Week 2

R Output.R

Student Name: Zihan Ma

Class NameALY6010: Probability Theory and Introductory Statistics

Instructor: Tom Breur

03/05/2023

Report part 1:

For the Dataset, I’m using the free Dataset that I found from Kaggle (https://www.kaggle.com/datasets/nicoelbert/aoe-matchups?resource=download); the chart below shows the first five observations of the data after cleaned.

It’s a data set which record 225.000 matches in the game Age of Empires 2: Definitive Edition.

The duration column means how long does the match goes.

The Elo column means the player’s skill ranking, p1&p2civ shows the civilization which player use, and if the winner = 0 then p1 won the match, otherwise p2 was the winner.



I picked Britons, Byzantines and Byzantines to focus on because in a famous youtuber.

Spirit Of The Law ’s video, he explains that these are the three civilizations that fit for low rank players. (Spirit Of The Law, 2020) But it’s just a conclude from his gaming experience. I want to use this dataset to prove wither he is right or wrong.

A picture containing table

Description automatically generated

The Table 1.1 show all dataset’s attributes. We can observe that p1 have about 51% win rate, which means we basically can ignore the different between the player is eighter in p1 or p2. The median of player’s elo is 1063, then we can treated the player under this elo are “low rank players”.

Graphical user interface

Description automatically generated with low confidence

Graphical user interface, application

Description automatically generated

A picture containing graphical user interface

Description automatically generated

The table 1.2, 1.3, 1.4 is showing the basic attributes of the player wins as the civilization on the top of the chart. There is a small amount of differences between the average elo, we can not find any valuable information from it. But the time Median shows these are all civilizations that always extend the length of the match.

Report part 2:

Chart, scatter chart

Description automatically generated

**Chart, scatter chart

Description automatically generated**

**Chart, scatter chart

Description automatically generated**

Chart, scatter chart

Description automatically generated

From plot 2.1, 2.2, 2.3 and 2.4 we can see that the low rank player usually make very long game time. These three civilization are have big advantage for late games, so that could a reason for they are good for new players cause they experience longer fights.

Chart, scatter chart

Description automatically generated

Chart

Description automatically generated

The table 3.1 and 3.2 shows these three new players civilizations have higher density than other civilizations on low rank distributions, which means they are used more on lower rank than most of civilizations but not higher ranks.

Chart, box and whisker chart

Description automatically generated

Timeline

Description automatically generated with medium confidence

The last Table 4.1 and 4.2 is only showing the winner results for all civilizations. From the table we can see clear they help lower rank people gained more victories more compared to higher elo players.

References:

Elbert, N. (2022, November 7). *Age of empires 2: De Match Data*. Kaggle. Retrieved March 5, 2023, from https://www.kaggle.com/datasets/nicoelbert/aoe-matchups?resource=download

Spirit Of The Law. (2020). *Best Civ for Beginners? [AoE2]*. *YouTube*. YouTube. Retrieved March 5, 2023, from https://www.youtube.com/watch?v=y0BA7iX14Wo&ab\_channel=SpiritOfTheLaw.