`ipeds`: R Package for Interfacing with the Integrated Postsecondary Education Data System

Jason M. Bryer

Excelsior College

jbryer@excelsior.edu

November 4, 2012

Abstract

This paper provides an overview of the `ipeds` package for R.

1   Introduction

Integrated Postsecondary Data System (IPEDS)

1.1   Installing the `ipeds` Package

The `ipeds` package is maintained on the R-Forge [**?**] website. This site allows for collaboration among researchers across the globe. The R-Forge website maintains nightly builds of all packages hosted and as such, the latest `ipeds` package can be retrieved directly from the R-Forge website. The following command will download and install the `ipeds` package along with all required package dependencies (note that the second `repos` parameter can be changed to a CRAN mirror closest to your location).

```
> require(devtools)
> install_github('ipeds','jbryer')
```

Once the package is installed it can be loaded. The `ls` command will print all the functions available in the package.

```
> library(ipeds)
> library(xtable)
> ls('package:ipeds')

 [1] "completions"            "downloadAllSurveys"
 [3] "downloadIPEDSSurvey"    "getIPEDSSurvey"
 [5] "ipedsHelp"              "mapIPEDSFields"
 [7] "recodeAwardLevel"       "recodeControl"
```

```
  [9] "recodeDirectory"             "recodeGeographicRegion"
 [11] "recodeGraduateOffering"      "recodeHighestDegreeOffered"
 [13] "recodeHighestLevelOfOffering" "recodeImputation"
 [15] "recodeInstitutionSize"       "recodeLevelOfInstitution"
 [17] "recodeOpenPublic"            "recodeSector"
 [19] "recodeTitleIVEligibility"    "recodeTitleIVIndicator"
 [21] "recodeUndergraduateOffering" "twelveMonthEnrollment"

> data(surveys)
```

2 Retrieving Data from IPEDS

The `ipeds` package includes a data frame that lists the data available from the IPEDS website. There are currently 36 data files that can be retrieved. Once the desired survey is identified, the value of the `SurveyID` column is used for most of the data retrieving functions.

```
> names(surveys)

[1] "SurveyID"    "Survey"     "Title"      "DataFilePre" "DataFilePost"
[6] "YearFormat"

> head(surveys[,c('SurveyID', 'Survey', 'Title')])

  SurveyID                      Survey
1       HD Institutional Characteristics
2       IC Institutional Characteristics
3    IC_AY Institutional Characteristics
4    IC_PY Institutional Characteristics
5    FLAGS Institutional Characteristics
6    EFEST                 Enrollments

                                                                  Title
1                                                 Directory information
2 Educational offerings, organization, admissions, services and athletic associations
3                                  Student charges for academic year programs
4                                  Student charges by program (vocational programs)
5                                  Response status for all survey components
6                                                        Estimated enrollment

> peerIds = c(105668,127918,128780,144777,163204,183257,187046,202806,206279,223816,260901,413413,
> directory = recodeDirectory(getIPEDSSurvey('HD', 2009))
> names(directory); nrow(directory)

 [1] "unitid"   "instnm"   "addr"     "city"     "stabbr"   "zip"
 [7] "fips"     "obereg"   "chfnm"    "chftitle" "gentele"  "ein"
[13] "opeid"    "opeflag"  "webaddr"  "adminurl" "faidurl"  "applurl"
[19] "sector"   "iclevel"  "control"  "hloffer"  "ugoffer"  "groffer"
[25] "hdegofr1" "deggrant" "hbcu"     "hospital" "medical"  "tribal"
```

```
[31] "locale"   "openpubl" "act"      "newid"    "deathyr" "closedat"
[37] "cyactive" "postsec"  "pseflag"  "pset4flg" "rptmth"  "ialias"
[43] "instcat"  "ccbasic"  "ccipug"   "ccipgrad" "ccugprof" "ccenrprf"
[49] "ccsizset" "carnegie" "tenursys" "landgrnt" "instsize" "cbsa"
[55] "cbsatype" "csa"      "necta"    "longitud" "latitude" "countycd"
[61] "countynm" "cngdstcd" "dfrcgid"

[1] 7316

> unitid = directory[which(directory$instnm == 'Excelsior College'),'unitid']
```

3   Student Charges

```
> studentCharges = getIPEDSSurvey('IC_AY', 2009)
> studentCharges = studentCharges[,c('unitid',
+          'tuition1', 'fee1', 'hrchg1', 'cmpfee1', #In-district average tuition for full-time unde
+          'tuition2', 'fee2', 'hrchg2', 'cmpfee2', #In-state average tuition for full-time undergr
+          'tuition3', 'fee3', 'hrchg3', 'cmpfee3', #Out-of-state average tuition for full-time und
+          'tuition5', 'fee5', 'hrchg5', #In-district average tuition full-time graduates
+          'tuition6', 'fee6', 'hrchg6', #In-state average tuition full-time graduates
+          'tuition7', 'fee7', 'hrchg7')] #Out-of-state average tuition full-time graduates
> studentCharges = rbind(studentCharges[which(studentCharges$unitid == unitid),], studentCharges[w.
> studentCharges = merge(studentCharges, directory[,c('unitid', 'instnm')], by='unitid', all.x=TRU.
> addtorow = list()
> addtorow$pos = list()
> addtorow$pos[[1]] = c(-1)
> addtorow$command = c("\\hline & \\multicolumn{3}{c}{In State} & \\multicolumn{3}{c}{Out-of-State,
> x = xtable(studentCharges[,c('instnm', 'tuition2', 'fee2', 'hrchg2', 'tuition3', 'fee3', 'hrchg3
> print(x, include.rownames=FALSE, include.colnames=FALSE, add.to.row=addtorow)
```

	In State			Out-of-State		
Institution	Tuition	Fees	Per Credit	Tuition	Fees	Per Credit

Table 1: Cost of Attendance - Undergraduate

4   Submitting Data to IPEDS

References

[1] R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org`.

[2] Stefan Theuβl and Achim Zeileis. Collaborative software development using R-Forge. *The R Journal*, 1(1):9-14, May 2009. URL `http://journal.R-project.org/`

[3] Duncan Temple Lang (duncan@wald.ucdavis.edu) (2010). XML: Tools for parsing and generating XML within R and S-Plus.. R package version 3.1-0. http://CRAN.R-project.org/package=XML

[4] National Center for Educational Statistics (2010). Integrated Postsecondary Data System. Washington D.C. URL `http://nces.ed.gov/ipeds/`

Appendices

A   Surveys Available from IPEDS

HD Directory information

IC Educational offerings, organization, admissions, services and athletic associations

IC_AY Student charges for academic year programs

IC_PY Student charges by program (vocational programs)

FLAGS Response status for all survey components

EFEST Estimated enrollment

EFFY 12-month unduplicated headcount

EFIA 12-month instructional activity

FLAGS Response status for all survey components

C_A Awards/degrees conferred by program (6-digit CIP code), award level, race/ethnicity, and gender

FLAGS Response status for all survey components

SAL_A Salaries of full-time instructional faculty, by contract length, gender, and academic rank

SAL_B Fringe benefits of full-time instructional faculty, by contract length

SAL_FACULTY Tenure status of full-time instructional faculty in 4-year institutions, by contract length, gender, and academic rank

SAL_A_LT9 Number of full-time instructional faculty with less than 9-month contracts, by gender and academic rank

FLAGS Response status for all survey components

S_ABD Employees by primary occupation, salary categories, race/ethnicity, and gender (Degree-granting institutions with 15 or more full-time employees)

S_F Full-time instruction/research/public service staff, by tenure status, academic rank, race/ethnicity, and gender (Degree-granting institutions with 15 or more full-time employees)

S_G New hires by primary occupation, race/ethnicity, and gender (Degree-granting institutions with 15 or more full-time employees)

S_CN Employees by primary occupation, race/ethnicity, and gender (Degree-granting institutions with less than 15 full-time employees and all nondegree-granting institutions)

FLAGS Response status for all survey components

EAP Employees by faculty status, primary function/occupational activity

FLAGS Response status for all survey components