# BOARD GAMES DATA ANALYSIS

**Dmitri Sajutinski, Taavi Aru, Elis Teemus**

Introduction

Our dataset from Kaggle contains information regarding around 20 000 board games. The columns include values such as min and max amount of players, average playtime, complexity and various other parameters that define a board game. Our goal was to find out how these parameters influence the avarage rating of a board game and also to predict the average rating based on those values.

Some interesting findings
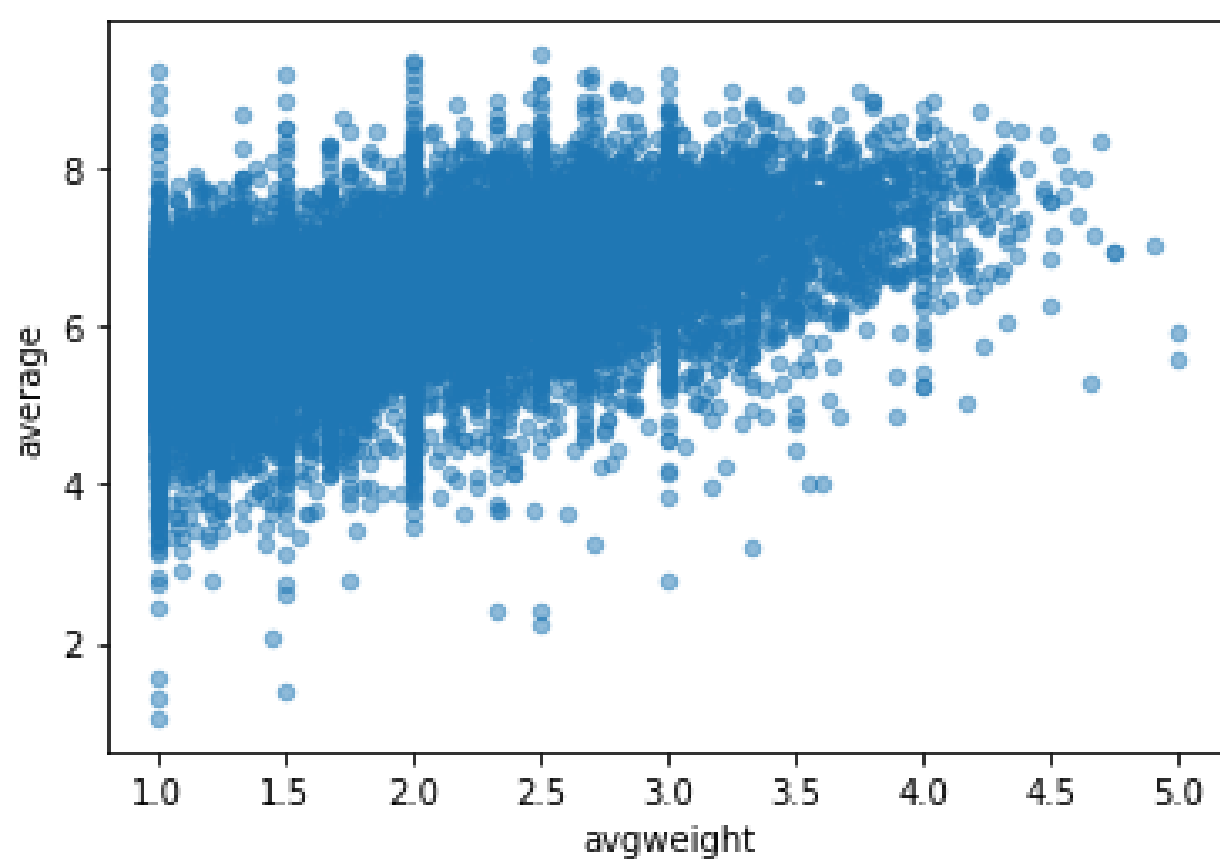
avgweight - game complexity rating from 1-5
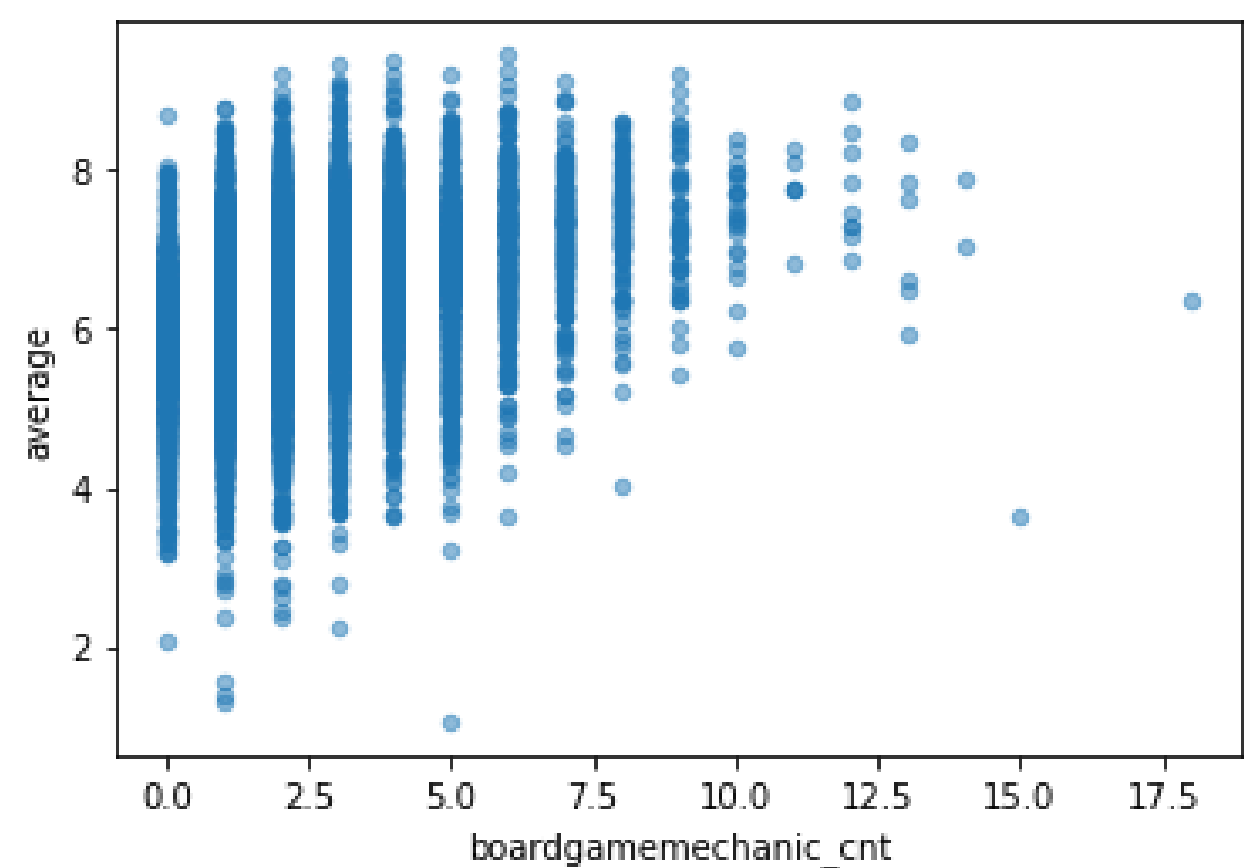
boardgamemechanic_cnt - number of different elements/types of gameplay

One of the findings that surprised us was that game complexity seemed to be positively correlated with the game rating, as seen on the above plots. The average rating was also positively correlated with the following attributes: the number of people who have rated the game, the number of times it has appeared on the news, blogs, podcasts or websites, the popularity of the game on the BoardGameGeek website (number of views, comments, plays, people wanting and trading the game etc), the number of expansion packs that have been released for the game. This was not as surprising for us however, since it seems reasonable that games that have gotten very popular are also very highly liked. What also surprised us was that there didn't seem to be much of a correlation between minimum and maximum playtime.

Predicting the average rating

To predict the average rating of a game we trained a neural net.
We succeeded with the mean absolute error of 0.19544... (aka on average the prediction is wrong by ± 0.19544...), which we consider a fairly satisfactory result.