**三（三）科举与官员的动态流动**

暂定

说明：动态流动情况需要结合官员的履历情况分析，即所用的数据是全部数据大约20000多条。

又因为数据库中履历信息是在“人物传记库”中检索“姓名”的结果，所以有可能实际履历信息与官员是不对应的。排除的依据可以用“任职时间”减去“生年”，如果数值为负或者大于90岁，可以认为其履历信息是错误的，因此分析中可以排除，**即排除履历不正确的数据（任职时间减去生年小于0或大于90的情况）。**

**“科举类别”：**进士（进士、翻译进士、武进士）；

举人（举人、翻译举人、武举人）；

举人以下（其他非空的类型，如贡生、监生等等）

1. **最初任职职位**

（1）目的：分析“科举类别”对不同“官职类别”的官员**初次任职职位**的影响。

（2）条件： 当“姓名内编码”=1或2（即时间最早的几条履历所对应的职位）

统计“任职职位”中各种职位出现的频次。

（3）结果：

不同“科举类别”下“任职职位或任职职位简写”及其出现的次数及比例。

按照不同的“官职类别”分别输出上面的结果。

1. **最初任职的地点**

（1）目的：分析“科举类别”对不同“官职类别”的官员初次**任职地点**的影响。

（2）条件： 当“姓名内编码”=1或2（即时间最早的几条履历所对应的职位）

统计“任职地”中各种任职地出现的频次。

（3）结果：

不同“科举类别”下“任职地”及其出现的次数及比例。

按照不同的“官职类别”分别输出上面的结果。

1. **上升空间**

（1）目的：分析“科举类别”对不同的 “官职类别”的官员履历中**最高职位**的影响。

（2）条件： 当“姓名内编码”最后几位，且“任职职位简写”为非空值（即时间最晚的几条履历所对应的职位，且是有效数据。有效数据是指“任职职位简写”中为非空值）

（3）结果：

不同“科举类别”下，统计“任职职位简写”中各种任职地及其出现的次数和比例。

按照不同的“官职类别”分别输出上面的结果。

1. **任职年限**

（1）目的：分析“科举类别”对不同的 “官职类别”的官员**任职年限（时间差）**的影响。

（2）条件：列C“姓名内编码”为0和1，保留按照任职时间排列的官员基本信息。

去掉“任职年限”为空的数据

计算所在职位任职的**时间差**。

**时间差**计算方法：同一“官职类别”中的同一“官职名称”内，

官员A的任职年限=b—a

官员B的任职年限=c—b

|  |  |
| --- | --- |
| 姓名 | 任职时间 |
| 官员A | a |
| 官员B | b |
| 官员C | c |

同一“官职名称”内的最末位的官员任职年限可以不用计算。

（3）结果：

不同“科举类别”下，统计时间差为“空值、0、1、2、3……”出现的次数和比例，即“时间差”+“次数”+“比例”

按照不同的“官职类别”分别输出上面的结果。

1. **区域间的流转**

（1）目的：为了统计官员从目前所在地到全国其他地方流动的特点。

（2）条件：不同“官职类别”下，不同“科举类别”下，履历信息中“任职地”出现的名称和次数。

（3）结果：

不同“科举类别”下， “任职地”及出现的次数

按照不同的“官职类别”分别输出上面的结果。

注意：

注：1、2、3项的统计条件待定，待定原因：

1. 不好确定条件中“任职地”“任职职位”“任职职位简写”是否选择“姓名内编号”的1或2或3？即最初任职地和职位的选择。
2. 涉及履历中“职位”的统计，应该选择“任职职位”还是“任职职位简写”。因为有的“任职职位”内容没办法简写。