

State Shares File Codebook

CollegeMarketShares_state.dta: Grads on the Go - State Shares File

Descriptives

Observations	135,200
variables	13
Last saved	12 May 2022 17:21
data signature	-

Variables

unitid: IPEDS ID

summary

type	numeric (long)
missing values	0
non-missing values	135,200
distinct non-missing values	2,600
minimum	100654
1st quartile	149811.5
median	183699
3rd quartile	215345
maximum	490805

instnm: Institution Name

summary

type	string (str75)
-------------	----------------

State Shares File Codebook

distinct non-missing values2,576

example values

Abilene Christian University
Abraham Baldwin Agricultural College
Adams State University
Adelphi University
Adirondack Community College
Adrian College
AdventHealth University
Agnes Scott College
Aiken Technical College
Aims Community College

stabbr: Institution State

summary

type	string (str2)
missing values	0
non-missing values	135,200
distinct non-missing values	51

example values

AK
AL
AR
AZ
CA

State Shares File Codebook

DC

DE

FL

geo_stabbr: State Geography Abbreviation

summary

type	string (str9)
missing values	0
non-missing values	135,200
distinct non-missing values	52

example values

AK
AL
AR
AZ
CA
CO
CT
DC
DE
FL

geo_fips: State Fips Code

summary

type	numeric (byte)
missing values	2,600

State Shares File Codebook

minimum	1
1st quartile	16
median	29
3rd quartile	42
maximum	56

alum_uscount: LinkedIn raw US count

summary

type	numeric (float)
missing values	0
non-missing values	135,200
distinct non-missing values	2,291
minimum	31
1st quartile	1364
median	3236.5
3rd quartile	7730
maximum	127859
note 1	Source: LinkedIn college alumni page data extracted in 2021 for all graduates with enrollment listed between 2010 and 2015. See associated paper text and data appendix for more details.

alum_uscovered: LinkedIn raw located US count

summary

State Shares File Codebook

non-missing values	135,200
distinct non-missing values	2,272
minimum	25
1st quartile	1146
median	2816
3rd quartile	6726
maximum	105235
note 1	Source: LinkedIn college alumni page data extracted in 2021 for all graduates with enrollment listed between 2010 and 2015. See associated paper text and data appendix for more details.

geo_count_state: LinkedIn count in state

summary

type	numeric (double)
missing values	0
non-missing values	135,200
distinct non-missing values	3,444
minimum	0
1st quartile	0
median	0

State Shares File Codebook

note 1 Derived by aggregating counts from substate LinkedIn geographies. The state associated with the LinkedIn substate geography text is generally listed explicitly (e.g., Fort Collins, Colorado Area), though in some instances it is inferred, (e.g., Greater Denver Area).

share_state: Share of Full US Count

summary

type	numeric (double)
missing values	0
non-missing values	135,200
distinct non-missing values	20,840
minimum	0
1st quartile	0
median	0
3rd quartile	0
maximum	.95664622
note 1	Derived by taking geo_count_state divided by alum_uscount

share_state_renorm: Share of US Count excl. Unlocated

summary

type	numeric (double)
missing values	2,600
non-missing values	132,600
distinct non-missing values	18,154
minimum	0
1st quartile	0

State Shares File Codebook

maximum 1

note 1 Derived by taking geo_count_state divided by alum_uscovered. This ignores unlocated US alumni (geo_stabbr = UNLOCATED).

geo_count_state_adj: Population Adjusted LinkedIn count in state

summary

type numeric (double)

missing values 0

non-missing values 135,200

distinct non-missing values 21,198

minimum 0

1st quartile 0

median 0

3rd quartile 2.0875131

maximum 80320.806

note 1 Derived by mapping substate LinkedIn geographies to CBSAs, then apportioning alumni counts to each state based on 2010 Census populations.

share_state_adj: Population Adjusted Share of Full US Count

summary

type numeric (double)

missing values 0

non-missing values 135,200

distinct non-missing values 39,764

minimum 0

1st quartile 0

State Shares File Codebook

maximum .95525622

note 1 Derived by taking geo_count_state_adj divided by alum_uscount

share_state_renorm_adj: Population Adjusted Share of US Count excl. Unlocated

summary

type	numeric (double)
missing values	2,600
non-missing values	132,600
distinct non-missing values	37,082
minimum	0
1st quartile	0
median	0
3rd quartile	.00050101
maximum	1
note 1	Derived by taking geo_count_state_adj divided by alum_uscovered. This ignores unlocated US alumni (geo_stabbr = UNLOCATED).

Codebook created on 12 May 2022 17:21:10