**说明文档**

**一、原始数据**

|  |  |  |  |
| --- | --- | --- | --- |
| 数据名称 | 数据来源 | 文件夹 | 具体文件**（未上传）** |
| **Panel A: 正文主要数据** | | | |
| 中国人口普查微观数据 | 复旦大学经济学院数据库（非公开） | Data | * census2000.dta * census2010.dta * census2015.dta |
| 美国O\*NET数据 | https://www.onetcenter.org/db\_releases.html | Appendix Data | * ONET N\_version\db\_25\_0\_excel.zip * ONET N\_version\db\_50.zip * ONET N\_version\onet151ac.zip * ONET N\_version\onet20ac.zip |
| ACS数据 | https://www.ipums.org/ | Appendix Data | * ACS\usa\_00003.dta * ACS\ACS-SOC2010.xlsx |
| **Panel B: 附录数据** | | | | |
| 中国女性社会地位调查 | 复旦大学经济学院数据库（非公开） | Appendix Data | * Women Status\wave3.dta * Women Status\wave2.dta |
| 中国家庭追踪调查数据 | https://www.isss.pku.edu.cn/cfps/index.htm | Appendix Data | * CFPS\paneldata\_person.dta * CFPS\cfps2018crossyearid\_202104.dta * CFPS2010职业编码对应表.xlsx |
| 中国动态劳动力调查数据 | http://css.sysu.edu.cn/ | Appendix Data | * CLDS\individual2014-clds.dta |
| 第五次和第六次人口普查统计年鉴 | * https://www.stats.gov.cn/sj/pcsj/rkpc/5rp/index.htm * https://www.stats.gov.cn/sj/pcsj/rkpc/6rp/indexch.htm | Appendix Data | * 职业和行业就业分布.xlsx |
| 2016年《中国劳动统计年鉴》 | https://www.mohrss.gov.cn/2016/indexch.htm | Appendix Data | * 01-37 2016中国劳动统计年鉴.xls |
| ChatGPT4.o | https://chat.openai.com/chat | Appendix Data | * ChatGPT Question.txt * Chatcheck.xlsx |
| 王亚菲等（2021） | https://publishmedia.cbpt.cnki.net/portal/minio/webs/tjyj/media/web/2024/06/12/%E7%8E%8B%E4%BA%9A%E8%8F%B2%E7%AD%89%EF%BC%9A%E4%B8%AD%E5%9B%BD%E8%A1%8C%E4%B8%9A%E5%B1%82%E9%9D%A2%E5%B0%B1%E4%B8%9A%E6%A0%B8%E7%AE%97%E7%A0%94%E7%A9%B6.pdf | Appendix Data | * 王亚菲行业就业测算结果.xlsx |
| 2022年中国职业技能发展数据库 | https://doc.weixin.qq.com/doc/w3\_AcgA8QasAHEenUZdg1jTwq0NFV5Tq?scode=AMgAPwf5AAgznsyo20AcgA8QasAHE | Appendix Data | * 香帅的职业数据库\香帅的职业发展指数.xlsx |
| **Panel C: 职业、行业分类标准** | | | |
| 《职业分类与代码》GB/T6565-2009  《职业分类与代码》GB/T6565-2015 | https://www.doc88.com/p-7919531862293.html  https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=26D453E518AFC7A826331A251E7EB1F9 | Classification | * occ2015\_match\_occ2010.xlsx * 《职业分类与代码》GB/T6565-2009.pdf * 《职业分类与代码》GB/T6565-2015.pdf |
| ONET SOC 职业分类 | https://www.onetcenter.org/taxonomy.html | Classification | * 2000\_to\_2006\_Crosswalk.xlsx * 2006\_to\_2009\_Crosswalk.xlsx * 2009\_to\_2010\_Crosswalk.xlsx |
| 中国行业分类  \*GBT4754-1994  \*GBT4754-2002  \*GBT4754-2002  \*GBT4754-2011 | https://international.ipums.org/international-action/variables/IND#codes\_section | Classification\行业标准转换 | * cic02\_3dnames\_new.dta * 2010industry\_分类.xlsx * 2015industry\_分类.xlsx * 国民经济行业分类和代码表\2002与1994版《国民经济行业分类》国家标准对照表.xlsx * 国民经济行业分类和代码表\2011与2002版《国民经济行业分类》国家标准对照表.xlsx |
| Classification\industry correspondance | * cic acsindustry correspondance.xlsx |
| 美国行业分类 | https://international.ipums.org/international-action/variables/IND#codes\_section | Classification\industry correspondance | * acs industry correspondence.xlsx |

**二、数据预处理**

|  |  |  |  |
| --- | --- | --- | --- |
| 原始数据**（未上传）** | 原始数据所在  文件夹 | 数据（基本在Lab文件夹） | Do file |
| * ONET N\_version\db\_25\_0\_excel.zip\Occupation Data * 《职业分类与代码》GB/T6565-2009 * 《职业分类与代码》GB/T6565-2015 | * Data | * occupation2009-ONET.xlsx * occ2015\_sem.xls | 手动匹配所得  （生成的数据文件位于Data文件夹） |
| * cic02\_3dnames\_new.dta * 2010industry\_分类.xlsx * 2015industry\_分类.xlsx * 国民经济行业分类和代码表\2002与1994版《国民经济行业分类》国家标准对照表.xlsx * 国民经济行业分类和代码表\2011与2002版《国民经济行业分类》国家标准对照表.xlsx | * Classification\行业标准转换 | * census2000-2015\_行业统一版本 | Pre-ind.do |
| * acs industry correspondence.xlsx * cic acsindustry correspondence.xlsx | * Lab | * cic\_acsind\_chn.dta * cic\_acsind\_usa2000.dta * cic\_acsind\_usa2010.dta * cic\_acsind\_usa2015.dta | Pre-ind.do |
| * CHN-consistent\_ONETSOC2010 | * Appendix Data | * process-CHN\_ACSSOC2010\_v4.xlsx | 手动整理所得  （生成的数据文件位于Appendix Data文件夹） |
| * occ2015\_match\_occ2010.xlsx | * Lab | * occ2010\_consistent.dta * occ2015\_consistent.dta | Pre-occ.do |
| * occupation2009-ONET.xlsx * occ2015\_sem.xls * 2000\_to\_2006\_Crosswalk.xlsx * 2006\_to\_2009\_Crosswalk.xlsx * 2009\_to\_2010\_Crosswalk.xlsx | * Lab | * CHN-consistent\_ONETSOC2010.dta * CHN-consistent\_ONETSOC2009.dta * CHN-consistent\_ONETSOC2006.dta * CHN-consistent\_ONETSOC2000.dta | Pre-occ.do |
| * ONET N\_version\db\_25\_0\_excel.zip * ONET N\_version\db\_50.zip * ONET N\_version\onet151ac.zip * ONET N\_version\onet20ac.zip * CHN-consistent\_ONETSOC2010.dta * CHN-consistent\_ONETSOC2000.dta | * Lab | * CHN\_cons\_to\_SOC2010\_v25.dta * CHN\_cons\_to\_SOC2010\_v20.dta * CHN\_cons\_to\_SOC2010\_v151.dta * CHN\_cons\_to\_SOC2000\_v5.dta | Pre-onet.do |
| * 2000census * 2010census * 2015census * CHN\_cons\_to\_SOC2010\_v25.dta * CHN\_cons\_to\_SOC2010\_v20.dta * CHN\_cons\_to\_SOC2010\_v151.dta * CHN\_cons\_to\_SOC2000\_v5.dta | * Lab | * CHN2000\_v5.dta * CHN2000\_v151.dta * CHN2000\_v20.dta * CHN2000\_v25.dta * CHN2010\_v5.dta * CHN2010\_v151.dta * CHN2010\_v20.dta * CHN2010\_v25.dta * CHN2015\_v5.dta * CHN2015\_v151.dta * CHN2015\_v20.dta * CHN2015\_v25.dta | Pre-onet.do |
| * CHN\_cons\_to\_SOC2010\_v25 * occ2010\_consistent.dta * occ2015\_consistent.dta | * Lab | * 2000\_occ\_consistent\_characteristics.dta * 2010\_occ\_consistent\_characteristics.dta * 2015\_occ\_consistent\_characteristics.dta | Pre-onet.do |
| * 2000census * 2010census * 2015census * 行业标准转换\census2000-2015\_行业统一版本 * 2000\_occ\_consistent\_characteristics.dta * 2010\_occ\_consistent\_characteristics.dta * 2015\_occ\_consistent\_characteristics.dta | * Lab | * whole\_occ2000.dta * whole\_occ2010.dta * whole\_occ2015.dta | * Pre-census.do |
| * Lab | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| * process-CHN\_ACSSOC2010\_v4.xlsx * ACS-SOC2010.xlsx * CHN-consistent\_ONETSOC2010.dta | * Lab | * CHN-SOC-consistent.dta * USA-SOC-consistent2010.dta * USA-SOC-consistent2000.dta * USA-SOC-consistent2015.dta | Pre-country.do |

**三、正文图表**

|  |  |  |
| --- | --- | --- |
| 编号 | Do file | 原始数据**（未上传）** |
| 图1 | * Fig1.do | * whole\_occ2000.dta * whole\_occ2010.dta * whole\_occ2015.dta |
| 图2 | * Fig2.do | * whole\_occ2000.dta * whole\_occ2010.dta * whole\_occ2015.dta |
| 图3 | * Fig3.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 图4 | * Fig4.do | * 2000census\_consistent.dta * 2015census\_consistent.dta * CHN\_cons\_to\_SOC2010\_v25.dta |
| 图5 | * Fig 5ab.do * Fig5c.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 表1 | * Tab1.do | * whole\_occ2000.dta * whole\_occ2010.dta * whole\_occ2015.dta * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta * CHN-SOC-consistent.dta * USA-SOC-consistent2000.dta * USA-SOC-consistent2010.dta * USA-SOC-consistent2015.dta * cic\_acsind\_chn.dta * cic\_acsind\_usa2000.dta * cic\_acsind\_usa2010.dta * cic\_acsind\_usa2015.dta * Appendix Data\ACS\usa\_00003.dta.dta |
| 表2 | * Tab2.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 表3 | * Tab3.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |

**四、附录图表**

|  |  |  |
| --- | --- | --- |
| 编号 | Do file | 原始数据**（未上传）** |
| 附表Ⅰ1 | * App1 Tab1 and App1 Fig1.do | * Census2000 * Census2010 * Census2015 * Appendix Data\职业和行业就业分布 * Appendix Data\王亚菲行业就业测算结果.xlsx |
| 附表Ⅰ1 | * App1 Tab2.do | * Census2000 * Census2010 * Census2015 * Appendix Data\职业和行业就业分布.xlsx * Appendix Data\01-37 2016中国劳动统计年鉴.xls |
| 附表Ⅰ3 |  | * 笔者根据文献整理所得 |
| 附表Ⅱ1 | * App2 Tab1 and App2 Fig2.do | * Appendix Data\CLDS\individual2014-clds.dta * Appendix Data\ONET N\_version\db\_25\_0\_excel * 香帅的职业数据库\香帅的职业发展指数.xlsx * occ2015\_sem.xls * Appendix Data\chatcheck.xlsx |
| 附表Ⅳ1 | * App4 Tab1.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 附表Ⅳ2 | * App4 Tab2.do | * Appendix Data\CFPS\paneldata\_person.dta * Appendix Data\CFPS\cfps2018crossyearid\_202104.dta * Appendix Data\ONET N\_version\onet20ac * occ2010\_consistent.dta * Appendix Data\CFPS\CFPS2010职业编码对应表.xlsx * Appendix Data\Women Status\wave3.dta * Appendix Data\Women Status\wave2.dta |
| 附图Ⅰ1 | * App1 Tab1 and App1 Fig1.do | * Census2000 * Census2010 * Census2015 * Appendix Data\职业和行业就业分布 * Appendix Data\王亚菲行业就业测算结果.xlsx |
| 附图Ⅰ2 | * App1 Fig2.do | * Census2000 * Census2010 * Appendix Data\职业和行业就业分布.xlsx * 2010\_occ\_consistent\_characteristics.dta |
| 附图Ⅱ1 |  | * 笔者根据匹配过程绘制所得 |
| 附图Ⅱ2系列 | * App2 Tab1 and App2 Fig2.do | * Appendix Data\CLDS\individual2014-clds.dta * Appendix Data\ONET N\_version\db\_25\_0\_excel * Appendix Data\香帅的职业数据库\香帅的职业发展指数.xlsx * Data\occ2015\_sem.xls * Data\chatcheck.xlsx |
| 附图Ⅲ 1 | * App3 Fig1 and App3 Fig2.do | * CHN2000\_v5.dta * CHN2000\_v151.dta * CHN2000\_v20.dta * CHN2000\_v25.dta * CHN2010\_v5.dta * CHN2010\_v151.dta * CHN2010\_v20.dta * CHN2010\_v25.dta * CHN2015\_v5.dta * CHN2015\_v151.dta * CHN2015\_v20.dta * CHN2015\_v25.dta |
| 附图Ⅲ 2 | * App3 Fig1 and App3 Fig2.do | * CHN2000\_v5.dta * CHN2000\_v151.dta * CHN2000\_v20.dta * CHN2000\_v25.dta * CHN2010\_v5.dta * CHN2010\_v151.dta * CHN2010\_v20.dta * CHN2010\_v25.dta * CHN2015\_v5.dta * CHN2015\_v151.dta * CHN2015\_v20.dta * CHN2015\_v25.dta |
| 附图A1 | * App5 FigA1.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 附图Ⅳ1 | * App4 Fig1.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta * Census2000 * Census2010 * Census2015 |
| 附图Ⅳ2 | * App4 Fig2.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 附图Ⅳ3 | * App4 Fig3 and App4 Fig4.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 附图Ⅳ4 | * App4 Fig3 and App4 Fig4.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 附图Ⅳ5 | * App4 Fig5.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |
| 附图Ⅳ6 | * App4 Fig6.do | * 2000census\_consistent.dta * 2010census\_consistent.dta * 2015census\_consistent.dta |