**物联网技术与大数据导论实验2上机要求**

**实验2根据给定的AXP\_NAN文件进行相应的操作，可参考视频**

1. 将AXP\_NAN数据读入，并赋值给变量pfnan
2. 利用isnull()函数判定pfnan是否包含缺失值。
3. 利用dropna()函数删除包含缺失值的行。
4. 再次将AXP\_NAN数据读入，并赋值给变量pfnan，然后利用SimpleImputer()函数填充数据，填充数据为中位数值(median()函数)。
5. 利用decribe()函数观测数据。
6. 利用loc()函数将pfnan的最后一条数据改为[3,200,1,2,3,4],并利用iloc选择前4列数据显示。
7. 利用drop(‘Volume’，axis=1)和boxplot显示箱图。
8. 利用3σ方法判定缺失值。公式如下:数据与平均值(mean())差的绝对值大于数据标准差(std())的3倍。

**实验3 根据数据集boston()完成以下操作,可参考**

1. 调用sklearn模块下的datasets数据集。
2. 将boston()数据库赋值给bo。(使用datasets.load\_boston)
3. 利用bo的4,5,6这3列构造dataframe类型数据df，并利用feature\_names()函数将bo的列名称给df。
4. 利用公式(df-df.min)/(df.max-df.min)来对df数据进行规范化,并赋值给df2。
5. 调用sklearn的preprocessing模块，并利用其中的minmax\_scale函数对df进行规范化，并赋值给df3。
6. 利用preprocessing模块中的scale()函数对df数据进行规范化，并赋值给df4。