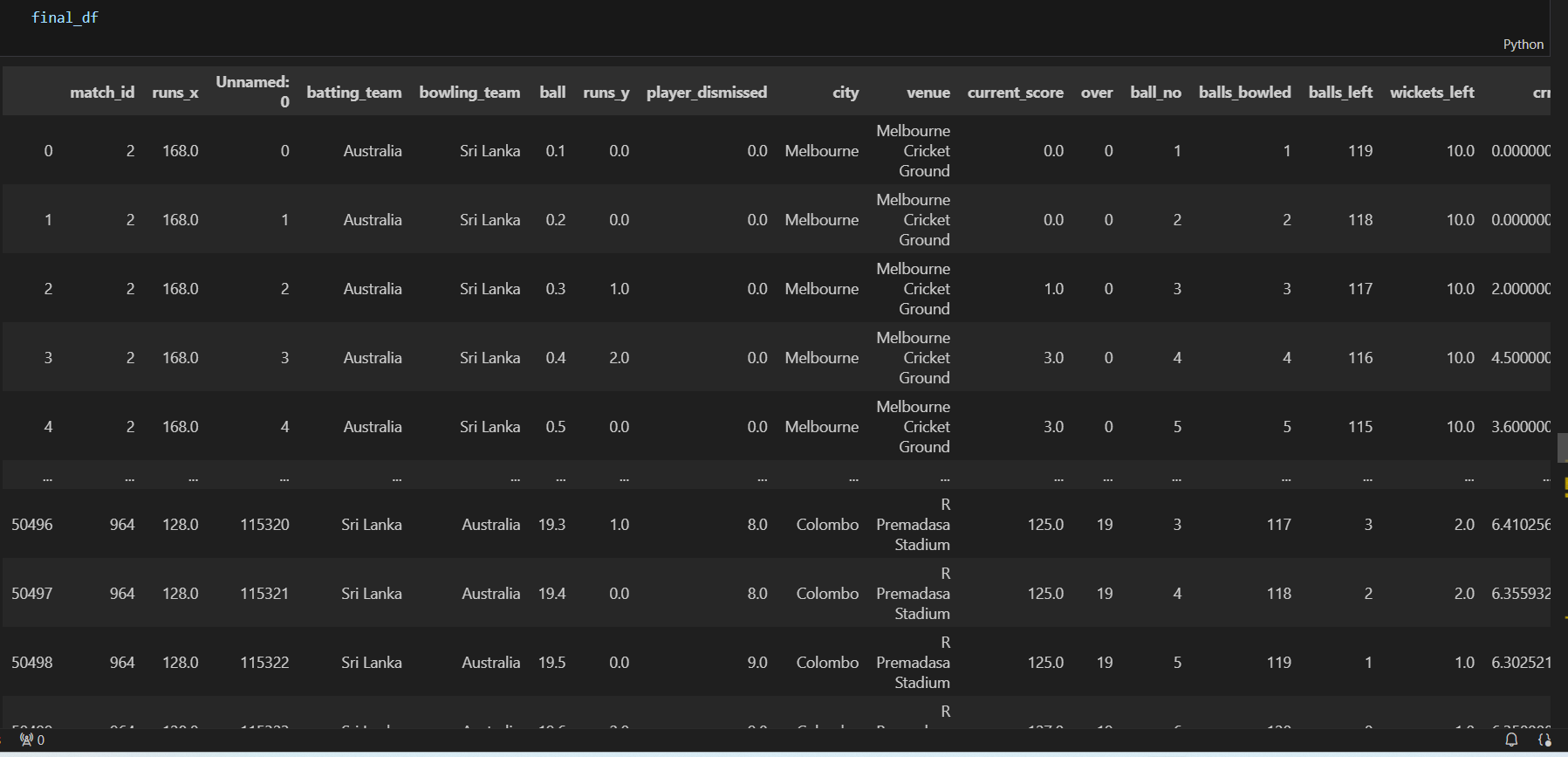


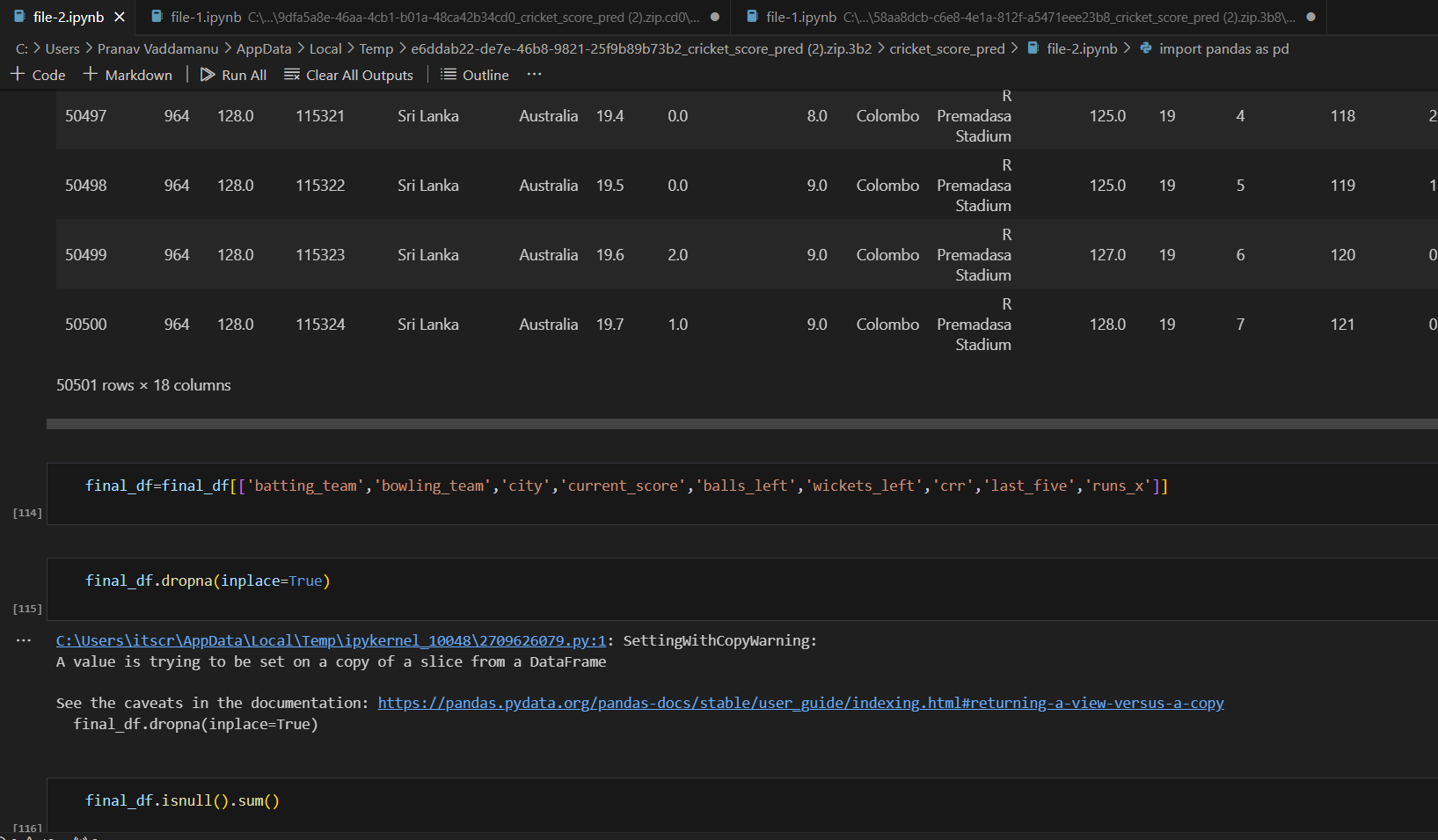


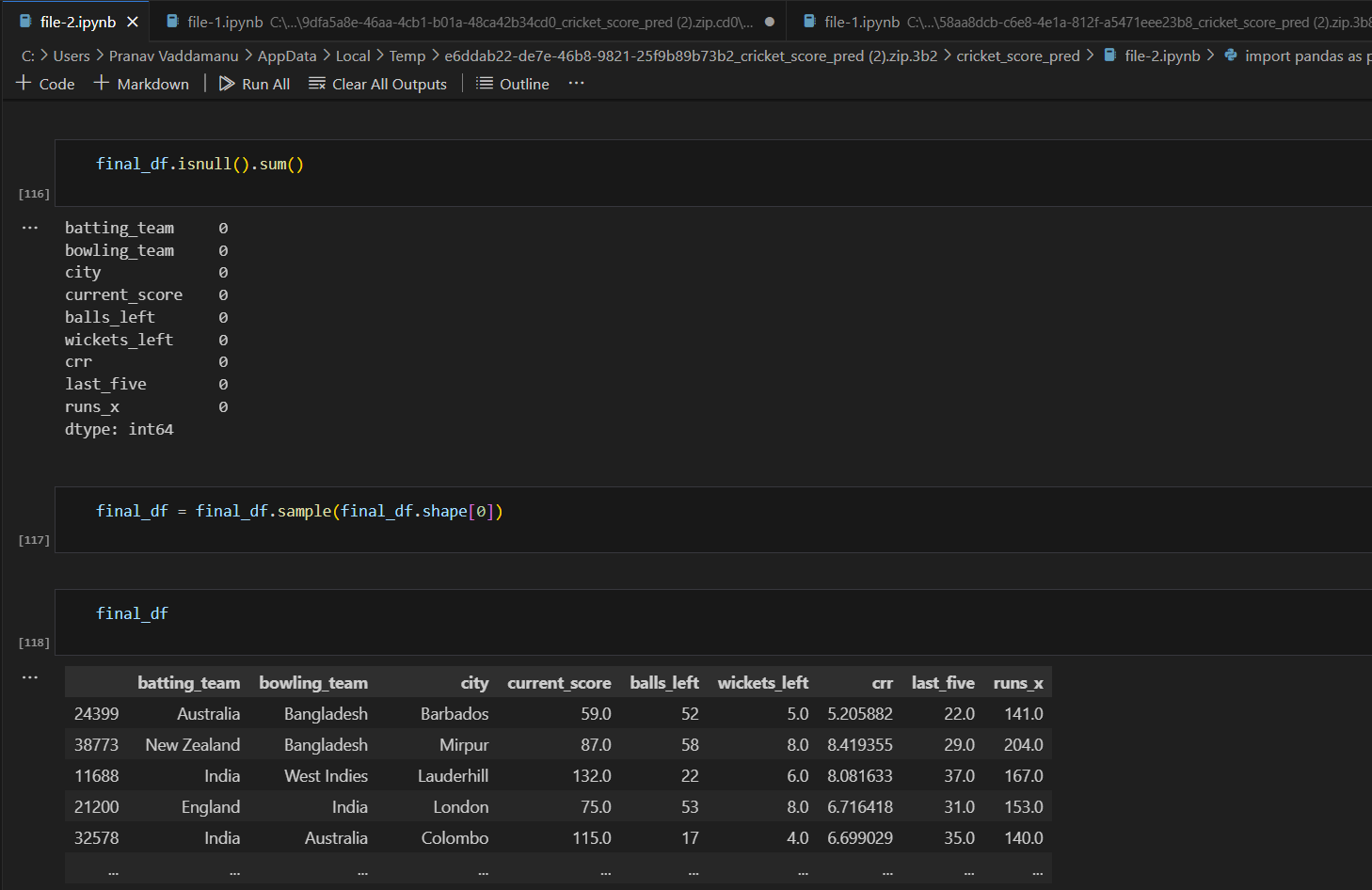


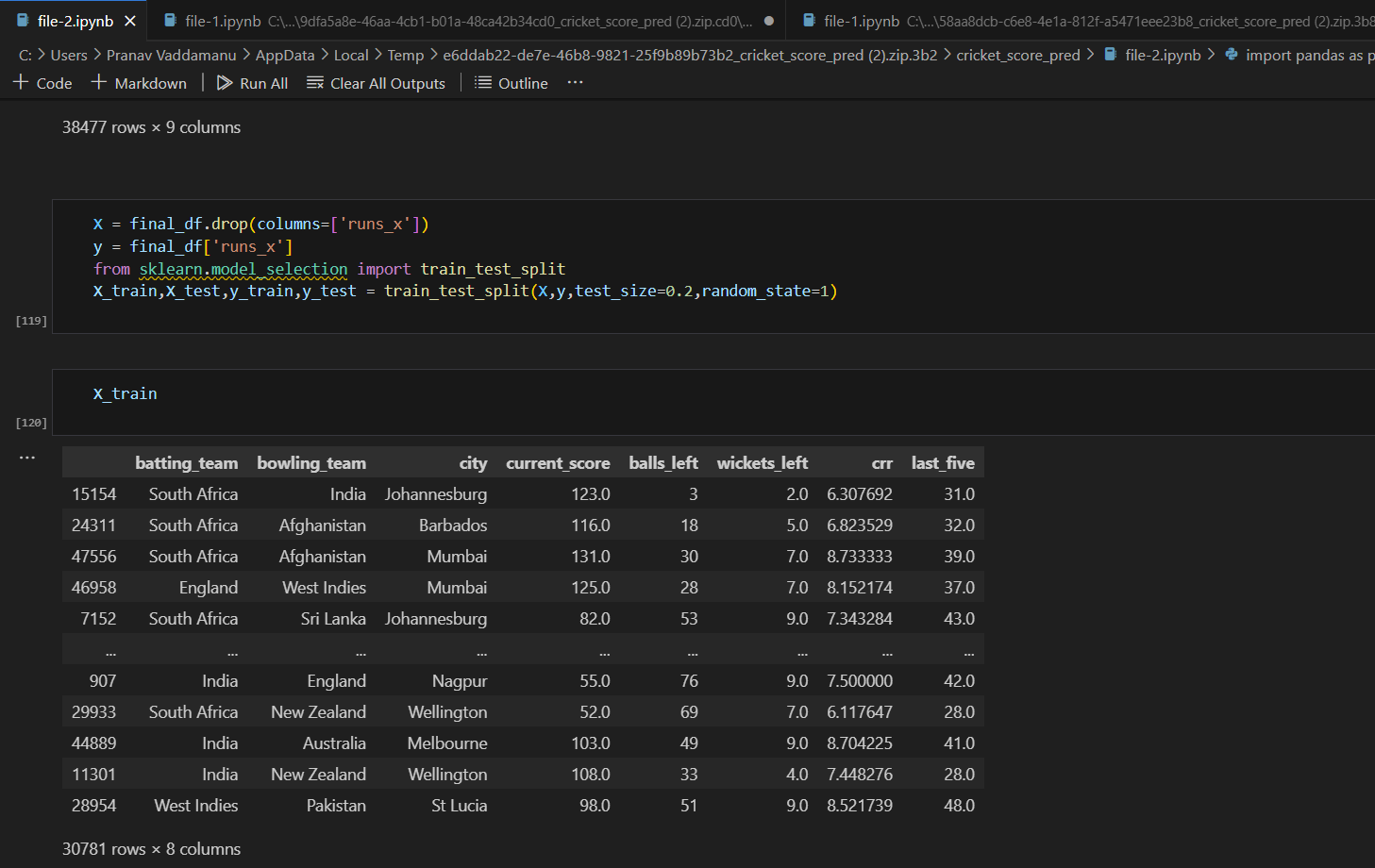


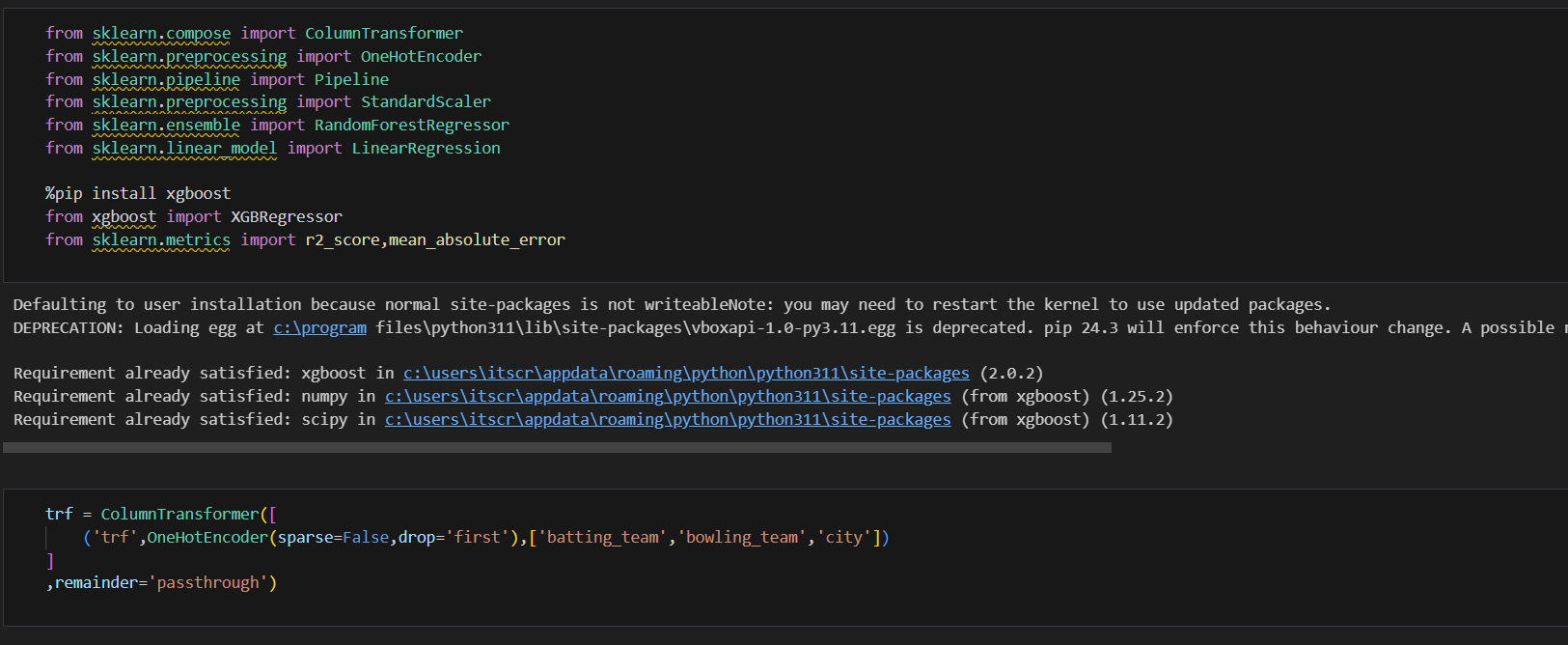


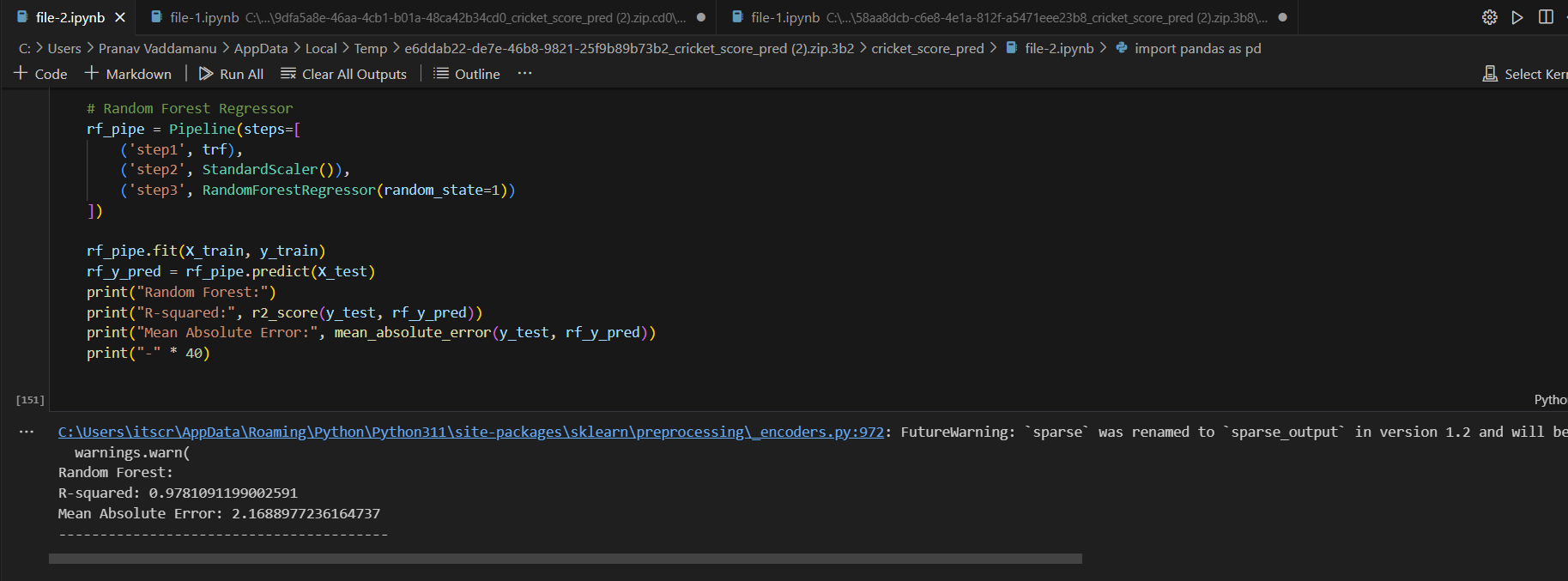




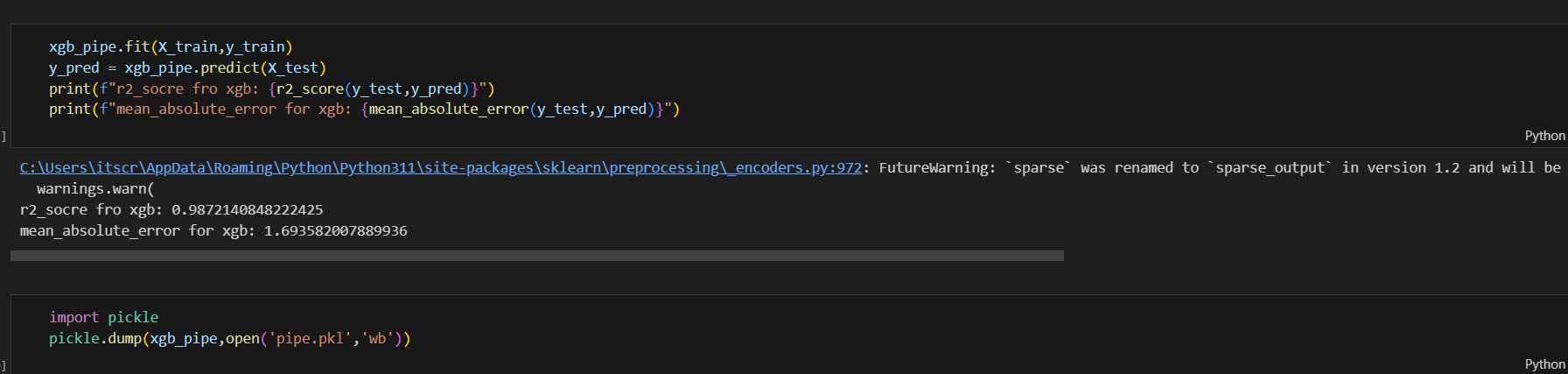


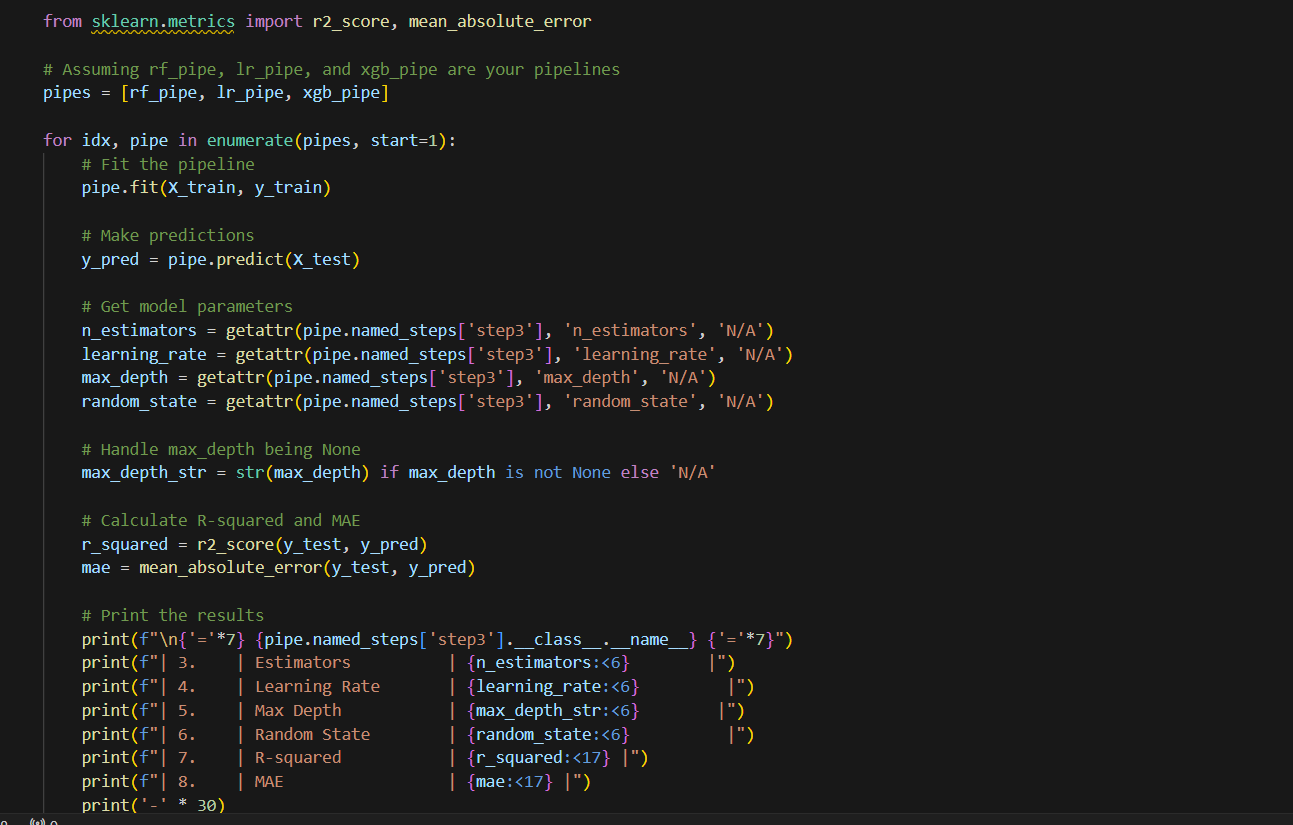


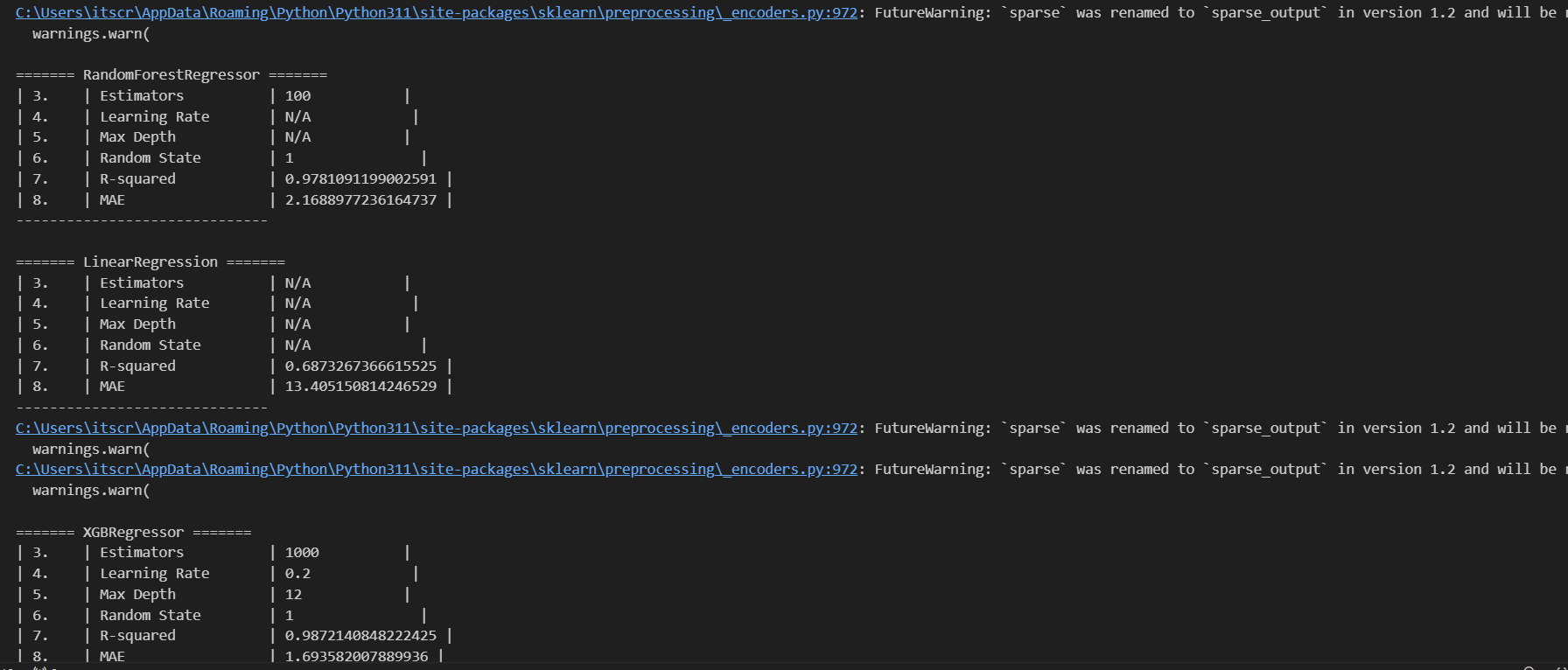




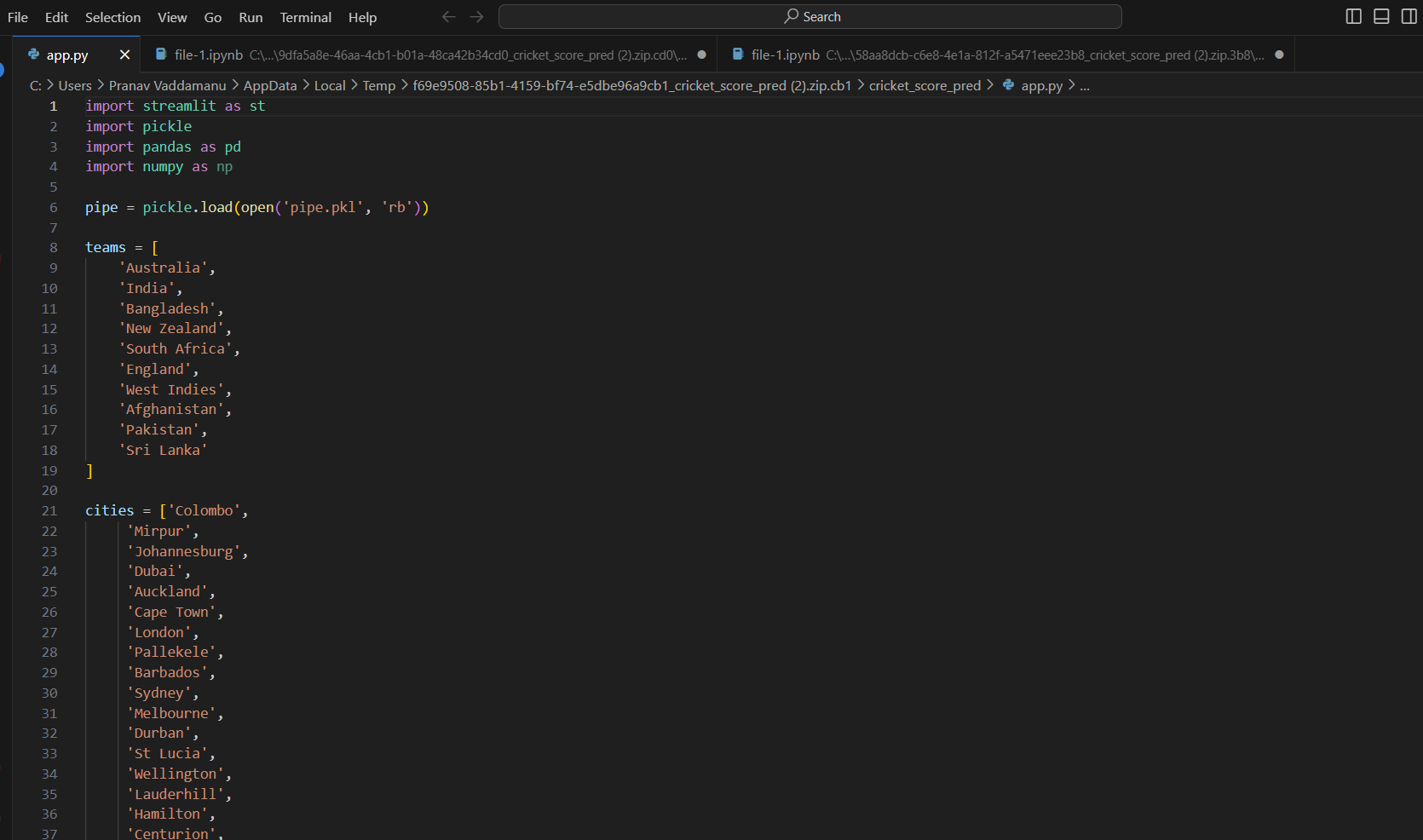


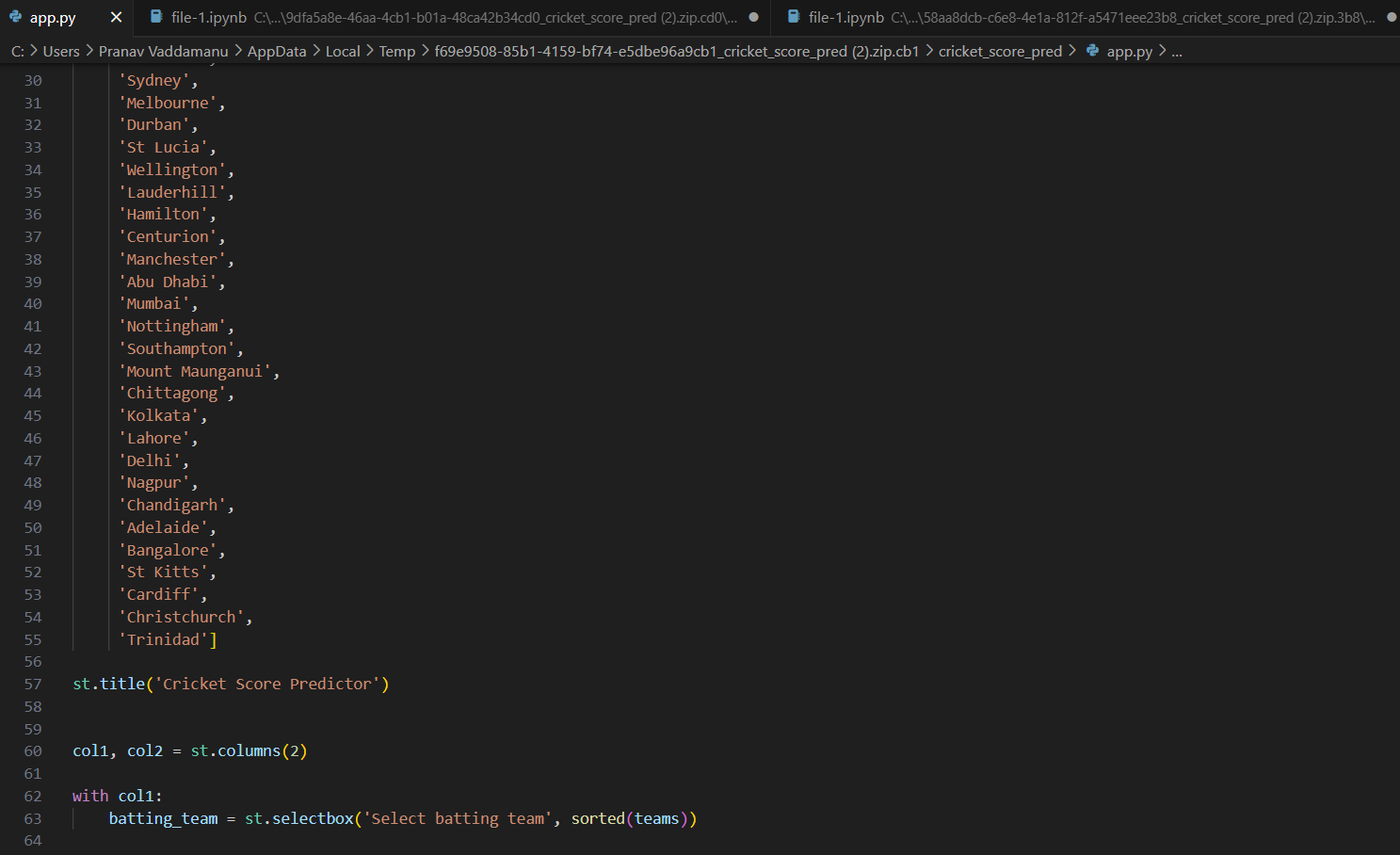


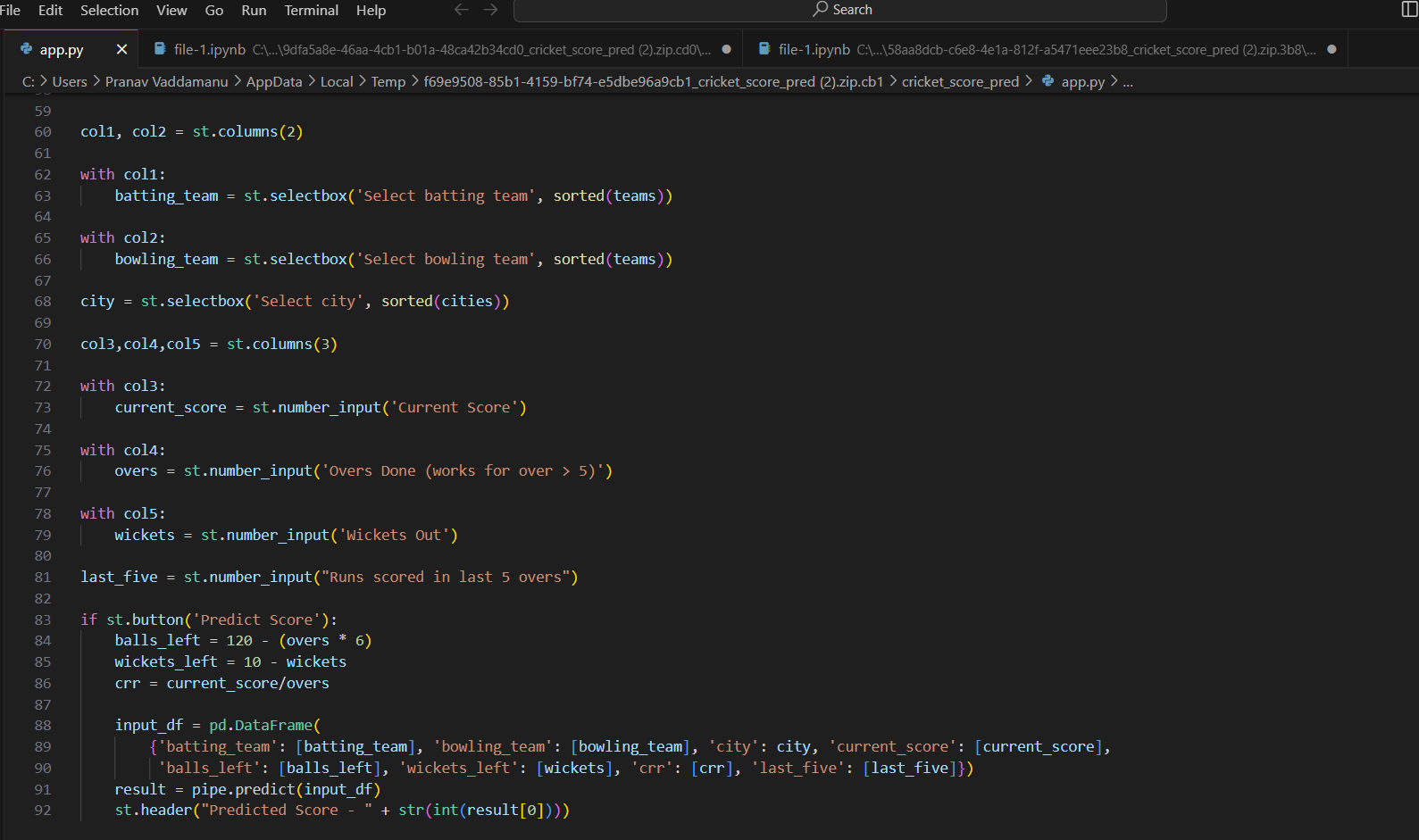




***Web Application Source Code:***

******

******

******

***OUTPUT***

******

******

### GitHub Repository

The complete project, including source code, datasets, and documentation, is hosted on GitHub. The repository can be accessed at the following link: https://github.com/smartinternz02/SI-GuidedProject-612871-1700582187

### Project Demo

For a live demonstration of the T20 Cricket Score Prediction application, please visit the following link:

### Data Dictionary

The data dictionary provides detailed information about the columns in the processed dataset, including their meanings and data types.

**Column Name Description Data Type**

batting\_team Team currently batting String

bowling\_team Team currently bowling String

city City where the match is played String current\_score Current total runs scored by the batting team Integer balls\_left Number of balls left in the match Integer wickets\_left Number of wickets left for the batting team Integer crr Current run rate Float

last\_five Runs scored in the last five overs Integer runs\_x Target runs for the batting team Integer

### Project Documentation

The project documentation, including detailed information on project overview, literature survey, requirements analysis, and more, can be accessed in the provided PDF [document: T20 Cricket Score Prediction Documentation (link-to- documentation.pdf)](file:///C:\tmp\MarkdownToPdfHandler\ucxongzg.qqz\link-to-documentation.pdf)

### Contact Information

For inquiries or additional information, please contact:

Project Team:

Pranav Vaddamanu: [pranav.vaddamanu2021@vitstudent.ac.in](mailto:pranav.vaddamanu2021@vitstudent.ac.in)

Sriram S: [sriram.s2021@vitstudent.ac.in](mailto:sriram.s2021@vitstudent.ac.in)

Lanka Vasanth: [lanka.vasanth2021@vitstudent.ac.in](mailto:lanka.vasanth2021@vitstudent.ac.in)

Note: Ensure that the GitHub repository is regularly updated with the latest code changes and documentation revisions.