

调查问卷数据

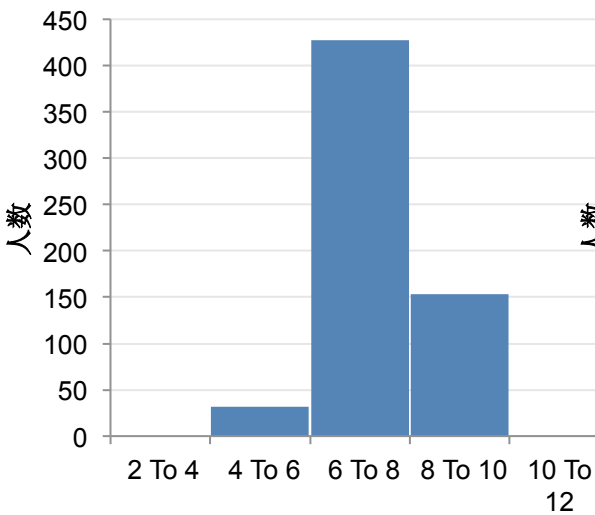
楚文亚

睡眠时间和是否工作之间存在什么关系？

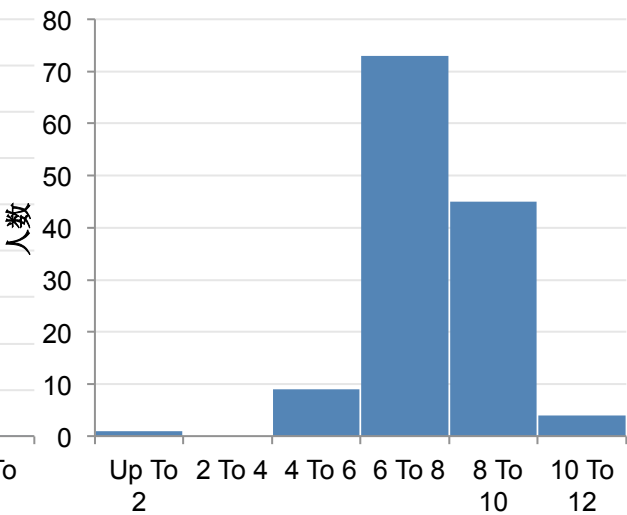
睡眠时间/h	工作	未工作
最小值	4	1
Q1	6	7
Q2	7	7
Q3	8	8
最大值	10	10
均值	6.9	7.1
异常值	无	有

从五数概括法可以看出，未工作人群的睡眠时间异常值较多，同时均值大于工作人群的睡眠时间。

工作睡眠时间直方图



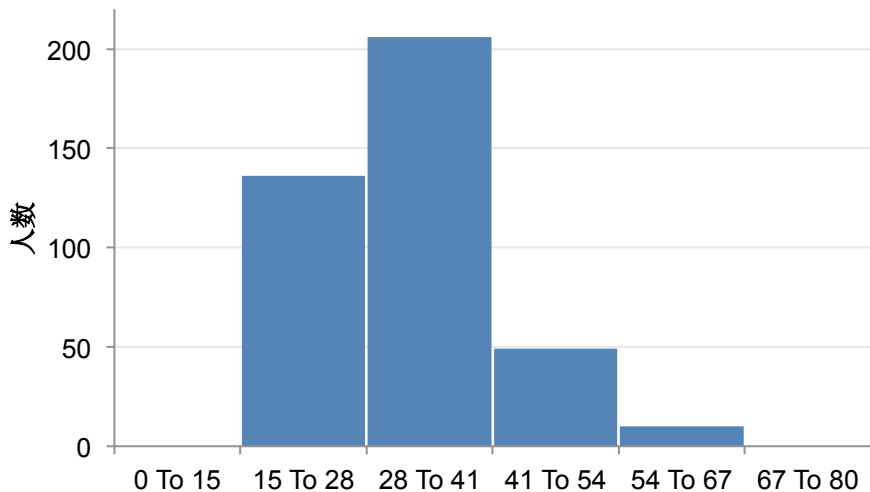
未工作睡眠时间直方图



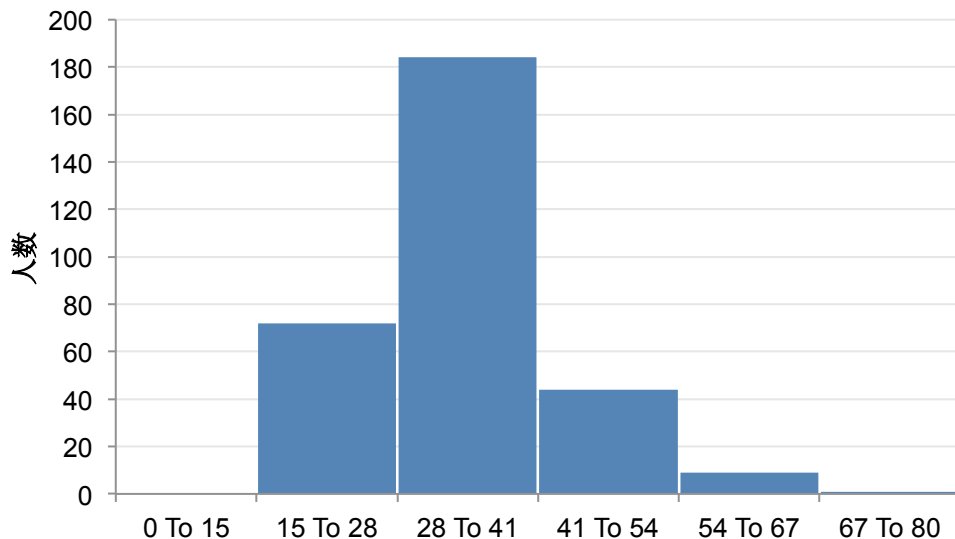
从直方图各组数据对比可以看出，是否工作对于平均睡眠时间有一定影响，在6~8小时睡眠时间中，工作人群比例（70%）高于未工作人群比例（55%）；在8~10小时睡眠时间中，未工作人群比例（34%）要高于工作人群比例（25%）；同时，未工作人群的睡眠时间数据分布较广

年龄与是否购买周边/购买周边种类的关系

想买的人的年龄直方图/岁



不想购买的人年龄直方图



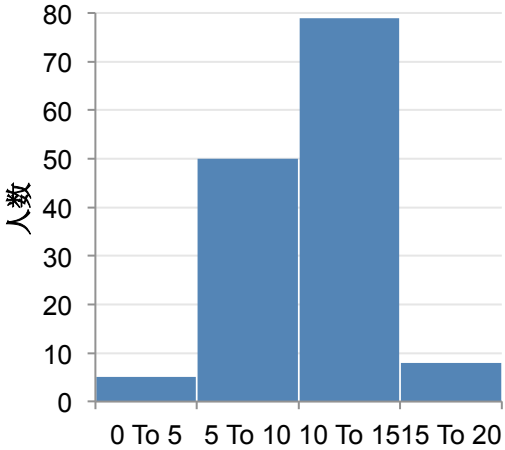
从上面两个直方图中可以看出，对于28~41岁和41~54岁年龄段的人来说，想要购买周边产品和不想购买周边产品的人在其各自的群体中基本持平，然而对于15~28岁的人群，想要购买周边的人数和所占比明显大于不想购买周边的人，这一点值得关注。

通勤时间与久坐时间之间的关系

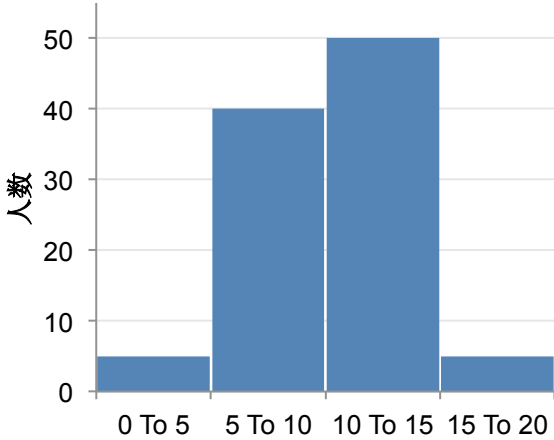
	通勤时间/ 小时	久坐时间/ 小时
值域	(0,10)	(1,800)
标准偏差	0.97	39.94

从上表可以看出，通勤时间的值域基本在正常时间范围内（小于24小时），同时标准偏差较小，说明数据离散程度较小；而久坐时间的值域很大，最大值甚至超过24小时，因此标准偏差也高达39.94，说明原始数据离散程度较大，有需要清洗的数据（大于24小时）

通勤时间为0—久坐时间直方图/h

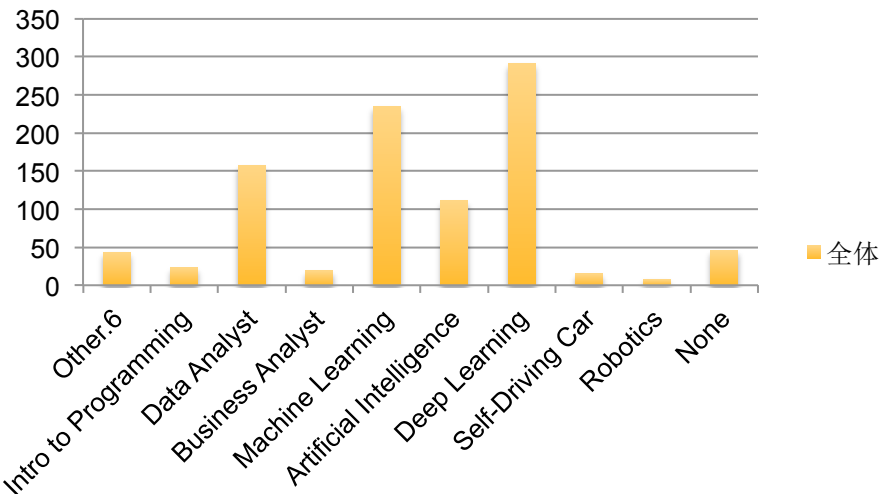


通勤时间大于2h—久坐时间直方图/h

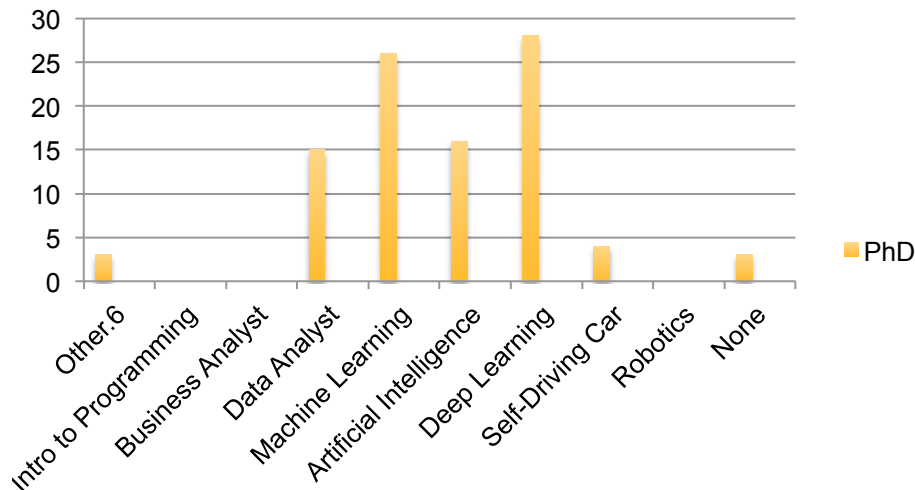


从上面两图中可以看出，通勤时间与久坐时间之间并无明显关系，不管是通勤时间为0（在家办公）或者通勤时间在2小时以上，最普遍的久坐时间都是10~15小时。不过值得注意的是，对于在家办公的人来说，5~10小时的久坐时间所占比例要小于通勤时间较长的人。

最高学历与选课项目之间是否存在一定关系？

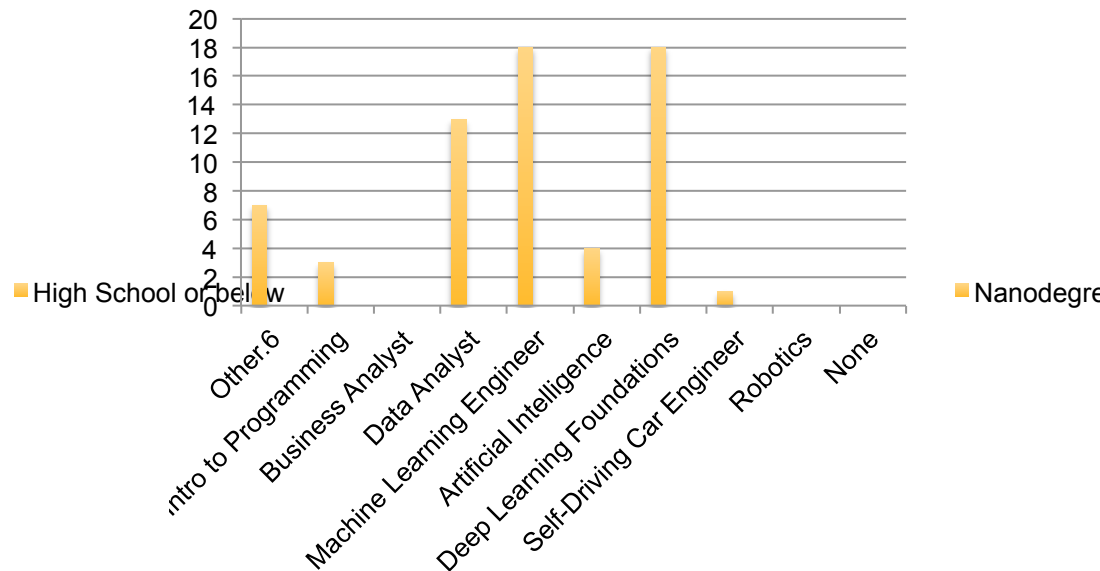
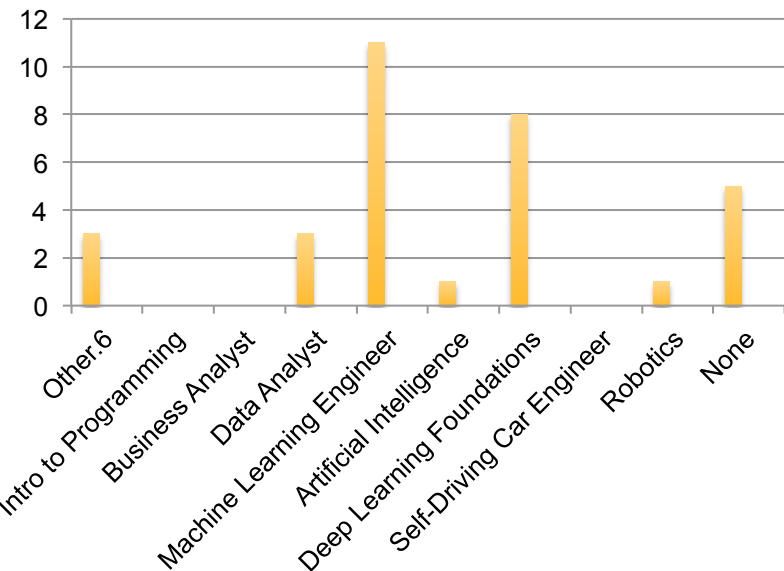


根据全体学员选择课程的柱形图来看，最受欢迎的四门课程分别是Data Analyst, Machine Learning Engineer, Artificial Intelligence, Deep Learning Foundations



根据最高学历为PhD的人群所选课程柱形图来看，最受欢迎的四门课程和全体情况相同，也是Data Analyst, Machine Learning Engineer, Artificial Intelligence, Deep Learning Foundations

最高学历与选课项目之间是否存在一定关系？



而从上面两个柱形图可看出，对于最高学历为High school or below和Nanodegree Program的学员来说，Artificial Intelligence课程逐渐失去了吸引力，但是Data Analyst, Machine Learning Engineer, Deep Learning Foundations三门课仍旧是这些人们最爱选择的三门课程