Dear Times puzzle editor

I am writing to you as MCM Team 2314475 to summarize our results on forecasting Wordle.

After analyzing various data sources and conducting extensive research, we have made some

interesting findings that I believe your readers may be interested in

First, we developed a model to explain the change in the number of results reported every

day. I found that using \* \* ARIMA model \* \* to establish a time series regression model

for the number of reported results can best explain this change. ARIMA is a time series

analysis method, which is a combination of autoregressive moving average model (ARMA)

and differential integration model (I). ARIMA model makes the time series data stable by

differential, and on this basis establishes an autoregressive moving average model to predict

the future value. Based on this model, we have created a forecast interval for the number of

reported results on March 1, 2023. My prediction is that the number of reported results on

this day will be in the range of 19368-20982

Secondly, we investigated whether any attribute of a given word would affect the percentage

of report scores played in difficult mode. Our analysis found that there was a significant

correlation between the frequency of words and the percentage of scores reported in the

difficult mode. Specifically, more common words tend to report higher scores in the difficult

mode. However, I did not find any significant correlation between other attributes of a given

word (such as the irregular spelling of the word) and the percentage of scores reported in the

difficult mode

Thirdly, we have developed a model to predict the distribution of report results for a given

solution word in the future. Our model uses the relevant attributes that are usually used to

describe the clustering results. Based on this model, we made a specific prediction for the

word "EERIE" on March 1, 2023. We predict that the reported result distribution of this word

will be (1=0%, 2=2%, 3=15%, 4=33%, 5=30%, 6=16%, X=4%)

However, there are also some uncertainties in this prediction, because the distribution of the

report results may vary according to the playerâ€™s choice and the uncertain factors of the day

Fourth, we have developed a model to classify solution words by difficulty. We use the central

decision cluster of the system cluster and the attributes of a given word to predict whether it

is classified as "difficult", "medium", and "easy". Our analysis found that words with fewer

occurrences are often classified as "difficult", while more common words are often classified

as "simple". In addition, words with high repetition times are often classified as "simple",

while words with high orthogonality are often classified as "difficult". Based on this model,

the word "EERIE" is classified as "easy"

Finally, there are other interesting features of this dataset that are worth mentioning. For

example, we note that the number of reported results will gradually decrease and stabilize

over time, which indicates that the popularity of Wordle games is gradually declining. In

addition, the data set contains a wide range of words, ranging from common English words

to more obscure terms.

In conclusion, we believe that our analysis provides valuable insights into the patterns and

trends in the data set provided. We hope that these findings will arouse the interest of your

readers, and we look forward to hearing any feedback or comments from you

Sincerely,

Sicerely yours,

Your friends