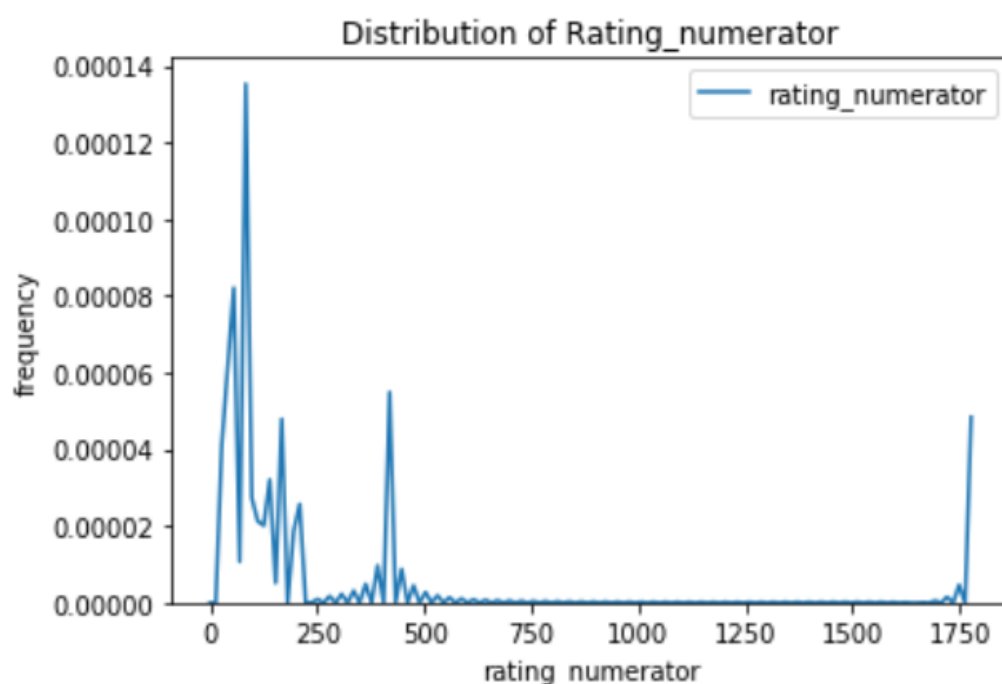


Act Report

本脚本程序通过诸如数据收集，数据评估，数据清理一系列步骤，最终得到了我们想要的整洁的数据，这份整洁且干净的数据将会是之后进一步分析的输入。

由于本项目侧重在收集评估清理这三部分，只是进行了简单的三四个分析。比如我们知道了评级为'pupper'的狗的数量最多，也知道了在用户发的一条 twitter 中，往往第一张图是狗的大头贴，因为其具有较高的置信度，并且还能得出‘金毛是全世界最受欢迎的狗’这样的结论。除此之外，还绘制了评分的分布图。因为分母都是统一的 10 分，我们只考虑分子。做图如下



，可以看到大部分的评分都小于 500 分，但有一少量极高的评分，考虑这是因为评分者对于某几张图片中的狗极度喜爱所致。