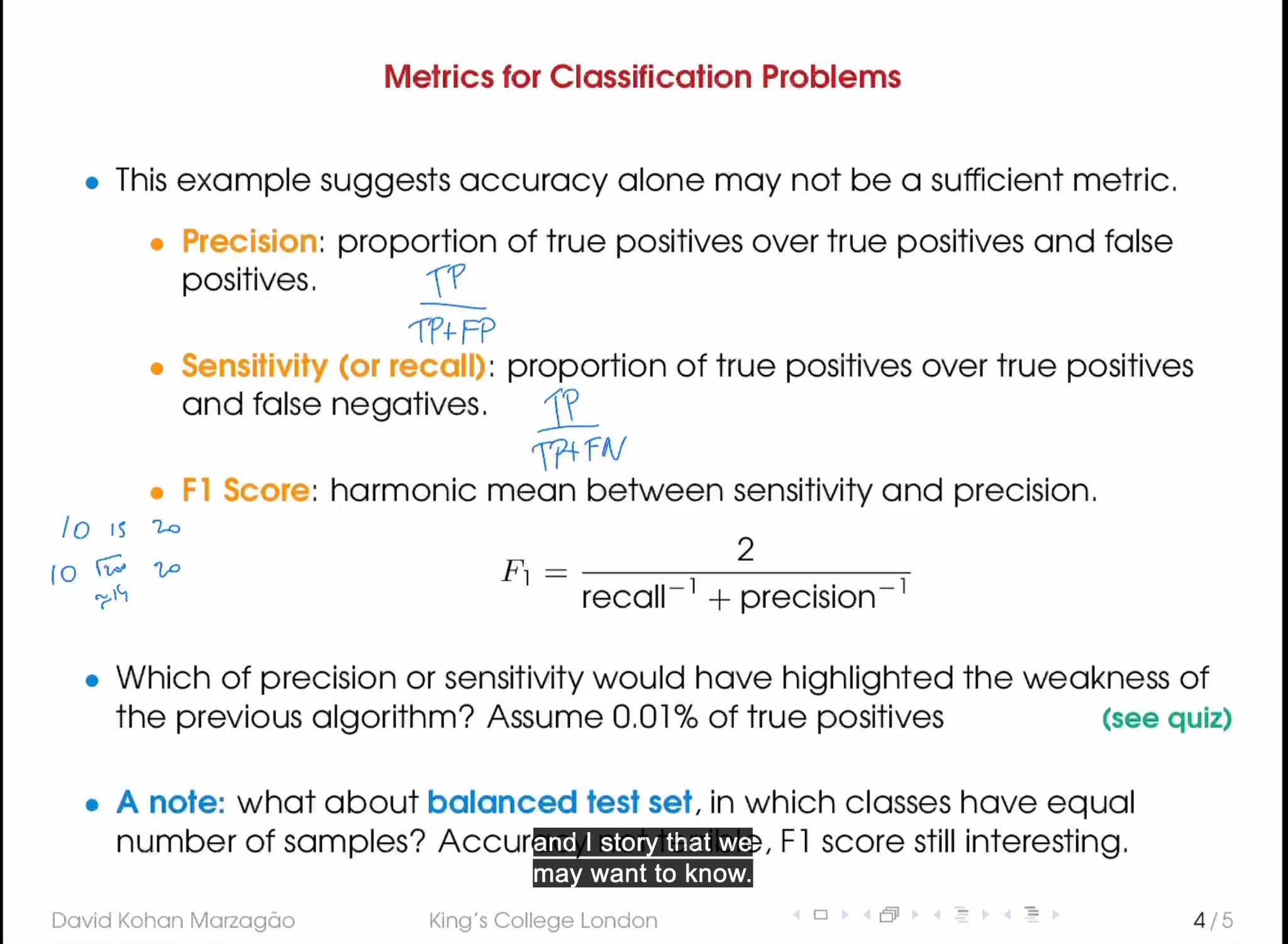


这部分实际上从第三张图可以看到概率是逐渐从1扩展开来的，比如最开始的0.72 = 80%概率去+1，10%概率去+0，10%概率去+0，综合下来就是+0.72。但是这只是第一个interation，到下一个iteration0.72隔壁的格子因为有了0.72这个参照物，他们也能根据自己的概率去计算当前这个格子的回报，以此类推。同时第三次interation的时候。0.72因为多了隔壁的格子的回报，参与到计算中，比如80%概率去+1，10%概率去+0.5，10%概率去+0.4这种，他自己的回报率也会上升，随着iteration的次数变多最终这个数字也会趋于稳定。

第二张途中最开始有些格子之所以是负return也是因为有-0.04的cost，并且随着周围的格子也有cost了，这个cost会随着iteration上升，但是等到一定的interation后，周围的格子突然有了正的回报率，这些离+1远的格子回报也会突然暴涨。

第二张图右图之所以discount factor越小需要的iteration越稳定是因为可能5个格子之后因为discount factor 比如说0.5，0.5的五次方就已经小到可以忽略不计了，所以容易趋于稳定

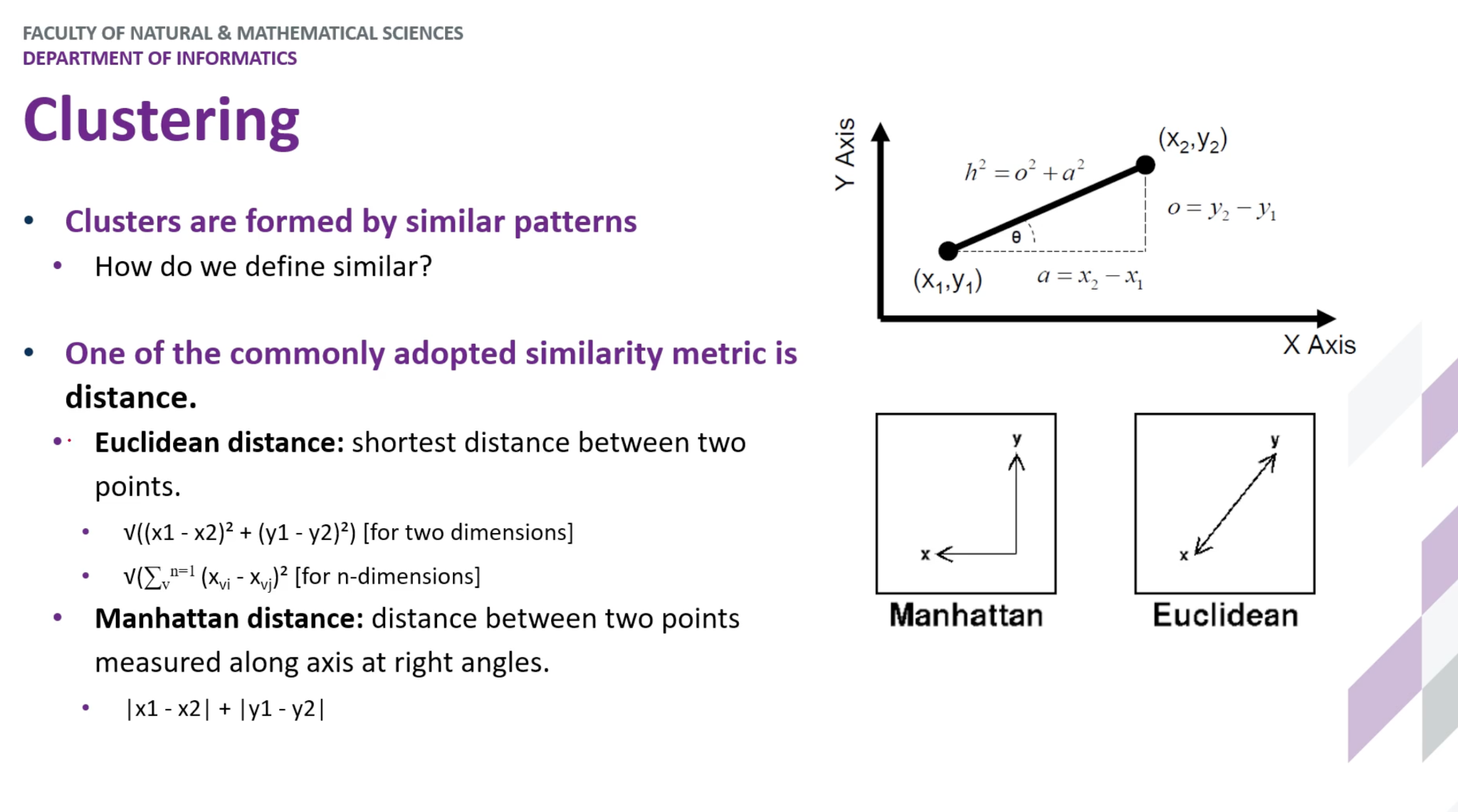


False Positive：比如一个人没病，但测试结果是有病，这就是FP，可以理解为乱说positive，negative说成positive

False Negative: 一个人有病，但是测试结果是没病，就是FN，可以理解为乱说negative，positive说成negative

Precision 用FP是因为比如，一个人没病，你硬说他有病，你这判的就不准了

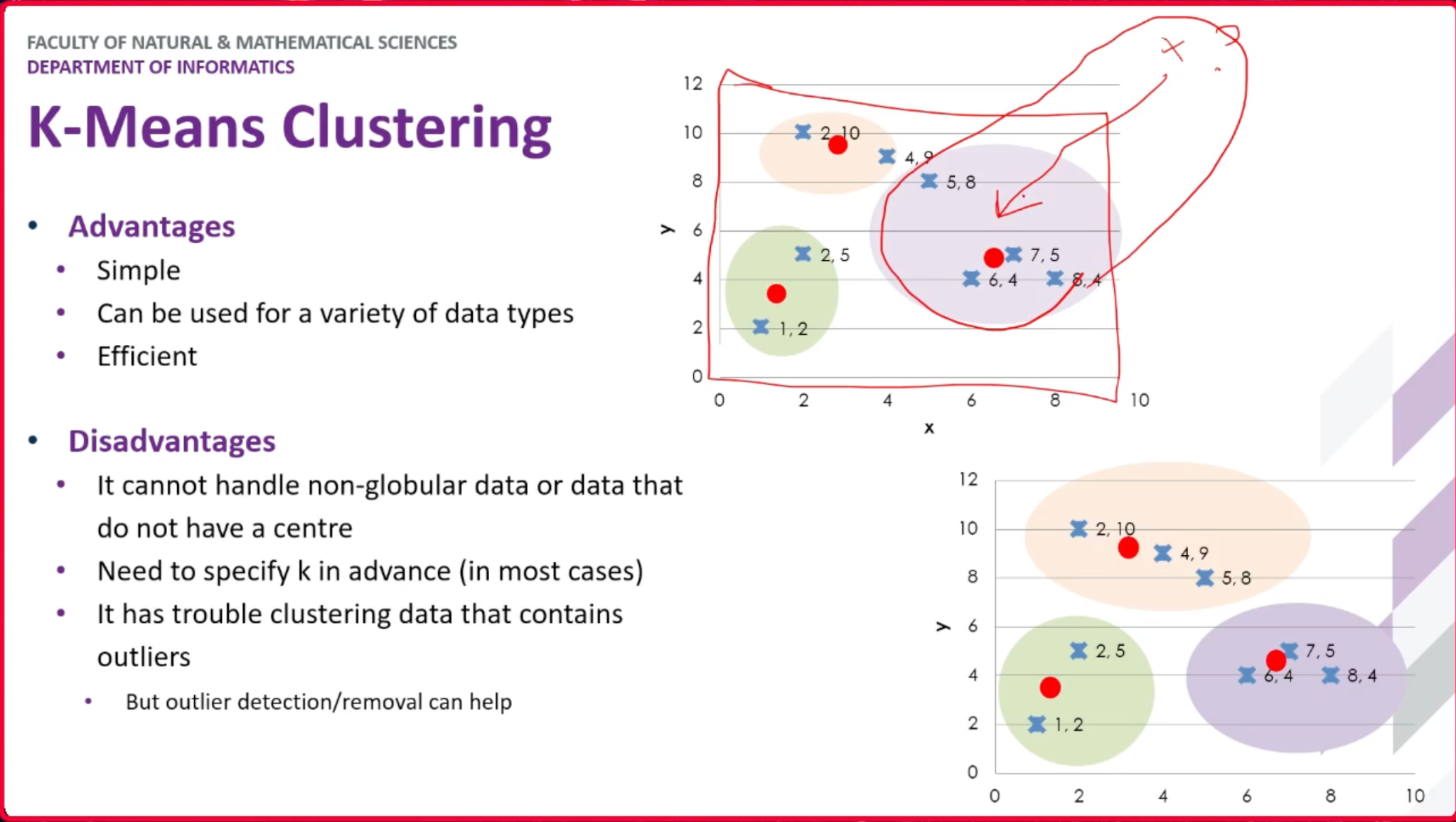
Sensitivity可以这么理解，一个非常敏感的新冠测试，只有你有一点点新冠病毒，他都会显示阳性，所以真的得病的人，一定会被检测出来，但是一个人可能直接呼吸到了一丁点有病毒的空气，但是也被检测出来了，证明敏感度非常高，这就是FN



如果是三维坐标系比如(x1,y1,z1),(x2,y2,z2)则等于**√((x1-x2)^2+(y1-y2)^2+(z1-z2)^2)**

**K means clustering**

Advantage and disadvantage



简单，好用，但是如果出现outlier，如图中所示，就会很大程度影响准确性。如果点很远那他还算这个组的吗？算的话会不会使质心偏离很多