ULMS861: Sports Economics and Analytics

Coursework

This coursework forms 100% of the final mark for the module. Submission deadline is 12 noon **Monday 11ᵗʰ March 2024**. This coursework requires **online submission.** You must submit via Turnitin. If a copy is not submitted to Turnitin the assessment will not be marked.

There are two sections to this assessment. Answers to Section A should be given using the template table provided below. Section B is an essay and your answer should be provided in the same document as Section A, starting on a new page. Include text, tables and figures (charts) in your document where and when appropriate.

Section A Assignment Project Exam Help

You are asked to obtain and analyse a dataset in Microsoft Excel (or other spreadsheet software of your choice), and provide insight and commentary on your analysis.
https://tutorcs.com

Download the csv file containing results of tennis matches from the men's ATP Tour during 2023 from http://tennis-data.co.uk/alldata.php
WeChat: cstutorcs

Note that the file is a csv (comma separated variable), and can be opened in Excel. Please search the internet for "*open a csv file in Excel*" if you are struggling. On the same website, there is a text file defining the variables in the dataset.

You will need to use a combination of formulas, functions, and pivot tables to answer these questions in Excel. All answers are based on the data provided in the spreadsheet.

Write your answers in the answer sheet provided.

Question 1

a)   How many matches did Novak Djokovic play in 2023?

(3 marks)

b)   Who played the most matches in 2023?

(2 marks)

c) In 2023, which player had the most losses and how many losses did they have?

(2 marks)

d) How many times did the underdog (player with the longest odds), according to Bet365, win?

(2 marks)

e) In what percentage of matches does the underdog win?

(1 mark)

Question 2

a) Copy and paste a bar chart showing how many times Novak Djokovic played in each round of a tournament (1st round, 2nd round, 3rd round, 4th round, quarter-finals, semi-finals, and finals). Exclude round-robin matches.

(5 marks)

b) Why has Novak Djokovic played more matches in quarter-finals than in the 4th round?

(1 mark)

c) What is Novak Djokovic's win percentage on each surface: clay, grass, hard?

(4 marks)

d) In all matches he played, how many games did Novak Djokovic win and lose in 2023? What is his game win percentage?

(3 marks)

Question 3

You are now going to build a forecasting model for tennis based on the Bradley-Terry model. To start with you will filter the data to exclude players that have played fewer than 20 matches in 2023.

a) Create a pivot table for the number of matches each player has played. The players will be order alphabetically. Paste top 10 rows of your pivot table into the answer sheet (use paste special and choose png).

(2 marks)

b) Use *vlookup()* to add two columns to the match result data for the number of matches the winning and losing player played in 2023. Copy the two formulae you have used into the answer sheet.

(2 marks)

c) Filter the match result data to only include players that played at least 20 matches in 2023. Copy and paste it to a new worksheet. How many matches does this leave?

(2 marks)

d) Using a Bradley-Terry model, estimate the player strengths using Carlos Alcaraz as the reference player with a strength of 1000. Note that the likelihood for player i beating player j is

$$L = \frac{\alpha_i}{\alpha_i + \alpha_j},$$

where α_i is the strength of player i.

Based on the estimated strengths, who are the top 10 players of 2023? Write your answer in the answer sheet.

(10 marks)

e) Based on this estimated model, what is the probability of Danill Medvedev beating Gregor Dimitrov?

Assignment Project Exam Help (2 marks)

Question 4

https://tutorcs.com

a) Extend the Bradley-Terry model you have fitted in Question 3 to use sets, not matches, as the unit of victory. Again, set Carlos Alcaraz as the reference player with a strength of 1000.

WeChat: cstutorcs

Note that the likelihood for player i winning s_i games and player j winning s_j games is

$$L \propto \frac{\alpha_i^{s_i} \alpha_j^{s_j}}{\left(\alpha_i + \alpha_j\right)^{s_i + s_j}}$$

and you should maximise the logarithm of this.

Based on this model, who are the top 5 players, and their estimated strengths?

(5 marks)

CONTINUED ON NEXT PAGE

Section B

In 1,000 words or less, please discuss **<u>one</u>** of the following.

(i)   In a team sport, explain why we cannot evaluate a managerial turnover decision by comparing team performance before and after such a decision. Choose one recent publication (not older than 10 years) to explain how to tackle this concern and discuss the main empirical results.

(ii)   Find an example where sports data can be used to investigate the presence of racial or gender discrimination. Discuss the employed analytical methods from a critical perspective.

Where appropriate you should support your work using relevant citations (news articles, websites, academic articles, and books are all acceptable forms of evidence), examples, and data.

Assignment Project Exam Help

https://tutorcs.com

WeChat: cstutorcs