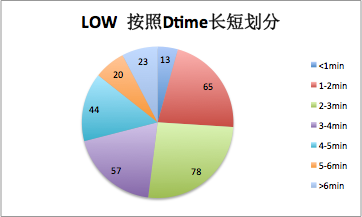
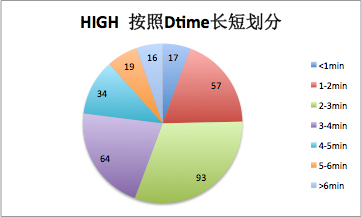
LOW temporal relevance下，根据Dtime长短划分区间，统计Dtime分布情况。

可见Dtime主要集中在1-5min。



HIGH temporal relevance下，根据Dtime长短划分区间，统计Dtime分布情况。

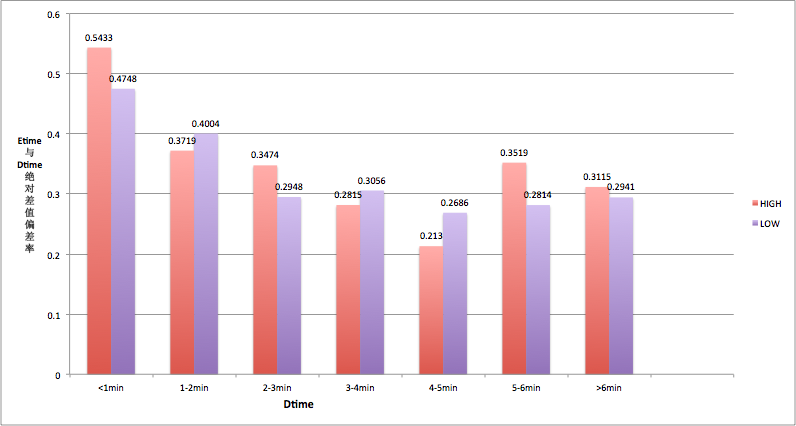
Dtime同样集中在1-5min。



在根据Dtime划分出的每个小区间中，计算Etime与Dtime二者的绝对差值偏差率（即二者差值的绝对值除以Dtime）的平均值。比较不同区间之间，以及HIGH、LOW两种设置下的偏差率。

基本可以看到查询时间越短，偏差率越大。尤其如果我们只关注Dtime主要分布于的1-5min，这个趋势还是比较明显的。

考虑HIGH与LOW两种设置之间的差异，从整体来看特征非常不明显。如果只考虑1-5min，则可以看到基本都是LOW的偏差率比HIGH高，即high temporal relevance下的查询时间估计更准确。



上面的对比是考虑了Etime与Dtime的绝对差值偏差率，也就是不能反映出用户的Etime是估计过高还是过低。如果我们不将两者的差值取绝对值，只计算它们的原始差值偏差率，则绘制出的图像如下图所示。只考虑分布集中的1-5min区间，可以明显看出随着Dtime增长偏差率逐渐减小，即时间越长估计越准确。而且可以看出随着Dtime增长，用户倾向于估计偏短。

但是在每个区间内，都是LOW比HIGH估计得更加准确一些，这和上面绝对差值偏差率比较的结论不尽相同。但是因为这张图的趋势很明显，而且数据处理方式更合理一些，所以还是应该以这里的结论为准。原因应该是HIGH条件下用户更容易估计出更长的时间。

再深入考虑，现在统计的偏差率都是计算的每个区间中的平均值，并不能准确反映整个区间的总体情况。所以这里采用箱体图来进行进一步的数据展示。

分别看HIGH条件下和LOW条件下的原始偏差率，在1-5min区间基本可以看出箱体的整体下移趋势。也就是说上面得出的随着Dtime增加，用户的时间感知越来越准确，这一结论是比较可靠的。

上面得出的第二个结论，LOW比HIGH估计得更加准确，这一点只能在1-4min区间中通过中位数看出来，并不十分显著。

还可以明显看出，LOW在每个区间的数据分布范围都比HIGH的要大很多，这比较容易理解，因为HIGH条件有一定的时间参考，个人估计时间的差异性会被削弱一些。

下面再附上两张在HIGH和LOW条件下，比较绝对差值偏差率的结果图。这两张图的特征并不十分明显，所以还是应该考虑分析原始偏差率更加合理。