MA22019 2025 - Coursework 2

Generative AI statement

Please comment on your use of generative AI for this coursework.

Question 1

Generative AI tools have gained a reputation for their ability to produce texts in the style of famous authors. You are asked to put this claim to the test using the work by Charles Dickens.

The file "Dickens_Fakes.csv" contains five pieces of text written by ChatGPT in the style of Charles Dickens, and "Dickens_Originals.csv" provides ten randomly selected chapters from the books "A Tale of Two Cities" and "Great Expectations". Explore the similarities and differences between the original and artificially produced pieces of text. Your analysis should consider both the words used and the emotional intent of the texts.

What do you conclude regarding ChatGPT's ability to produce work in the style of Charles Dickens? How confident should we be about your conclusion?

Answer

Please present your analysis for Question 1 here.

Question 2

The region of Tesremos is one of Utopia's key sources for freshwater. Unfortunately, Tesremos has seen lower than expected rainfall over the past years, which in the longer term may threaten the country's capacity to deliver clean water to its citizens. Due to its importance, the government of Utopia has been monitoring ground water levels at several sites across Tesremos for decades.

The government of Utopia has now approached you to analyse some of last year's data to get a better understanding of the current situation. Specifically, they provided you with daily observations from their 102 groundwater monitoring sites across Tesremos for March and August 2024, which correspond to the beginning and end of Utopia's dry season respectively. Instead of the actual values, the government of Utopia has decided to provide you with the differences (in percent) to historical averages. As such, a value of x for a day and site in 2024 corresponds to ground water levels being x% lower (for x negative), or x% higher (for x positive), than the groundwater levels which were observed historically for that day and site.

You are also provided with a shapefile for Tesremos, and the locations of the 102 monitoring sites. To hide Tesremos's location, constants have been added to the latitude and longitude coordinates, but the shapes they define are correct.

There are two tasks the government of Utopia asks you to complete using the provided data:

i. *Explore the spatial pattern/structure and dependence of the monthly average deficit in groundwater levels for the 102 monitoring sites separately for March and August. Discuss all the assumptions you make for the analysis.*

ii. *Produce maps which provide estimates for the monthly average deficit in groundwater levels across the region for March and August 2024. Discuss the reliability of your estimates.*

Answer

Please present your analysis for Question 2 here.

Question 3

The Isles of Sofara are one of the most beautiful spots across Utopia and very popular with tourists. There are a total of five islands: Calmorra, Ecliptria, Justitia, Paxora and Solvenya. You have been provided with a shapefile of the islands; to hide their location, constants have been added to the latitude and longitude coordinates, but the shapes they define are correct.

Over recent decades, the islands have been frequently hit by devastating tropical storms and the government of Utopia was repeatedly blamed for their poor disaster response. To address this issue, the government of Utopia set up a helpline in 2020 - every person in distress due to a tropical storm should send a text message to the number 971 with the hashtag #IoSTSHelp and, if possible, provide their location information. Since 2020, a few tropical storms hit the islands and, while the introduction of the helpline and hashtag was generally successful, the current setup requires a government employee to manually evaluate the text messages received, which leads to delays in the disaster response.

The government of Utopia has now approached you for help with the analysis of the text messages sent to 971. Specifically, for the messages with the hashtag #IoSTSHelp, they ask you to:

i. *Visualize the number of messages per island and identify, as precisely as possible, the areas from which the most distress messages were sent.*

ii. *Develop an approach to identify the kind of help required. The four key categories are (a) lack of food, (b) lack of fuel, (c) lack of shelter and (d) need for medical assistance. Use your approach to provide estimates for the number of people who required help related to the categories (a)-(d).*

The government of Utopia is planning to use your analysis approach for any future tropical storms. Specifically, they want to use your approach to identify the areas that need help and the kind of help required, without a government employee having to manually evaluate the text messages. Discuss the strengths and limitations of your analysis approach regarding this aspect.

Answer

Please present your analysis for Question 3 here.