mapstu

Jian Lu

January 23, 2017

Abstract The **mapstu** package uses the geographical distribution data published in the Williams Course Catalogs each year to compare which states/ countries students are accepted from each year. This is done for fifteen years from 2000 to 2015 by pulling data from text versions of the geographical distribution data. The core component of the package creates a cloropleth map of the difference of students accepted in each area between any two years of the dataset. A printed comparison of the two years is also produced to understand the vizualization of the data.

Introduction

The geographical distribution of students from Williams College is an interesting dataset that can show the diversity of the student population. A visualization of the dataset can easily present the student diversity of the college, and the changes of where students are accepted from each year. The package **mapstu** pulls and cleans data from the yearly Williams College Course Catalogs and organizes the data into a readable csv file. Then the package reads the csv files into a dataframe and combines all the dataframes from all years merging into a complete dataframe which can be easily manipulated.
The vectors of the dataframe, or years, can be called on and manipulated into cloropleth maps with the help of the S4 classes in R. There are two diffferent S4 datasets in the package. One includes all of the countries in the world, and the other includes all U.S. owned territories. The years can be combined with either of the S4 datasets, and then plotted. The package also includes another function which plots the change between any two years, and creates a plot where gradients of red and green represent decreases and increases.

Data

Tools to visualize data

Creates the desired maps from year to year

countrymap2000.2001 <- countrymap(yearsdata$X2001, yearsdata$X2000, "countrymap2000.2001.png", "Change in Students 2000-2001") countrymap2001.2002 <- countrymap(yearsdata$X2002, yearsdata$X2001, "countrymap2001.2002.png", "Change in Students 2001-2002") countrymap2002.2003 <- countrymap(yearsdata$X2003, yearsdat "countrymap2002.2003.png", "Change in Students 2002-2003") countrymap2003.2004 <- countrymap(yearsdata$X2004, yearsdat "countrymap2003.2004.png", "Change in Students 2003-2004") countrymap2004.2005 <- countrymap(yearsdata$X2005, yearsdat "countrymap2004.2005.png", "Change in Students 2004-2005") countrymap2005.2006 <- countrymap(yearsdata$X2006, yearsdat "countrymap2005.2006.png", "Change in Students 2005-2006") countrymap2006.2007 <- countrymap(yearsdata$X2007, yearsdat "countrymap2006.2007.png", "Change in Students 2006-2007") countrymap2007.2008 <- countrymap(yearsdata$X2008, yearsdat "countrymap2007.2008.png", "Change in Students 2007-2008") countrymap2008.2009 <- countrymap(yearsdata$X2009, yearsdat "countrymap2008.2009.png", "Change in Students 2008-2009") countrymap2009.2010 <- countrymap(yearsdata$X2010, yearsdat "countrymap2009.2010.png", "Change in Students 2009-2010") countrymap2010.2011 <- countrymap(yearsdata$X2011, yearsdat "countrymap2010.2011.png", "Change in Students 2010-2011") countrymap2011.2012 <- countrymap(yearsdata$X2012, yearsdat "countrymap2011.2012.png", "Change in Students 2011-2012") countrymap2012.2013 <- countrymap(yearsdata$X2013, yearsdat "countrymap2012.2013.png", "Change in Students 2012-2013") countrymap2013.2014 <- countrymap(yearsdata$X2014, yearsdat "countrymap2013.2014.png", "Change in Students 2013-2014") countrymap2014.2015 <- countrymap(yearsdata$X2015, yearsdat "countrymap2014.2015.png", "Change in Students 2014-2015") countrymap2015.2016 <- countrymap(yearsdata$X2016, yearsdat "countrymap2015.2016.png", "Change in Students 2015-2016") #Calls the maps

Creating the maps for year to year

usmap2000.2002 <- usmap(yearsdata$X2001, yearsdataX2000, "usmap2000.2001.png", "Change in Students 2000-2001") usmap2001.2003 <- usmap(yearsdata$X2002, yearsdataX2001, "usmap2001.2002.png", "Change in Students 2001-2002") usmap2002.2004 <- usmap(yearsdata$X2003, yearsdataX2002, "usmap2002.2003.png", "Change in Students 2002-2003") usmap2003.2005 <- usmap(yearsdata$X2004, yearsdataX2003, "usmap2003.2004.png", "Change in Students 2003-2004") usmap2004.2006 <- usmap(yearsdata$X2005, yearsdataX2004, "usmap2004.2005.png", "Change in Students 2004-2005") usmap2005.2007 <- usmap(yearsdata$X2006, yearsdataX2005, "usmap2005.2006.png", "Change in Students 2005-2006") usmap2006.2008 <- usmap(yearsdata$X2007, yearsdataX2006, "usmap2006.2007.png", "Change in Students 2006-2007") usmap2007.2009 <- usmap(yearsdata$X2008, yearsdataX2007, "usmap2007.2008.png", "Change in Students 2007-2008") usmap2008.2010 <- usmap(yearsdata$X2009, yearsdataX2008, "usmap2008.2009.png", "Change in Students 2008-2009") usmap2009.2011 <- usmap(yearsdata$X2010, yearsdataX2009, "usmap2009.2010.png", "Change in Students 2009-2010") usmap2010.2012 <- usmap(yearsdata$X2011, yearsdataX2010, "usmap2010.2011.png", "Change in Students 2010-2011") usmap2011.2013 <- usmap(yearsdata$X2012, yearsdataX2011, "usmap2011.2012.png", "Change in Students 2011-2012") usmap2012.2014 <- usmap(yearsdata$X2013, yearsdataX2012, "usmap2012.2013.png", "Change in Students 2012-2013") usmap2013.2015 <- usmap(yearsdata$X2014, yearsdataX2013, "usmap2013.2014.png", "Change in Students 2013-2014") usmap2014.2016 <- usmap(yearsdata$X2015, yearsdataX2014, "usmap2014.2015.png", "Change in Students 2014-2015")

loading the maps

usmap2000.2002 usmap2001.2003 usmap2002.2004 usmap2003.2005 usmap2004.2006 usmap2005.2007 usmap2006.2008 usmap2007.2009 usmap2008.2010 usmap2009.2011 usmap2010.2012 usmap2011.2013 usmap2012.2014 usmap2013.2015 usmap2014.2016

Creates the desired map of the countries from year to year

countrymap2000.2001 <- countrymap(yearsdata$X2001, yearsdataX2000, "countrymap2000.2001.png", "Change in Students 2000-2001") countrymap2001.2002 <- countrymap(yearsdata$X2002, yearsdataX2001, "countrymap2001.2002.png", "Change in Students 2001-2002") countrymap2002.2003 <- countrymap(yearsdata$X2003, yearsdat "countrymap2002.2003.png", "Change in Students 2002-2003") countrymap2003.2004 <- countrymap(yearsdata$X2004, yearsdat "countrymap2003.2004.png", "Change in Students 2003-2004") countrymap2004.2005 <- countrymap(yearsdata$X2005, yearsdat "countrymap2004.2005.png", "Change in Students 2004-2005") countrymap2005.2006 <- countrymap(yearsdata$X2006, yearsdat "countrymap2005.2006.png", "Change in Students 2005-2006") countrymap2006.2007 <- countrymap(yearsdata$X2007, yearsdat "countrymap2006.2007.png", "Change in Students 2006-2007") countrymap2007.2008 <- countrymap(yearsdata$X2008, yearsdat "countrymap2007.2008.png", "Change in Students 2007-2008") countrymap2008.2009 <- countrymap(yearsdata$X2009, yearsdat "countrymap2008.2009.png", "Change in Students 2008-2009") countrymap2009.2010 <- countrymap(yearsdata$X2010, yearsdat "countrymap2009.2010.png", "Change in Students 2009-2010") countrymap2010.2011 <- countrymap(yearsdata$X2011, yearsdat "countrymap2010.2011.png", "Change in Students 2010-2011") countrymap2011.2012 <- countrymap(yearsdata$X2012, yearsdat "countrymap2011.2012.png", "Change in Students 2011-2012") countrymap2012.2013 <- countrymap(yearsdata$X2013, yearsdat "countrymap2012.2013.png", "Change in Students 2012-2013") countrymap2013.2014 <- countrymap(yearsdata$X2014, yearsdat "countrymap2013.2014.png", "Change in Students 2013-2014") countrymap2014.2015 <- countrymap(yearsdata$X2015, yearsdat "countrymap2014.2015.png", "Change in Students 2014-2015") countrymap2015.2016 <- countrymap(yearsdata$X2016, yearsdat "countrymap2015.2016.png", "Change in Students 2015-2016") ## Data analysis

Conclusion

Appendix