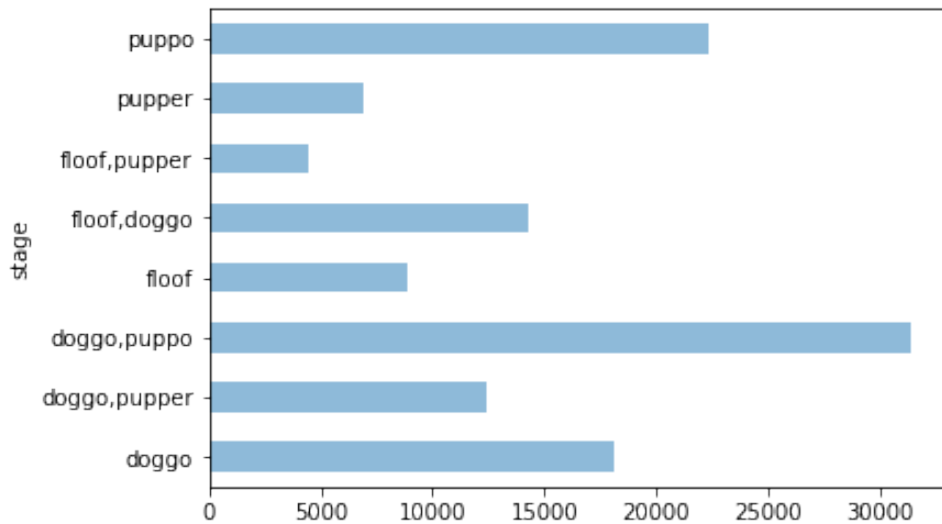
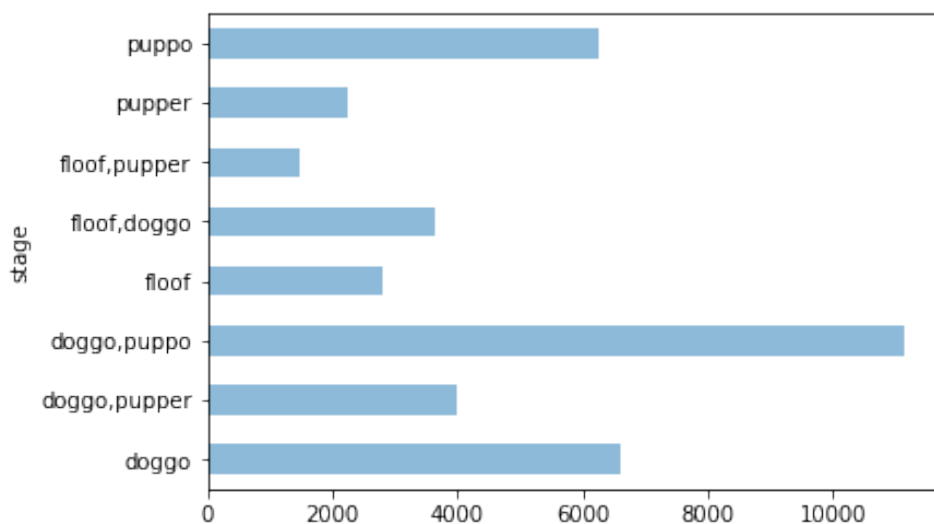


- 在最后的可视化分析中，按照 stage（狗狗的地位）这一列进行分组，分别求出 floofer, doggo, pupper, puppo 的收到的喜欢数的平均值，由作出的横向柱形图可以看出 doggo, puppo 两种最受人喜爱，最后是 floof, pupper。



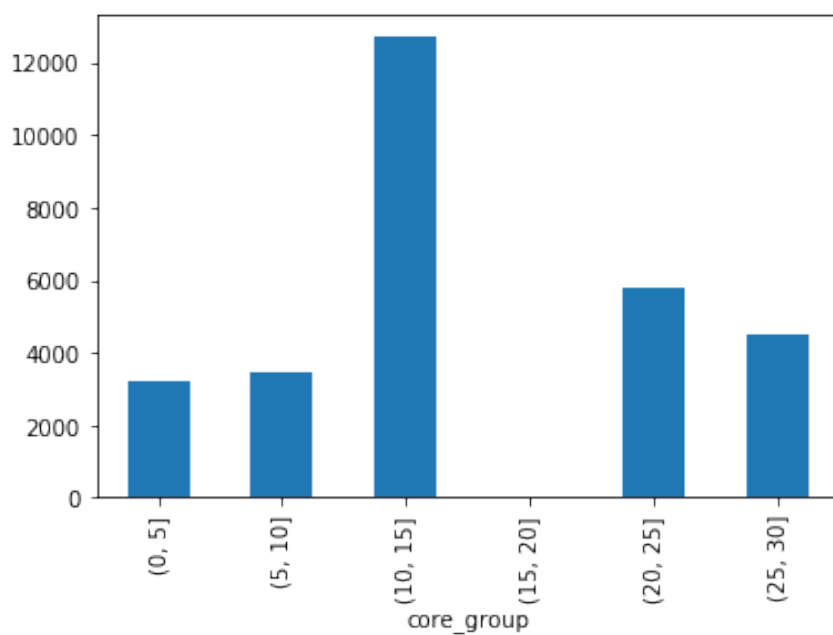
- 按照 stage（狗狗的地位）这一列进行分组，分别求出 floofer, doggo, pupper, puppo 的转发数的平均值，由作出的横向柱形图可以看出 doggo, puppo 的转发数最高，最后是 floof, pupper。



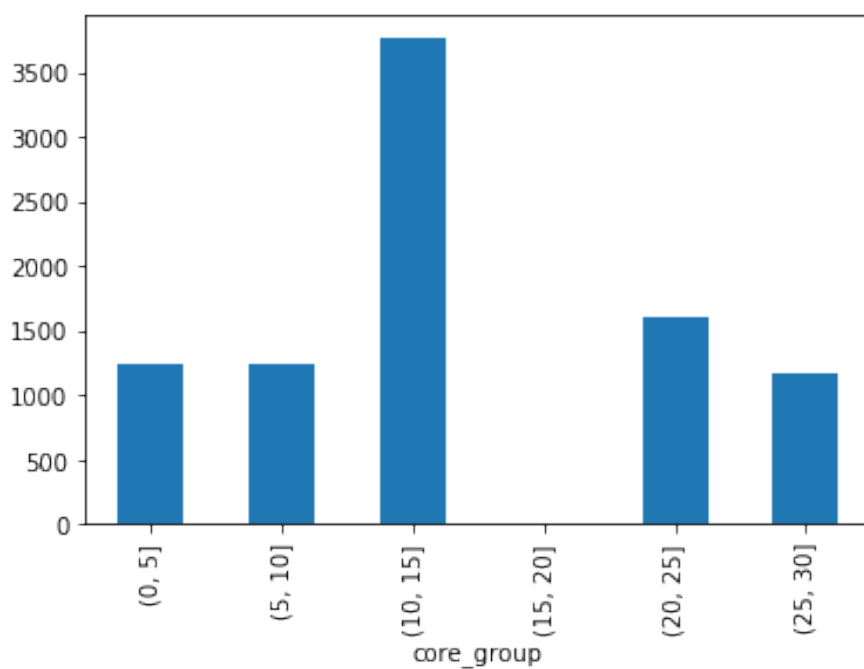
由于 stage 列的缺失值很多，所以数据的说服力很弱。

- 将评分的分子 rating\_numerator 按照数值进行分组，0~5, 6~10, 11~15, 15~20, 25~30 共五组，然后求得每一组的受喜爱数和转发数的平均值，可以发现，评分为 10~15 这一组的平均受喜爱数和平均转发数都是最高的，然而评分比较高的几组和评分比较低的几组的平均受喜爱数和平均转发数相差不是很大，说明评分特别高的可能存在数据录入误差，

我们期望的评分越高，平均受喜爱数和平均转发数就越大，然而事实不是这样。



【dog favorite】



【dog retweet】