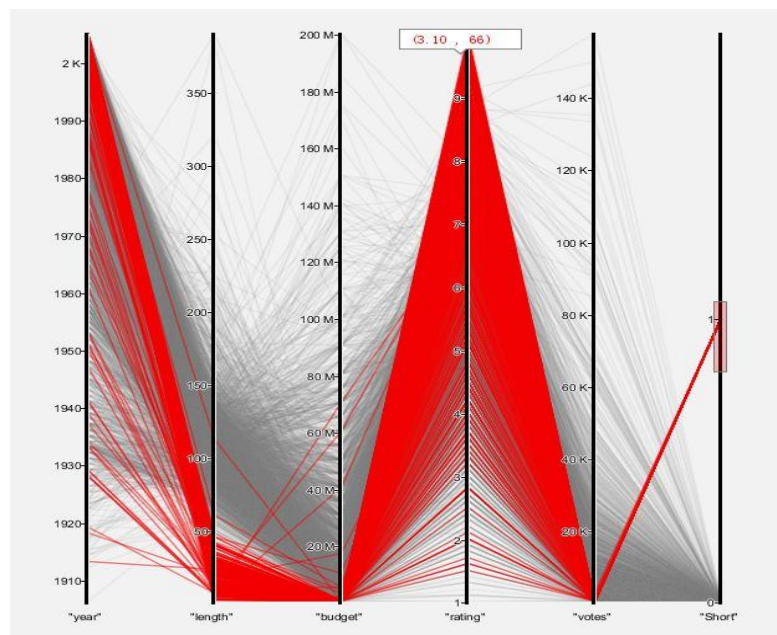
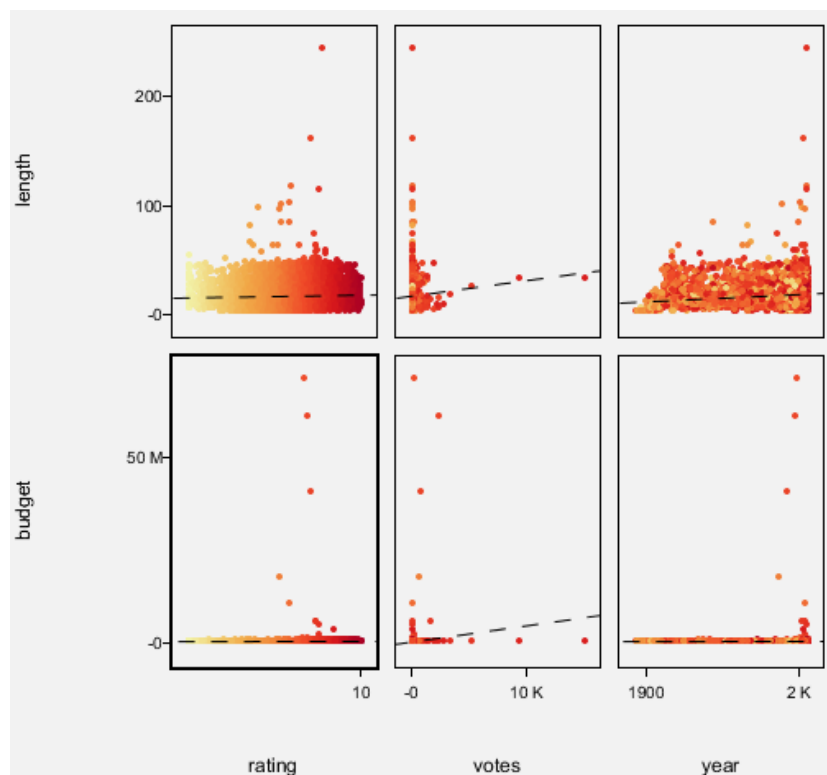


Through observing the table, we know that most logs of budget are missed, which value 0. In this case I select the part whose budget > 0.



Use parallel coordinate system to show the new data set, select the movies whose types value "short". We can see short movies often get rating scores higher than average. Also there were few short movies in the past compare to the amount of such in recent years. Besides, most short movies cost less.



This scatter plot creates mental image to show short movies get less attention as fewer people gave votes to them. In addition, we can notice an exceptional dot which shows the "short" movie covers 112 minutes. So I consider it a mistake among the data.