Scatterplot based on how num\_voted\_users relates to score

```{r}

# scatterplot using ggplot2

movie\_df %>% ggplot(aes(x = num\_voted\_users, y = imdb\_score)) +

geom\_point(shape = 1) +

ggtitle("Voted User & Score Linear Scatterplot")+

ylab("IMDB Score") +

xlab("Voted Users")

```

Based on this scatterplot, it appears that there is a moderate positive correlation between the amount of users who have voted on a movie and its overall score. There is a cluster where presumable niche or indie movies with a smaller amount of user votes have a range of varied scores, but as the number of user votes increases, the IMDb score approaches 10 (the maximum possible score). This may actually be because better movies would be seen by larger audiences over time and would therefore receive more votes.

Number of movie facebook likes seems closely related with # of critic reviews, let's take a look

```{r}

# Scatterplot using ggplot 2

ggplot(small\_movie\_df, aes(x = movie\_facebook\_likes, y = num\_critic\_for\_reviews)) +

geom\_point(shape = 1) +

geom\_smooth(method = lm, color = "Green") +

ggtitle("Facebook likes and # of critc review scatterplot") +

ylab("# of Critic Reviews") +

xlab("Move facebook likes")

```

Similar to the previous scatterplot of user votes and IMDb score, there appears to be a cluster with a large spread initially where the number of movie likes on Facebook isn't a strong indicator of the number of critic reviews, but a moderate positive correlation overall between the two exists. This may be because a given movie's overall popularity could attract more of both critics and the general audience alike.