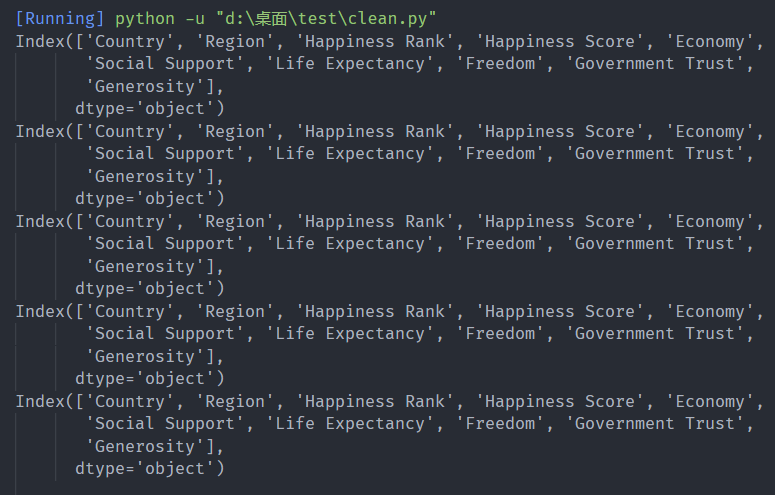
《世界幸福报告》是对全球幸福状况的一次里程碑式的调查。第一份报告于2012年发布，第二份于2013年发布，第三份于2015年发布，第四份报告于2016年更新。3月20日，在联合国举行的庆祝国际幸福日的活动上，根据155个国家的幸福水平，发布了《2017年世界幸福》。随着各国政府、组织和民间社会越来越多地使用幸福指数来指导其决策，该报告继续得到全球的认可。经济学、心理学、调查分析、国家统计、卫生、公共政策等领域的权威专家描述了如何有效地利用幸福感来评估国家的进步。这些报告回顾了当今世界的幸福状况，并展示了新的幸福科学如何解释个人和国家在幸福方面的差异。



实现原理：  
首先将五个数据集全部放到一个数组中。事先观察每个数据集的特点，我们发现某些数据集多出了一些列信息，比如2015年的数据集多出了标准差，反乌托邦残余两列信息。  
为了能够使所有数据集拥有相同的数据集列属性，决定对这些多出的不具备比较分析意义的属性进行删除，主要是用到了drop方法。  
drop方法接受三个参数，  
第一个参数指定的是要删除的属性，它是由一个数组构成的。  
第二个参数设置为1的时候即表明需要删除的是列属性，因为drop方法默认删除的是行属性，需要指定axis=1即可。  
第三个参数指定是否修改原数剧集，默认是false，我们这里不做任何修改即可。

同时我们还需要对列属性进行具体的修改，因为原数剧集的列属性过长，例如Economy (GDP per Capita)。

这里使用的是一个函数。

匹配（），首先我们寻找（出现的索引值，然后用到了切片的思想，使用：匹配（后续的部分字符串，将它们全部替换成空

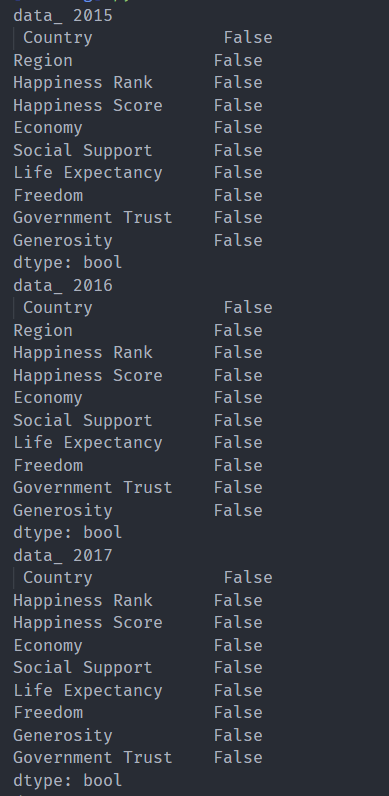
同时因为存在列属性由多个单词构成，我们需要将他们其中的空格替换成\_以便于更好的阅读

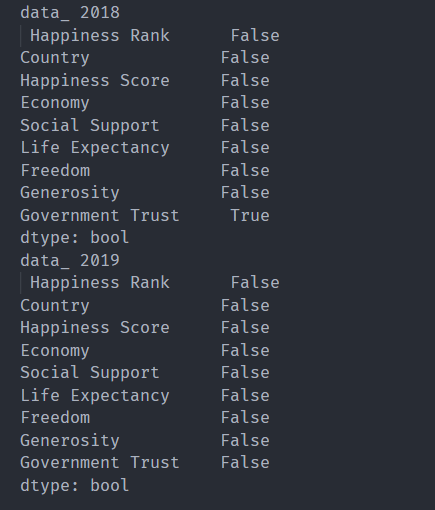
空值处理：

循环遍历data中每项数据集，通过isnull方法可以判断是否为空，而当我们加上了any后显示的结果会更改为判断哪些项有空值。

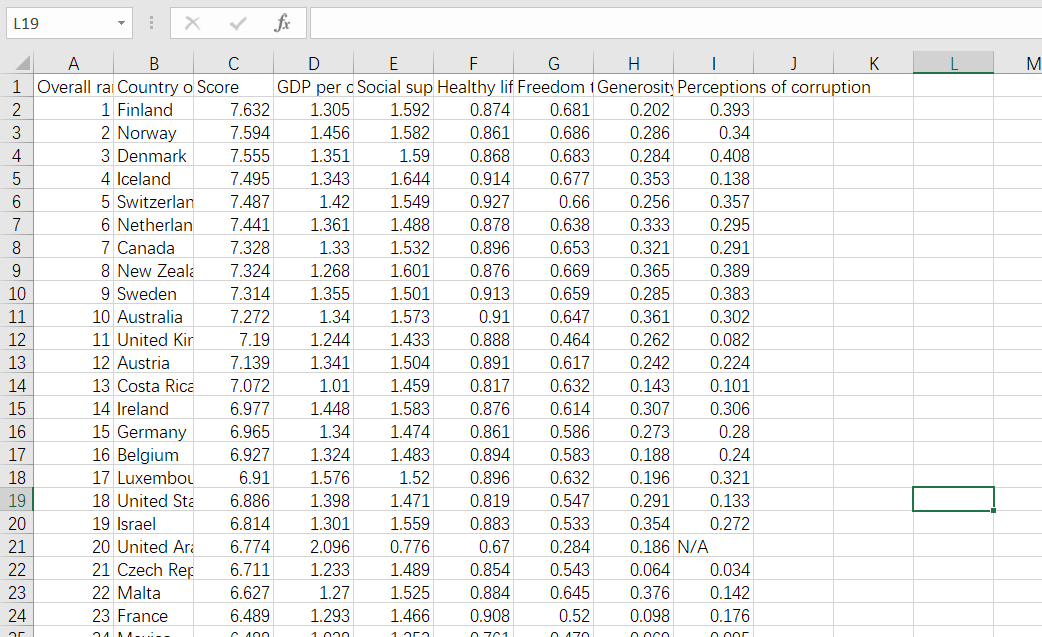
[print('data\_', i+2015, '\n', data[i].isnull().any())

 for i in range(len(list(data)))]





我们可以通过输出结果准确定位到2018年的数据集中Government Trust(Perceptions of corruption)列下存在空属性，直接查看源数据集，可以查看到确实存在空值。



空值的处理：

因为这个空值的存在，导致阿拉伯联合酋长国关于政府自信度的值得到了空值，这里我们可以通过遍历历年来阿拉伯联合酋长国此属性的值来求证。



最后选择了将2018年的值设置为2017年和2019年的平均值，也就是0.253245。