Natality Models

DATA 621: Business Analytics and Data Mining

Daniel Dittenhafer & Justin Hink

April 24, 2016

1 Abstract

2 Keywords

Natality, Births

3 Literature Review

4 Methodology

4.1 Data Preparation

Data sets from the Census Bureau, Centers for Disease Control, and Bureau of Labor Statistics were joined together in order to provide a unified data set for analysis and modeling.

4.1.1 Natality Data

4.1.2 Census Data

For the period of May 2010 - Decemeber 2015, the Census Bureau's census data was available as monthly population estimates broken down by age and gender (Census Bureau, 2015). The age data was in whole year granulatity and we created 10 year buckets for the female population by age: 15-24, 25-34, and 35-44.

For the period of 2000 - April 2010, monthly population estimates were only available for the total population (Census Bureau, 2010). We used annual age and gender estimates from the Census Bureau 2000 - 2010 time period, converted to ratio's, to divide the monthly total population into age and gender bins as shown in the following expressions:

Gender Bins

$$G_{year} = \frac{F_{year}}{P_{year}}$$

$$F_{month} = P_{month} * G_{year}$$

$$M_{month} = P_{month} - F_{month}$$

Where:

G Gender Ratio

F Total females, TOT_FEMALE

M Total males, TOT_MALE

P Total population, TOT_POP

Age Bins

$$F_{year\_x\_y} = \sum_{i=x}^{y-1} F_{year\_i}$$

$$A_{year\_x\_y} = \frac{F_{year\_x\_y}}{F_{year}}$$

$$F_{month\_x\_y} = F_{month} * A_{year\_x\_y}$$

Where:

x Lower age bound of bin

y Upper age bound of bin

A Age bin's ratio

4.1.3   Earnings Data

The earnings data was acquired from the Bureau of Labor Statistics and specifically covers women's weekly earnings from 2003 - 2015 (Bureau of Labor Statistics, 2015). The acquired data was at a quarter year granularity and was transformed to a monthly granularity for use in this study where a given quarter's weekly earnings were assigned to each of the 3 months in the 12 month annual period.

4.1.4   Unemployment Data

Unemployment data (U3) was acquired from the Bureau of Labor Statistics. The data was at a monthly granularity with no transformations applied before use in the study (Bureau of Labor Statistics, 2015).

5   Results

6   Summary

7   Appendix: Supplemental Tables & Figures

8   Appendix: Raw Code

9   References

Bureau of Labor Statistics. Labor Force Statistics from the Current Population Survey 2003-2015 - LNS14000000. Accessed: April 24, 2016. 2015. URL: http://data.bls.gov/timeseries/LNS14000000.

— Median wkly earnings, Emp FT, Wage & sal wrkrs, Women - LEU0252882700. Accessed: March 10, 2016. 2015. URL: http://data.bls.gov/cgi-bin/surveymost?le.

Census Bureau. Monthly Postcensal Resident Population, by single year of age, sex, race, and Hispanic origin. Accessed: April 24, 2016. 2015. URL: http://www.census.gov/popest/data/national/asrh/2014/2014-nat-res.html.

— National Intercensal Estimates (2000-2010). Accessed: April 24, 2016. 2010. URL: http://www.census.gov/popest/data/intercens