

我国老年群体医疗情况研究

首都经贸大学

程义淇 22019212043

李 康 220192120xx

2020.1.5

一、 研究问题说明

本篇报告的分析基于北大国家发展中心《中国健康与养老追踪调查试调查追踪问卷（2012）》中所得出的相关数据。小组成员在通览数据后一致决定选择“医疗保健与保险”这一篇调查内容为研究主体，这一篇内容包含医疗保险和医疗成本与使用情况这两大部分，其中医疗成本与使用情况又可以细分到最近一次就医以及最近一次住院等相关信息，较有研究价值。此外，为使研究更具实用性、更能清晰具体地关联不同人群特征，我们加入了基本信息表、工作、退休和养老金表以及收入、支出与资产表中的变量，经过整合后对问卷结果进行可视化分析，所有数据处理及绘图工作使用 python 语言完成。分析思路如下：

- 概览受访者总体背景情况，包含户口类型、最高学历、目前婚姻状况、性别、参保情况、收入及支出情况等
- 具体分析参保问题，包含各因素对是否参保、何种医保以及保费的影响等
- 具体分析最近一次就医情况，包含各因素对医疗成本以及使用情况的影响等，并重点关注不同情况下药费的差异。
- 具体分析最近一次住院情况，包含各因素对是否住院、住院花销、住院报销情况的影响等，并重点关注需要住院但没有选择住院这类群体的深层原因。

二、 数据处理

(一) 基本信息表

1. 字段筛选

筛选后保留字段数据字典如下：

B 基本信息表数据字典-demographic_background	
字段名称	含义
ID	ID
householdID	householdID
ba002_1	出生年份
ba002_2	出生月份
ba002_3	出生日
ba003	公历或阴历
bc001	目前户口类型
bd001	最高学历
be001	目前婚姻状况
rgender	性别

(二) 医疗保险与保健表

1. 字段筛选

筛选后保留字段数据字典如下：

E 医疗保健与保险-health_care_and_insurance	
ID	ID
householdID	householdID
ea001	参加何种保险
ea006	每年自己缴纳多少保费
ec001_1	最近一次体检的年份
ec002	最近一次体检谁付的体检费
ed001	过去一个月里，您是否去医疗机构看过门诊或者接受过上门医疗服务
ED002	您在过去一个月内生过病吗
ED004	过去一个月，您去过哪些医疗机构接受门诊治疗

ED008	过去一个月，您最近一次去了哪种医疗机构就诊
ED009	最近去的医疗机构是公立的还是私立的
ED013	从您家到这家医疗机构有多远
ED014	您上次去这家医疗机构单程花多少时间
ED014_1	分钟或小时
ED014_2	交通方式
ED015	您上次去这家医疗机构的单程交通成本是多少
ED022	最近一次就诊，您排队花了多长时间
ED023	这次就诊(包括药费和诊疗费，药费包括在这家医疗机构或其他药店购买医生开的处方药)总共花了多少钱
ED024	这次就诊所花的费用里，您自己支付了多少
ED025	自付部分谁给您交得最多
ED026_1	这次就诊所付的药费(包括在这家医疗机构或其他药店购买医生开的处方药)一共是多少钱
ED027_1	这次就诊所花的药费(包括在这家医疗机构或其他药店购买医生开的处方药)里，您自己支付了多少钱
ed028s1~s12	这次就诊的费用(包含诊疗费和药费)您使用或将使用哪种医疗保险报销
ED029	您最近一次就诊有没有给医生送红包
EE001	过去一年，是否有医生说您应该住院而没有住院的情况
EE002	您没去住院的主要原因是什么
EE003	过去一年内，您住过院吗
EE004	过去一年，您接受过几次住院治疗
EE005_1	过去一年住院的总费用大概是多少?(只包括付给医院的费用，不包括陪护的工资、自己或家人的交通费和住宿费，但包括医院病房费。
EE006_1	其中自付部分
EE008	过去一年中，您最后一次接受住院治疗所在的医疗机构的类型是什么
EE009	这家医疗机构是公立的还是私立的
EE010	这家医疗机构是什么级别的
EE013	从您家到这家医疗机构有几公里
EE015	您上次去这家医疗机构接受住院治疗的单程交通成本是多少钱
EE016	您在那里住了几晚
EE019	您住院的原因是什么
ee021s1~s10	在最近的一次住院期间，您接受了哪些治疗
ee022	您这次是在什么情况下出院的
EE023	您要求提前出院的原因是什么
EE024_1	这次住院一共花了多少钱?(只包括付给医院的费用，不包括陪护的工资、自己或家人的交通费和住宿费，但包括医院病房费
EE025_1	雇佣陪护照顾病人花了多少钱
EE026_1	自己和家人往返医院交通费、伙食费以及家人或陪护的住宿费一共花了多少钱
EE027_1	这次住院，您最后自己总共付了(将要付)多少钱
EE028	自付部分谁给您交得最多

EE029_1	这次住院药费一共花了多少钱
EE030_1	这次住院，您自己一共付了多少钱的药费
EE031s1~s12	您使用或将使用哪种医疗保险报销
EE032	这次住院期间有没有给医生送红包

2. 字段合并

a) 字符串形式的合并

问卷所得数据中部分问题用哑变量形式分列展示，故将其拼接为一列，涉及到的变量包括：医疗报销途径、住院期间接受何种治疗等

b) 数值形式的合并

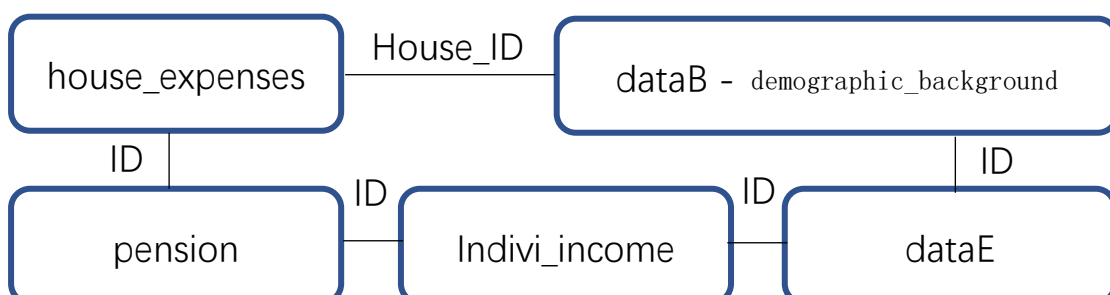
问卷中存在循环提问的情况，故同一指标的真实值由每部分具体值相加得出，如每年自己缴纳多少保费这一指标，就是将受访者所有参保种类的缴纳保费相加得到最终结果。

(三) 工作、退休和养老金表

(四) 收入、支出与资产表

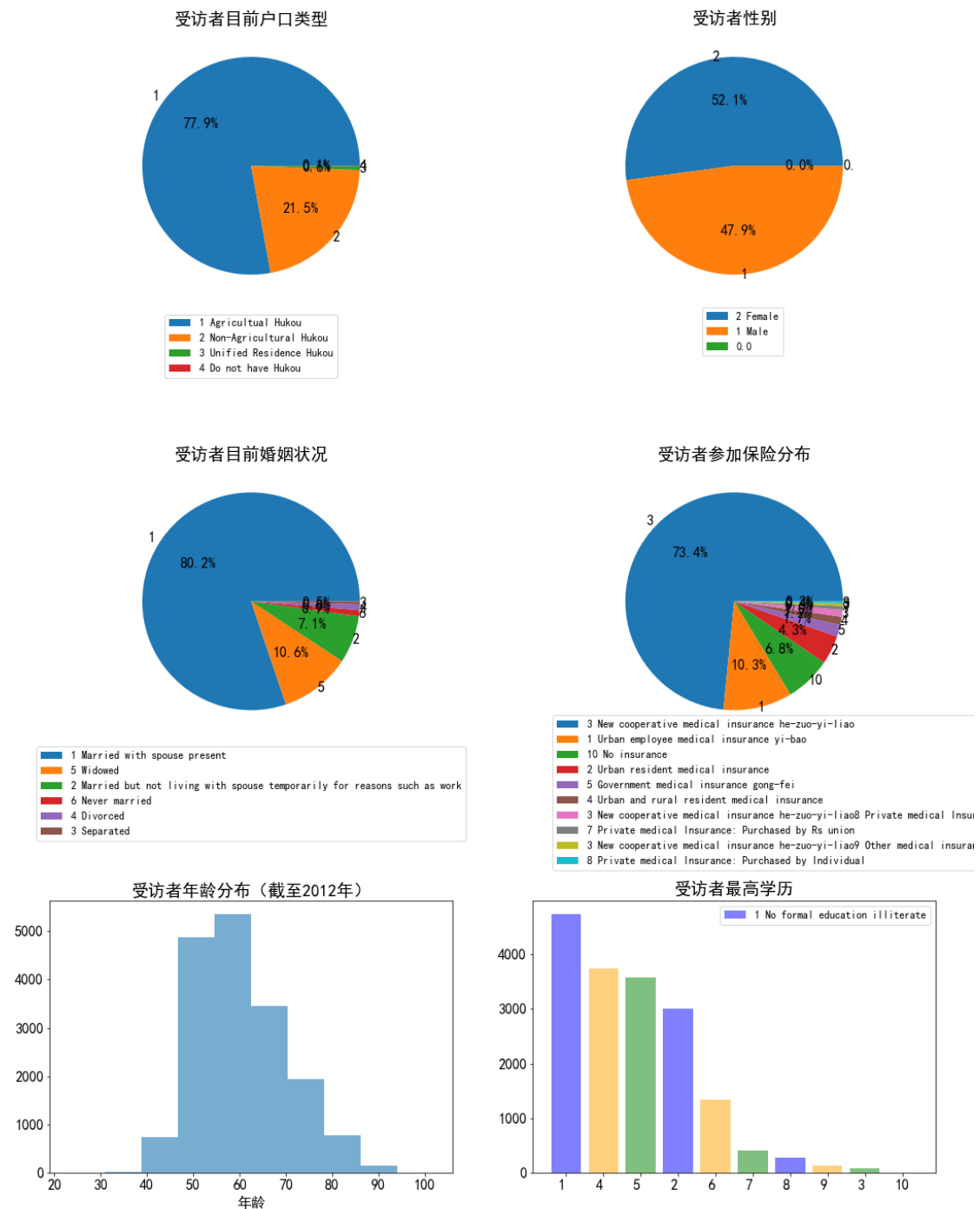
(五) 合并表

合并规则表示如下：



三、 数据可视化

(一) 受访者基本信息



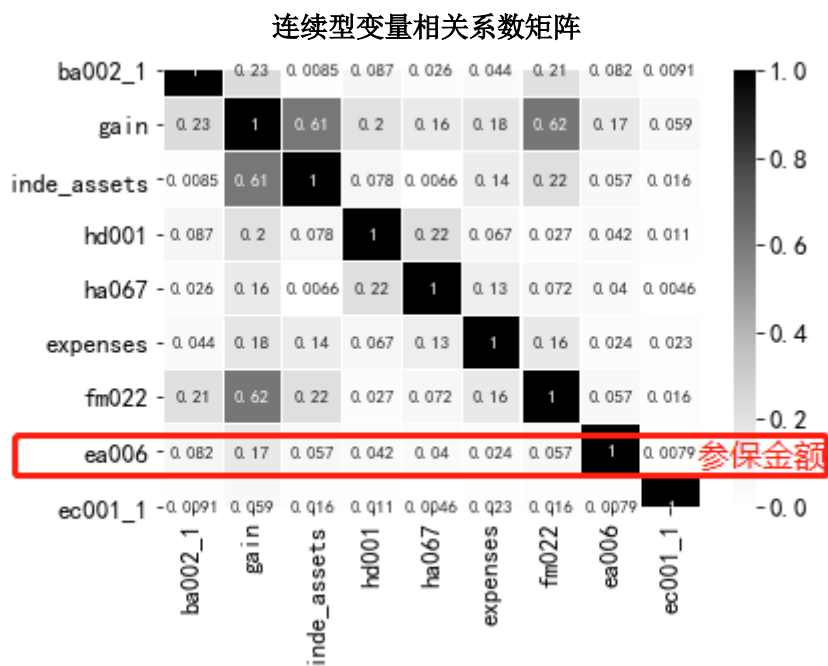
这一部分采用 Python 的子图功能，在一张大图中分别采用饼图、

直方图以及柱状图描述受访者的各项基本信息，结果清晰明了可让分析人员对受访者有一个初步的认识：这是年龄大多在 50~80 岁之间、农村户口较多、性别比例均衡、非正式学历较多且大部分参加了农村联合医保的一个信息采集群体。

(二) 受访者医疗保险情况分析

1. 各连续型变量相关性

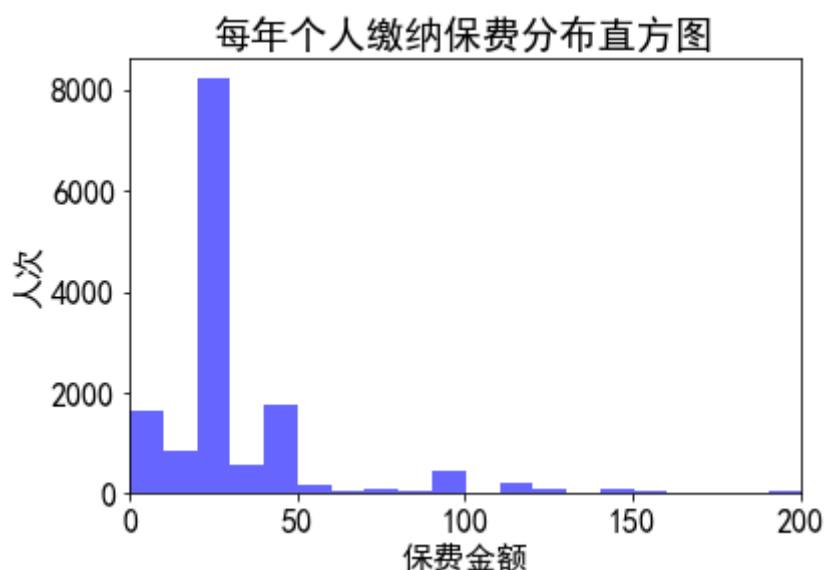
计算受访者收入、不定资产、其他固定资产、养老金领取金额、最近一次体检的年份 与每年缴纳保费的相关系数，得出下图。



从图中可看出，受访者每年缴纳的保费金额与其自己的收入存在较弱的相关关系，可见对于这部分受访者而言，无论收入高低都可以获得一定的医疗保障。

2. 个人缴纳保费金额

为分析该群体每年负担保费金额多少，绘制个人缴纳部分的直方图，由于个人缴纳保费金额这一指标存在极端值，我们将其隐去后展示如下：

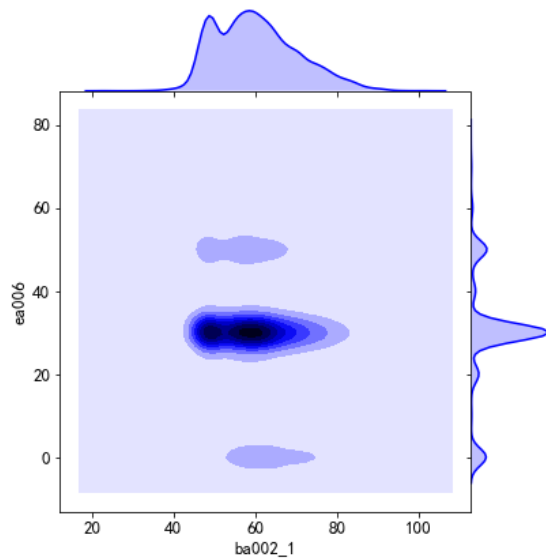


从图中可以看出，大部分的受访者每年的保费支出均为 30 元。其次是 0 元和 50 元的两种情况。几乎没有人的每年缴纳保费金额超过 100 元。

3. 各年龄段人群缴纳保费情况

因为我们了解到，受访者的年龄分布在 40 以上，跨度较大，且很可能属于不同人生阶段（如是否退休等），故对不同年龄段的人每一年的医保费用支出进行描述如下：

年龄-保费 热力图

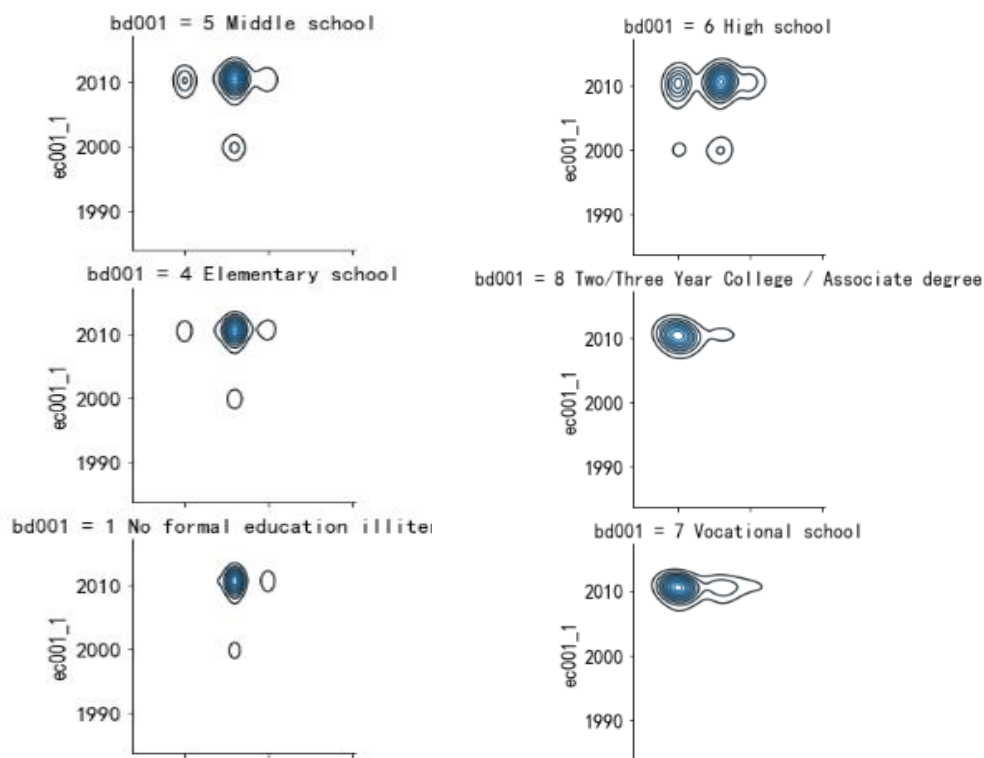


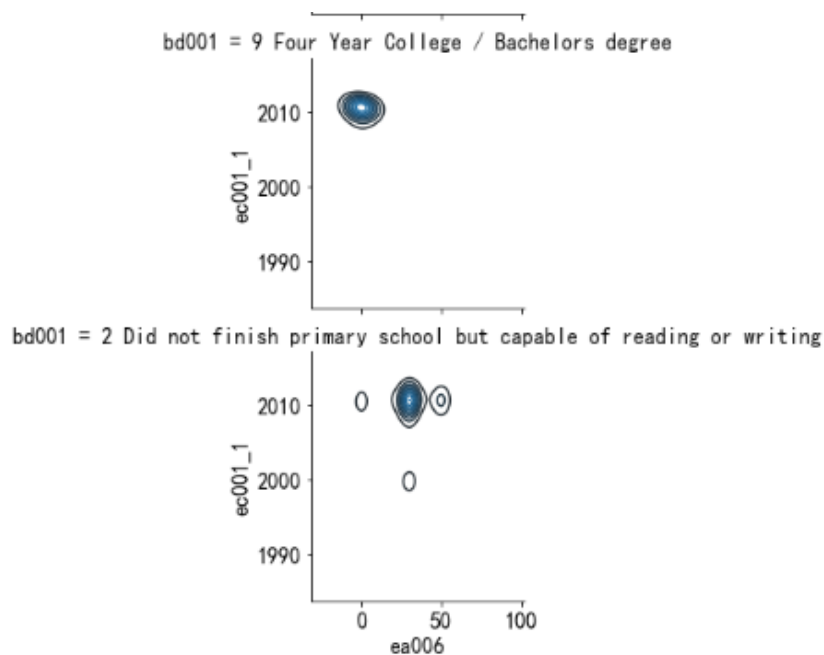
从左图可以看出缴纳 50 元保费的人群在 50 岁左右，无缴纳保费的人群在 70 岁左右，存在年龄上的差异。

4. 不同学历背景下健康关注度与每年缴纳保费是否有关

为分析不同学历背景下，健康关注度与缴纳保费的关系选取受访者最近一次体检的年份来表示其健康关注度。

学历-保费-体检时间组图

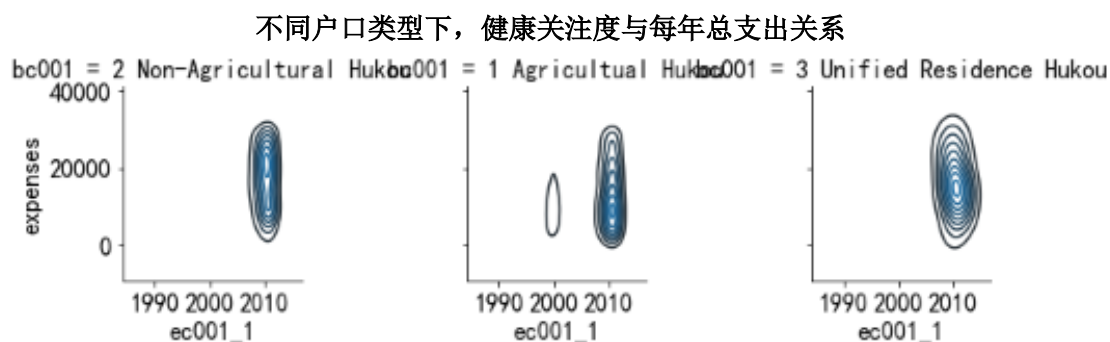




从上图我们可以发现，学历越高的受访者最近一次体检的时间越近，且他们往往自身不缴纳医保费用。

5. 不同户口类型下，健康关注度与每年总支出关系

此处仍以最近一次体检年份来表示受访者的健康关注度，以受访者的户口类型作为第一级区分，绘制支出和健康关注度的密度图，表示二者相关关系。

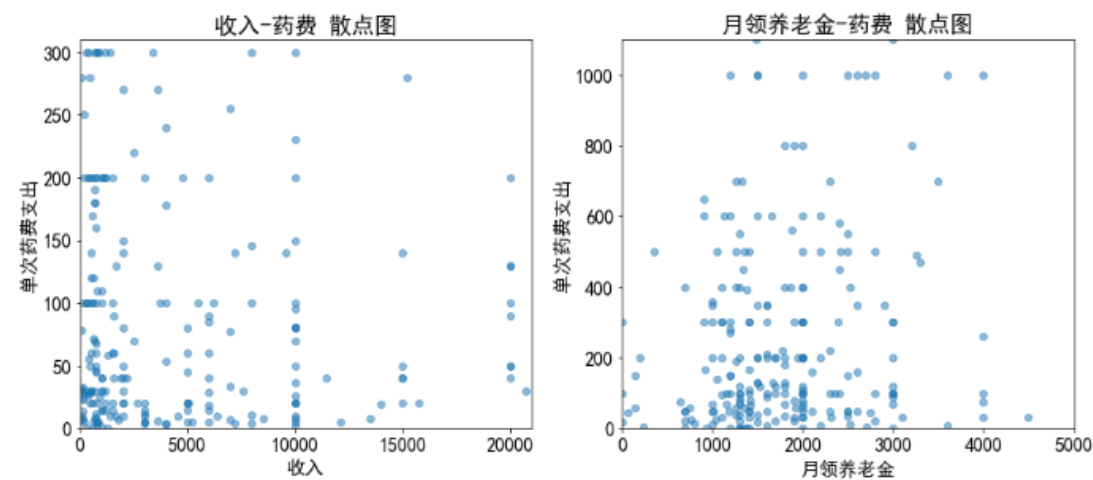


从上图我们可以看出，只有农村户口的受访者存在最近一次体检在十年前的情况，非农村户口的受访者体检时间与每年开销基本无关，而农业户口的受访者密度图像下半部分略微左偏，可见在开销小的这

部分群体中健康关注度较低，而集体户口的这部分受访者密度图像下半部分略微右偏，可见在开销小的这部分群体中健康关注度反而更高。

(三) 受访者一个月内最近一次就医相关情况分析

1. 收入与药费关联情况

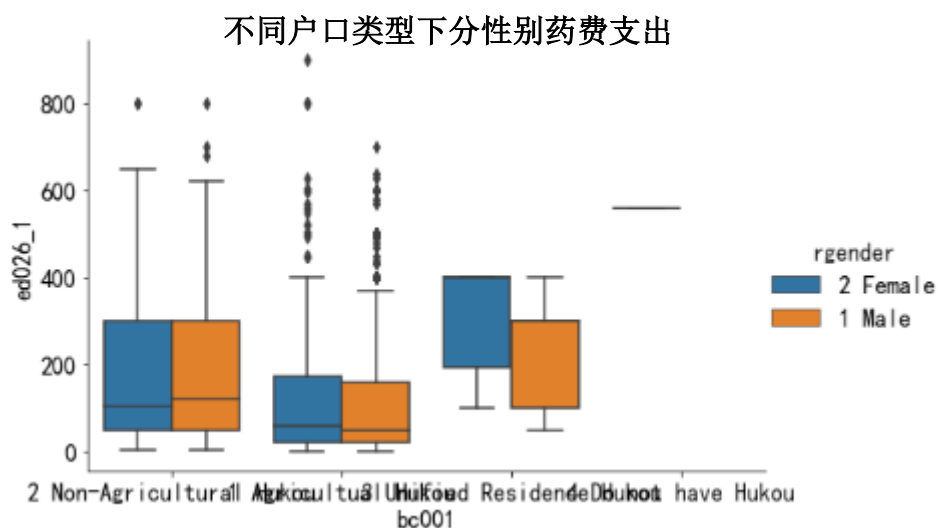


分别观察收入和月领养老金与药费的关系，两图对比可见领养老金的受访者的平均单次药费支出比正常收入的受访者要多出许多，可见老年人的药费需求更大，基本要占每月养老金的十分之一。

且从两图中均可看出，两种收入与药费都没有明显的相关关系，这从侧面表明，受访者购买的药品均为刚需商品，没有收入越多药品越贵的情况。

2. 不同户口类型下，不同性别受访者药费区间

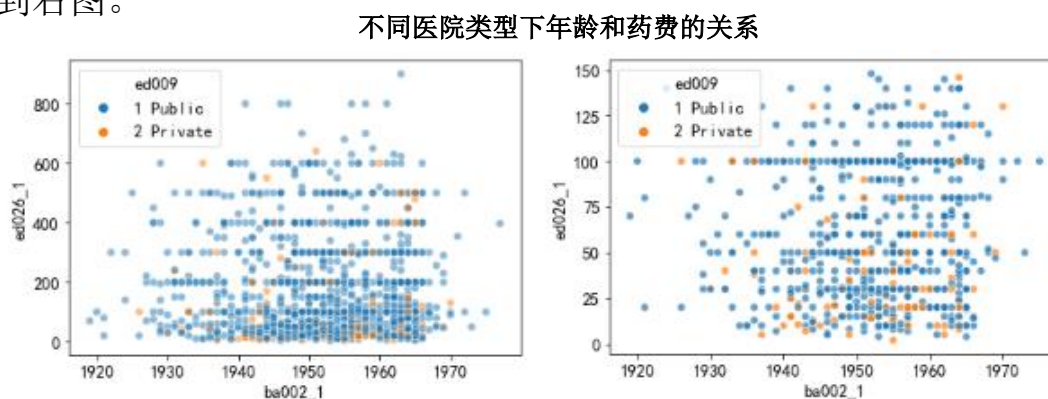
选取户口类型作为箱线图的一轴，选取药费作为另一轴，以性别作为箱线图的分箱指标，得出图形如下：



从上图可以看出，农村户口的受访者花费的药费要整体低于其他两种情况。仅在集体户口的受访者中男女支出药品金额差异明显。

3. 不同医院类型下年龄和药费的关系

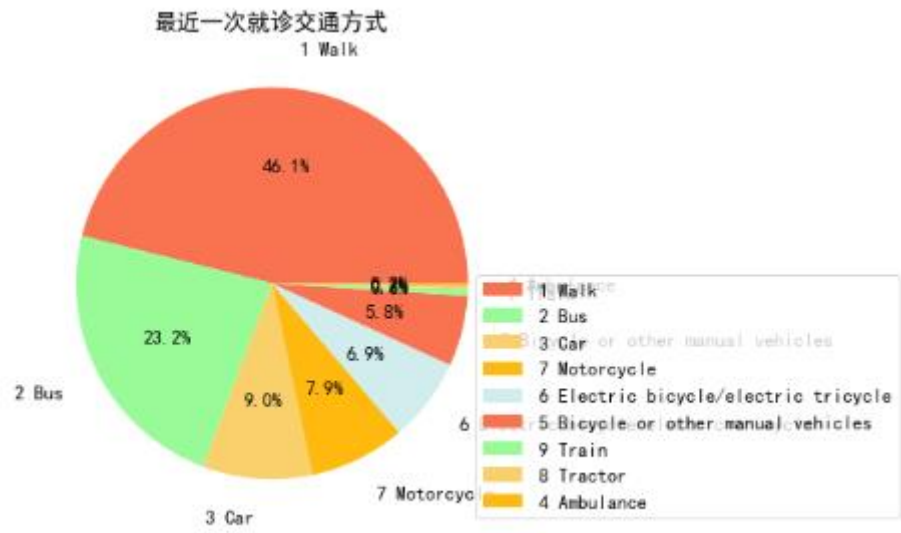
选取医院的类型即公立或者私立作为散点图的颜色区分，以受访者出生年份作为一轴，支出药费作为另一轴分析，最初得到左图，因去私立医院的受访者较少，信息表述不明，故对密集区域另外作图得到右图。



从图中来看，私立医院和公立医院的医药费没有拉开差距，病人年龄也没有拉开差距。

4. 最近一次就诊交通方式

选取受访者最近一次就诊时的交通方式绘制饼图



这部分受访者最近一次就诊时有将近一半的人选择步行，公交占比约五分之一，考虑到他们的年龄，可以认为他们就诊的交通便捷性较差。

(四) 受访者一年内住院情况分析

四、 研究发现

1. 对于此次问卷采集的受访者而言，无论收入高低都可以获得一定的医疗保障
2. 大部分的受访者每年的保费支出均为 30 元
3. 学历越高的受访者健康关注度更高，且他们往往自身不缴纳医保费用
4. 只有农村户口的受访者存在最近一次体检在十年前的情况，且其

中开销小的一部分群体中健康关注度较低

5. 公立医院与私立医院的药费支出无明显差距
6. 近半数受访者去医院就诊的便捷性有待改进