**Changes made to data**

**Splitting ethnic groups**

| Initial | Change |
| --- | --- |
| pole and italian | + polish +italian – pole and italian |
| english and german | + english + german – english and german |
| roumanian and turkis | + turkish + roumanian – roumanian and turkis |
| spanish, mexican | + spanish + mexican – spnaish, mexican |
| jwish | + jewish – jwish |
| polish, slavic | + polish + slavic -polish, slavic |
| armenian,syrian,turkish,greek | + armenian + syrian + turkish + greek - armenian,syrian,turkish,greek |
| norwegian and swedish | + norwegian + swedish – norwegian and swedish |
| frnch | + french - frnch |
| portugaise | + portuguese – portugaise |
| german-jewish | + jewish – german-jewish |
| armenian, persian, turkish, greek | + armenian + persian + turkish + greek - armenian, persian, turkish, greek |
| germany and german | replace all germany for german |
| pole and polish | replace all pole for polish |
| french and france | replace all for french |
| portuguese and portugal | replace all for portugese |
| slovenian, slavic | + slavic – slovenian, slavic |
| spanish and portuguese | + spanish + portugese – spanish and portuguese |
| spanish/mexican | + spanish + mexican – spanish/mexican |
| french and austrian | + french + austrian – french and austrian |
| hamburg | - hamburg + german |
| french and german | + french + german – french and german |
| turkish and greek | + turkish + greek – turkish and greek |
| greek and turkish | + greek + turkish – greek and turkish |
| american, english | + american + english – american, english |
| londoner | + english – londoner |
| sicilian | + italian – sicilian |
| parisian | + french – parisian |
| venetian | + italian – venetian |
| bavarian | + german – bavarian |
| berlin | + german – berlin |
| boulogne | + french – boulogne |
| britanny | + french – britanny |
| milanese | + italian – milanese |
| vienna | + austrian – vienna |
| genevan | + swiss – genevan |
| flanders | + belgian – flanders |
| genoese | + italian – genoese |
| mirabeau | + french - mirabeau |
| montpellier | + french – montpellier |
| neapolitan | + italian – neapolitan |
| norman | + french – norman |
| rouen | + french – rouen |
| oiental | + oriental – oiental |
| russia | + russian – russia |
| saxony | + german – saxony |
| saxon | + german – saxony |
| scotch | - scotch + scottish |
| tyrolian | + tyrolean - tyrolian |
| suabian | + german – suabian |
| roman, napolishon, mazarin, circassian, coblentz, alsatian |  |

**Removal of categories with not enough data: every ethnical group with entries < 10**

| Old list | New list | |
| --- | --- | --- |
| | arabian | | --- | | african | | mazarin | | napolishon | | coblentz | | polish | | puerto rican | | tyrolean | | egyptian | | syrian | | japanese | | portuguese | | polish | | oriental | | french | | moorish | | circassian | | brazilian | | mexican | | hispanic/mexican | | austrian | | roman | | alsatian | | dutch | | greek | | (blank) | | norwegian | | persian | | creole | | armenian | | italian | | slavic | | scottish | | russian | | jewish | | american | | turkish | | roumanian | | belgian | | english | | spanish | | chinese | | swiss | | german | | danish | | swedish | | hungarian | | indian | | irish | | cuban | | havanese | | french | | purpose | | | japanese | | --- | | portuguese | | polish | | oriental | | french | | mexican | | greek | | persian | | creole | | armenian | | italian | | slavic | | scottish | | russian | | jewish | | turkish | | english | | spanish | | chinese | | swiss | | german | | swedish | | hungarian | | indian | | irish | |

**List of removed outliers analysis relationship year and unhealthy score**

Removed: 1827,1846.1867.1912 and 1920. because n<90