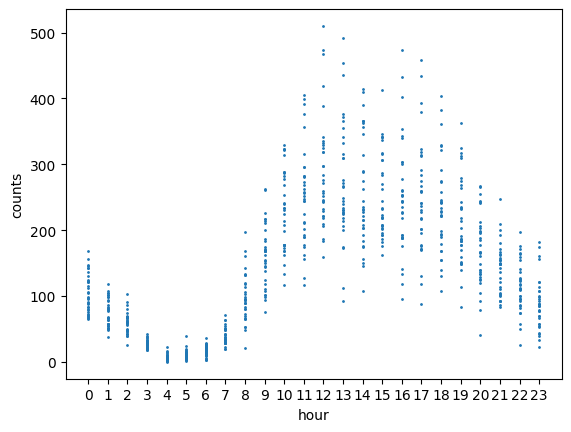
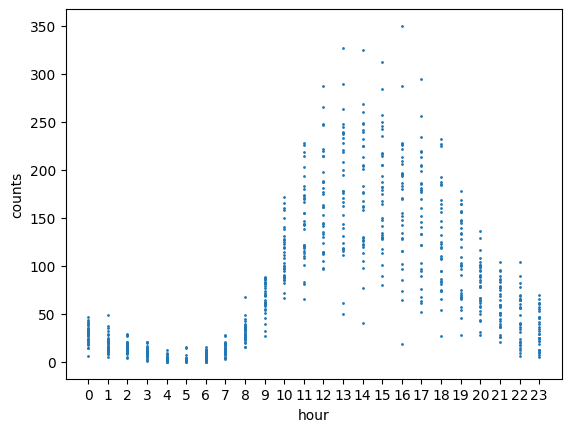
代码对原始的数据进行了准备，包括简单筛选了假期信息 – 将所有周末划入了假期并将所有周1-5的日期调整为工作日。同时，对存在0值的风速进行了矫正，对于每一个0值，将其取为0值前后两个风速的平均值。

由于共享单车使用集中在白天，晚上22点之后直到次日4点间使用人数较少，故将此时间段排除在分析范围之外,而更关注其他时间（5至21点）。

**时间和用户数量的关系**

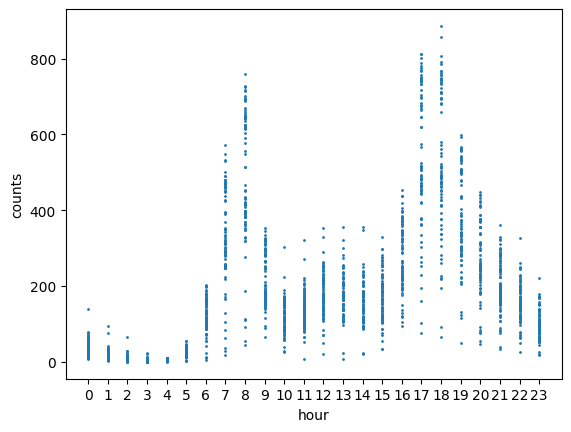
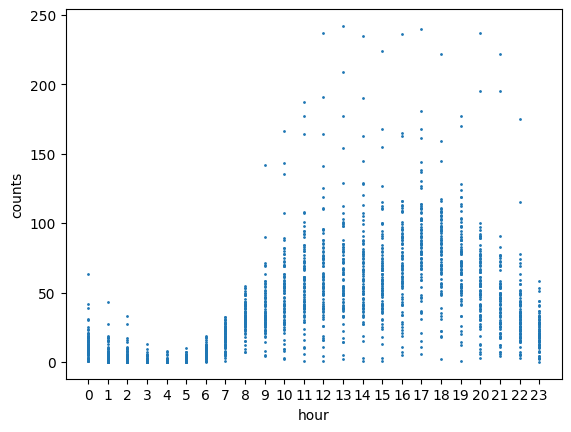
可以观察到，在假期时间，用车人群集中在下午12-18点之间

Casual, Holiday Registered, Holiday



工作日则存在明显区别。注册用户使用集中在8点周边和17，18点，对应上下班高峰，可见注册用户多为有固定上班需求。未注册用户则与假期分布保持一致，可见未注册用户多数为非上班族。

Casual, Workday Registered, workday

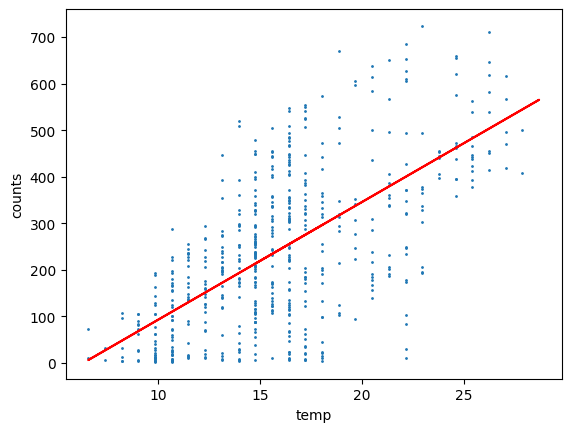
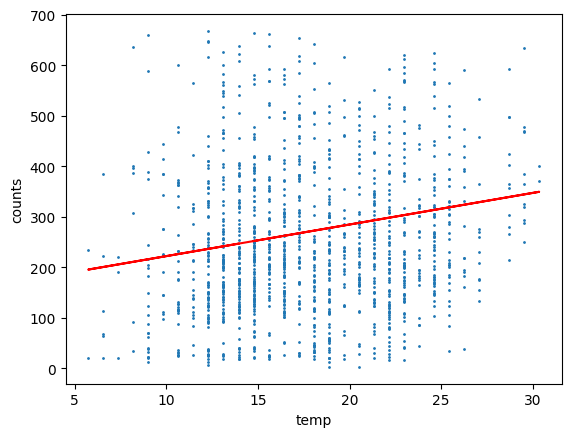


**温度，体感温度对于用户数量的影响**

由于同年月份不同会带来温度的较大差距，在每次分析时仅选择一个季节（Season），即三个月为范围进行分析。

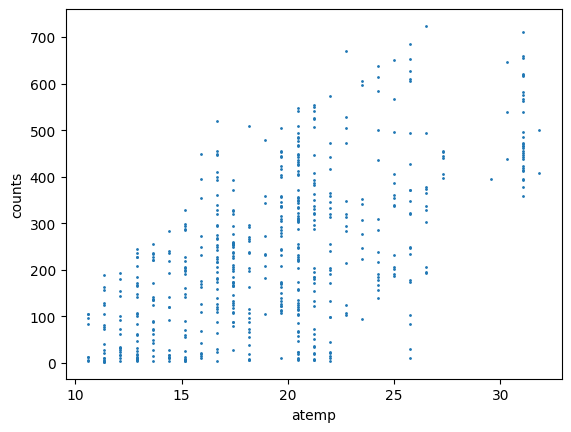
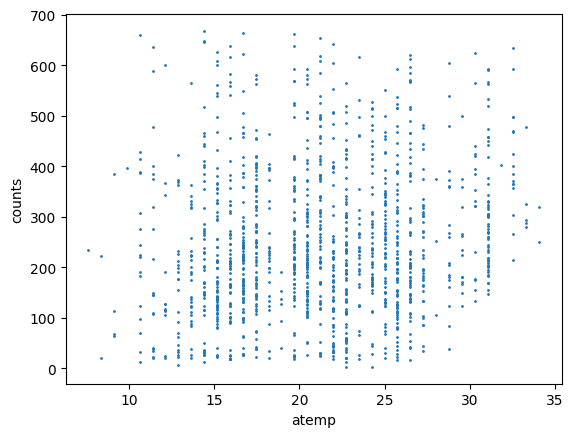
此处选择温差较大的Season4(10-12月) 为样例进行分析。

Workday Holiday



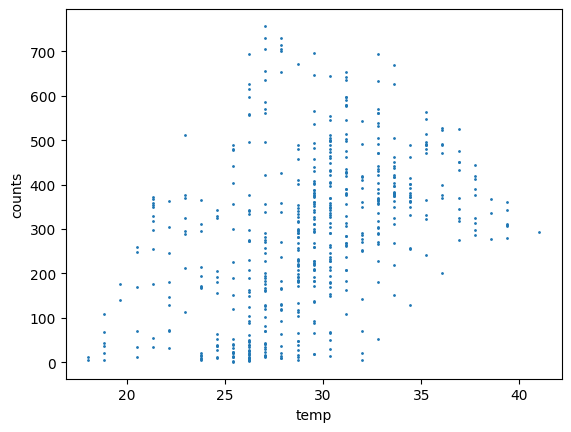
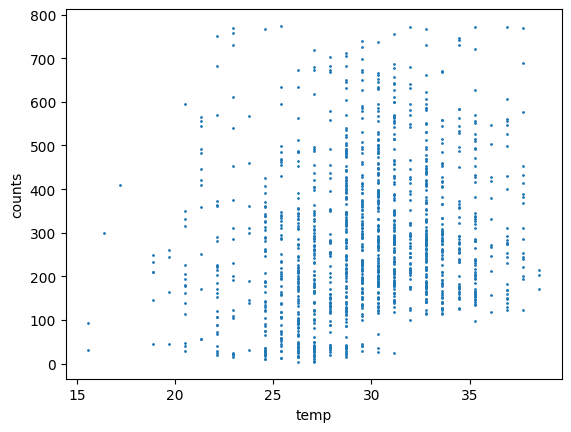
在工作日，气温与共享单车的使用情况没有明显的线性关系。由于上班的需求相对固定，气温对人们的骑行选择没有显著的影响。在非工作日，气温的影响则更加明显。随着气温的上升，更多人会在外出时选择骑行。用户数量与体感温度的图像同样符合上述关系。

Workday Holiday



不过，当气温进一步上升时，例如在Season 3(7,8,9)这三个较热的月份，情况会与较冷的月份相比不同。

Season 3, Holiday Season 3, workday

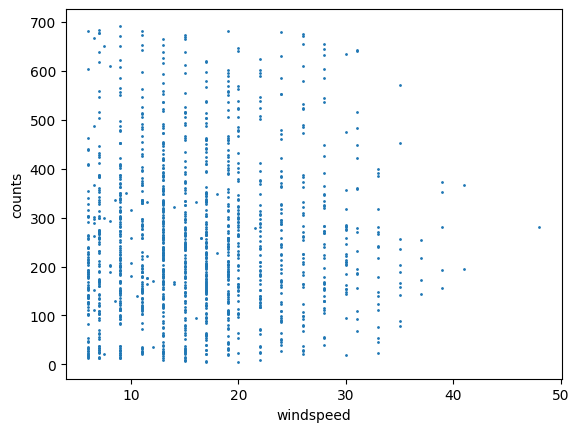
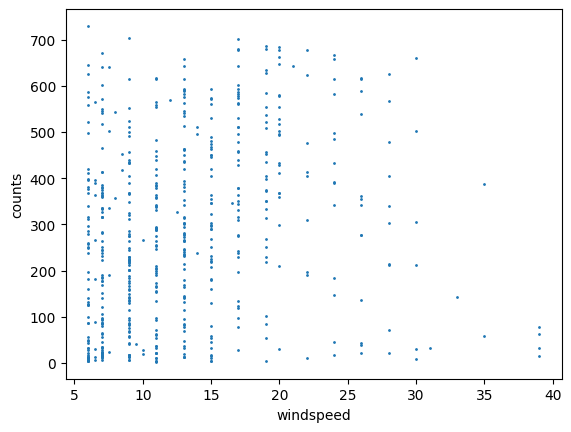
 

可以看到，随着气温进一步上升（>35°C），非工作日的单车用户数量出现了下降。工作日的单车用户数量变化不明显，不过依然可以看出，当气温大于35°时，用户数量大于400的分布较30°-35°有所减少。

**风速与天气类型**

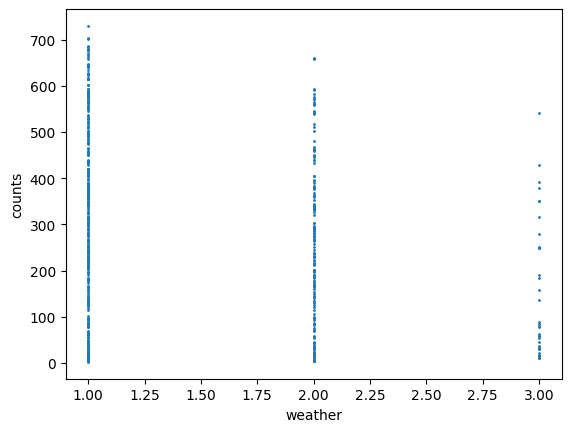
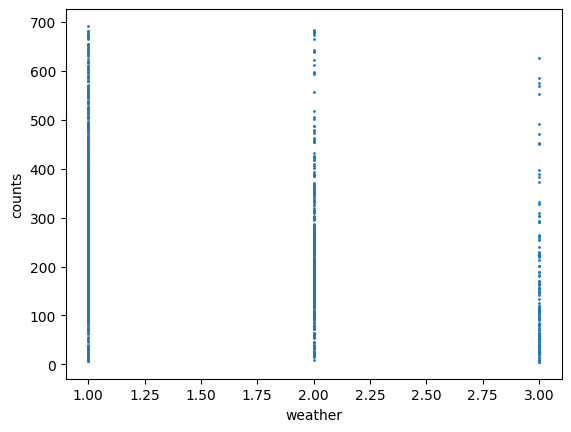
从图表中可以看出，没有明显的关于风速与用户数量的线性关系。不过，从大致趋势上来说，当风速逐渐变大时，用户的数量逐渐减少。

Season 2, holiday Season 2, workday



在天气上同样有类似的趋势出现。随着天气条件的不断恶化（从1到3），用户的数量在不断地减少。

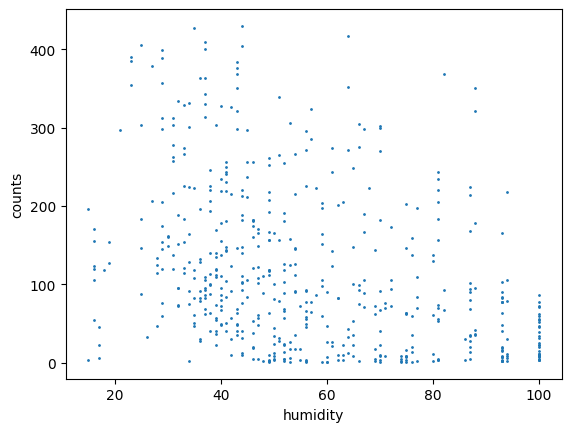
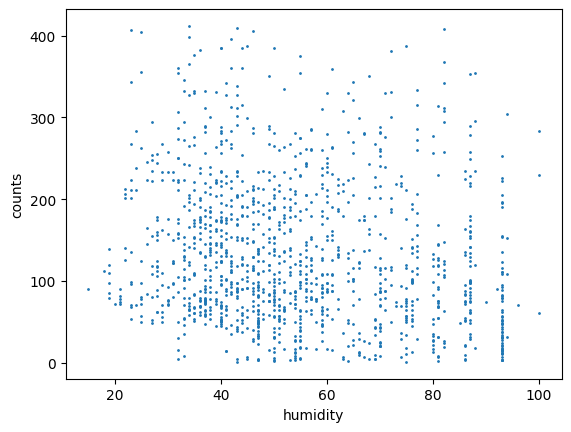
Season 2, Holiday Season 2, Workday

**湿度对用户数量的影响**

湿度同样决定了用户使用共享单车的舒适程度。此处选择Season 1的数据为样例进行分析。

Holiday Workday

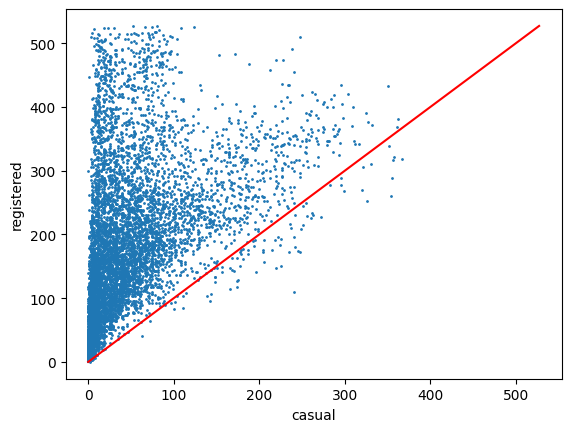
 

工作日和非工作日的图表展示出了相似的趋势。在非工作日，随着湿度的上升，用户数量出现了明显的下降。在工作日，尽管趋势并不如在非工作日那么显著，还是可以观察到在湿度较大的时候（80%-100%），代表较多用户数量的高位数据点减少，而数据点更多集中在低位。

**用户分布情况**

通过观察图表可知，已注册用户占据用户的绝大多数。在每日的用户分布图中，绝大多数的数据点位于平分线上。

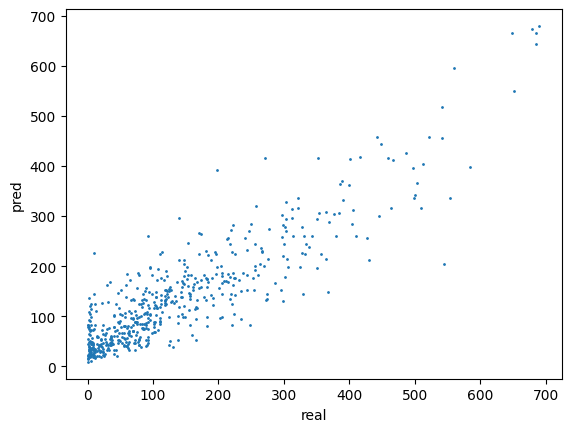
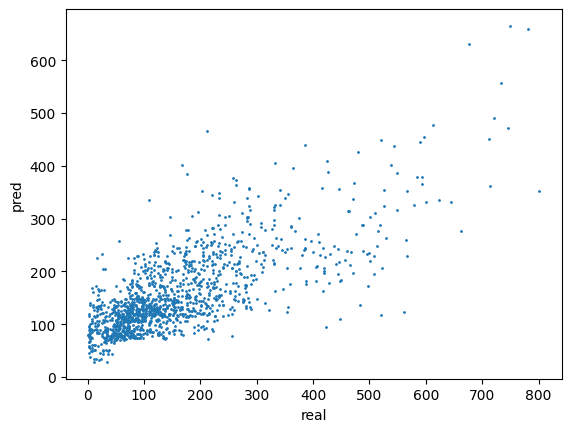
All season, holiday and workday



**机器学习- 回归**

运用机器学习对数据进行了回归计算，通过Random Forest算法对天气，温度，体感温度，湿度，风速进行线性回归拟合，得出以下结论。

Season 1, Holiday, R^2:0.51 Season 1, Workday, R^2: 0.22

**** ****

在节假日，拟合的模型对于用户数量有较好的估计效果，而工作日则效果较差。这很大程度是由于工作日骑行需求弹性较差，骑行用户数量与各项数据关联性较低导致的。

**总结**

通过对共享单车租赁数据的分析，我们可以发现，当天气条件比较恶劣时（湿度较大，温度较高较低，风速较大）时，用户数量相较其他时间有所减少。其中，非工作日的用户租赁数量对这些因素更加敏感，变化幅度也大于工作日。同时，工作日和非工作日的用户使用时间分布存在明显差异。工作日的单车租赁时间集中在早上8点和下午6点两个时段，可以理解为上下班的高峰期。而在非工作日则观察到下午12至16点为租赁高峰，可以认为是用户出行游玩，逛街的需求高峰。由于非工作日用户对天气等各项因素反应更强烈和明显，其数据拟合模型的决定系数(R square)高于工作日用户。

由此，可以对单车的租赁提出以下建议：

1. 针对较好的天气加大投放力度，吸引用户出行。
2. 在上下班高峰期时于各种交通枢纽，大型写字楼处加大投放数量。
3. 晚上为单车租赁的空窗期，可以用于单车的部署和维护。