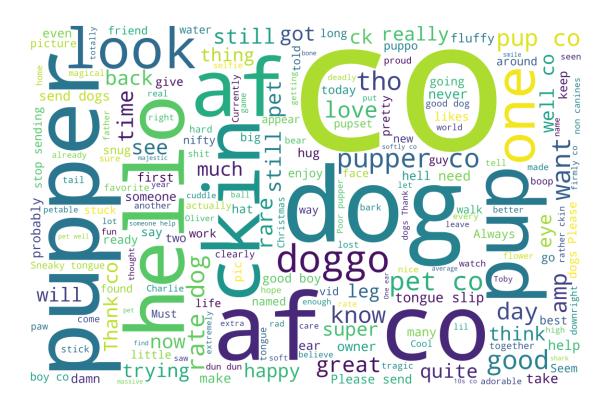
在本项目中,为实践和展示数据处理技巧,数据集并没有全部处理干净。还可以找出很多问题来,也能找出一些问题来进行分析。这就取两个可视化是图形结果与大家分享。

1. 图一是针对 WeRateDogs 推特档案中的 text 生成的词云图,通过这个词云图可以一目了然的看到出现频率最高的关键词。从当前图看出,狗狗的地位出现最多的是 doggo、pupper,对狗狗描述最多的关键词是 hello、dog、look。



(图一)

2. 图 二 是 对 参 加 图 片 预 测 的 狗 狗 进 行 分 类 , 运 用 了 df2_clean.p1.value_counts().plot(kind='bar',figsize=(20,6)),绘画条形图,明显看出参与图片预测狗狗什么种类最多。可看出,品种数量最多分别为

golden_retriever(金毛)、Labrador_retriever(拉布拉多)、Pembroke(彭布罗克柯基犬)、Chihuahua(吉娃娃)、pug(哈巴狗)

