

# Where did we come from? Where might we go?

Tidewater Big Data Enthusiasts

Chuck Cartledge

Developer

September 27, 2016 at 1:33pm

## Contents

|                            |    |
|----------------------------|----|
| <b>List of Figures</b>     | i  |
| <b>1 Introduction</b>      | 1  |
| <b>2 Data sources</b>      | 1  |
| <b>3 Data manipulation</b> | 1  |
| <b>4 Results</b>           | 6  |
| <b>5 Conclusion</b>        | 10 |
| <b>A Misc. files</b>       | 10 |

## List of Figures

|   |                                     |   |
|---|-------------------------------------|---|
| 1 | Pre-1990 raw census file.           | 3 |
| 2 | 2000 raw census file.               | 4 |
| 3 | 2010 raw census file.               | 5 |
| 4 | Normalized population distribution. | 7 |
| 5 | Raw population data.                | 8 |
| 6 | Normalized growth by county.        | 9 |

# 1 Introduction

The US population is constantly moving. From the East westward. From the North southward, and then back. Is there a way to visualize these movements over time?

We'll dive into the US Census databases looking for population density data. We expect that we'll easily find data for recent censuses, earlier ones may be a challenge. The data should have at least these attributes: location, magnitude, and time. It may also have these: age, gender, race, or type of household. We'll see what data is available, and how decide on how to visualize the changes over time.

## 2 Data sources

We are looking at US census data as a function of time; so the United States Census Bureau<sup>1</sup> is the source for all data, including the geographic data. Surprisingly enough the data took a little bit of digging to find, and ultimately I was not able to locate all the data I had hoped to find.

My initial goal was to have a fine grained look at population movements as a function of time, where fine grained meant age, gender, race, and income down to the county level. I was finally able to locate some of the data in three different files, and three different formats. Copies of all files are listed in the Misc. files section (see Section A on page 10).

The files were located at:

Pre 2000 <https://www.census.gov/population/www/censusdata/pop1790-1990.html>

2000 <https://www.census.gov/geo/maps-data/data/gazetteer2000.html>

2010 <https://www.census.gov/geo/maps-data/data/gazetteer2010.html>

I believe that the data exists somewhere on the US Census site, but it was too hard to find in the time I allocated.

Having the census data was part of the problem, the other was how to map the county population to a geographic location. Again, the US Census has a partial solution by providing cartographic shapefiles to the county level for different years. The shapefiles are located at:

[https://www.census.gov/geo/maps-data/data/cbf/cbf\\_counties.html](https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html)

The 2015 shapefiles were used.

## 3 Data manipulation

A Postgres database as created to store the geographic and population data. The data is very simple in format and could have been managed by an R data frame. During software

---

<sup>1</sup><http://www.census.gov/>

development, there were many different iterations of the software and it just seemed to make more sense to have the data persisted in a database rather than re-create it each time. At the end, an R data frame would have had faster execution, but the database was fast enough.

Each data file had to be handled differently. The pre-2000 file (see Figure 1 on the next page) required the most hands on tweaking, and resulted in a standardized file with consistent columns and format. The 2000 file (see Figure 2 on page 4) is a simple column oriented flat file. Data fields are not delimited by any characters. The 2010 file (see Figure 3 on page 5) has one line per county, and data fields are delimited by tab characters.

The county shapefiles were used without change or modification. The “labpt” associated with the county was used as the geographic centroid of the county. The “labpt” was used as the geographic point associated with the population data.

The various data files list approximately 3,300 counties in the US. This does not include those in Puerto Rico, Guam, Samoa, or other territories or possessions. Nor does the current location necessarily represent where the county was in the past. More time related data is available from the Census website, but was not used because of time limitations. There are approximately 150 county name mismatches between the census data and the geographic data.

Once the data was stored in the database, various queries and analyzes were conducted. Each of these efforts resulted in a single PNG image. These images were consolidated into a GIF using this command:

```
convert -delay 100 -loop 0 *png pop.gif
```

in the directory where the images were created.

| Population_Part.xls - LibreOffice Calc                    |   |
|---|---|
| Area  | W279  |
| A   | B   |
| B   | C   |
| C   | D   |
| D   | E   |
| E   | F   |
| F   | G   |
| G   | H   |
| H   | I   |
| I   | J   |
| J   | K   |
| K   | L   |
| L   | M   |
| M   | N   |
| N   | O   |
| O   | P   |
| P   | Q   |
| Q   | R   |
| R   | S   |
| S   | T   |
| T   | U   |
| U   | V   |
| V   | W   |
| W   | X   |
| X   | Y   |
| Y   | Z   |
| Z   |   |
| Part III. Population of Counties, Earliest Census to 1990 |   |
| 1   | 2   |
| 3   | POPULATION OF COUNTIES - ALABAMA, 1880-1990 |
| 4   | 1880  |
| 5   | 1890  |
| 6   | 1900  |
| 7   | 1910  |
| 8   | 1920  |
| 9   | 1930  |
| 10  | 1940  |
| 11  | 1950  |
| 12  | 1960  |
| 13  | 1970  |
| 14  | 1980  |
| 15  | 1990  |
| 16  | 1990  |
| 17  | 1990  |
| 18  | 1990  |
| 19  | 1990  |
| 20  | 1990  |
| 21  | 1990  |
| 22  | 1990  |
| 23  | 1990  |
| 24  | 1990  |
| 25  | 1990  |
| 26  | 1990  |
| 27  | 1990  |
| 28  | 1990  |
| 29  | 1990  |
| 30  | 1990  |
| 31  | 1990  |
| 32  | 1990  |
| 33  | 1990  |
| 34  | 1990  |
| 35  | 1990  |
| 36  | 1990  |
| 37  | 1990  |
| 38  | 1990  |
| 39  | 1990  |
| 40  | 1990  |
| 41  | 1990  |
| 42  | 1990  |
| 43  | 1990  |
| 44  | 1990  |
| 45  | 1990  |
| 46  | 1990  |
| 47  | 1990  |
| 48  | 1990  |
| 49  | 1990  |
| 50  | 1990  |
| 51  | 1990  |
| 52  | 1990  |
| 53  | 1990  |
| 54  | 1990  |
| 55  | 1990  |
| 56  | 1990  |
| 57  | 1990  |
| 58  | 1990  |
| 59  | 1990  |
| 60  | 1990  |
| 61  | 1990  |
| 62  | 1990  |
| 63  | 1990  |
| 64  | 1990  |
| 65  | 1990  |
| 66  | 1990  |
| 67  | 1990  |
| 68  | 1990  |
| 69  | 1990  |
| 70  | 1990  |
| 71  | 1990  |
| 72  | 1990  |
| 73  | 1990  |
| 74  | 1990  |
| 75  | 1990  |
| 76  | 1990  |
| 77  | 1990  |
| 78  | 1990  |
| 79  | 1990  |
| 80  | 1990  |
| 81  | 1990  |
| 82  | 1990  |
| 83  | 1990  |
| 84  | 1990  |
| 85  | 1990  |
| 86  | 1990  |
| 87  | 1990  |
| 88  | 1990  |
| 89  | 1990  |
| 90  | 1990  |
| 91  | 1990  |
| 92  | 1990  |
| 93  | 1990  |
| 94  | 1990  |
| 95  | 1990  |
| 96  | 1990  |
| 97  | 1990  |
| 98  | 1990  |
| 99  | 1990  |
| 100   | 1990  |
| 101   | 1990  |
| 102   | 1990  |
| 103   | 1990  |
| 104   | 1990  |
| 105   | 1990  |
| 106   | 1990  |
| 107   | 1990  |
| 108   | 1990  |
| 109   | 1990  |
| 110   | 1990  |
| 111   | 1990  |
| 112   | 1990  |
| 113   | 1990  |
| 114   | 1990  |
| 115   | 1990  |
| 116   | 1990  |
| 117   | 1990  |
| 118   | 1990  |
| 119   | 1990  |
| 120   | 1990  |
| 121   | 1990  |
| 122   | 1990  |
| 123   | 1990  |
| 124   | 1990  |
| 125   | 1990  |
| 126   | 1990  |
| 127   | 1990  |
| 128   | 1990  |
| 129   | 1990  |
| 130   | 1990  |
| 131   | 1990  |
| 132   | 1990  |
| 133   | 1990  |
| 134   | 1990  |
| 135   | 1990  |
| 136   | 1990  |
| 137   | 1990  |
| 138   | 1990  |
| 139   | 1990  |
| 140   | 1990  |
| 141   | 1990  |
| 142   | 1990  |
| 143   | 1990  |
| 144   | 1990  |
| 145   | 1990  |
| 146   | 1990  |
| 147   | 1990  |
| 148   | 1990  |
| 149   | 1990  |
| 150   | 1990  |
| 151   | 1990  |
| 152   | 1990  |
| 153   | 1990  |
| 154   | 1990  |
| 155   | 1990  |
| 156   | 1990  |
| 157   | 1990  |
| 158   | 1990  |
| 159   | 1990  |
| 160   | 1990  |
| 161   | 1990  |
| 162   | 1990  |
| 163   | 1990  |
| 164   | 1990  |
| 165   | 1990  |
| 166   | 1990  |
| 167   | 1990  |
| 168   | 1990  |
| 169   | 1990  |
| 170   | 1990  |
| 171   | 1990  |
| 172   | 1990  |
| 173   | 1990  |
| 174   | 1990  |
| 175   | 1990  |
| 176   | 1990  |
| 177   | 1990  |
| 178   | 1990  |
| 179   | 1990  |
| 180   | 1990  |
| 181   | 1990  |
| 182   | 1990  |
| 183   | 1990  |
| 184   | 1990  |
| 185   | 1990  |
| 186   | 1990  |
| 187   | 1990  |
| 188   | 1990  |
| 189   | 1990  |
| 190   | 1990  |
| 191   | 1990  |
| 192   | 1990  |
| 193   | 1990  |
| 194   | 1990  |
| 195   | 1990  |
| 196   | 1990  |
| 197   | 1990  |
| 198   | 1990  |
| 199   | 1990  |
| 200   | 1990  |
| 201   | 1990  |
| 202   | 1990  |
| 203   | 1990  |
| 204   | 1990  |
| 205   | 1990  |
| 206   | 1990  |
| 207   | 1990  |
| 208   | 1990  |
| 209   | 1990  |
| 210   | 1990  |
| 211   | 1990  |
| 212   | 1990  |
| 213   | 1990  |
| 214   | 1990  |
| 215   | 1990  |
| 216   | 1990  |
| 217   | 1990  |
| 218   | 1990  |
| 219   | 1990  |
| 220   | 1990  |
| 221   | 1990  |
| 222   | 1990  |
| 223   | 1990  |
| 224   | 1990  |
| 225   | 1990  |
| 226   | 1990  |
| 227   | 1990  |
| 228   | 1990  |
| 229   | 1990  |
| 230   | 1990  |
| 231   | 1990  |
| 232   | 1990  |
| 233   | 1990  |
| 234   | 1990  |
| 235   | 1990  |
| 236   | 1990  |
| 237   | 1990  |
| 238   | 1990  |
| 239   | 1990  |
| 240   | 1990  |
| 241   | 1990  |
| 242   | 1990  |
| 243   | 1990  |
| 244   | 1990  |
| 245   | 1990  |
| 246   | 1990  |
| 247   | 1990  |
| 248   | 1990  |
| 249   | 1990  |
| 250   | 1990  |
| 251   | 1990  |
| 252   | 1990  |
| 253   | 1990  |
| 254   | 1990  |
| 255   | 1990  |
| 256   | 1990  |
| 257   | 1990  |
| 258   | 1990  |
| 259   | 1990  |
| 260   | 1990  |
| 261   | 1990  |
| 262   | 1990  |
| 263   | 1990  |
| 264   | 1990  |
| 265   | 1990  |
| 266   | 1990  |
| 267   | 1990  |
| 268   | 1990  |
| 269   | 1990  |
| 270   | 1990  |
| 271   | 1990  |
| 272   | 1990  |
| 273   | 1990  |
| 274   | 1990  |
| 275   | 1990  |
| 276   | 1990  |
| 277   | 1990  |
| 278   | 1990  |
| 279   | 1990  |
| 280   | 1990  |
| 281   | 1990  |
| 282   | 1990  |
| 283   | 1990  |
| 284   | 1990  |
| 285   | 1990  |
| 286   | 1990  |
| 287   | 1990  |
| 288   | 1990  |
| 289   | 1990  |
| 290   | 1990  |
| 291   | 1990  |
| 292   | 1990  |
| 293   | 1990  |
| 294   | 1990  |
| 295   | 1990  |
| 296   | 1990  |
| 297   | 1990  |
| 298   | 1990  |
| 299   | 1990  |
| 300   | 1990  |
| 301   | 1990  |
| 302   | 1990  |
| 303   | 1990  |
| 304   | 1990  |
| 305   | 1990  |
| 306   | 1990  |
| 307   | 1990  |
| 308   | 1990  |
| 309   | 1990  |
| 310   | 1990  |
| 311   | 1990  |
| 312   | 1990  |
| 313   | 1990  |
| 314   | 1990  |
| 315   | 1990  |
| 316   | 1990  |
| 317   | 1990  |
| 318   | 1990  |
| 319   | 1990  |
| 320   | 1990  |
| 321   | 1990  |
| 322   | 1990  |
| 323   | 1990  |
| 324   | 1990  |
| 325   | 1990  |
| 326   | 1990  |
| 327   | 1990  |
| 328   | 1990  |
| 329   | 1990  |
| 330   | 1990  |
| 331   | 1990  |
| 332   | 1990  |
| 333   | 1990  |
| 334   | 1990  |
| 335   | 1990  |
| 336   | 1990  |
| 337   | 1990  |
| 338   | 1990  |
| 339   | 1990  |
| 340   | 1990  |
| 341   | 1990  |
| 342   | 1990  |
| 343   | 1990  |
| 344   | 1990  |
| 345   | 1990  |
| 346   | 1990  |
| 347   | 1990  |
| 348   | 1990  |
| 349   | 1990  |
| 350   | 1990  |
| 351   | 1990  |
| 352   | 1990  |
| 353   | 1990  |
| 354   | 1990  |
| 355   | 1990  |
| 356   | 1990  |
| 357   | 1990  |
| 358   | 1990  |
| 359   | 1990  |
| 360   | 1990  |
| 361   | 1990  |
| 362   | 1990  |
| 363   | 1990  |
| 364   | 1990  |
| 365   | 1990  |
| 366   | 1990  |
| 367   | 1990  |
| 368   | 1990  |
| 369   | 1990  |
| 370   | 1990  |
| 371   | 1990  |
| 372   | 1990  |
| 373   | 1990  |
| 374   | 1990  |
| 375   | 1990  |
| 376   | 1990  |
| 377   | 1990  |
| 378   | 1990  |
| 379   | 1990  |
| 380   | 1990  |
| 381   | 1990  |
| 382   | 1990  |
| 383   | 1990  |
| 384   | 1990  |
| 385   | 1990  |
| 386   | 1990  |
| 387   | 1990  |
| 388   | 1990  |
| 389   | 1990  |
| 390   | 1990  |
| 391   | 1990  |
| 392   | 1990  |
| 393   | 1990  |
| 394   | 1990  |
| 395   | 1990  |
| 396   | 1990  |
| 397   | 1990  |
| 398   | 1990  |
| 399   | 1990  |
| 400   | 1990  |
| 401   | 1990  |
| 402   | 1990  |
| 403   | 1990  |
| 404   | 1990  |
| 405   | 1990  |
| 406   | 1990  |
| 407   | 1990  |
| 408   | 1990  |
| 409   | 1990  |
| 410   | 1990  |
| 411   | 1990  |
| 412   | 1990  |
| 413   | 1990  |
| 414   | 1990  |
| 415   | 1990  |
| 416   | 1990  |
| 417   | 1990  |
| 418   | 1990  |
| 419   | 1990  |
| 420   | 1990  |
| 421   | 1990  |
| 422   | 1990  |
| 423   | 1990  |
| 424   | 1990  |
| 425   | 1990  |
| 426   | 1990  |
| 427   | 1990  |
| 428   | 1990  |
| 429   | 1990  |
| 430   | 1990  |
| 431   | 1990  |
| 432   | 1990  |
| 433   | 1990  |
| 434   | 1990  |
| 435   | 1990  |
| 436   | 1990  |
| 437   | 1990  |
| 438   | 1990  |
| 439   | 1990  |
| 440   | 1990  |
| 441   | 1990  |
| 442   | 1990  |
| 443   | 1990  |
| 444   | 1990  |
| 445   | 1990  |
| 446   | 1990  |
| 447   | 1990  |
| 448   | 1990  |
| 449   | 1990  |
| 450   | 1990  |
| 451   | 1990  |
| 452   | 1990  |
| 453   | 1990  |
| 454   | 1990  |
| 455   | 1990  |
| 456   | 1990  |
| 457   | 1990  |
| 458   | 1990  |
| 459   | 1990  |
| 460   | 1990  |
| 461   | 1990  |
| 462   | 1990  |
| 463   | 1990  |
| 464   | 1990  |
| 465   | 1990  |
| 466   | 1990  |
| 467   | 1990  |
| 468   | 1990  |
| 469   | 1990  |
| 470   | 1990  |
| 471   | 1990  |
| 472   | 1990  |
| 473   | 1990  |
| 474   | 1990  |
| 475   | 1990  |
| 476   | 1990  |
| 477   | 1990  |
| 478   | 1990  |
| 479   | 1990  |
| 480   | 1990  |
| 481   | 1990  |
| 482   | 1990  |
| 483   | 1990  |
| 484   | 1990  |
| 485   | 1990  |
| 486   | 1990  |
| 487   | 1990  |
| 488   | 1990  |
| 489   | 1990  |
| 490   | 1990  |
| 491   | 1990  |
| 492   | 1990  |
| 493   | 1990  |
| 494   | 1990  |
| 495   | 1990  |
| 496   | 1990  |
| 497   | 1990  |
| 498   | 1990  |
| 499   | 1990  |
| 500   | 1990  |
| 501   | 1990  |
| 502   | 1990  |
| 503   | 1990  |
| 504   | 1990  |
| 505   | 1990  |
| 506   | 1990  |
| 507   | 1990  |
| 508   | 1990  |
| 509   | 1990  |
| 510   | 1990  |
| 511   | 1990  |
| 512   | 1990  |
| 513   | 1990  |
| 514   | 1990  |
| 515   | 1990  |
| 516   | 1990  |
| 517   | 1990  |
| 518   | 1990  |
| 519   | 1990  |
| 520   | 1990  |
| 521   | 1990  |
| 522   | 1990  |
| 523   | 1990  |
| 524   | 1990  |
| 525   | 1990  |
| 526   | 1990  |
| 527   | 1990  |
| 528   | 1990  |
| 529   | 1990  |
| 530   | 1990  |
| 531   | 1990  |
| 532   | 1990  |
| 533   | 1990  |
| 534   | 1990  |
| 535   | 1990  |
| 536   | 1990  |
| 537   | 1990  |
| 538   | 1990  |
| 539   | 1990  |
| 540   | 1990  |
| 541   | 1990  |
| 542   | 1990  |
| 543   | 1990  |
| 544   | 1990  |
| 545   | 1990  |
| 546   | 1990  |
| 547   | 1990  |
| 548   | 1990  |
| 549   | 1990  |
| 550   | 1990  |
| 551   | 1990  |
| 552   | 1990  |
| 553   | 1990  |
| 554   | 1990  |
| 555   | 1990  |
| 556   | 1990  |
| 557   | 1990  |
| 558   | 1990  |
| 559   | 1990  |
| 560   | 1990  |
| 561   | 1990  |
| 562   | 199   |

Figure 1: Pre-1990 raw census file. The file contains a lot of interesting, but not useful information. Things like: a Notes column to the right, when the state was not part of the Union, triple dashes when no data was available (because no one was there, or the county didn't exist), a textual section after the state that addresses the notes. Each of these non-useful items had to be removed by hand for smooth processing.

| File Edit Options Buffers Tools Text VirtualEnv Help |  |
|--|--|
| AL190938A Baldwin County                             | 1136158094 1596.346694 430.562346 36.592775 -87.748266   |
| AL190938Barbour County                               | 229188972 1596.983349 884.983349 31.546553 -85.331312  |
| AL190938Bibb County                                  | 1613627425 8345.093585 623.593585 33.546989 -87.12243  |
| AL190938Blount County                                | 211582 1672657889 1285.585188 645.585188 5.018892 3.377040 -86.584768                                    |
| AL190938Butler County                                | 1618775183 4727 12998871 655.093582 2.998925 -86.589165  |
| AL190938Cahaba County                                | 112399 9557 2012676972 1.643659 776.867795 1.654861 1.738884 -86.682222                                  |
| AL190938Cullman County                               | 15758906972 51522 159569154 3.859609 33.745939 1.654861 1.738884 -85.81544                               |
| AL190938Chambers County                              | 365838 16556 14865221 15388288 597.18613 1.654861 1.738884 -85.288548                                    |
| AL190938Clayton County                               | 112249 1153 16965221 117552 675.9877 1.654861 1.738884 -85.288548  |
| AL190938Cleburne County                              | 30593 70593 17553 17975964 46.7729 29.18413 1.654861 1.738884 -85.288548                                 |
| AL190938Crittenden County                            | 159222 28626 17553 28626 675.9877 1.654861 1.738884 -85.288548   |
| AL190938Cochran County                               | 186999651 1596.580217 913.513822 7.335575 31.227683 -87.834164   |
| AL190938Clarke County                                | 27867 12331 328737186 14.129916 31.688155 609.580217 1.6.392822 34.155683 -86.839179                     |
| AL190938Clay County                                  | 14254 6612 1567134311 1458946389 605.093586 6.927817 33.265880 -85.848258                                |
| AL190938Cleburne County                              | 6189 14254 1567134311 1458946389 605.093586 6.927817 33.265880 -85.848258                                |
| AL190938Coffee County                                | 43615 19837 20450 2348933739 1.689814 6.868682 33.547988 -85.51937                                       |
| AL190938Cohart County                                | 14215 24980 1593831753 175876346 38546633 678.998152 1.4.488381 31.739887 -85.958641                     |
| AL190938Conklin County                               | 14089 24980 1593831753 175876346 38546633 678.998152 1.4.488381 31.739887 -85.958641                     |
| AL190938Conecuh County                               | 14889 2765 2283546316 4436304 856.734625 1.7.72880 31.416613 -87.082168                                  |
| AL190938Cosa County                                  | 12202 6142 1689828597 13.914884 32.595511 -86.236319   |
| AL190938Covington County                             | 37631 18378 2677571159 10.846616 31.251036 -86.416195  |
| AL190938Crenshaw County                              | 13665 6644 1578805448 1.295137 31.10305 -86.289299   |
| AL190938Dallas County                                | 77483 35233 1912519740 42457211 738.428831 1.6.392822 34.155683 -86.839179                               |
| AL190938DeKalb County                                | 49129 21179 1453174585 14.6847 6.1515013   |
| AL190938Dade County                                  | 46365 20450 2348933739 1.689814 6.868682 33.547988 -85.848258  |
| AL190938Dekalb County                                | 42851 169984562 169984562 1.689814 6.868682 33.547988 -85.848258   |
| AL190938Dent County                                  | 65874 25733 169984562 1.689814 6.868682 33.547988 -85.848258   |
| AL190938Fayette County                               | 38448 16544 2433716269 147.322948 621.255085 35.546706 1.4.59133 1.7.72880 31.416613 -87.082168          |
| AL190938Franklin County                              | 103459 45359 138518532 148659152 148659152 1.6.392822 34.155683 -86.839179                               |
| AL190938Gadsden County                               | 18495 8472 1625332174 1.689814 6.868682 33.547988 -85.848258   |
| AL190938Hale County                                  | 31223 13149 1625332174 1.689814 6.868682 33.547988 -85.848258  |
| AL190938Henry County                                 | 2954 145584452 145584452 1.6.392822 34.155683 -86.839179   |
| AL190938Hickman County                               | 9974 5111 1627278267 1.689814 6.868682 33.547988 -85.848258  |
| AL190938Hinds County                                 | 17185 8037 145584452 1.6.392822 34.155683 -86.839179   |
| AL190938Houston County                               | 16310 8037 145584452 1.6.392822 34.155683 -86.839179   |
| AL190938Jackson County                               | 88787 3937 1583184582 135124582 1.689814 6.868682 33.547988 -85.848258                                   |
| AL190938Jefferson County                             | 53927 14168 2793932065 1471241120 161969592 1.689814 6.868682 33.547988 -85.848258                       |
| AL190938Jeffersonson County                          | 662467 288162 2881633946 1566546589 1566546589 1.6.392822 34.155683 -86.839179                           |
| AL190938Jasper County                                | 797944 7517 1566546589 1566546589 1.6.392822 34.155683 -86.839179  |
| AL190938Lauderdale County                            | 879766 40424 1733895504 127745460 669.468025 49.617974 33.227683 -87.685706                              |
| AL190938Lawrence County                              | 34803 15099 1795852273 639328499 693.312469 7.157412 32.332124 1.6.686987 31.739887 -85.51937            |
| AL190938Madison County                               | 115992 50329 1576552257 17223894 688.710555 6.842925 32.332124 -87.382184                                |
| AL190938Marion County                                | 65676 26897 1468677998 106498676 1233.98096 416.93196 30.722236 -88.139667                               |
| AL190938Morgan County                                | 13473 5801 1859446669 18378683 717.936488 7.095347 2.171280 -86.659917                                   |
| AL190938Montgomery County                            | 24165 10627 1581247728 7042132 610.523384 2.718979 32.421555 -85.721158                                  |
| AL190938Nashville County                             | 2767608 18288 2084739566 263352053 684.922458 49.617974 33.227683 -86.588226                             |
| AL190938Perry County                                 | 252539 10127 25385233970 15031202 977.044677 5.886761 32.332124 -87.788789                               |
| AL190938Perdido County                               | 31214 14416 192853331 5.886761 32.332124 -87.788789  |
| AL190938Shelby County                                | 82231 16331 1468677998 106498676 1233.98096 416.93196 30.722236 -88.139667                               |
| AL190938St. Louis County                             | 399843 165101 31393698623 2257598 1055.845959 8.695924 31.55536 87.359837                                |
| AL190938Tipton County                                | 235310 225310 204547293 2588019 789.734625 0.1.... county2.txt (comtypes.zip) Top L1 (Text Archive Path) |

Figure 2: 2000 raw census file. The file has one line per county, and is column oriented (i.e., there are no delimiters between data fields).

| File | Edit  | Buffers  | Tools  | Text   | VirtualEnv | Help          |
|------|-------|----------|--|--------|------------|---------------|
| USPS | AL    | ANISCODE | NAME   | POP10  | HU19       | ALAND         |
| AL   | 01001 | 00161526 | Autauga County                                     | 54271  | 22135      | 134952276     |
| AL   | 01003 | 00161527 | Baldwin County                                     | 182825 | 104861     | 257175735     |
| AL   | 01005 | 00161528 | Baldwin County                                     | 27457  | 11829      | 11133190211   |
| AL   | 01007 | 00161529 | Bibb County  | 22915  | 8981       | 508647216     |
| AL   | 01009 | 00161530 | Blount County                                      | 57212  | 23887      | 92989057      |
| AL   | 01011 | 00161531 | Bullcock County                                    | 16914  | 4493       | 15157440      |
| AL   | 01013 | 00161532 | Butler County                                      | 20847  | 9564       | 6866228       |
| AL   | 01015 | 00161533 | Cahoon County                                      | 118572 | 5289       | 26726134      |
| AL   | 01017 | 00161534 | Chambers County                                    | 34215  | 17084      | 16624267      |
| AL   | 01019 | 00161535 | Cherokee County                                    | 15598  | 16267      | 10481412      |
| AL   | 01021 | 00161536 | Chilton County                                     | 13043  | 19278      | 119883898     |
| AL   | 01023 | 00161537 | Choctaw County                                     | 13059  | 19278      | 19052531      |
| AL   | 01025 | 00161538 | Clarke County                                      | 25833  | 12638      | 362659488     |
| AL   | 01027 | 00161539 | Clay County  | 13932  | 776        | 5234628       |
| AL   | 01029 | 00161540 | Cleburne County                                    | 44912  | 6118       | 156222280     |
| AL   | 01031 | 00161541 | Coffee County                                      | 2739   | 1586522167 | 2358581       |
| AL   | 01033 | 00161542 | Colbert County                                     | 54278  | 7899       | 12389733      |
| AL   | 01035 | 00161543 | Conecuh County                                     | 33226  | 23884769   | 6644946       |
| AL   | 01037 | 00161544 | Cosa County  | 11539  | 6478       | 399465646     |
| AL   | 01039 | 00161545 | Covington County                                   | 37765  | 18829      | 40686955      |
| AL   | 01041 | 00161546 | Crenshaw County                                    | 13986  | 6735       | 266886866     |
| AL   | 01043 | 00161547 | Dale County  | 80466  | 37654      | 106222833     |
| AL   | 01045 | 00161548 | Dallas County                                      | 43820  | 20268      | 23348807386   |
| AL   | 01047 | 00161549 | DeKalb County                                      | 71169  | 31189      | 4013826       |
| AL   | 01049 | 00161550 | Elmore County                                      | 11938  | 32657      | 1901867593    |
| AL   | 01051 | 00161551 | Escambia County                                    | 38319  | 64648      | 244774632     |
| AL   | 01053 | 00161552 | Eufaula County                                     | 104430 | 47454      | 1385618894    |
| AL   | 01055 | 00161553 | Fayette County                                     | 17241  | 8437       | 4326531609    |
| AL   | 01057 | 00161554 | Franklin County                                    | 31704  | 14022      | 164587713     |
| AL   | 01059 | 00161555 | Geneva County                                      | 9045   | 5987       | 1070807032    |
| AL   | 01061 | 00161556 | Hale County  | 15760  | 6551       | 1687803514    |
| AL   | 01063 | 00161557 | Henry County                                       | 19303  | 16267      | 1459676558    |
| AL   | 01065 | 00161558 | Houston County                                     | 101547 | 8891       | 1459770720    |
| AL   | 01067 | 00161559 | Ivyton County                                      | 53327  | 104430     | 1759851794    |
| AL   | 01069 | 00161560 | Jefferson County                                   | 15686  | 2586       | 2078192269    |
| AL   | 01071 | 00161561 | Kosciusko County                                   | 16156  | 168455     | 16066379      |
| AL   | 01073 | 00161562 | Lake County  | 14564  | 7354       | 1465546454    |
| AL   | 01075 | 00161563 | Lauderdale County                                  | 92799  | 43751      | 17932791      |
| AL   | 01077 | 00161564 | Lawrence County                                    | 44339  | 15229      | 1788646568    |
| AL   | 01079 | 00161565 | Lee County   | 146247 | 63391      | 1737510228    |
| AL   | 01081 | 00161566 | Limestone County                                   | 82782  | 30977      | 145622336     |
| AL   | 01083 | 00161567 | Lowndes County                                     | 11299  | 5149       | 105219648     |
| AL   | 01085 | 00161568 | Macon County                                       | 21452  | 10259      | 1377065215    |
| AL   | 01087 | 00161569 | Madison County                                     | 334811 | 164447     | 207116529     |
| AL   | 01089 | 00161570 | Marengo County                                     | 21027  | 10237      | 2303112461    |
| AL   | 01091 | 00161571 | Morgan County                                      | 30776  | 14737      | 1822556826    |
| AL   | 01093 | 00161572 | Marshall County                                    | 33019  | 40342      | 146533556     |
| AL   | 01095 | 00161573 | Mobile County                                      | 112992 | 178196     | 318422378     |
| AL   | 01097 | 00161574 | Monroe County                                      | 23068  | 11333      | 2056485275    |
| AL   | 01099 | 00161575 | Truncate Long Lines enabled                        |        |            | 22643193      |
| AL   |       |          | U:---- Gz_counties_national.txt Top L1 (Text) pair |        |            | -87.383266    |
| AL   |       |          | Truncate Long Lines enabled                        |        |            | -87.31.588332 |

Figure 3: 2010 raw census file. The file has one line per county with tab characters between data fields.

## 4 Results

Census data was analyzed to answer three different questions.

1. What was the relative distribution of population as a function of time? (see Figure 4 on the following page)
2. What was the absolute distribution of population as a function of time? (see Figure 5 on page 8)
3. What areas had the greatest percentage of change in population as a function of time? (see Figure 6 on page 9)

The static images do not convey how national and international events affect the movement of population across the continental United States. Things like the Westward Expansion, World War 2, rise of the West Coast are readily evident in the animated versions of the data (see Section A on page 10).

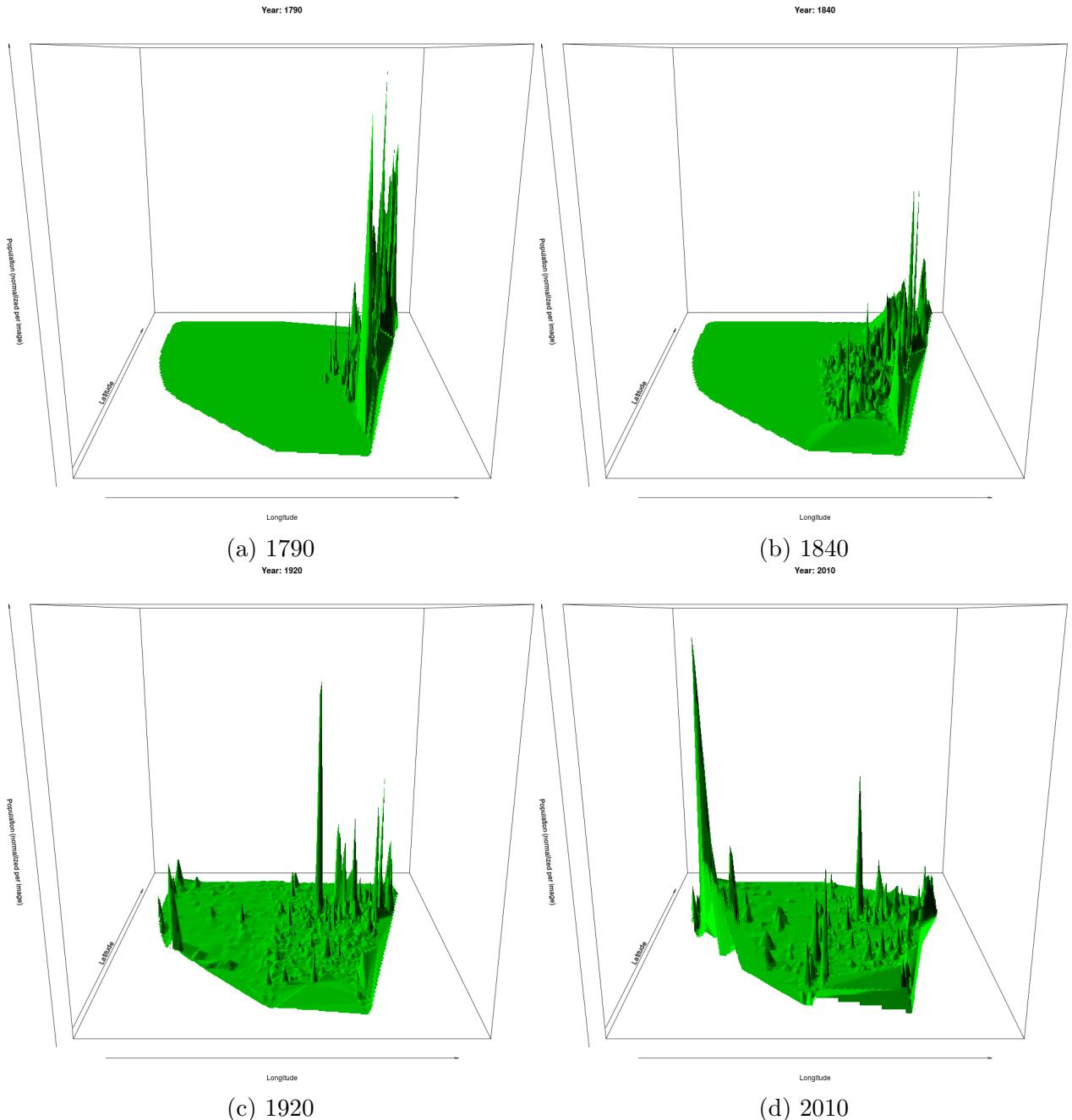


Figure 4: Normalized population distribution. The normalized population starts off on the East Coast, and then around 1950 shifts to the West Coast.

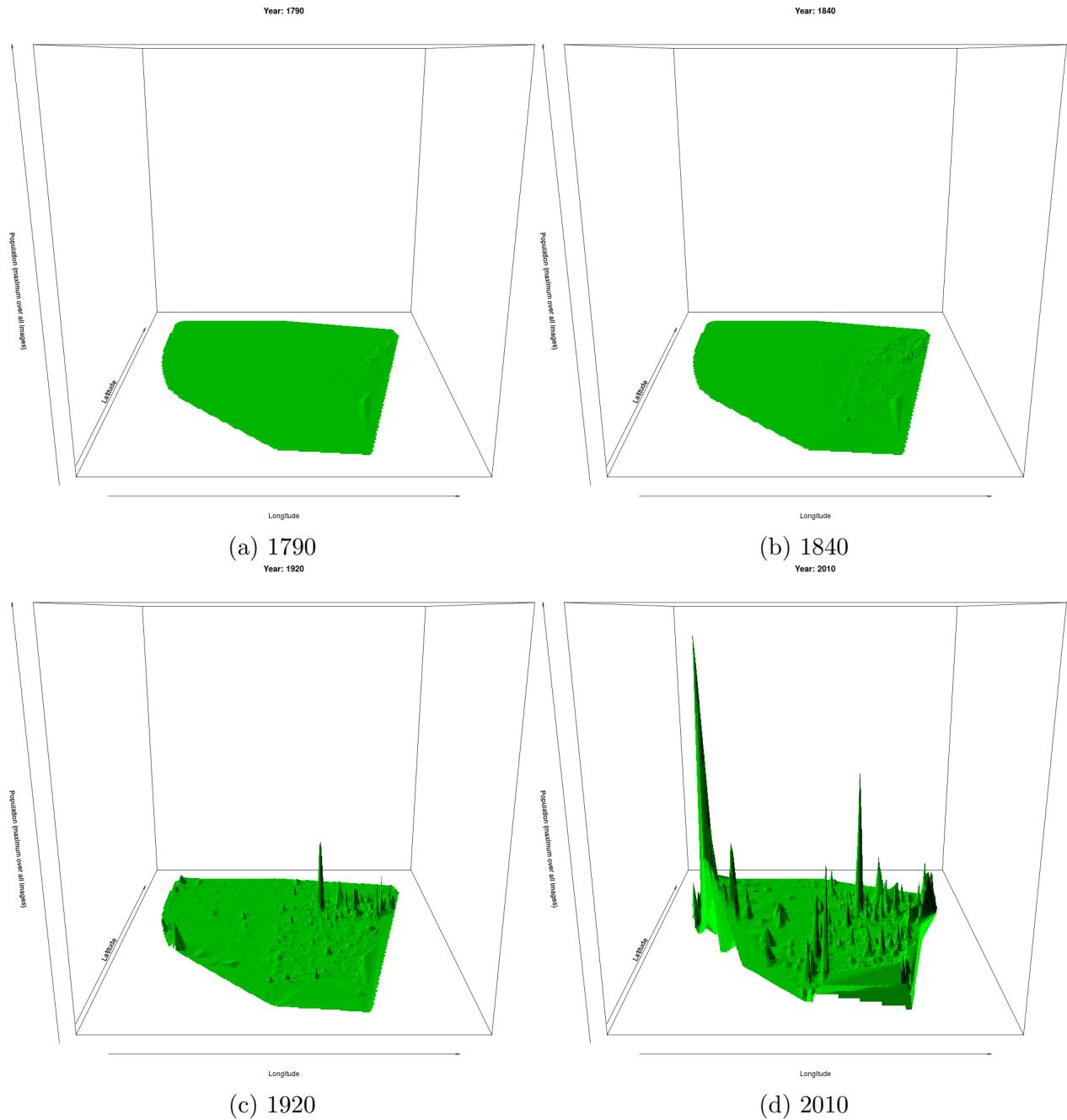


Figure 5: Raw population data. Z axis set to the highest population value.

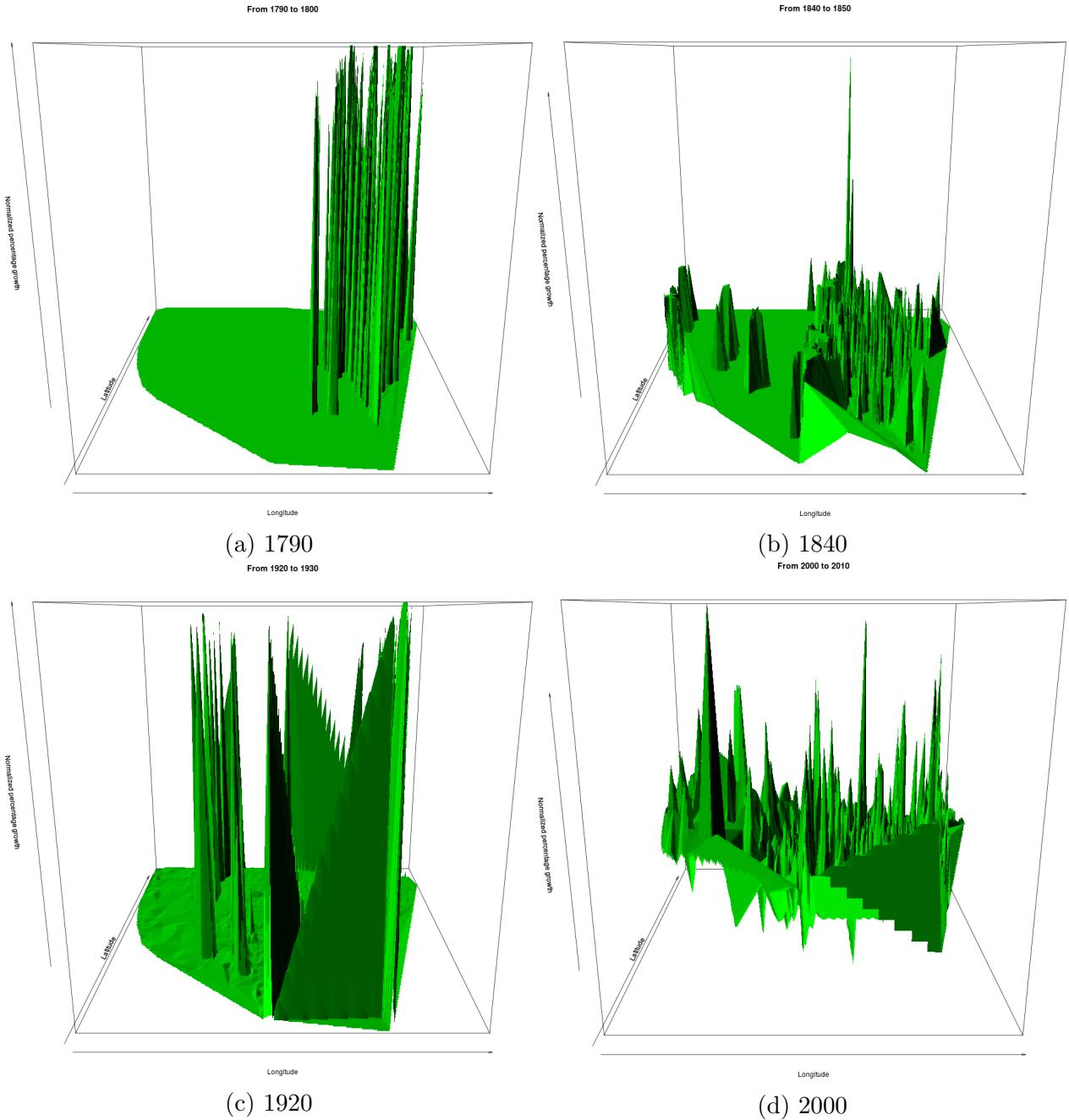


Figure 6: Normalized growth by county. The percentage of growth/shrinkage by county can be a misleading value. Small population changes in counties with small populations can result in large percentage changes, and vice versa.

## 5 Conclusion

Publicly available data exists within the Census Bureau to support asking and answering interesting questions about US demographics. In this exercise, we looked at population movements down to the county level from the first census to the most recent. In doing that we found that counties are not fixed in time or space, but come into existence, possibly change shape, and in some cases die as they are absorbed into other entities. Nor do their names remain the same over time. The changes in county designations add interesting complexity to what would initially seem to be an easily answered question about population in a county over time.

Looking at the available data, and discounting problems associated with changing boundaries; it is easy to see population shifts at specific intervals. These intervals can be tied to national and international events.

Given more time, more refined questions about changes in demographics could be asked including looking at racial changes, gender changes, and education changes. These analyzes could be used to predict where and how the population will change be the next census.

## A Misc. files

The files used to create all these figures are attached to this report. They are:

1. Population\_PartIII.xls - raw 1790 - 1990 census county population 
2. populationWrangled.xls - cleaned 1790 - 1990 census county population 
3. cb\_2015\_us\_county\_500k.zip - shape files for all US state counties 
4. county2k.zip - 2000 census county population 
5. Gaz\_counties\_national.txt - 2010 census county population 
6. pop.gif - percentage population growth by county 1790 - 2010 
7. pop.gif - normalized population by county 1790 - 2010 
8. pop.gif - absolute population growth by county 1790 - 2010 
9. viewRGL.R - an R script to view RGL data of the last year in the gif files 

10. census.R - an R script to load the census data into a PostGres database and create  
images

11. rglData.dat - an R data file with normalized data to be viewed with viewRGL.R

12. rglData.dat - an R data file with raw data to be viewed with viewRGL.R

13. rglData.dat - an R data file with growth data to be viewed with viewRGL.R