叶、林-数据分析工作

【34家电厂-42个数据集】

1、数据缺失情况统计（日期纬度）【列表格说明-说明哪几个数据集是存在日期缺失的】【需要全部数据集的统计情况】

1-1 行缺失数据补齐方法【需要找一下】

如果缺失月级别的用什么方法（需要找一下）

如果缺失天级别的用什么方法（平均值相关方法）

2、数据列维度转化检验（根据转换公式图）【煤质数据自洽性检验】【5个数据集即可】

直接检验【收到基】列就行

收到基 灰分 Aar(%)

收到基 挥发分 Var(%)

收到基固定碳 FCar(%)

收到基全硫St,ar%

相当于去通过公式去计算一列新的收到基，（新-旧）/旧【真实数据和换算数据差】

3、基于均值的数据异常检测 【5个数据集即可】

收到基低位发热量 Qnet,ar(MJ/kg)

全水分/收到基水分 Mt/Mar (%)

收到基固定碳 FCar(%)

4、基于分布的数据异常检测 【5个数据集即可】

收到基低位发热量 Qnet,ar(MJ/kg)

全水分/收到基水分 Mt/Mar (%)

收到基固定碳 FCar(%)

5、基于模型的数据异常检测 【5个数据集即可】

收到基低位发热量 Qnet,ar(MJ/kg)

全水分/收到基水分 Mt/Mar (%)

收到基固定碳 FCar(%)

主要检测的检测异常的是

发热量、煤量【这个可待讨论】

如果是基于模型的数据检验

【X】

全水分/收到基水分 Mt/Mar (%)

收到基灰分 Aar(%)

收到基挥发分 Var(%)

收到基固定碳 FCar(%)

收到基全硫St,ar%

【Y】

收到基低位发热量 Qnet,ar(MJ/kg)

PPT-P14—P28作为参考