

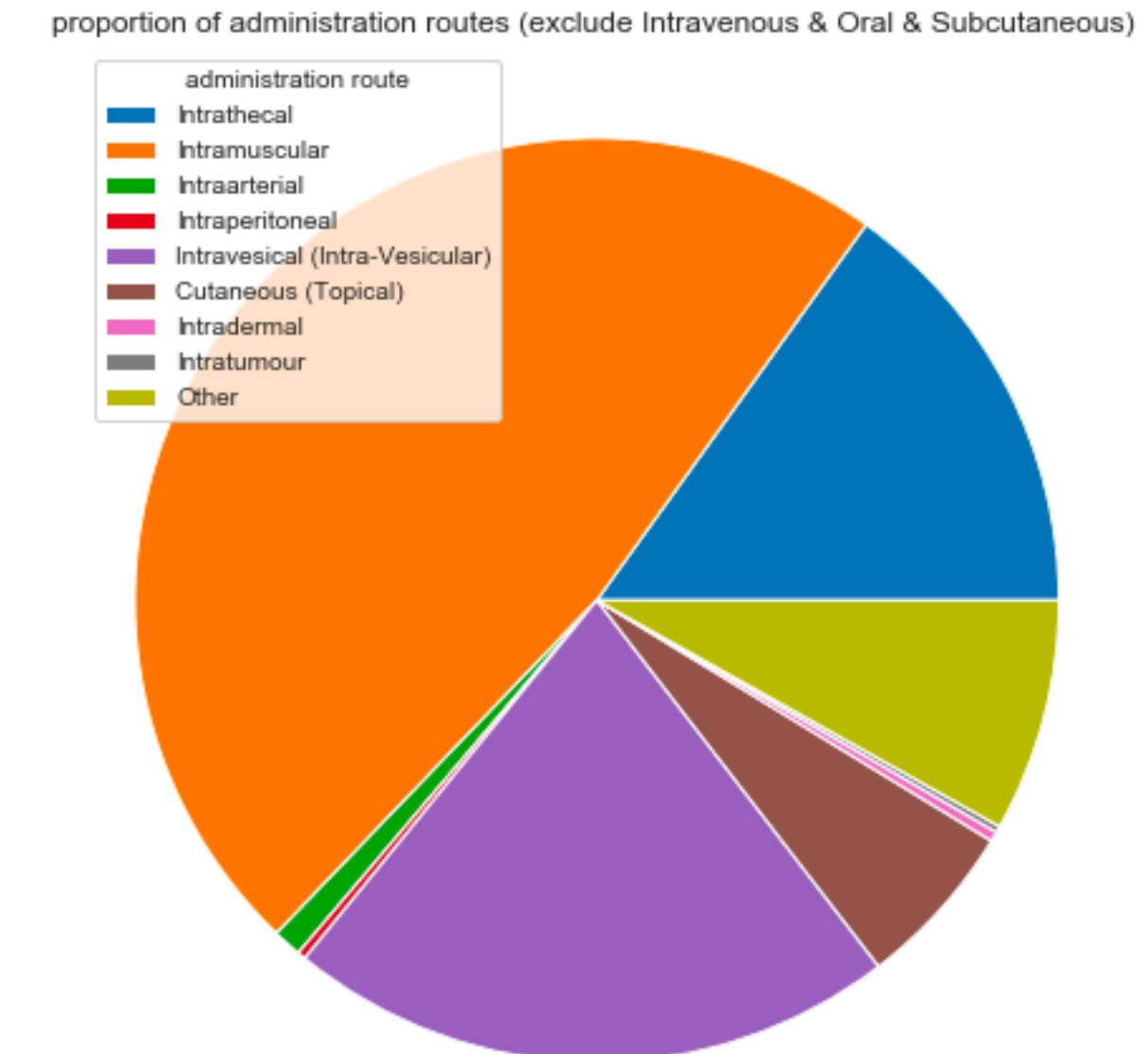
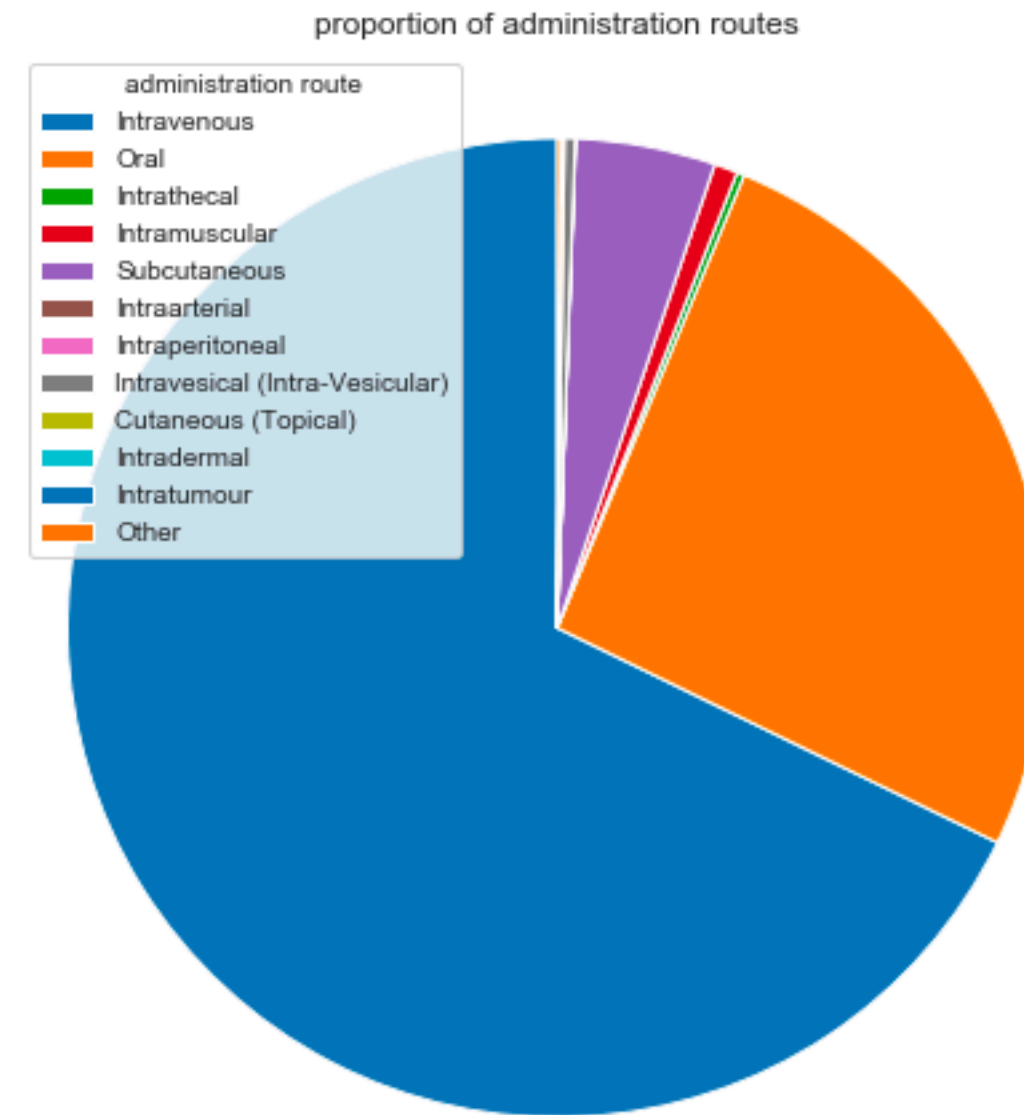
数据分析1

数据清理

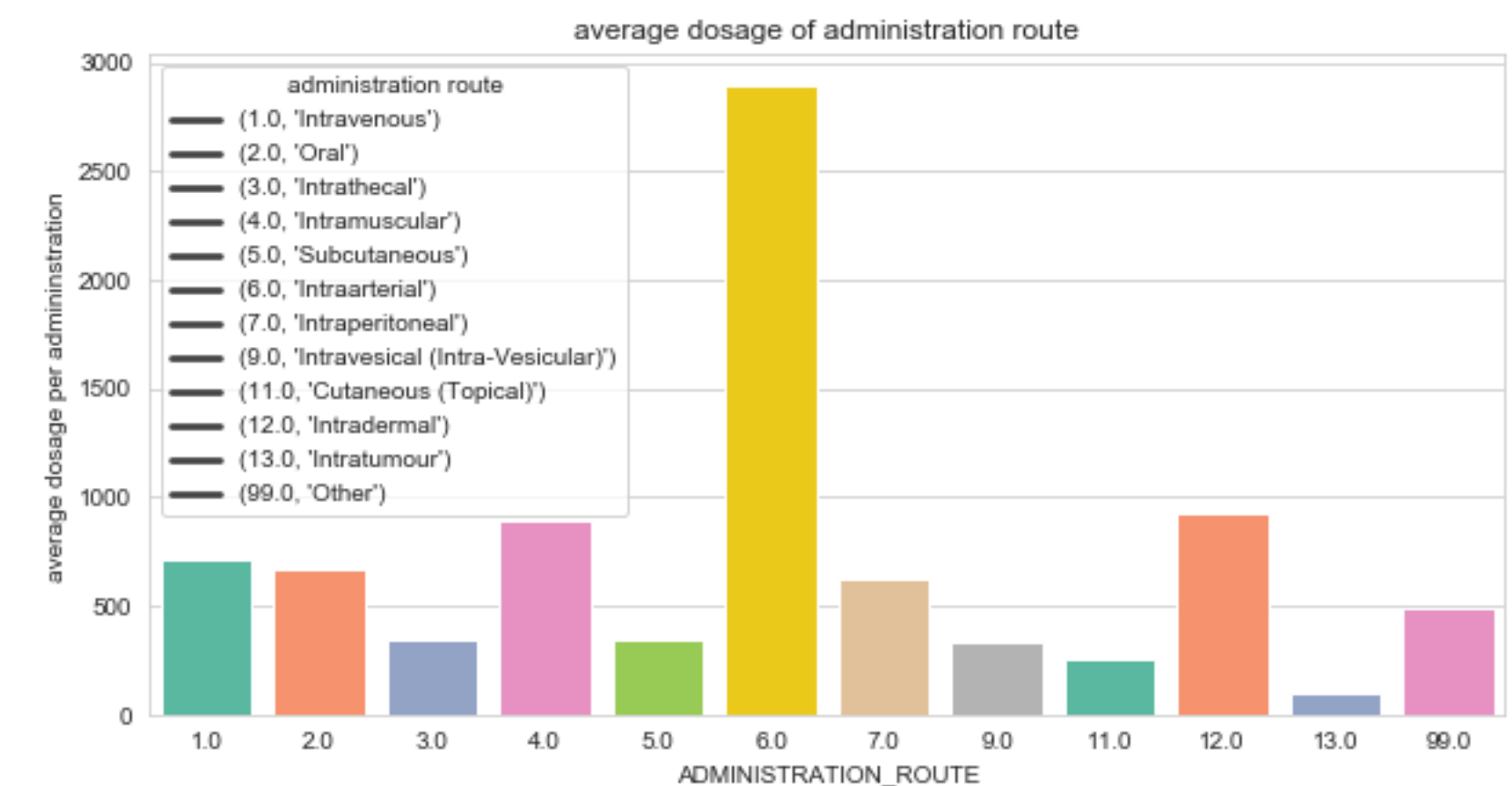
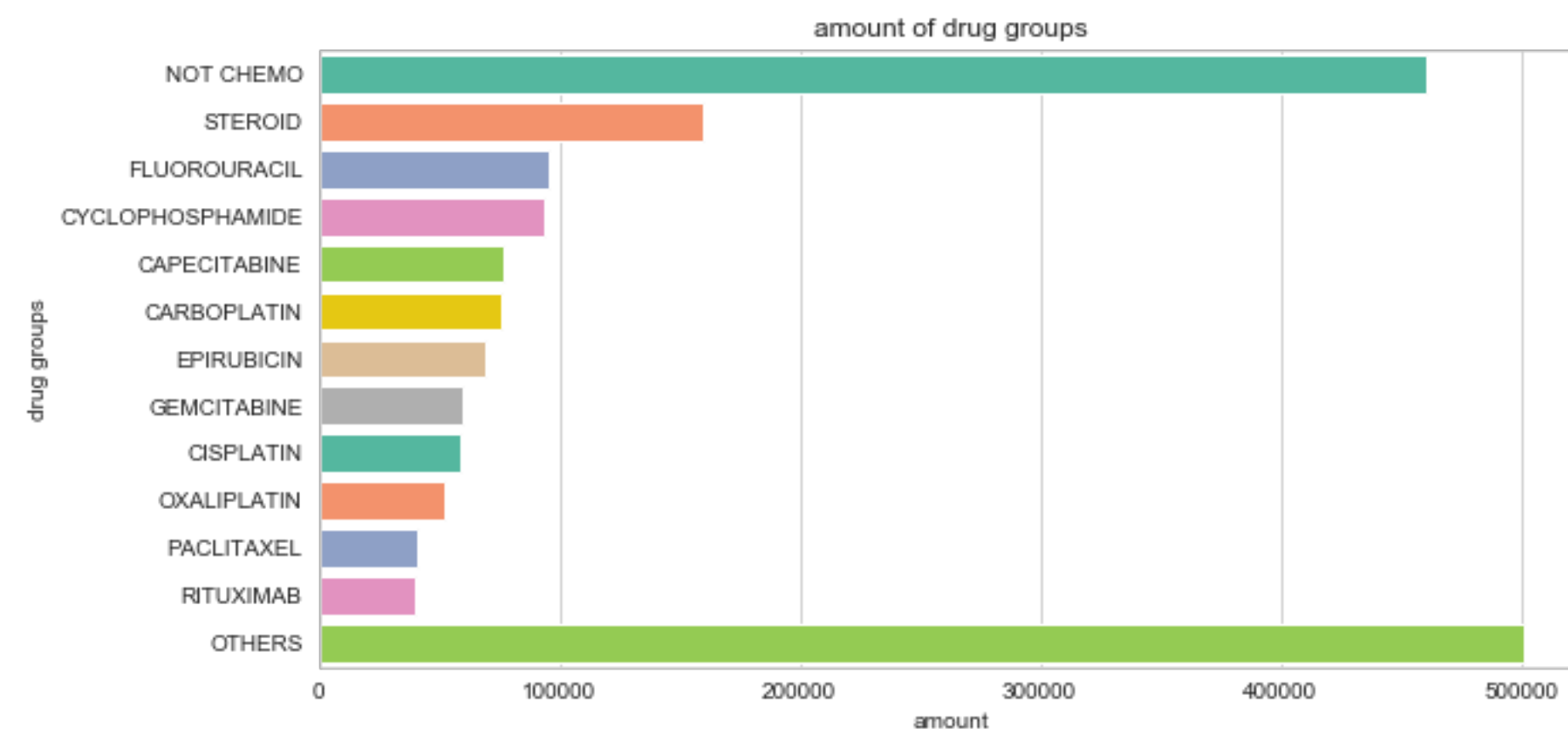
- 分析数据集大小：分别为2,000,000和3,000,000条数据
- 清除无效数据、错误数据及空值，规范化数据等

简单数据分析

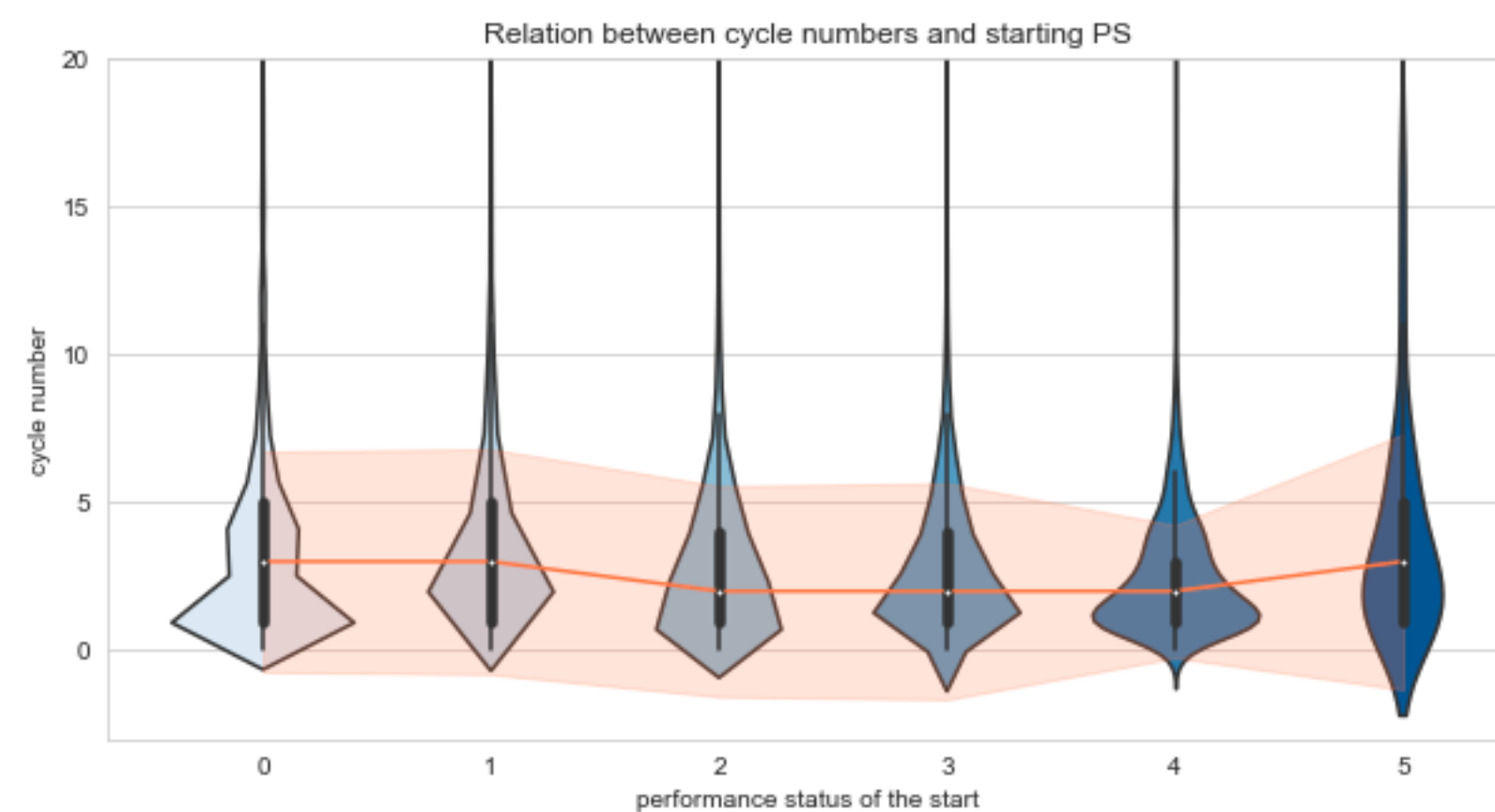
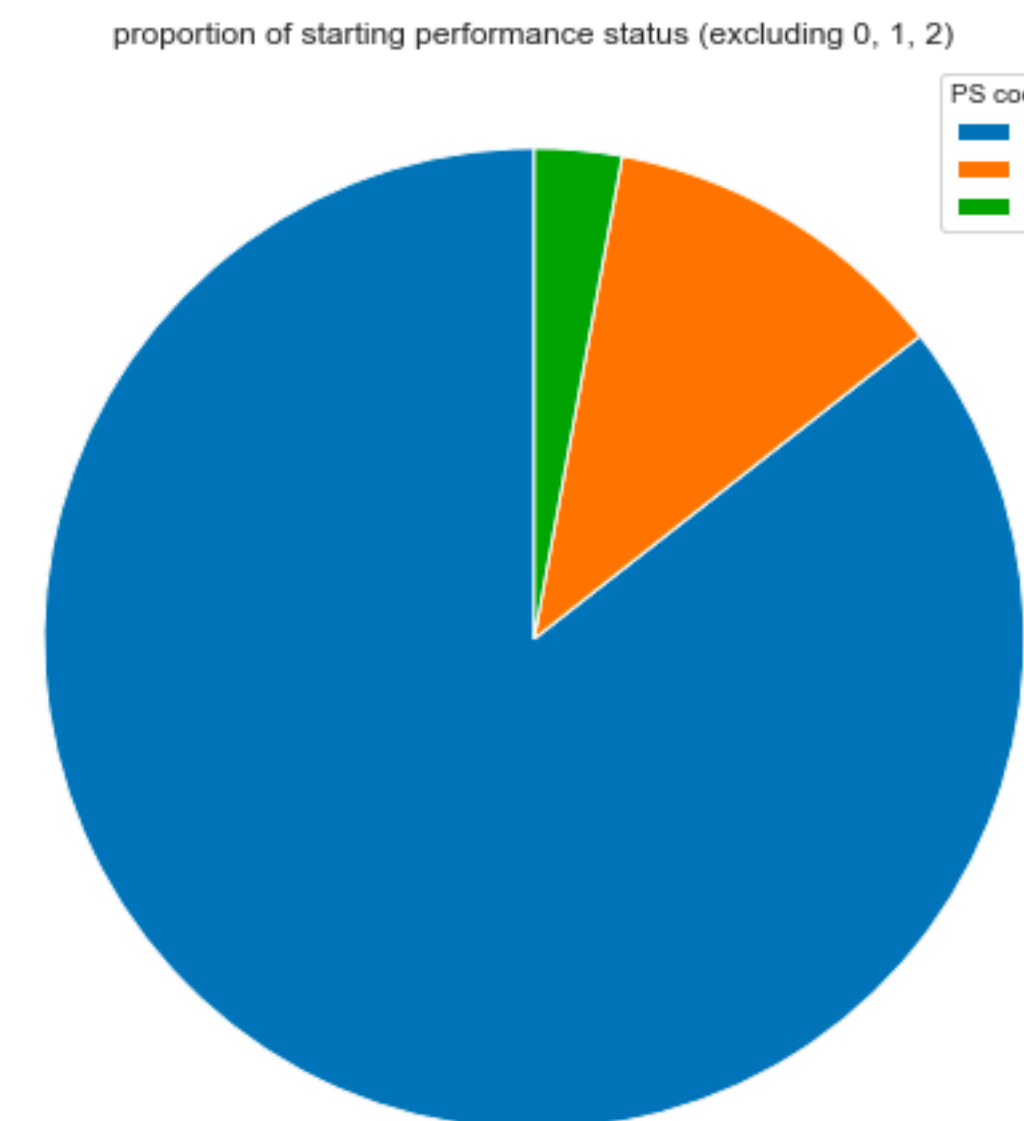
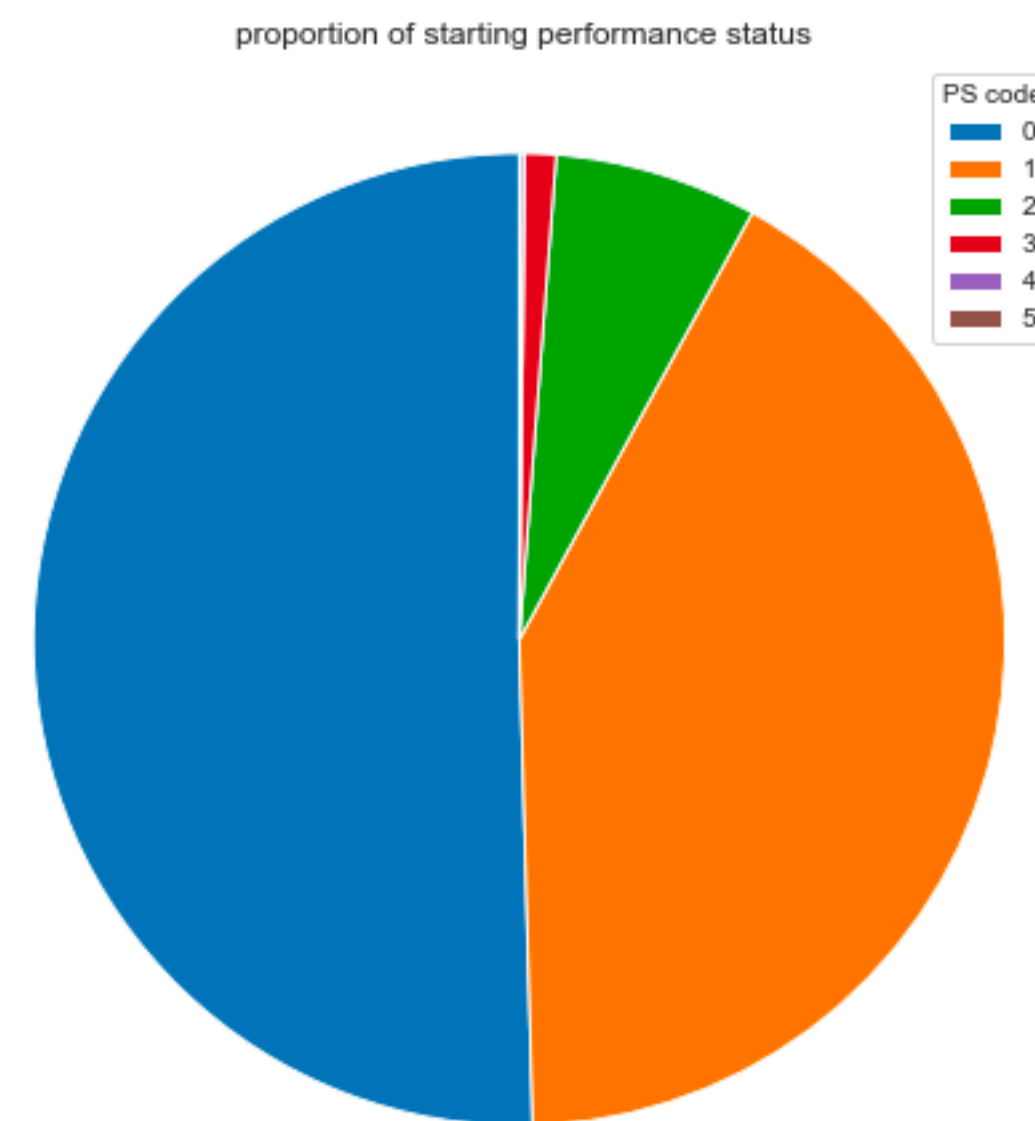
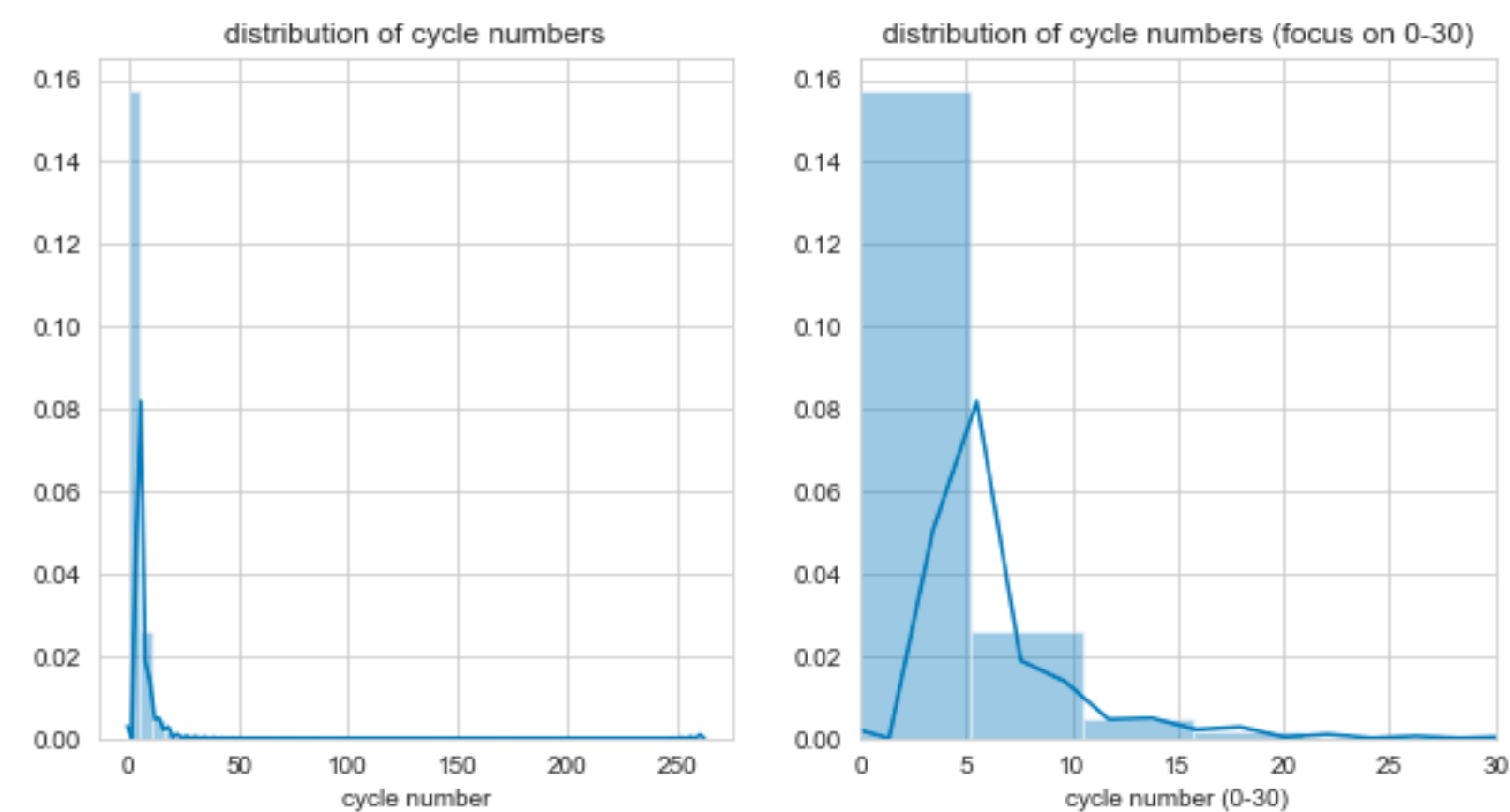
进行最值及高频值分析：
每次给药的剂量值在0-1801201之间



数据可视化2



*此页仅列出我负责分析及可视化的数据



结果分析及假设

- 疗程结束后，患者的表现状态会比开始时更差。经过一定的周期后，可能会出现较差的表现，给患者的身体带来负担。
- 根据疗程数分布的柱状图，通常情况下，最佳循环次数是5次左右（大多数患者接受了5个疗程）。
- 由于平均给药路径的柱状图显示内材料的值最高，动脉给药是需要大量剂量的给药途径。
- 根据给药途径比例的饼状图，静脉注射占比最大，猜测静脉注射是治疗癌症效果最好的一种给药途径。

*此页仅列出我负责分析及可视化的数据