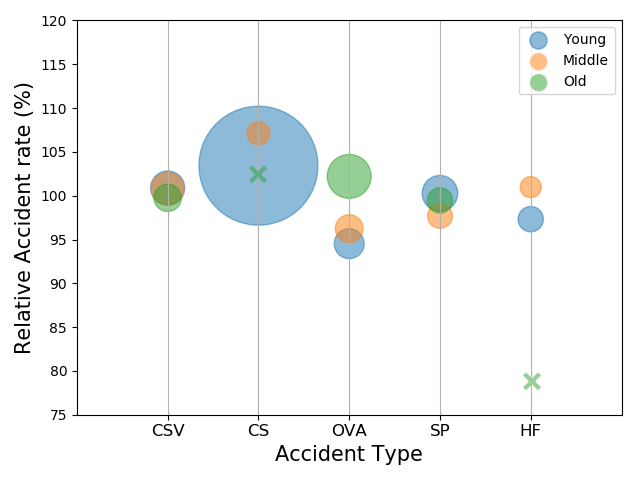
老师好，下面分别介绍【散点图】、【折线图】和【调和曲线】的结果：

1. **散点图**

之前的数据星期那一栏有些问题，所以重新整合了数据，散点图看起来和上次有点不同，但是散点图的画法是相同的

* 纵坐标：雨天事故率/晴天事故率
* 圆的大小：雨天死亡率/晴天死亡率
* 交叉：代表相对死亡率为0（若NAN，也当成0）

****

1. **折线图**

按照上次讨论的方法划分样本，每周作为一个样本

（1）先分别求出【晴天平均每天死亡人数】和【雨天平均每天死亡人数】

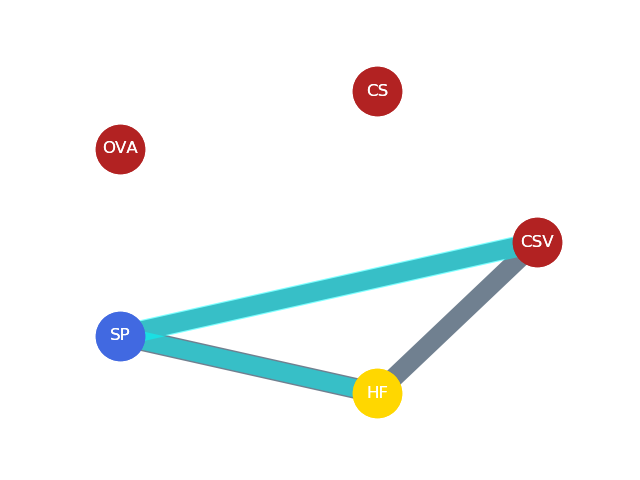
（2）分别计算晴天和雨天三种年龄段各自的显著性和相关性

（3）分别绘制晴天和雨天不同年龄段的折线图（灰色代表晴天，蓝色代表雨天）

（4）将相同年龄组的晴天和雨天两张图叠加在一起

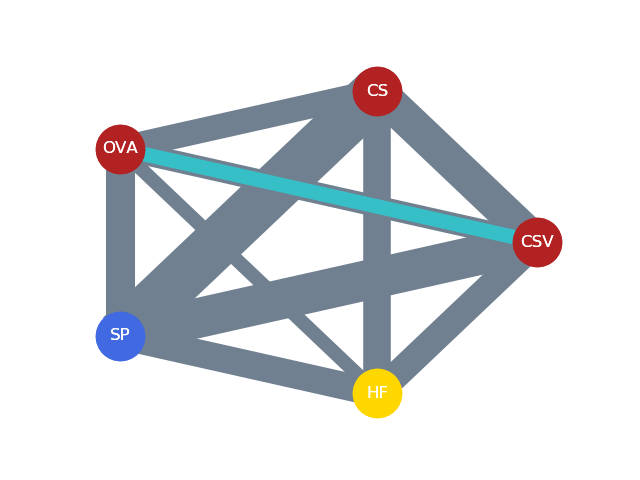
【图中的折线均为正相关】

1. **青年组：**



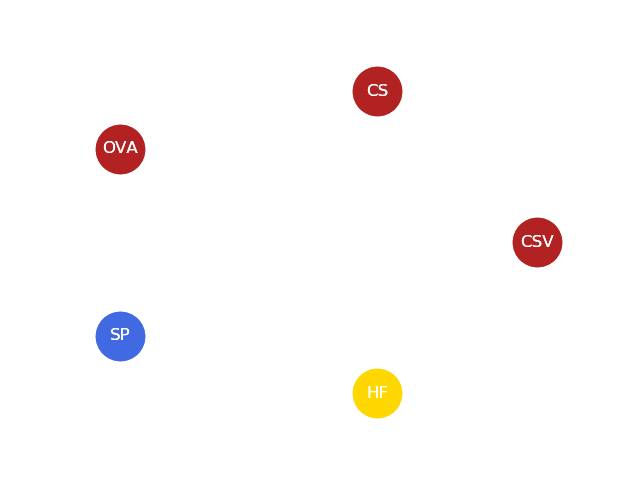
**SP和HF之间也有灰色的折线，但是被蓝色覆盖了，放大后能看到**

1. **中年组：**



**该种情况下雨天只有一对组合有显著性**

1. **老年组：**



**晴天和雨天均无显著性**

1. **调和曲线**

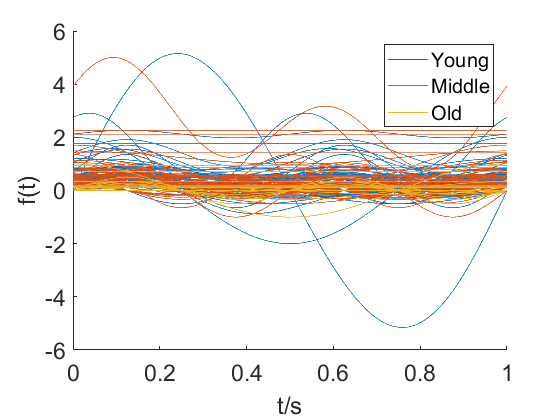
调和曲线依然非常诡异，原因为主要有两个：

* 虽然样本有157个（157周），但是有很多样本的全部特征为都为0（例如雨天老年组，实际上不全为零的样本只有9个，所以老年组的曲线非常稀疏）
* 三个年龄组可能本身差异不大

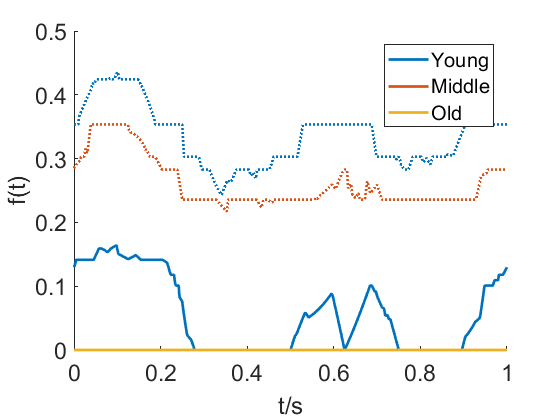
同时发现

* 晴天和雨天的调和曲线差不多
* 数据在归一化前和归一化后，曲线变化不大

1. **不分晴雨天：**

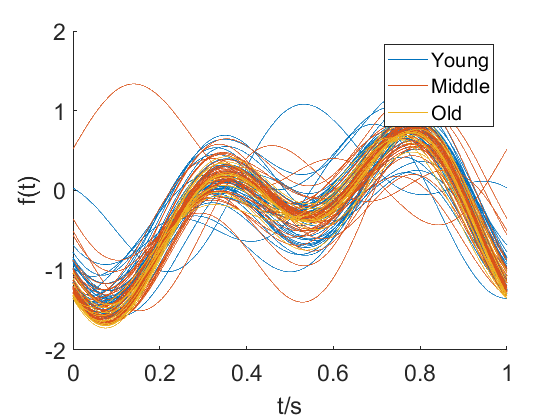


【显示所有样本曲线】

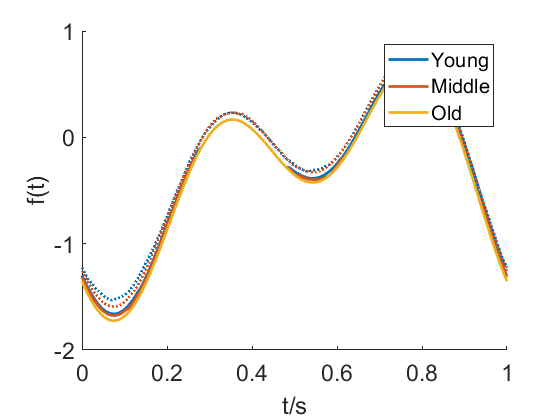


【显示中位数和四分位数】

1. **晴天：**

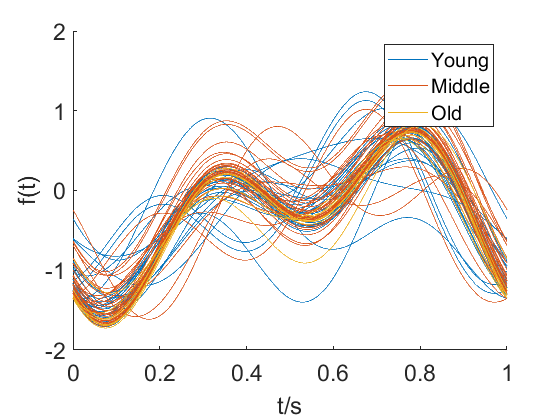


【显示所有样本曲线】

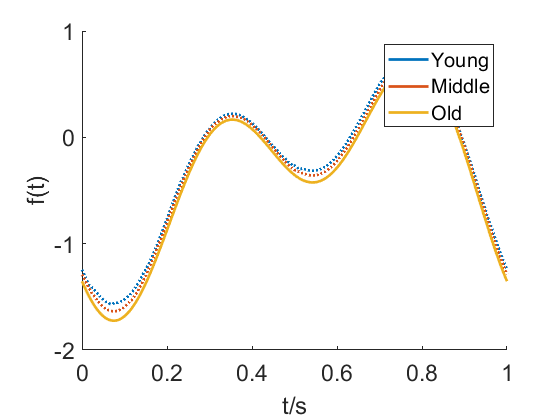


【显示中位数和四分位数】

1. **雨天：**



【显示所有样本曲线】



【显示中位数和四分位数】

1. **问题**现在的主要问题还是虽然样本有了，但是样本的特征很多为全0行，这一点在调和曲线的时候表现的尤为显著，如果将一个月作为一个样本的话，会将星期的数据合并在一起，那这样的话不为零的样本会更少