

Data Analysis and Simulation for Stratified Random Sampling

Longhai Li

September 2024

Contents

1	Functions and packages for Analyzing Data	1
2	Analysis of “agstrat.csv” data	2
2.1	Importing agstrat.csv data	2
2.2	Spreadsheet calculation	19
2.3	Using the function “str_mean_estimate”	20
2.4	Using the function “str_mean_estimate_data”	20
2.5	Comparing with SRS estimate	21
3	Allocation of stratum sample size	21
3.1	Analyzing seals.csv collected with stratified sampling	21
3.2	Neyman allocation of stratum sample size	22
4	Simulation to Study the Efficiency of Stratified Sampling with Different Allocation	23
4.1	Using “Region” to Stratify	23
4.2	Using “acres82” to Define Strata	33

1 Functions and packages for Analyzing Data

```
library(latex2exp)
library(sampling)
library(plyr)
## this function finds statistical estimates given a dataset with sampling weight
# stratdata --- data.frame containing stratified sample
# y --- name of variable for which we want to estimate population mean
# stratum --- name of variable that will be used as stratum variable
# weight --- name of variable indicating sampling weight
# post = TRUE is to indicate that we will do post-stratification analysis (ignoring it for the moment)
# note: from weights we can find Nh (see the code for formula)
str_mean_estimate_data <- function (stratdata, y, stratum, weight)
{
  ## compute stratum-wise data
  n <- nrow (stratdata)
  sh <- tapply (stratdata[, y], stratdata[,stratum], sd)
  ybarh <- tapply (stratdata[, y], stratdata[,stratum], mean)
  ## find population stratum size using sampling weight included in the data set
  Nh <- tapply (stratdata[, weight], stratdata[,stratum], sum)
  nh <- tapply (1:nrow(stratdata), stratdata[,stratum], length)
  ## find mean estimates
```

```

N <- sum (Nh)
Pi_h <- Nh/N
ybar <- sum(ybarh * Pi_h)
seybar <- sqrt(sum((1-nh/Nh)*Pi_h^2*sh^2/nh))
mem <- 1.96 * seybar
c(Est. = ybar, S.E. = seybar, ci.low = ybar - mem, ci.upp = ybar + mem)
}

## to find total, multiply N to the estimate returned by this function
## for poststratification, use nh = n * Nh/N
str_mean_estimate <- function (ybarh, sh, nh, Nh)
{
  N <- sum (Nh)
  Pi_h <- Nh/N
  ybar <- sum(ybarh * Pi_h)
  seybar <- sqrt(sum((1-nh/Nh)*Pi_h^2*sh^2/nh))
  mem <- 1.96 * seybar
  c(Est. = ybar, S.E. = seybar, ci.low = ybar - mem, ci.upp = ybar + mem)
}

## sdata --- a vector of original survey data
## N --- population size
## to find total, multiply N to the estimate returned by this function

srs_mean_est <- function (sdata, N = Inf)
{
  n <- length (sdata)
  ybar <- mean (sdata)
  se.ybar <- sqrt((1 - n / N)) * sd (sdata) / sqrt(n)
  mem <- qt (0.975, df = n - 1) * se.ybar
  c (Est. = ybar, S.E. = se.ybar, ci.low = ybar - mem, ci.upp = ybar + mem)
}

```

2 Analysis of “agstrat.csv” data

2.1 Importing agstrat.csv data

```

agstrat <- read.csv("data/agstrat.csv")
agstrat

```

```

##           county state acres92 acres87 acres82 farms92 farms87
## 1      PIERCE COUNTY   NE  297326  332862  319619    725    857
## 2    JENNINGS COUNTY   IN  124694  131481  139111    658    671
## 3      WAYNE COUNTY   OH  246938  263457  268434   1582   1734
## 4    VAN BUREN COUNTY  MI  206781  190251  197055   1164   1278
## 5      OZAUKEE COUNTY  WI   78772   85201   89331    448    483
## 6  CLEARWATER COUNTY  MN  210897  229537  213105    583    699
## 7      POTTER COUNTY  SD  507101  552844  541015    321    371
## 8      HARDIN COUNTY  IA  332358  337990  355823    986   1065

```

## 9	VERNON COUNTY	MO	402202	396638	400466	1249	1251
## 10	SHERIDAN COUNTY	KS	535359	503582	513458	488	518
## 11	MACOUPIN COUNTY	IL	402310	444816	467453	1308	1509
## 12	CARROLL COUNTY	IL	238906	259634	270129	657	775
## 13	HOUSTON COUNTY	MN	272049	285056	286561	974	1073
## 14	CRAWFORD COUNTY	KS	302849	283589	325082	780	809
## 15	MISSISSIPPI COUNTY	MO	265245	259207	253150	293	340
## 16	PUTNAM COUNTY	IL	78081	85254	84630	201	242
## 17	STARK COUNTY	IL	169622	179267	179044	362	434
## 18	PIKE COUNTY	IN	85366	98958	96821	309	382
## 19	SHERMAN COUNTY	KS	620144	625942	672254	500	524
## 20	SHERBURNE COUNTY	MN	117701	124288	134960	530	604
## 21	MOULTRIE COUNTY	IL	184599	184566	183803	491	561
## 22	VANDERBURGH COUNTY	IN	80958	85852	81779	305	378
## 23	SAUK COUNTY	WI	335517	370141	388255	1383	1502
## 24	DODGE COUNTY	MN	241148	239443	269247	740	830
## 25	EDWARDS COUNTY	KS	403375	387309	382477	325	361
## 26	GREENE COUNTY	IL	303715	296591	312325	783	827
## 27	NANCE COUNTY	NE	236950	248639	245898	440	508
## 28	POLK COUNTY	IA	229818	238256	247261	832	1001
## 29	MONROE COUNTY	IA	223638	225145	222200	682	736
## 30	JONES COUNTY	SD	584231	594236	540344	198	213
## 31	WALLACE COUNTY	KS	471658	529749	505916	283	330
## 32	ALLAMAKEE COUNTY	IA	321728	321226	351756	1000	1062
## 33	THOMAS COUNTY	KS	702549	677199	658558	547	644
## 34	STODDARD COUNTY	MO	438142	434403	423972	953	1159
## 35	MIAMI COUNTY	IN	188843	196019	198468	771	818
## 36	CHEROKEE COUNTY	IA	336254	338708	340594	979	1091
## 37	HETTINGER COUNTY	ND	688468	724825	768099	427	525
## 38	RIPLEY COUNTY	IN	164025	173795	173605	963	1071
## 39	SCOTT COUNTY	MO	219042	241410	229535	545	642
## 40	CLINTON COUNTY	IL	229120	241466	246185	942	1115
## 41	MARSHALL COUNTY	IA	312858	330012	337775	949	1073
## 42	SULLIVAN COUNTY	IN	181020	188948	184890	544	599
## 43	MONONA COUNTY	IA	392835	377030	378621	822	854
## 44	JO DAVIESS COUNTY	IL	290454	313934	313436	955	1070
## 45	SEDGWICK COUNTY	KS	510319	523580	522674	1421	1589
## 46	HICKORY COUNTY	MO	174314	174007	155011	532	587
## 47	LAWRENCE COUNTY	SD	195077	207123	196321	272	253
## 48	LA SALLE COUNTY	IL	612112	641835	637108	1669	1978
## 49	DUNDY COUNTY	NE	528731	536373	557748	308	389
## 50	MASSAC COUNTY	IL	98838	103802	104233	401	414
## 51	ALEXANDER COUNTY	IL	69354	73076	78693	171	184
## 52	PENNINGTON COUNTY	MN	280089	305784	313450	480	585
## 53	SANILAC COUNTY	MI	444407	431199	444294	1433	1559
## 54	MONTCALM COUNTY	MI	224030	238387	240591	900	980
## 55	WALWORTH COUNTY	SD	448834	410375	421727	378	360
## 56	ALGER COUNTY	MI	16099	16140	17700	58	63
## 57	DOOR COUNTY	WI	130051	147860	155318	760	911
## 58	WARREN COUNTY	IA	302487	303783	320840	1216	1310
## 59	KNOX COUNTY	NE	612694	614959	641105	1086	1212
## 60	BARRY COUNTY	MI	165371	168493	188096	833	908
## 61	DOUGLAS COUNTY	IL	259498	271098	276455	682	807
## 62	JEFFERSON COUNTY	WI	232591	256282	269791	1280	1440

## 63	MINNEHAHA COUNTY	SD	425288	432472	425971	1262	1382
## 64	WOOD COUNTY	WI	221357	236904	242248	1029	1157
## 65	WAUPACA COUNTY	WI	241778	272429	274368	1190	1365
## 66	GRUNDY COUNTY	MO	226336	223823	245418	618	650
## 67	LOUISA COUNTY	IA	191291	198855	208670	554	634
## 68	GRAY COUNTY	KS	517623	532284	536969	497	547
## 69	HALL COUNTY	NE	316551	306869	304925	744	788
## 70	STONE COUNTY	MO	137747	148207	160233	698	782
## 71	BOLLINGER COUNTY	MO	197530	193547	210806	786	860
## 72	SHELBY COUNTY	IL	402212	422071	420047	1305	1431
## 73	FAIRFIELD COUNTY	OH	197736	214316	221802	1058	1217
## 74	DODGE COUNTY	NE	298854	321996	310513	855	962
## 75	TRIPP COUNTY	SD	1006831	960092	937667	741	770
## 76	WAYNE COUNTY	IN	189467	196898	200759	828	888
## 77	EFFINGHAM COUNTY	IL	257761	258305	255895	1140	1228
## 78	ATCHISON COUNTY	MO	304032	321878	306986	509	622
## 79	WASECA COUNTY	MN	237239	231788	247571	759	813
## 80	CALHOUN COUNTY	MI	244927	253383	266680	1080	1166
## 81	CODINGTON COUNTY	SD	392935	341159	352025	658	636
## 82	WATONWAN COUNTY	MN	249731	252824	260395	663	748
## 83	ELKHART COUNTY	IN	192311	204547	213225	1447	1556
## 84	ROCK COUNTY	WI	343115	357751	362206	1398	1518
## 85	NEMAHA COUNTY	NE	226042	241335	239728	511	591
## 86	LINN COUNTY	IA	349252	356000	368525	1529	1690
## 87	THAYER COUNTY	NE	347598	380147	362375	623	744
## 88	SALINE COUNTY	NE	312079	342898	331520	742	881
## 89	THURSTON COUNTY	NE	193556	193365	204344	386	462
## 90	HENDERSON COUNTY	IL	203974	212513	211574	468	522
## 91	DEARBORN COUNTY	IN	86236	90024	99239	738	796
## 92	EMMONS COUNTY	ND	834293	866121	813645	759	868
## 93	REPUBLIC COUNTY	KS	443290	440215	441118	746	833
## 94	ISANTI COUNTY	MN	131563	142998	153003	680	817
## 95	WARREN COUNTY	OH	129416	136623	154744	792	920
## 96	WORTH COUNTY	MO	134028	144572	155282	327	367
## 97	POTTAWATOMIE COUNTY	KS	451362	443660	464739	777	790
## 98	VINTON COUNTY	OH	41666	43303	50514	205	217
## 99	BROWN COUNTY	MN	347420	335559	339037	1190	1317
## 100	BOND COUNTY	IL	182572	186777	193237	629	714
## 101	CHEYENNE COUNTY	KS	592207	619870	602726	426	493
## 102	BOYD COUNTY	NE	296164	293720	282981	395	443
## 103	CASS COUNTY	NE	296016	338390	316564	721	913
## 104	CHESTER COUNTY	PA	176643	189943	219980	1367	1573
## 105	OSWEGO COUNTY	NY	112334	122648	139440	659	749
## 106	MONTOUR COUNTY	PA	41347	41870	48815	249	270
## 107	NEW LONDON COUNTY	CT	65987	74063	82709	550	556
## 108	CUMBERLAND COUNTY	ME	53893	57745	62096	440	456
## 109	FULTON COUNTY	PA	88982	99920	104793	438	467
## 110	ESSEX COUNTY	VT	17710	22237	26028	74	81
## 111	PLYMOUTH COUNTY	MA	72247	77140	80392	668	775
## 112	MONROE COUNTY	PA	20777	26898	29597	147	192
## 113	CLINTON COUNTY	NY	158392	172734	206396	488	591
## 114	ST LAWRENCE COUNTY	NY	396721	456497	493073	1367	1602
## 115	MORRIS COUNTY	NJ	23915	27086	25576	395	430
## 116	CUMBERLAND COUNTY	PA	141919	153746	163186	940	1100

## 117	YATES COUNTY	NY	102024	113922	112344	602	619
## 118	DELAWARE COUNTY	NY	192116	225899	297071	716	883
## 119	NEW YORK COUNTY	NY	0	0	0	0	0
## 120	ULSTER COUNTY	NY	69643	78437	85203	433	539
## 121	UNION COUNTY	PA	63159	64622	72935	451	505
## 122	SULLIVAN COUNTY	PA	30613	30496	35654	142	165
## 123	MONTGOMERY COUNTY	NY	138822	156368	164000	537	616
## 124	LEHIGH COUNTY	PA	82982	96931	95302	427	541
## 125	LAMPASAS COUNTY	TX	432379	404876	431276	689	615
## 126	MONROE COUNTY	GA	44599	39407	58630	179	160
## 127	LOVING COUNTY	TX	346653	415540	350350	14	17
## 128	FISHER COUNTY	TX	545666	504096	507523	547	629
## 129	SEMINOLE COUNTY	OK	250958	256310	256094	872	990
## 130	LINCOLN COUNTY	KY	173892	184377	181006	1444	1475
## 131	SEVIER COUNTY	AR	131353	126457	122126	549	558
## 132	GWINNETT COUNTY	GA	24239	29435	33762	345	441
## 133	FULTON COUNTY	GA	21975	32832	42527	235	344
## 134	SUMTER COUNTY	SC	138573	152452	177061	406	488
## 135	CHATHAM COUNTY	GA	8518	10641	12715	40	51
## 136	GREENBRIER COUNTY	WV	179736	191525	200127	705	729
## 137	TUCKER COUNTY	WV	32093	32123	36723	169	160
## 138	CARROLL COUNTY	MS	151743	173064	188333	394	431
## 139	QUITMAN COUNTY	MS	186297	199767	212667	219	256
## 140	HALE COUNTY	TX	560355	582208	610359	774	818
## 141	STERLING COUNTY	TX	835337	675977	726837	74	75
## 142	MONTGOMERY COUNTY	MD	82470	103377	106157	561	669
## 143	MONTGOMERY COUNTY	MS	80272	97454	110766	290	307
## 144	SABINE PARISH	LA	57789	58112	65071	414	468
## 145	CRAIGHEAD COUNTY	AR	350402	351106	345804	781	902
## 146	UNION COUNTY	NC	167379	170378	202639	1037	1086
## 147	CARTER COUNTY	OK	372901	402260	386472	992	1056
## 148	DECATUR COUNTY	GA	168593	163114	194568	342	376
## 149	CHOWAN COUNTY	NC	53902	50446	54749	179	220
## 150	HILLSBOROUGH COUNTY	FL	265443	287951	329293	2760	2754
## 151	CLAYTON COUNTY	GA	4519	8028	7137	56	73
## 152	LOWNDES COUNTY	AL	199714	207753	226744	315	378
## 153	LIBERTY COUNTY	GA	15583	18248	19965	49	41
## 154	RUSK COUNTY	TX	268058	271230	281569	1226	1327
## 155	TROUSDALE COUNTY	TN	55097	58550	58921	389	439
## 156	PERQUIMANS COUNTY	NC	68736	75808	86376	226	272
## 157	JEFFERSON COUNTY	WV	74268	83079	87648	334	363
## 158	OGLETHORPE COUNTY	GA	55310	63352	69948	303	314
## 159	GA RFIELD COUNTY	OK	662121	633271	659095	1152	1182
## 160	DALE COUNTY	AL	134555	129105	134743	403	490
## 161	CULLMAN COUNTY	AL	196859	193771	206050	2086	2210
## 162	DUVAL COUNTY	TX	801159	996776	970827	946	1151
## 163	SAMPSON COUNTY	NC	266067	263626	263007	1342	1477
## 164	CAMPBELL COUNTY	VA	134474	134093	143127	612	628
## 165	GREENE COUNTY	NC	112291	117207	107523	407	510
## 166	CARTER COUNTY	KY	112831	118340	118009	986	1025
## 167	GA TES COUNTY	NC	64532	68992	67985	199	251
## 168	ORLEANS PARISH	LA	100	11	13	17	7
## 169	HAYS COUNTY	TX	463450	297443	210668	704	701
## 170	EASTLAND COUNTY	TX	493227	433691	469113	1120	1085

## 171	TALBOT COUNTY	GA	38313	38854	41585	127	128
## 172	DODDRIDGE COUNTY	WV	59184	57795	59608	261	272
## 173	HICKMAN COUNTY	TN	130167	129661	130045	642	650
## 174	BURKE COUNTY	NC	31671	34833	35776	348	366
## 175	HOUSTON COUNTY	TX	417187	422172	450995	1360	1421
## 176	HALL COUNTY	TX	443027	393949	458988	297	296
## 177	MCCRACKEN COUNTY	KY	62766	70148	72377	404	434
## 178	MCCREARY COUNTY	KY	13887	11584	10946	114	116
## 179	HART COUNTY	KY	200455	194172	189316	1582	1518
## 180	SARASOTA COUNTY	FL	151242	166766	206976	328	352
## 181	PICKENS COUNTY	AL	106206	108861	133757	404	417
## 182	KING COUNTY	TX	436040	409706	418003	34	50
## 183	SALINE COUNTY	AR	45609	55253	58961	330	427
## 184	CALHOUN COUNTY	FL	43314	48166	55986	132	159
## 185	PRESIDIO COUNTY	TX	1695484	1890612	1981461	151	139
## 186	CLARK COUNTY	KY	144904	155437	147236	966	947
## 187	JASPER COUNTY	SC	72500	102205	103145	146	150
## 188	OCHILTREE COUNTY	TX	593819	607038	612836	374	391
## 189	CALLOWAY COUNTY	KY	137337	137781	152270	694	749
## 190	HABERSHAM COUNTY	GA	36074	39886	42024	455	452
## 191	COFFEE COUNTY	TN	132388	143496	152296	838	887
## 192	TAYLOR COUNTY	TX	509017	431408	435628	915	903
## 193	HAYWOOD COUNTY	NC	69961	79672	81070	812	912
## 194	BECKHAM COUNTY	OK	493631	495415	498195	732	815
## 195	COTTON COUNTY	OK	358446	347784	362443	540	513
## 196	DALLAS COUNTY	AR	20589	18918	25706	108	128
## 197	ORANGE COUNTY	VA	107700	113175	118613	419	424
## 198	SMITH COUNTY	TX	247626	249326	246909	1609	1701
## 199	ITAWAMBA COUNTY	MS	76673	78932	98384	417	452
## 200	NASH COUNTY	NC	179051	184304	214759	560	692
## 201	WASHINGTON COUNTY	TX	328367	338840	305119	1903	1983
## 202	WARREN COUNTY	KY	252817	239462	245457	1956	1866
## 203	DICKENS COUNTY	TX	561521	439124	513862	270	285
## 204	ORANGE COUNTY	NC	67491	81108	90575	433	522
## 205	HOPKINS COUNTY	KY	144828	144862	158726	617	620
## 206	ORANGEBURG COUNTY	SC	262093	292177	340090	910	961
## 207	HIGHLAND COUNTY	VA	96910	94880	103470	298	303
## 208	CITRUS COUNTY	FL	70672	74264	93183	288	331
## 209	WOLFE COUNTY	KY	61145	59856	67007	456	453
## 210	PASQUOTANK COUNTY	NC	83218	81626	73766	199	225
## 211	CATAWBA COUNTY	NC	62854	67408	81452	507	567
## 212	MONTGOMERY COUNTY	KY	113383	115897	121161	772	793
## 213	KIMBLE COUNTY	TX	774804	781013	720012	476	442
## 214	TWIGGS COUNTY	GA	31161	31693	40169	113	112
## 215	DAVIDSON COUNTY	TN	47319	57917	74003	440	561
## 216	SURRY COUNTY	VA	52770	46030	55565	110	126
## 217	WASHITA COUNTY	OK	577693	589015	586142	1004	1089
## 218	WETZEL COUNTY	WV	37130	36217	45877	199	203
## 219	OKTIBBEHA COUNTY	MS	80761	91819	124895	339	355
## 220	OVERTON COUNTY	TN	105519	110079	126718	818	842
## 221	SEBASTIAN COUNTY	AR	115019	118946	137686	689	738
## 222	CHESTERFIELD COUNTY	SC	109652	109613	139303	491	429
## 223	HOOD COUNTY	TX	225852	212741	238234	659	638
## 224	REAL COUNTY	TX	362642	318164	311411	215	189

## 225	LEE COUNTY	NC	37434	42636	52614	301	345
## 226	MONROE COUNTY	WV	148842	143762	153723	606	610
## 227	POCAHONTAS COUNTY	WV	115487	118540	129717	355	379
## 228	EDGECOMBE COUNTY	NC	180400	182498	219242	376	449
## 229	NOWATA COUNTY	OK	282659	231409	274430	695	679
## 230	WAYNE COUNTY	KY	135850	136970	140014	889	979
## 231	JEFFERSON COUNTY	TN	98669	109592	115196	1234	1326
## 232	JASPER COUNTY	TX	70165	101968	76097	532	602
## 233	CHOCTAW COUNTY	MS	42712	47224	76105	195	228
## 234	COFFEE COUNTY	GA	178861	178875	200598	711	649
## 235	POLK COUNTY	TX	141215	144390	161772	515	559
## 236	COLORADO COUNTY	TX	549167	559698	596786	1547	1589
## 237	WEST BATON ROUGE PARISH	LA	38566	42488	42970	90	95
## 238	KERR COUNTY	TX	531206	539371	589223	651	584
## 239	WASHINGTON COUNTY	AL	85086	86532	96955	361	416
## 240	GRAHAM COUNTY	NC	8882	7533	10507	147	147
## 241	CLINTON COUNTY	KY	75409	86085	87629	747	777
## 242	MCINTOSH COUNTY	GA	8003	5071	5107	33	23
## 243	TOWNS COUNTY	GA	9910	10638	13613	128	149
## 244	PULASKI COUNTY	GA	80396	86400	89515	137	163
## 245	LOGAN COUNTY	AR	186829	187992	187551	940	995
## 246	ANNE ARUNDEL COUNTY	MD	43320	42413	44722	477	567
## 247	WHEELER COUNTY	TX	501692	486321	492597	445	467
## 248	MONROE COUNTY	AR	219444	234605	235409	278	306
## 249	BEAUFORT COUNTY	NC	144529	156433	158281	447	630
## 250	NORTHAMPTON COUNTY	VA	52469	50530	62283	162	180
## 251	MENARD COUNTY	TX	487573	501761	463848	280	261
## 252	SAN SABA COUNTY	TX	743638	741678	708937	640	609
## 253	HENDERSON COUNTY	TX	356170	333190	323887	1579	1577
## 254	KENEDY COUNTY	TX	553226	621878	569640	29	24
## 255	CARROLL COUNTY	AR	246184	240838	252536	1031	1045
## 256	BRAZOS COUNTY	TX	295601	271421	256837	1006	971
## 257	COLUMBIA COUNTY	AR	57253	66305	80909	320	411
## 258	GREENUP COUNTY	KY	100468	93353	92487	849	824
## 259	FLEMING COUNTY	KY	193859	204660	177444	1232	1308
## 260	CUSTER COUNTY	ID	140701	137022	148063	267	261
## 261	UMATILLA COUNTY	OR	1466580	1451108	1407279	1441	1453
## 262	ESMERALDA COUNTY	NV	1949420	1798823	2133372	23	28
## 263	MAUI COUNTY	HI	355786	359310	403140	850	722
## 264	MILLARD COUNTY	UT	484156	480195	487961	612	630
## 265	DUCHESNE COUNTY	UT	399011	366471	315061	733	753
## 266	DAWSON COUNTY	MT	1334041	1326655	1353774	451	499
## 267	SHERMAN COUNTY	OR	487534	462424	431172	179	187
## 268	HILL COUNTY	MT	1644001	1722206	1734585	641	713
## 269	POWELL COUNTY	MT	675569	670508	742622	233	237
## 270	COLUMBIA COUNTY	OR	71839	73949	77182	661	695
## 271	EDDY COUNTY	NM	1138681	967816	957155	495	503
## 272	BENEWAH COUNTY	ID	111510	115100	124190	195	205
## 273	KLICKITAT COUNTY	WA	689639	698453	725048	508	545
## 274	HAWAII COUNTY	HI	926607	1007287	1172448	3157	2810
## 275	CROOK COUNTY	WY	1542262	1509320	1548500	442	461
## 276	TOOLE COUNTY	MT	1063086	1131519	1056130	358	393
## 277	GREENLEE COUNTY	AZ	137834	139840	141595	107	102
## 278	GRAND COUNTY	CO	299142	319578	284306	149	163

## 279	UTAH COUNTY	UT	450315	493902	432326	1696	1723		
## 280	POLK COUNTY	OR	167880	176178	179518	1027	1072		
## 281	BEAR LAKE COUNTY	ID	269435	269211	292783	415	446		
## 282	LAS ANIMAS COUNTY	CO	2286947	2149828	2137550	490	481		
## 283	MADISON COUNTY	MT	1271160	1195898	1196626	418	453		
## 284	EAGLE COUNTY	CO	213004	213441	201657	134	147		
## 285	TETON COUNTY	ID	134788	148908	156160	257	268		
## 286	ROOSEVELT COUNTY	MT	1414415	1364020	1371234	525	598		
## 287	CARBON COUNTY	MT	598694	536553	628680	599	635		
## 288	KITSAP COUNTY	WA	10302	9576	10974	366	404		
## 289	HINSDALE COUNTY	CO	9021	9899	10920	16	16		
## 290	HUMBOLDT COUNTY	CA	597766	616267	648820	874	890		
## 291	SUMMIT COUNTY	UT	373582	348827	339347	419	439		
## 292	STEVENS COUNTY	WA	546303	525783	578060	1054	1073		
## 293	CLARK COUNTY	WA	82967	94646	101660	1257	1428		
## 294	WAHIAKUM COUNTY	WA	12611	14616	15915	110	135		
## 295	FRANKLIN COUNTY	WA	670149	660813	632519	857	894		
## 296	LEA COUNTY	NM	2149450	2220431	2178568	544	561		
## 297	THURSTON COUNTY	WA	59890	56799	67628	811	806		
## 298	CARSON CITY (IC)	NV	5361	17859	18780	28	37		
## 299	BANNOCK COUNTY	ID	325338	358189	352306	588	655		
## 300	LA PLATA COUNTY	CO	587339	613579	589167	709	682		
##	farms82	largef92	largef87	largef82	smallf92	smallf87	smallf82	region	rn
## 1	865	54	54	42	58	67	48	NC	805
## 2	751	14	13	14	42	36	38	NC	241
## 3	1866	20	19	16	175	186	184	NC	913
## 4	1464	23	17	9	56	66	55	NC	478
## 5	527	6	5	5	56	49	48	NC	1028
## 6	693	34	32	23	8	19	13	NC	496
## 7	341	163	180	176	10	24	16	NC	969
## 8	1208	56	36	42	90	115	132	NC	42
## 9	1320	86	78	69	42	38	28	NC	676
## 10	571	216	204	193	16	37	24	NC	383
## 11	1754	80	77	56	71	108	117	NC	158
## 12	846	48	45	40	39	57	65	NC	107
## 13	1128	22	13	15	49	50	44	NC	509
## 14	915	73	61	65	34	40	44	NC	312
## 15	361	104	96	85	14	10	8	NC	635
## 16	261	12	16	16	12	12	18	NC	177
## 17	474	44	38	29	24	27	30	NC	187
## 18	413	12	14	11	14	24	13	NC	264
## 19	550	211	224	225	12	25	14	NC	384
## 20	684	20	19	14	20	22	18	NC	552
## 21	597	54	44	32	44	44	47	NC	169
## 22	408	15	9	11	26	40	41	NC	283
## 23	1635	30	24	19	60	54	58	NC	1039
## 24	933	47	43	49	44	63	76	NC	501
## 25	402	129	144	133	11	10	14	NC	317
## 26	912	71	66	58	56	62	55	NC	130
## 27	478	67	71	61	24	52	43	NC	798
## 28	1110	48	48	39	70	108	100	NC	77
## 29	761	31	25	24	26	34	28	NC	68
## 30	222	122	137	137	12	17	8	NC	953
## 31	338	142	161	148	6	30	27	NC	393

## 32	1134	46	37	37	55	61	56	NC	3
## 33	628	236	248	221	12	20	18	NC	390
## 34	1338	132	109	94	21	45	62	NC	671
## 35	924	33	21	21	75	58	77	NC	253
## 36	1214	41	31	18	64	86	94	NC	18
## 37	502	275	301	295	14	14	14	NC	703
## 38	1165	18	15	15	55	55	67	NC	270
## 39	687	70	55	51	36	46	40	NC	668
## 40	1154	20	16	11	75	73	69	NC	113
## 41	1201	55	46	39	72	94	110	NC	64
## 42	660	52	48	36	49	43	29	NC	278
## 43	933	104	93	71	59	50	53	NC	67
## 44	1142	43	39	33	49	61	64	NC	142
## 45	1665	133	116	125	102	135	135	NC	380
## 46	562	25	32	22	13	20	13	NC	611
## 47	245	55	50	51	15	22	22	NC	956
## 48	2146	97	77	52	84	106	119	NC	149
## 49	382	150	171	178	7	26	17	NC	764
## 50	480	23	24	18	11	18	9	NC	163
## 51	246	20	16	19	2	2	9	NC	101
## 52	655	93	91	80	14	23	23	NC	538
## 53	1846	89	71	52	39	63	47	NC	474
## 54	1132	37	36	28	30	48	33	NC	457
## 55	405	158	163	176	26	22	21	NC	980
## 56	76	2	1	1	2	4	1	NC	400
## 57	1032	5	3	3	37	26	31	NC	997
## 58	1394	54	47	44	68	86	91	NC	91
## 59	1280	156	134	142	69	91	81	NC	789
## 60	1061	21	18	11	30	35	29	NC	406
## 61	845	74	59	47	69	74	56	NC	120
## 62	1602	20	15	16	79	84	101	NC	1010
## 63	1490	78	62	41	73	88	96	NC	965
## 64	1253	18	18	16	39	42	52	NC	1054
## 65	1488	12	17	13	53	69	67	NC	1051
## 66	740	54	52	51	19	30	35	NC	608
## 67	699	30	28	21	29	43	43	NC	58
## 68	552	181	183	170	20	37	23	NC	328
## 69	794	66	59	54	71	70	76	NC	775
## 70	824	16	15	13	21	41	35	NC	672
## 71	994	33	19	22	18	33	25	NC	577
## 72	1599	73	61	43	76	93	81	NC	186
## 73	1322	31	30	25	84	79	88	NC	851
## 74	977	61	53	38	88	98	75	NC	762
## 75	713	329	338	336	38	33	34	NC	977
## 76	1000	32	26	13	57	59	73	NC	290
## 77	1261	23	16	15	75	85	71	NC	124
## 78	669	88	80	61	22	36	19	NC	571
## 79	878	34	23	25	41	63	65	NC	562
## 80	1281	27	30	27	59	34	41	NC	411
## 81	675	121	100	95	33	27	47	NC	930
## 82	885	38	30	25	30	44	58	NC	564
## 83	1678	20	22	14	236	234	239	NC	221
## 84	1680	64	66	55	138	160	165	NC	1036
## 85	614	56	55	36	26	28	27	NC	799

## 86	1810	44	35	27	134	135	128	NC	57
## 87	707	100	120	91	32	44	32	NC	820
## 88	905	63	58	38	56	47	39	NC	811
## 89	535	41	33	24	17	31	33	NC	822
## 90	562	48	41	36	35	22	26	NC	135
## 91	905	2	1	0	34	33	42	NC	216
## 92	849	352	355	328	15	6	24	NC	697
## 93	923	140	135	112	28	42	43	NC	372
## 94	947	18	14	8	14	26	34	NC	511
## 95	1048	24	22	20	61	73	76	NC	911
## 96	389	27	26	25	11	11	13	NC	681
## 97	869	129	124	122	26	50	48	NC	368
## 98	249	3	4	3	10	6	5	NC	910
## 99	1390	33	16	15	60	88	96	NC	489
## 100	824	28	35	30	32	50	35	NC	102
## 101	546	189	222	221	10	17	33	NC	305
## 102	442	99	91	93	28	39	35	NC	743
## 103	915	88	91	59	56	72	56	NC	748
## 104	1825	21	18	20	172	211	263	NE	149
## 105	826	3	4	4	40	35	32	NE	110
## 106	314	2	4	4	8	15	22	NE	181
## 107	573	1	1	2	66	55	52	NE	6
## 108	507	2	3	3	84	59	54	NE	28
## 109	491	5	5	7	10	9	8	NE	163
## 110	90	3	2	4	8	2	0	NE	211
## 111	649	12	10	6	153	158	121	NE	23
## 112	207	3	4	1	14	26	13	NE	179
## 113	705	14	9	14	19	25	23	NE	82
## 114	1807	41	30	20	37	46	35	NE	117
## 115	385	1	1	1	116	134	104	NE	65
## 116	1174	7	7	6	60	86	84	NE	155
## 117	617	3	9	8	18	29	32	NE	134
## 118	1067	15	14	21	28	56	40	NE	85
## 119	0	0	0	0	0	0	0	NE	103
## 120	570	9	6	8	54	84	57	NE	128
## 121	524	3	0	1	16	33	36	NE	194
## 122	187	3	1	1	7	9	3	NE	191
## 123	657	9	7	6	16	21	21	NE	101
## 124	574	15	16	11	66	74	61	NE	173
## 125	611	110	107	101	21	25	24	S	1116
## 126	172	5	7	9	12	9	11	S	311
## 127	15	13	12	10	0	1	1	S	1126
## 128	688	153	142	134	17	28	23	S	1051
## 129	1030	29	35	33	34	31	34	S	824
## 130	1611	11	7	8	185	181	236	S	437
## 131	567	11	12	16	28	30	23	S	134
## 132	492	1	1	1	51	60	44	S	276
## 133	379	3	5	6	40	58	61	S	269
## 134	567	40	45	45	45	54	56	S	877
## 135	59	4	4	5	10	8	15	S	234
## 136	790	29	33	31	31	23	23	S	1340
## 137	175	3	4	6	3	7	2	S	1374
## 138	490	32	34	33	8	12	11	S	583
## 139	323	57	66	73	6	7	6	S	635

## 140	832	193	160	172	21	49	29	S 1070
## 141	78	54	54	58	4	5	2	S 1191
## 142	675	15	17	20	83	90	90	S 567
## 143	374	15	18	21	13	10	19	S 624
## 144	527	1	2	4	15	20	28	S 531
## 145	1020	94	78	75	24	36	42	S 83
## 146	1263	30	31	32	72	93	79	S 747
## 147	1063	70	75	73	28	48	53	S 767
## 148	484	47	45	53	17	27	30	S 252
## 149	259	14	12	8	10	20	26	S 678
## 150	2748	45	51	48	812	786	741	S 171
## 151	81	0	1	0	12	12	14	S 240
## 152	440	56	57	51	7	25	24	S 43
## 153	61	3	4	4	7	1	5	S 298
## 154	1372	36	35	46	62	60	52	S 1176
## 155	501	3	3	3	50	65	85	S 965
## 156	343	19	11	9	23	25	21	S 729
## 157	398	13	16	15	23	26	17	S 1346
## 158	351	9	10	12	19	9	13	S 318
## 159	1261	205	186	192	66	66	62	S 781
## 160	582	30	23	21	22	26	38	S 23
## 161	2303	5	6	9	122	165	176	S 22
## 162	1074	149	172	158	19	33	32	S 1041
## 163	1818	52	31	19	103	102	151	S 739
## 164	740	18	17	16	31	33	59	S 1245
## 165	670	20	9	5	17	27	43	S 697
## 166	1088	3	4	4	108	136	135	S 390
## 167	282	16	15	6	21	27	23	S 694
## 168	8	0	0	0	14	7	8	S 524
## 169	643	64	57	37	47	55	48	S 1080
## 170	1111	97	78	82	41	46	28	S 1042
## 171	144	6	6	7	4	3	3	S 339
## 172	279	6	3	4	6	3	0	S 1336
## 173	704	12	13	8	25	29	38	S 921
## 174	409	1	0	1	30	41	45	S 669
## 175	1387	77	75	73	53	56	30	S 1088
## 176	341	107	103	127	12	13	12	S 1071
## 177	555	12	14	7	37	43	59	S 441
## 178	125	0	0	0	3	8	7	S 442
## 179	1607	5	8	4	194	195	247	S 418
## 180	317	23	21	26	107	82	90	S 200
## 181	525	18	19	20	16	17	24	S 54
## 182	45	20	21	20	1	4	1	S 1110
## 183	421	2	0	5	27	35	24	S 130
## 184	169	13	12	17	8	10	13	S 149
## 185	137	97	95	80	12	9	9	S 1164
## 186	1018	22	23	19	121	103	115	S 393
## 187	189	13	13	15	16	12	13	S 861
## 188	394	188	186	190	6	22	12	S 1154
## 189	949	32	21	21	108	95	133	S 386
## 190	409	1	2	2	58	62	51	S 277
## 191	979	19	15	18	50	61	61	S 896
## 192	911	141	114	111	73	66	50	S 1196
## 193	1043	5	6	6	155	166	224	S 701

## 194	780	150	137	135	36	51	39	S 762
## 195	560	120	115	114	22	25	27	S 774
## 196	143	0	1	1	6	5	9	S 87
## 197	461	20	22	22	20	17	15	S 1295
## 198	1689	26	30	35	119	113	102	S 1187
## 199	600	10	6	6	18	17	28	S 604
## 200	997	44	41	38	62	66	72	S 721
## 201	1839	43	44	27	104	91	98	S 1214
## 202	2113	20	19	16	205	176	282	S 482
## 203	330	93	86	85	5	12	11	S 1038
## 204	599	6	6	4	27	31	45	S 725
## 205	688	34	28	27	30	34	33	S 422
## 206	1240	67	72	82	66	68	61	S 872
## 207	320	11	16	19	17	16	9	S 1274
## 208	293	11	9	14	55	46	45	S 151
## 209	510	3	3	5	70	63	78	S 487
## 210	253	26	20	15	14	9	15	S 727
## 211	669	5	4	5	27	32	40	S 675
## 212	874	11	9	12	119	114	122	S 455
## 213	423	202	184	188	18	19	11	S 1109
## 214	149	9	6	10	7	4	7	S 352
## 215	693	1	1	4	52	47	66	S 899
## 216	186	14	11	13	6	8	8	S 1316
## 217	1135	196	181	152	55	67	46	S 832
## 218	269	2	2	2	3	6	3	S 1379
## 219	454	11	15	23	26	24	12	S 628
## 220	991	7	10	9	56	51	79	S 947
## 221	814	14	12	15	41	50	54	S 133
## 222	518	18	17	29	20	22	29	S 847
## 223	593	48	46	34	51	38	41	S 1086
## 224	177	85	83	81	11	11	9	S 1168
## 225	457	5	4	4	27	33	39	S 710
## 226	720	11	9	13	22	14	22	S 1359
## 227	430	17	16	17	9	9	9	S 1365
## 228	615	57	48	50	30	33	40	S 690
## 229	696	58	46	58	30	22	23	S 810
## 230	1082	12	9	8	124	130	146	S 484
## 231	1456	1	1	3	153	149	183	S 925
## 232	659	1	11	13	35	58	56	S 1096
## 233	332	7	6	9	5	7	5	S 585
## 234	833	26	29	24	53	47	55	S 243
## 235	530	24	28	36	32	32	31	S 1162
## 236	1424	95	106	109	64	70	48	S 1020
## 237	93	14	13	12	13	6	11	S 549
## 238	529	122	117	115	44	33	29	S 1108
## 239	466	13	15	15	24	29	42	S 65
## 240	197	0	0	0	39	33	52	S 695
## 241	848	5	6	5	120	103	103	S 395
## 242	29	2	1	1	6	2	1	S 304
## 243	176	0	0	0	6	11	11	S 348
## 244	188	21	26	26	7	5	7	S 325
## 245	1058	21	16	14	30	30	42	S 109
## 246	604	0	0	0	78	101	127	S 554
## 247	507	131	141	136	12	22	24	S 1217

## 248	369	87	86	72	8	6	4	S	115
## 249	815	39	37	29	20	37	43	S	664
## 250	246	11	8	10	11	14	18	S	1292
## 251	268	113	109	111	13	11	9	S	1139
## 252	643	164	169	153	23	21	17	S	1181
## 253	1509	51	51	38	120	100	75	S	1082
## 254	25	14	13	15	0	2	0	S	1106
## 255	1081	33	29	29	36	52	47	S	75
## 256	874	66	61	51	70	82	65	S	996
## 257	477	6	4	9	17	12	27	S	81
## 258	857	3	3	2	90	111	125	S	413
## 259	1295	9	8	3	132	154	193	S	403
## 260	240	36	32	38	24	30	16	W	164
## 261	1441	300	322	290	330	327	330	W	325
## 262	29	9	8	9	0	1	1	W	283
## 263	670	26	25	24	505	365	316	W	145
## 264	612	108	97	89	41	43	32	W	345
## 265	677	65	59	48	37	56	36	W	338
## 266	499	289	308	310	9	41	32	W	200
## 267	198	124	124	123	9	15	10	W	323
## 268	675	443	467	443	12	17	18	W	210
## 269	210	99	100	97	13	29	17	W	228
## 270	781	11	9	10	121	106	114	W	300
## 271	465	133	108	96	88	87	66	W	254
## 272	222	39	38	40	10	7	11	W	150
## 273	575	124	129	134	39	40	43	W	380
## 274	2539	55	60	58	1960	1602	1468	W	142
## 275	442	271	292	293	10	31	27	W	405
## 276	354	255	287	273	6	20	13	W	240
## 277	121	8	15	12	11	11	13	W	11
## 278	138	56	75	52	13	6	8	W	103
## 279	1848	64	53	62	475	475	542	W	356
## 280	1196	36	30	30	163	150	180	W	322
## 281	451	65	72	71	20	43	30	W	149
## 282	484	214	216	209	13	31	13	W	114
## 283	439	179	181	194	13	37	32	W	218
## 284	124	29	35	30	9	7	12	W	97
## 285	270	44	44	40	8	8	9	W	186
## 286	584	335	353	350	9	20	18	W	232
## 287	597	128	141	137	23	47	43	W	194
## 288	422	1	0	0	143	164	137	W	378
## 289	19	2	2	3	2	0	0	W	105
## 290	876	114	109	109	139	134	120	W	32
## 291	417	65	64	66	47	69	55	W	353
## 292	1191	80	87	95	47	58	50	W	393
## 293	1618	8	5	7	271	274	317	W	366
## 294	145	0	0	0	5	12	8	W	395
## 295	856	127	140	120	107	109	101	W	371
## 296	534	208	205	191	59	67	63	W	259
## 297	856	5	4	4	171	143	151	W	394
## 298	34	3	2	2	15	15	17	W	295
## 299	617	79	81	83	98	112	106	W	148
## 300	625	67	79	66	25	39	33	W	112
##	weight								

1 10.23301
2 10.23301
3 10.23301
4 10.23301
5 10.23301
6 10.23301
7 10.23301
8 10.23301
9 10.23301
10 10.23301
11 10.23301
12 10.23301
13 10.23301
14 10.23301
15 10.23301
16 10.23301
17 10.23301
18 10.23301
19 10.23301
20 10.23301
21 10.23301
22 10.23301
23 10.23301
24 10.23301
25 10.23301
26 10.23301
27 10.23301
28 10.23301
29 10.23301
30 10.23301
31 10.23301
32 10.23301
33 10.23301
34 10.23301
35 10.23301
36 10.23301
37 10.23301
38 10.23301
39 10.23301
40 10.23301
41 10.23301
42 10.23301
43 10.23301
44 10.23301
45 10.23301
46 10.23301
47 10.23301
48 10.23301
49 10.23301
50 10.23301
51 10.23301
52 10.23301
53 10.23301
54 10.23301

55 10.23301
56 10.23301
57 10.23301
58 10.23301
59 10.23301
60 10.23301
61 10.23301
62 10.23301
63 10.23301
64 10.23301
65 10.23301
66 10.23301
67 10.23301
68 10.23301
69 10.23301
70 10.23301
71 10.23301
72 10.23301
73 10.23301
74 10.23301
75 10.23301
76 10.23301
77 10.23301
78 10.23301
79 10.23301
80 10.23301
81 10.23301
82 10.23301
83 10.23301
84 10.23301
85 10.23301
86 10.23301
87 10.23301
88 10.23301
89 10.23301
90 10.23301
91 10.23301
92 10.23301
93 10.23301
94 10.23301
95 10.23301
96 10.23301
97 10.23301
98 10.23301
99 10.23301
100 10.23301
101 10.23301
102 10.23301
103 10.23301
104 10.47619
105 10.47619
106 10.47619
107 10.47619
108 10.47619

109 10.47619
110 10.47619
111 10.47619
112 10.47619
113 10.47619
114 10.47619
115 10.47619
116 10.47619
117 10.47619
118 10.47619
119 10.47619
120 10.47619
121 10.47619
122 10.47619
123 10.47619
124 10.47619
125 10.23704
126 10.23704
127 10.23704
128 10.23704
129 10.23704
130 10.23704
131 10.23704
132 10.23704
133 10.23704
134 10.23704
135 10.23704
136 10.23704
137 10.23704
138 10.23704
139 10.23704
140 10.23704
141 10.23704
142 10.23704
143 10.23704
144 10.23704
145 10.23704
146 10.23704
147 10.23704
148 10.23704
149 10.23704
150 10.23704
151 10.23704
152 10.23704
153 10.23704
154 10.23704
155 10.23704
156 10.23704
157 10.23704
158 10.23704
159 10.23704
160 10.23704
161 10.23704
162 10.23704

163 10.23704
164 10.23704
165 10.23704
166 10.23704
167 10.23704
168 10.23704
169 10.23704
170 10.23704
171 10.23704
172 10.23704
173 10.23704
174 10.23704
175 10.23704
176 10.23704
177 10.23704
178 10.23704
179 10.23704
180 10.23704
181 10.23704
182 10.23704
183 10.23704
184 10.23704
185 10.23704
186 10.23704
187 10.23704
188 10.23704
189 10.23704
190 10.23704
191 10.23704
192 10.23704
193 10.23704
194 10.23704
195 10.23704
196 10.23704
197 10.23704
198 10.23704
199 10.23704
200 10.23704
201 10.23704
202 10.23704
203 10.23704
204 10.23704
205 10.23704
206 10.23704
207 10.23704
208 10.23704
209 10.23704
210 10.23704
211 10.23704
212 10.23704
213 10.23704
214 10.23704
215 10.23704
216 10.23704

217 10.23704
218 10.23704
219 10.23704
220 10.23704
221 10.23704
222 10.23704
223 10.23704
224 10.23704
225 10.23704
226 10.23704
227 10.23704
228 10.23704
229 10.23704
230 10.23704
231 10.23704
232 10.23704
233 10.23704
234 10.23704
235 10.23704
236 10.23704
237 10.23704
238 10.23704
239 10.23704
240 10.23704
241 10.23704
242 10.23704
243 10.23704
244 10.23704
245 10.23704
246 10.23704
247 10.23704
248 10.23704
249 10.23704
250 10.23704
251 10.23704
252 10.23704
253 10.23704
254 10.23704
255 10.23704
256 10.23704
257 10.23704
258 10.23704
259 10.23704
260 10.29268
261 10.29268
262 10.29268
263 10.29268
264 10.29268
265 10.29268
266 10.29268
267 10.29268
268 10.29268
269 10.29268
270 10.29268

```
## 271 10.29268
## 272 10.29268
## 273 10.29268
## 274 10.29268
## 275 10.29268
## 276 10.29268
## 277 10.29268
## 278 10.29268
## 279 10.29268
## 280 10.29268
## 281 10.29268
## 282 10.29268
## 283 10.29268
## 284 10.29268
## 285 10.29268
## 286 10.29268
## 287 10.29268
## 288 10.29268
## 289 10.29268
## 290 10.29268
## 291 10.29268
## 292 10.29268
## 293 10.29268
## 294 10.29268
## 295 10.29268
## 296 10.29268
## 297 10.29268
## 298 10.29268
## 299 10.29268
## 300 10.29268
```

2.2 Spreadsheet calculation

2.2.1 Summarizing acre92 in each stratum

```
nh <- tapply (agstrat[, "acres92"], agstrat[, "region"], length)
sh <- tapply (agstrat[, "acres92"], agstrat[, "region"], sd)
ybarh <- tapply (agstrat[, "acres92"], agstrat[, "region"], mean)
# create a vector with external information
Nh <- c(NC = 1054, NE = 220, S = 1382, W = 422)

data.frame (Nh, nh, ybarh, sh)
```

```
##      Nh  nh    ybarh    sh
## NC 1054 103 300504.16 172099.34
## NE  220  21  97629.81  87449.83
## S   1382 135 211315.04 231489.71
## W    422  41 662295.51 629433.04
```

2.2.2 Estimates

```
N <- sum (Nh)
pi_h <- Nh/N
weighted_ybar_h <- ybarh * pi_h
```

```

ybar <- sum(weighted_ybar_h)
var_h=(1-nh/Nh)*pi_h^2*sh^2/nh
library ("tibble")
tibble (`$N_h$`=Nh, nh=nh,pi_h, ybarh, sh=sh, var_h)

## # A tibble: 4 x 6
##   `$N_h$`      nh    pi_h      ybarh      sh      var_h
##   <dbl> <int[1d]> <dbl> <dbl[1d]> <dbl[1d]> <dbl[1d]>
## 1    1054      103 0.342   300504.   172099.  30423214.
## 2     220       21 0.0715   97630.    87450.   1682818.
## 3    1382      135 0.449   211315.   231490.  72204937.
## 4     422       41 0.137   662296.   629433. 163989261.

seybar <- sqrt(sum((1-nh/Nh)*pi_h^2*sh^2/nh))
mem <- 1.96 * seybar
c(Est. = ybar, S.E. = seybar, ci.low = ybar - mem, ci.upp = ybar + mem)

##      Est.      S.E.    ci.low    ci.upp
## 295560.77 16379.87 263456.21 327665.32

```

2.3 Using the function “str_mean_estimate”

```

nh <- tapply (agstrat[, "acres92"], agstrat[, "region"], length)
sh <- tapply (agstrat[, "acres92"], agstrat[, "region"], sd)
ybarh <- tapply (agstrat[, "acres92"], agstrat[, "region"], mean)
# create a vector with external information
Nh <- c(NC = 1054, NE = 220, S= 1382, W = 422)

## strata mean estimate
str_mean_estimate (ybarh, sh, nh, Nh)

##      Est.      S.E.    ci.low    ci.upp
## 295560.77 16379.87 263456.21 327665.32

## population total estimate
str_mean_estimate (ybarh, sh, nh, Nh) * sum (Nh)

##      Est.      S.E.    ci.low    ci.upp
## 909736035 50417248 810918229 1008553842

```

2.4 Using the function “str_mean_estimate_data”

```

## In the fucntion "str_mean_estimate_data", we can find the stratum size with the variable "weights":
tapply (agstrat$weight, agstrat$region, sum)

##    NC    NE     S     W
## 1054   220  1382   422

## if the dataset contains a variable "weight"
str_mean_estimate_data (agstrat, y="acres92", stratum="region", weight="weight")

##      Est.      S.E.    ci.low    ci.upp
## 295560.77 16379.87 263456.21 327665.32

## estimating the mean of the number of small farms in 1992
str_mean_estimate_data (agstrat, y="smallf92", stratum="region", weight="weight")

```

```
##      Est.      S.E.    ci.low  ci.upp
## 56.862794  7.201417 42.748016 70.977572
```

2.5 Comparing with SRS estimate

```
agsrs <- read.csv ("data/agsrs.csv")
srs_mean_est(agsrs[, "acres92"], N=3078)
```

```
##      Est.      S.E.    ci.low  ci.upp
## 297897.05 18898.43 260706.26 335087.84
```

```
## the ratio of estimated variance
(16379.87/18898.43)^2
```

```
## [1] 0.751224
```

```
## percentage of reduction of variance of str estimates from that of SRS estimate
1- (16379.87/18898.43)^2
```

```
## [1] 0.248776
```

3 Allocation of stratum sample size

3.1 Analyzing seals.csv collected with stratified sampling

Lydersen and Ryg (1991) used stratification techniques to estimate ringed seal populations in Svalbard fjords. The 200 km² study area was divided into three zones: Zone 1, outer Sassenfjorden, was covered with relatively new ice during the study period in March, 1990, and had little snow cover; Zone 3, Tempelfjorden, had a stable ice cover throughout the year; Zone 2, inner Sassenfjorden, was intermediate between the stable Zone 3 and the unstable Zone 1. Ringed seals need good ice to establish territories with breathing holes, and snow cover enables females to dig out birth lairs. Thus, it was thought that the three zones would have different seal densities. The investigators took a stratified random sample of 20% of the 200 1-km² areas.

```
#####
```

```
## load data
seals <- read.csv("data/seals.csv")
seals
```

```
##      zone holes
## 1      1      0
## 2      1      0
## 3      1      0
## 4      1      0
## 5      1      1
## 6      1      1
## 7      1      1
## 8      1      1
## 9      1      1
## 10     1      1
## 11     1      2
## 12     1      2
## 13     1      3
## 14     1      3
## 15     1      3
## 16     1      4
```

```
## 17    1    7
## 18    2    0
## 19    2    2
## 20    2    2
## 21    2    2
## 22    2    3
## 23    2    3
## 24    2    4
## 25    2    4
## 26    2    5
## 27    2    7
## 28    2    9
## 29    2   12
## 30    3    0
## 31    3    1
## 32    3    5
## 33    3    8
## 34    3   10
## 35    3   10
## 36    3   12
## 37    3   14
## 38    3   17
## 39    3   18
## 40    3   21
```

```
# survey data summary in each stratum
nh <- as.vector(table(seals[, "zone"]))
sh <- tapply(seals[, "holes"], seals[, "zone"], sd)
ybarh <- tapply(seals[, "holes"], seals[, "zone"], mean)
Nh <- c(68, 84, 48)
data.frame(Nh, nh, ybarh, sh)
```

```
##   Nh nh   ybarh   sh
## 1 68 17  1.764706 1.821037
## 2 84 12  4.416667 3.396745
## 3 48 11 10.545455 6.787689
```

```
str_mean_estimate(ybarh, sh, nh, Nh)
```

```
##      Est.      S.E.    ci.low  ci.upp
## 4.9859091 0.5901322 3.8292499 6.1425683
```

3.2 Neyman allocation of stratum sample size

```
## assuming the cost of surveying each zone is the same
Ch <- rep(1, 3)
Sh <- sh ## estimate Sh with sh
NhShDCh <- Nh * Sh / sqrt(Ch)

## the optimal allocation scheme for estimating the population total or means holes:
Lh <- NhShDCh / sum(NhShDCh)
data.frame(Ch, Nh, Sh, NhShDCh, Lh)
```

```
##   Ch Nh   Sh NhShDCh   Lh
## 1  1 68 1.821037 123.8305 0.1684847
## 2  1 84 3.396745 285.3266 0.3882173
```

```
## 3 1 48 6.787689 325.8091 0.4432980
```

4 Simulation to Study the Efficiency of Stratified Sampling with Different Allocation

4.1 Using “Region” to Stratify

4.1.1 Read Population Data

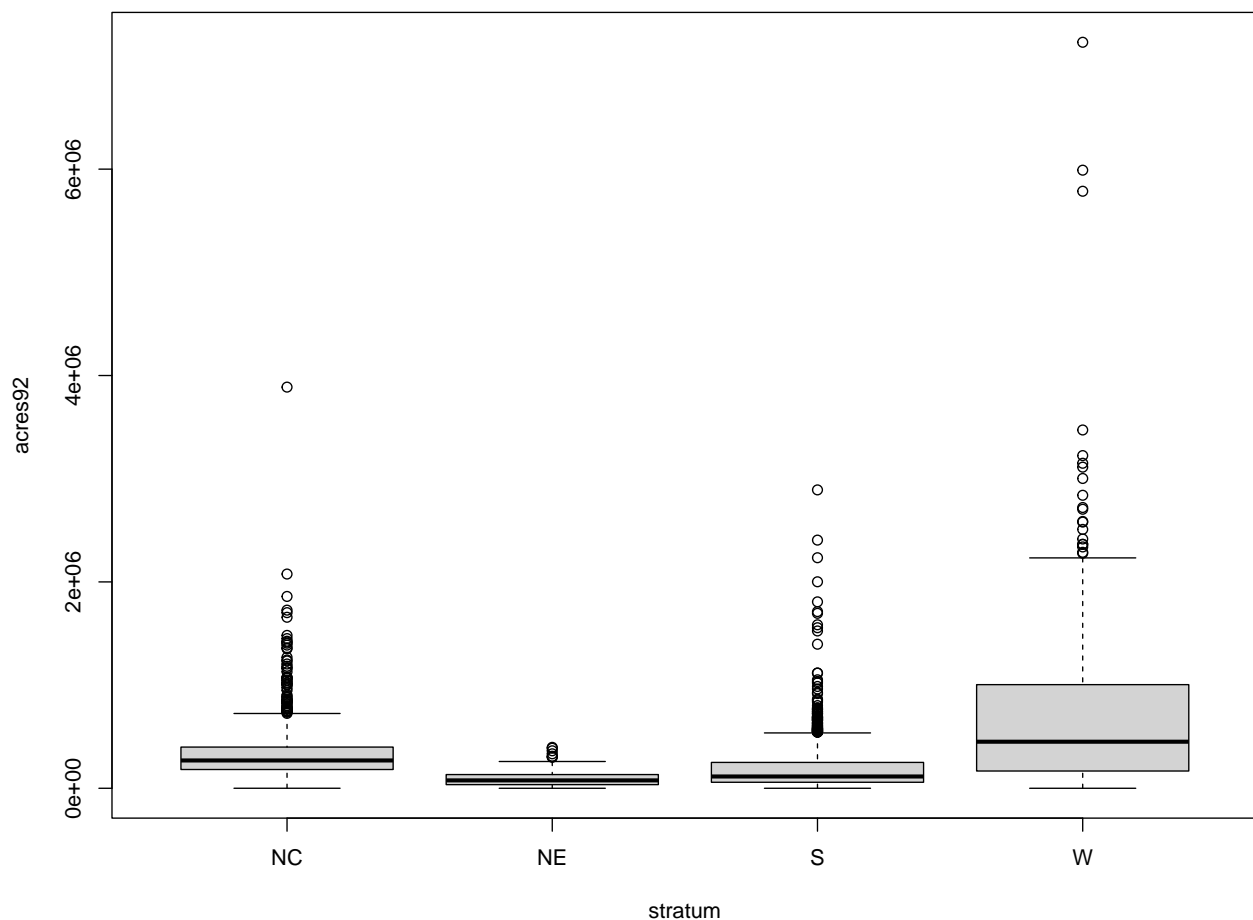
```
agpop <- read.csv ("data/agpop.csv")
agpop <- agpop[agpop$acres92 != -99, ] ## remove those counties with na
N <- nrow(agpop)
```

4.1.2 Define Stratum Variable

```
agpop$stratum <- agpop$region
## reorder agpop for the ease of using strata of "sampling" (very important)
agpop <- agpop[order (agpop$stratum), ]
```

Look at variance decomposition

```
boxplot (acres92 ~ stratum, data = agpop)
```



```
anova(lm (agpop$acres92~agpop$stratum))
```

```
## Analysis of Variance Table
##
## Response: agpop$acres92
##           Df      Sum Sq   Mean Sq F value    Pr(>F)
## agpop$stratum    3 1.0073e+14 3.3578e+13  226.73 < 2.2e-16 ***
## Residuals      3055 4.5243e+14 1.4810e+11
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Look at the R^2 of predicting “acres92” with “region”.

```
summary(lm (agpop$acres92~agpop$stratum))$r.squared
```

```
## [1] 0.1821031
```

4.1.3 Stratified Sampling with P2S allocation

```
# doing one stratified sampling
Nh <- tapply (1:nrow(agpop), agpop$stratum, length)
nh <- round(Nh/sum (Nh)*300)
data.frame(Nh,nh)
```

```
##      Nh  nh
## NC 1052 103
## NE  213  21
## S  1376 135
## W   418  41
```

```
nh
```

```
##  NC  NE   S   W
## 103  21 135  41
```

```
## doing stratified sampling
strsample <- strata (agpop, "stratum", size = nh, method = "srswor")
strsample
```

```
##      stratum ID_unit      Prob Stratum
## 12         NC      12 0.09790875      1
## 15         NC      15 0.09790875      1
## 23         NC      23 0.09790875      1
## 29         NC      29 0.09790875      1
## 36         NC      36 0.09790875      1
## 41         NC      41 0.09790875      1
## 53         NC      53 0.09790875      1
## 79         NC      79 0.09790875      1
## 93         NC      93 0.09790875      1
## 96         NC      96 0.09790875      1
## 104        NC     104 0.09790875      1
## 116        NC     116 0.09790875      1
## 118        NC     118 0.09790875      1
## 119        NC     119 0.09790875      1
## 133        NC     133 0.09790875      1
## 135        NC     135 0.09790875      1
## 145        NC     145 0.09790875      1
```


## 146	NC	146 0.09790875	1
## 150	NC	150 0.09790875	1
## 151	NC	151 0.09790875	1
## 155	NC	155 0.09790875	1
## 157	NC	157 0.09790875	1
## 160	NC	160 0.09790875	1
## 169	NC	169 0.09790875	1
## 181	NC	181 0.09790875	1
## 194	NC	194 0.09790875	1
## 195	NC	195 0.09790875	1
## 209	NC	209 0.09790875	1
## 211	NC	211 0.09790875	1
## 212	NC	212 0.09790875	1
## 220	NC	220 0.09790875	1
## 240	NC	240 0.09790875	1
## 245	NC	245 0.09790875	1
## 256	NC	256 0.09790875	1
## 266	NC	266 0.09790875	1
## 280	NC	280 0.09790875	1
## 285	NC	285 0.09790875	1
## 287	NC	287 0.09790875	1
## 289	NC	289 0.09790875	1
## 291	NC	291 0.09790875	1
## 301	NC	301 0.09790875	1
## 311	NC	311 0.09790875	1
## 314	NC	314 0.09790875	1
## 331	NC	331 0.09790875	1
## 335	NC	335 0.09790875	1
## 343	NC	343 0.09790875	1
## 359	NC	359 0.09790875	1
## 375	NC	375 0.09790875	1
## 377	NC	377 0.09790875	1
## 401	NC	401 0.09790875	1
## 408	NC	408 0.09790875	1
## 410	NC	410 0.09790875	1
## 415	NC	415 0.09790875	1
## 428	NC	428 0.09790875	1
## 449	NC	449 0.09790875	1
## 451	NC	451 0.09790875	1
## 455	NC	455 0.09790875	1
## 465	NC	465 0.09790875	1
## 466	NC	466 0.09790875	1
## 477	NC	477 0.09790875	1
## 479	NC	479 0.09790875	1
## 518	NC	518 0.09790875	1
## 557	NC	557 0.09790875	1
## 562	NC	562 0.09790875	1
## 572	NC	572 0.09790875	1
## 580	NC	580 0.09790875	1
## 598	NC	598 0.09790875	1
## 610	NC	610 0.09790875	1
## 624	NC	624 0.09790875	1
## 651	NC	651 0.09790875	1
## 666	NC	666 0.09790875	1

## 669	NC	669 0.09790875	1
## 677	NC	677 0.09790875	1
## 721	NC	721 0.09790875	1
## 729	NC	729 0.09790875	1
## 737	NC	737 0.09790875	1
## 756	NC	756 0.09790875	1
## 779	NC	779 0.09790875	1
## 798	NC	798 0.09790875	1
## 818	NC	818 0.09790875	1
## 840	NC	840 0.09790875	1
## 851	NC	851 0.09790875	1
## 853	NC	853 0.09790875	1
## 858	NC	858 0.09790875	1
## 859	NC	859 0.09790875	1
## 868	NC	868 0.09790875	1
## 886	NC	886 0.09790875	1
## 920	NC	920 0.09790875	1
## 924	NC	924 0.09790875	1
## 935	NC	935 0.09790875	1
## 943	NC	943 0.09790875	1
## 952	NC	952 0.09790875	1
## 957	NC	957 0.09790875	1
## 976	NC	976 0.09790875	1
## 979	NC	979 0.09790875	1
## 983	NC	983 0.09790875	1
## 985	NC	985 0.09790875	1
## 991	NC	991 0.09790875	1
## 993	NC	993 0.09790875	1
## 1007	NC	1007 0.09790875	1
## 1017	NC	1017 0.09790875	1
## 1038	NC	1038 0.09790875	1
## 1042	NC	1042 0.09790875	1
## 1054	NE	1054 0.09859155	2
## 1057	NE	1057 0.09859155	2
## 1058	NE	1058 0.09859155	2
## 1066	NE	1066 0.09859155	2
## 1075	NE	1075 0.09859155	2
## 1124	NE	1124 0.09859155	2
## 1136	NE	1136 0.09859155	2
## 1137	NE	1137 0.09859155	2
## 1138	NE	1138 0.09859155	2
## 1150	NE	1150 0.09859155	2
## 1156	NE	1156 0.09859155	2
## 1160	NE	1160 0.09859155	2
## 1188	NE	1188 0.09859155	2
## 1195	NE	1195 0.09859155	2
## 1207	NE	1207 0.09859155	2
## 1208	NE	1208 0.09859155	2
## 1224	NE	1224 0.09859155	2
## 1230	NE	1230 0.09859155	2
## 1245	NE	1245 0.09859155	2
## 1249	NE	1249 0.09859155	2
## 1250	NE	1250 0.09859155	2
## 1270	S	1270 0.09811047	3

## 1275	S	1275 0.09811047	3
## 1278	S	1278 0.09811047	3
## 1289	S	1289 0.09811047	3
## 1312	S	1312 0.09811047	3
## 1319	S	1319 0.09811047	3
## 1337	S	1337 0.09811047	3
## 1338	S	1338 0.09811047	3
## 1341	S	1341 0.09811047	3
## 1347	S	1347 0.09811047	3
## 1351	S	1351 0.09811047	3
## 1352	S	1352 0.09811047	3
## 1353	S	1353 0.09811047	3
## 1377	S	1377 0.09811047	3
## 1380	S	1380 0.09811047	3
## 1383	S	1383 0.09811047	3
## 1390	S	1390 0.09811047	3
## 1392	S	1392 0.09811047	3
## 1397	S	1397 0.09811047	3
## 1399	S	1399 0.09811047	3
## 1401	S	1401 0.09811047	3
## 1404	S	1404 0.09811047	3
## 1448	S	1448 0.09811047	3
## 1460	S	1460 0.09811047	3
## 1462	S	1462 0.09811047	3
## 1482	S	1482 0.09811047	3
## 1503	S	1503 0.09811047	3
## 1520	S	1520 0.09811047	3
## 1521	S	1521 0.09811047	3
## 1526	S	1526 0.09811047	3
## 1528	S	1528 0.09811047	3
## 1554	S	1554 0.09811047	3
## 1555	S	1555 0.09811047	3
## 1567	S	1567 0.09811047	3
## 1606	S	1606 0.09811047	3
## 1612	S	1612 0.09811047	3
## 1619	S	1619 0.09811047	3
## 1633	S	1633 0.09811047	3
## 1635	S	1635 0.09811047	3
## 1652	S	1652 0.09811047	3
## 1668	S	1668 0.09811047	3
## 1669	S	1669 0.09811047	3
## 1674	S	1674 0.09811047	3
## 1692	S	1692 0.09811047	3
## 1709	S	1709 0.09811047	3
## 1725	S	1725 0.09811047	3
## 1730	S	1730 0.09811047	3
## 1731	S	1731 0.09811047	3
## 1736	S	1736 0.09811047	3
## 1738	S	1738 0.09811047	3
## 1781	S	1781 0.09811047	3
## 1791	S	1791 0.09811047	3
## 1796	S	1796 0.09811047	3
## 1798	S	1798 0.09811047	3
## 1825	S	1825 0.09811047	3

## 1833	S	1833 0.09811047	3
## 1854	S	1854 0.09811047	3
## 1874	S	1874 0.09811047	3
## 1879	S	1879 0.09811047	3
## 1886	S	1886 0.09811047	3
## 1901	S	1901 0.09811047	3
## 1904	S	1904 0.09811047	3
## 1906	S	1906 0.09811047	3
## 1911	S	1911 0.09811047	3
## 1932	S	1932 0.09811047	3
## 1942	S	1942 0.09811047	3
## 1944	S	1944 0.09811047	3
## 1947	S	1947 0.09811047	3
## 1948	S	1948 0.09811047	3
## 1963	S	1963 0.09811047	3
## 1968	S	1968 0.09811047	3
## 1984	S	1984 0.09811047	3
## 1988	S	1988 0.09811047	3
## 1991	S	1991 0.09811047	3
## 1999	S	1999 0.09811047	3
## 2004	S	2004 0.09811047	3
## 2010	S	2010 0.09811047	3
## 2030	S	2030 0.09811047	3
## 2040	S	2040 0.09811047	3
## 2044	S	2044 0.09811047	3
## 2058	S	2058 0.09811047	3
## 2070	S	2070 0.09811047	3
## 2072	S	2072 0.09811047	3
## 2079	S	2079 0.09811047	3
## 2082	S	2082 0.09811047	3
## 2086	S	2086 0.09811047	3
## 2103	S	2103 0.09811047	3
## 2114	S	2114 0.09811047	3
## 2124	S	2124 0.09811047	3
## 2155	S	2155 0.09811047	3
## 2160	S	2160 0.09811047	3
## 2166	S	2166 0.09811047	3
## 2175	S	2175 0.09811047	3
## 2193	S	2193 0.09811047	3
## 2195	S	2195 0.09811047	3
## 2204	S	2204 0.09811047	3
## 2223	S	2223 0.09811047	3
## 2233	S	2233 0.09811047	3
## 2235	S	2235 0.09811047	3
## 2238	S	2238 0.09811047	3
## 2253	S	2253 0.09811047	3
## 2261	S	2261 0.09811047	3
## 2271	S	2271 0.09811047	3
## 2275	S	2275 0.09811047	3
## 2280	S	2280 0.09811047	3
## 2284	S	2284 0.09811047	3
## 2303	S	2303 0.09811047	3
## 2349	S	2349 0.09811047	3
## 2372	S	2372 0.09811047	3

##	2376	S	2376 0.09811047	3
##	2382	S	2382 0.09811047	3
##	2390	S	2390 0.09811047	3
##	2392	S	2392 0.09811047	3
##	2400	S	2400 0.09811047	3
##	2422	S	2422 0.09811047	3
##	2429	S	2429 0.09811047	3
##	2433	S	2433 0.09811047	3
##	2438	S	2438 0.09811047	3
##	2456	S	2456 0.09811047	3
##	2482	S	2482 0.09811047	3
##	2483	S	2483 0.09811047	3
##	2486	S	2486 0.09811047	3
##	2490	S	2490 0.09811047	3
##	2499	S	2499 0.09811047	3
##	2518	S	2518 0.09811047	3
##	2526	S	2526 0.09811047	3
##	2534	S	2534 0.09811047	3
##	2550	S	2550 0.09811047	3
##	2551	S	2551 0.09811047	3
##	2571	S	2571 0.09811047	3
##	2573	S	2573 0.09811047	3
##	2579	S	2579 0.09811047	3
##	2587	S	2587 0.09811047	3
##	2598	S	2598 0.09811047	3
##	2627	S	2627 0.09811047	3
##	2651	W	2651 0.09808612	4
##	2656	W	2656 0.09808612	4
##	2662	W	2662 0.09808612	4
##	2669	W	2669 0.09808612	4
##	2670	W	2670 0.09808612	4
##	2685	W	2685 0.09808612	4
##	2707	W	2707 0.09808612	4
##	2711	W	2711 0.09808612	4
##	2726	W	2726 0.09808612	4
##	2733	W	2733 0.09808612	4
##	2757	W	2757 0.09808612	4
##	2762	W	2762 0.09808612	4
##	2765	W	2765 0.09808612	4
##	2773	W	2773 0.09808612	4
##	2781	W	2781 0.09808612	4
##	2805	W	2805 0.09808612	4
##	2811	W	2811 0.09808612	4
##	2812	W	2812 0.09808612	4
##	2822	W	2822 0.09808612	4
##	2867	W	2867 0.09808612	4
##	2878	W	2878 0.09808612	4
##	2881	W	2881 0.09808612	4
##	2886	W	2886 0.09808612	4
##	2901	W	2901 0.09808612	4
##	2902	W	2902 0.09808612	4
##	2913	W	2913 0.09808612	4
##	2922	W	2922 0.09808612	4
##	2934	W	2934 0.09808612	4

```
## 2937      W      2937 0.09808612      4
## 2938      W      2938 0.09808612      4
## 2957      W      2957 0.09808612      4
## 2968      W      2968 0.09808612      4
## 2970      W      2970 0.09808612      4
## 2977      W      2977 0.09808612      4
## 2985      W      2985 0.09808612      4
## 3008      W      3008 0.09808612      4
## 3016      W      3016 0.09808612      4
## 3022      W      3022 0.09808612      4
## 3043      W      3043 0.09808612      4
## 3048      W      3048 0.09808612      4
## 3055      W      3055 0.09808612      4
```

```
## checking sampling results
```

```
table (strsample [,1])
```

```
##
```

```
## NC NE S W
```

```
## 103 21 135 41
```

```
# collecting data on sampled counties
```

```
agstrat <- agpop [strsample$ID_unit, ]
```

```
agstrat$weight <- 1/strsample$Prob
```

```
str_mean_estimate_data (agstrat, "acres92", "stratum", "weight")
```

```
##      Est.      S.E.      ci.low      ci.upp
```

```
## 337500.44  28795.76 281060.75 393940.12
```

Estimate with spreadsheet method for a double checking

```
## double checking the estimates
```

```
nh <- tapply (agstrat[, "acres92"], agstrat[, "region"], length)
```

```
sh <- tapply (agstrat[, "acres92"], agstrat[, "region"], sd)
```

```
ybarh <- tapply (agstrat[, "acres92"], agstrat[, "region"], mean)
```

```
# create a vector with external information
```

```
Nh <- tapply (1:nrow(agpop), agpop[, "region"], length)
```

```
## strata mean estimate
```

```
str_mean_estimate (ybarh, sh, nh, Nh)
```

```
##      Est.      S.E.      ci.low      ci.upp
```

```
## 337500.44  28795.76 281060.75 393940.12
```

4.1.4 Repeat stratified sampling with P2S allocation 2000 times

```
Nh <- tapply (1:nrow (agpop), agpop$stratum, length)
```

```
nh <- round(Nh/sum (Nh)*300) ## make sure the order matches Nh
```

```
nres <- 2000
```

```
str_p2s_simulated <- matrix (0, nres, 4)
```

```
for (i in 1:nres)
```

```
{
```

```
  ## doing stratified sampling
```

```
  strsample <- strata (agpop, "stratum", size = nh, method = "srswor")
```

```
  agstrat <- agpop [strsample$ID_unit, ]
```

```
  agstrat$weight <- 1/strsample$Prob
```

```

    ## analyzing data
    str_p2s_simulated[i,] <- str_mean_estimate_data (agstrat, "acres92", "stratum", "weight")
  }

```

4.1.5 Repeat stratified sampling with optimal allocation 2000 times

```

Nh <- tapply (agpop$acres92, agpop$stratum, length)

Sh <- tapply (agpop$acres92, agpop$stratum, sd)
nh_opt <- round((Nh*Sh)/sum (Nh*Sh) * 300);nh_opt

##  NC  NE   S   W
##  87   5 102 106

data.frame(Nh,Sh,nh_opt)

##      Nh      Sh nh_opt
## NC 1052 271188.0     87
## NE  213  78906.2      5
## S  1376 244132.0    102
## W   418 836613.6    106

nres <- 2000
str_neyman_simulated <- matrix (0, nres, 4)
for (i in 1:nres)
{
  ## doing stratified sampling
  strsample <- strata (agpop, "stratum", size = nh_opt, method = "srswor")
  # collecting data on sampled counties
  agstrat <- agpop [strsample$ID_unit, ]
  agstrat$weight <- 1/strsample$Prob

  str_neyman_simulated[i,] <- str_mean_estimate_data (agstrat, "acres92", "stratum", "weight")
}

```

4.1.6 Repeat simple random sampling 2000 times

```

nres <- 2000
srs_simulated <- matrix (0, nres, 4)
for (i in 1:nres)
{
  ## doing stratified sampling
  srs <- sample (sum(Nh), sum (nh))
  # collecting data on sampled counties
  agsrs <- agpop [srs, ]
  srs_simulated[i,] <- srs_mean_est (agsrs[, "acres92"], N= sum(Nh))
}

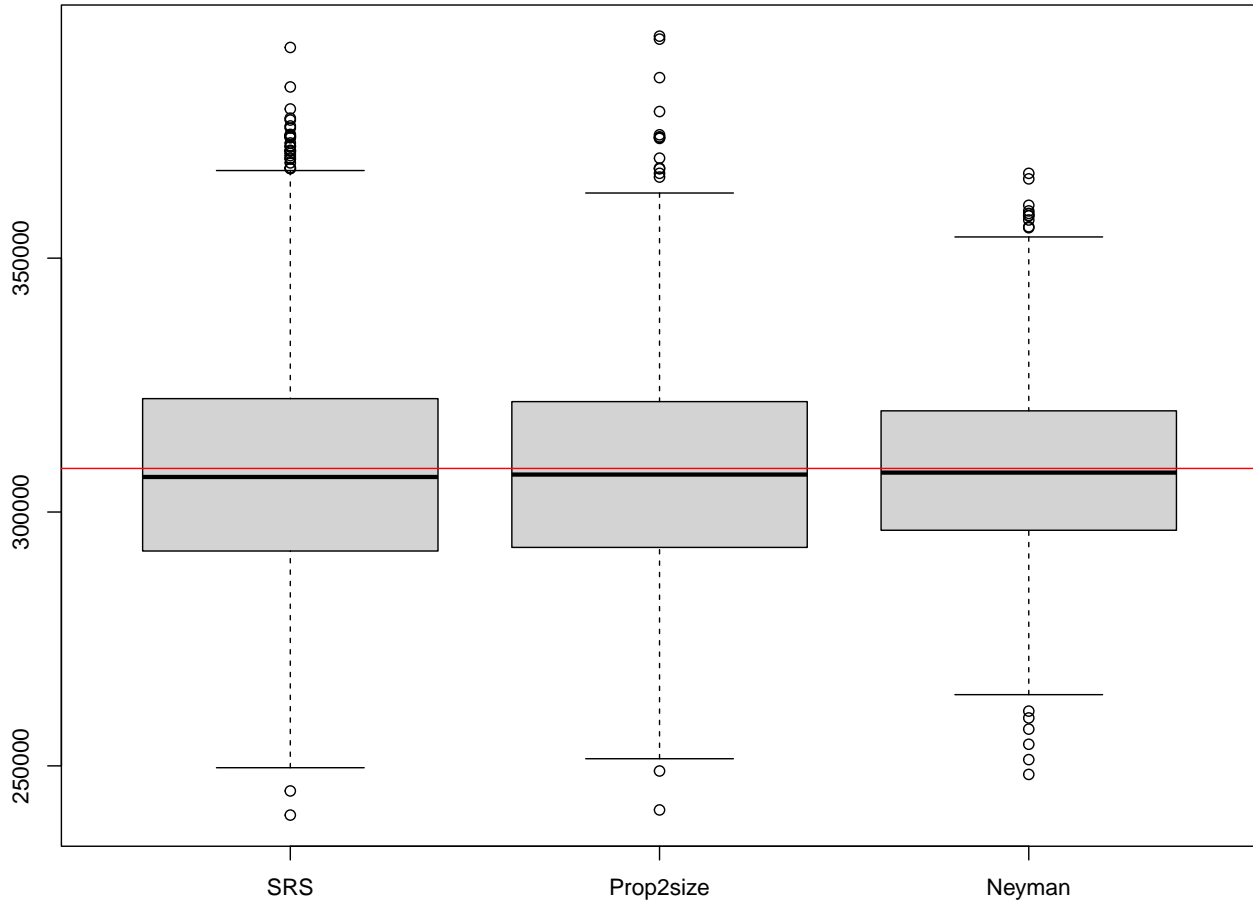
```

4.1.7 Compare the efficiency of different methods

```

sim_results_str_region <- data.frame("SRS"=srs_simulated[,1],
                                     "Prop2size"=str_p2s_simulated[,1],
                                     "Neyman"=str_neyman_simulated[,1])
boxplot (sim_results_str_region)
abline (h = mean (agpop$acres92), col = "red")

```



```

sapply (sim_results_str_region, mean) -> sim_means
sapply (sim_results_str_region, var) -> sim_var

sim_var/sim_var[1] -> sim_var_relative

#Percentage of reduction of variances of estimates compared to SRS

1-sim_var/sim_var[1] -> sim_var_reduction

cbind("Mean"=sim_means, "Variance"=sim_var, "Relative Variance"= sim_var_relative, "Percentage of Variance Reduction"=sim_var_reduction)

```

	Mean	Variance	Relative Variance	Percentage of Variance Reduction
## SRS	308075.0	530066896	1.0000000	0.0000000
## Prop2size	308182.8	447482662	0.8442004	0.1557996
## Neyman	308180.5	297403782	0.5610684	0.4389316

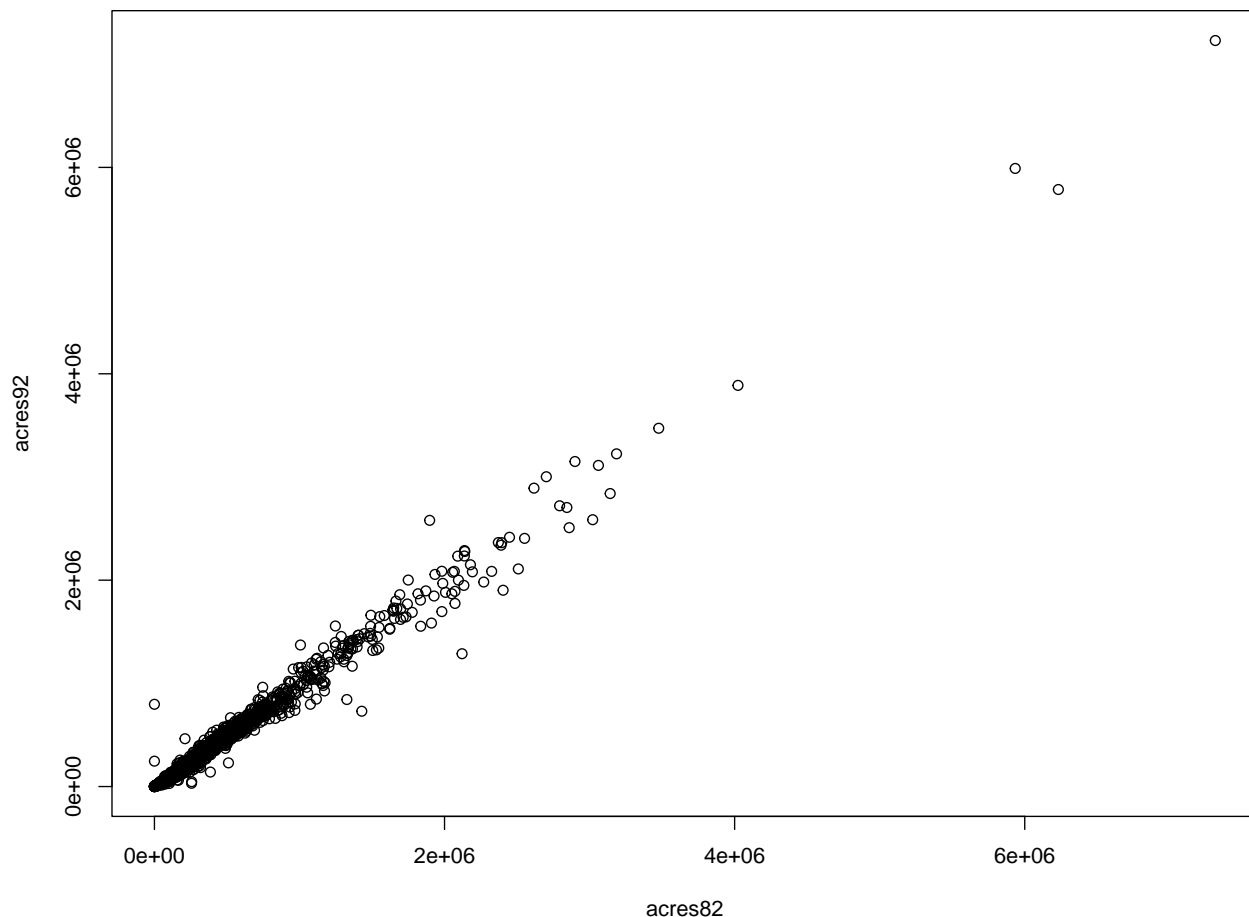
4.2 Using “acres82” to Define Strata

4.2.1 Read Population Data

```
agpop <- read.csv ("data/agpop.csv")
agpop <- agpop[agpop$acres92 != -99, ] ## remove those counties with na
N <- nrow(agpop)
```

4.2.2 Define Stratum Variable with Quantiles of “acres82”

```
plot (acres92~acres82, data = agpop)
```



```
summary (lm(acres92~acres82, data = agpop))
```

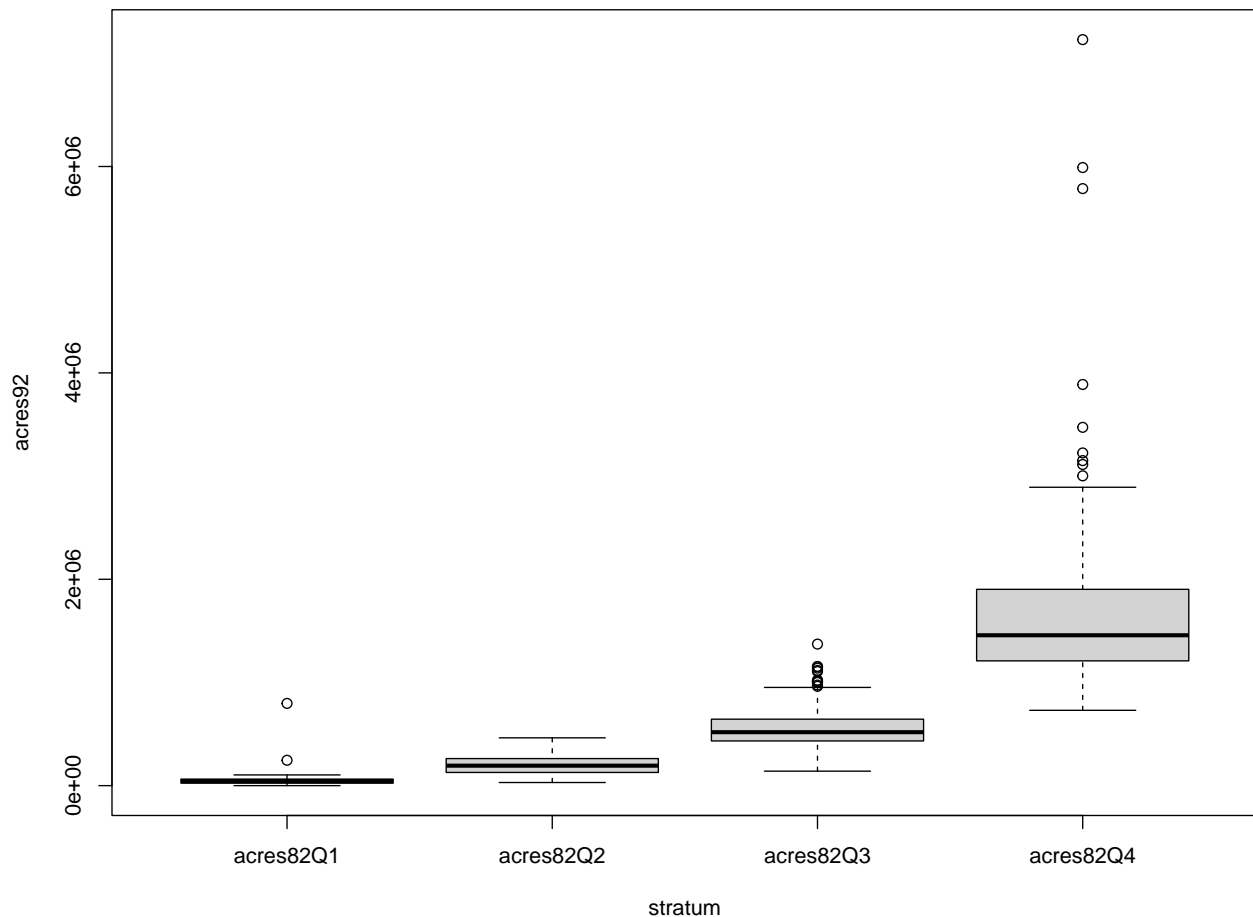
```
##
## Call:
## lm(formula = acres92 ~ acres82, data = agpop)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -786951  -11680    -548     8907   804356
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -7.142e+03  1.159e+03  -6.165 7.99e-10 ***
```

```
## acres82      9.813e-01  2.157e-03 455.015 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 51310 on 3057 degrees of freedom
## Multiple R-squared:  0.9854, Adjusted R-squared:  0.9854
## F-statistic: 2.07e+05 on 1 and 3057 DF,  p-value: < 2.2e-16

agpop$stratum <- cut(agpop$acres82,
                     breaks = quantile (agpop$acres82,
                                         probs = c(0,0.25,0.75,0.95,1)),
                     include.lowest = T)
agpop$stratum <- mapvalues(agpop$stratum,
                          from = (sort(unique (agpop$stratum))),
                          to= paste0("acres82",c("Q1", "Q2", "Q3", "Q4")))
agpop <- agpop[order(agpop$stratum), ]
```

Look at variance decomposition

```
boxplot (acres92 ~ stratum, data = agpop)
```



```
anova(lm (agpop$acres92~agpop$stratum))
```

```
## Analysis of Variance Table
##
## Response: agpop$acres92
##      Df      Sum Sq   Mean Sq F value    Pr(>F)
```

```
## agpop$stratum      3 4.0896e+14 1.3632e+14 2887.9 < 2.2e-16 ***
## Residuals         3055 1.4421e+14 4.7204e+10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Look at the R^2 of predicting “acres92” with “acres82” quantile.

```
summary(lm (agpop$acres92~agpop$stratum))$r.squared
```

```
## [1] 0.7393047
```

4.2.3 Stratified sampling with P2S allocation

```
# doing one stratified sampling
Nh <- tapply (1:nrow(agpop), agpop$stratum, length)
nh <- round(Nh/sum (Nh)*300)
nh
```

```
## acres82Q1 acres82Q2 acres82Q3 acres82Q4
##          75         150          60         15
```

```
## doing stratified sampling
strsample <- strata (agpop, "stratum", size = nh, method = "srswor")
strsample
```

```
##      stratum ID_unit      Prob Stratum
## 4    acres82Q1      4 0.09803922      1
## 12   acres82Q1     12 0.09803922      1
## 31   acres82Q1     31 0.09803922      1
## 38   acres82Q1     38 0.09803922      1
## 46   acres82Q1     46 0.09803922      1
## 51   acres82Q1     51 0.09803922      1
## 55   acres82Q1     55 0.09803922      1
## 56   acres82Q1     56 0.09803922      1
## 62   acres82Q1     62 0.09803922      1
## 70   acres82Q1     70 0.09803922      1
## 72   acres82Q1     72 0.09803922      1
## 73   acres82Q1     73 0.09803922      1
## 93   acres82Q1     93 0.09803922      1
## 94   acres82Q1     94 0.09803922      1
## 106  acres82Q1    106 0.09803922      1
## 120  acres82Q1    120 0.09803922      1
## 132  acres82Q1    132 0.09803922      1
## 136  acres82Q1    136 0.09803922      1
## 145  acres82Q1    145 0.09803922      1
## 146  acres82Q1    146 0.09803922      1
## 150  acres82Q1    150 0.09803922      1
## 152  acres82Q1    152 0.09803922      1
## 156  acres82Q1    156 0.09803922      1
## 163  acres82Q1    163 0.09803922      1
## 167  acres82Q1    167 0.09803922      1
## 180  acres82Q1    180 0.09803922      1
## 185  acres82Q1    185 0.09803922      1
## 187  acres82Q1    187 0.09803922      1
## 213  acres82Q1    213 0.09803922      1
## 218  acres82Q1    218 0.09803922      1
```

## 226	acres82Q1	226	0.09803922	1
## 230	acres82Q1	230	0.09803922	1
## 231	acres82Q1	231	0.09803922	1
## 232	acres82Q1	232	0.09803922	1
## 257	acres82Q1	257	0.09803922	1
## 267	acres82Q1	267	0.09803922	1
## 284	acres82Q1	284	0.09803922	1
## 297	acres82Q1	297	0.09803922	1
## 302	acres82Q1	302	0.09803922	1
## 309	acres82Q1	309	0.09803922	1
## 327	acres82Q1	327	0.09803922	1
## 339	acres82Q1	339	0.09803922	1
## 343	acres82Q1	343	0.09803922	1
## 351	acres82Q1	351	0.09803922	1
## 367	acres82Q1	367	0.09803922	1
## 379	acres82Q1	379	0.09803922	1
## 384	acres82Q1	384	0.09803922	1
## 389	acres82Q1	389	0.09803922	1
## 397	acres82Q1	397	0.09803922	1
## 410	acres82Q1	410	0.09803922	1
## 415	acres82Q1	415	0.09803922	1
## 443	acres82Q1	443	0.09803922	1
## 465	acres82Q1	465	0.09803922	1
## 470	acres82Q1	470	0.09803922	1
## 472	acres82Q1	472	0.09803922	1
## 493	acres82Q1	493	0.09803922	1
## 519	acres82Q1	519	0.09803922	1
## 523	acres82Q1	523	0.09803922	1
## 533	acres82Q1	533	0.09803922	1
## 548	acres82Q1	548	0.09803922	1
## 557	acres82Q1	557	0.09803922	1
## 563	acres82Q1	563	0.09803922	1
## 579	acres82Q1	579	0.09803922	1
## 590	acres82Q1	590	0.09803922	1
## 599	acres82Q1	599	0.09803922	1
## 611	acres82Q1	611	0.09803922	1
## 617	acres82Q1	617	0.09803922	1
## 629	acres82Q1	629	0.09803922	1
## 650	acres82Q1	650	0.09803922	1
## 654	acres82Q1	654	0.09803922	1
## 681	acres82Q1	681	0.09803922	1
## 687	acres82Q1	687	0.09803922	1
## 716	acres82Q1	716	0.09803922	1
## 724	acres82Q1	724	0.09803922	1
## 727	acres82Q1	727	0.09803922	1
## 801	acres82Q2	801	0.09810334	2
## 813	acres82Q2	813	0.09810334	2
## 845	acres82Q2	845	0.09810334	2
## 886	acres82Q2	886	0.09810334	2
## 894	acres82Q2	894	0.09810334	2
## 900	acres82Q2	900	0.09810334	2
## 901	acres82Q2	901	0.09810334	2
## 903	acres82Q2	903	0.09810334	2
## 906	acres82Q2	906	0.09810334	2

## 920	acres82Q2	920	0.09810334	2
## 932	acres82Q2	932	0.09810334	2
## 966	acres82Q2	966	0.09810334	2
## 972	acres82Q2	972	0.09810334	2
## 989	acres82Q2	989	0.09810334	2
## 995	acres82Q2	995	0.09810334	2
## 1019	acres82Q2	1019	0.09810334	2
## 1024	acres82Q2	1024	0.09810334	2
## 1031	acres82Q2	1031	0.09810334	2
## 1052	acres82Q2	1052	0.09810334	2
## 1064	acres82Q2	1064	0.09810334	2
## 1083	acres82Q2	1083	0.09810334	2
## 1086	acres82Q2	1086	0.09810334	2
## 1100	acres82Q2	1100	0.09810334	2
## 1106	acres82Q2	1106	0.09810334	2
## 1108	acres82Q2	1108	0.09810334	2
## 1112	acres82Q2	1112	0.09810334	2
## 1130	acres82Q2	1130	0.09810334	2
## 1134	acres82Q2	1134	0.09810334	2
## 1145	acres82Q2	1145	0.09810334	2
## 1157	acres82Q2	1157	0.09810334	2
## 1158	acres82Q2	1158	0.09810334	2
## 1179	acres82Q2	1179	0.09810334	2
## 1182	acres82Q2	1182	0.09810334	2
## 1193	acres82Q2	1193	0.09810334	2
## 1194	acres82Q2	1194	0.09810334	2
## 1198	acres82Q2	1198	0.09810334	2
## 1202	acres82Q2	1202	0.09810334	2
## 1218	acres82Q2	1218	0.09810334	2
## 1220	acres82Q2	1220	0.09810334	2
## 1228	acres82Q2	1228	0.09810334	2
## 1234	acres82Q2	1234	0.09810334	2
## 1235	acres82Q2	1235	0.09810334	2
## 1245	acres82Q2	1245	0.09810334	2
## 1247	acres82Q2	1247	0.09810334	2
## 1248	acres82Q2	1248	0.09810334	2
## 1262	acres82Q2	1262	0.09810334	2
## 1263	acres82Q2	1263	0.09810334	2
## 1264	acres82Q2	1264	0.09810334	2
## 1265	acres82Q2	1265	0.09810334	2
## 1296	acres82Q2	1296	0.09810334	2
## 1303	acres82Q2	1303	0.09810334	2
## 1317	acres82Q2	1317	0.09810334	2
## 1324	acres82Q2	1324	0.09810334	2
## 1332	acres82Q2	1332	0.09810334	2
## 1367	acres82Q2	1367	0.09810334	2
## 1381	acres82Q2	1381	0.09810334	2
## 1382	acres82Q2	1382	0.09810334	2
## 1386	acres82Q2	1386	0.09810334	2
## 1390	acres82Q2	1390	0.09810334	2
## 1392	acres82Q2	1392	0.09810334	2
## 1395	acres82Q2	1395	0.09810334	2
## 1427	acres82Q2	1427	0.09810334	2
## 1430	acres82Q2	1430	0.09810334	2

## 1437 acres82Q2	1437	0.09810334	2
## 1447 acres82Q2	1447	0.09810334	2
## 1451 acres82Q2	1451	0.09810334	2
## 1464 acres82Q2	1464	0.09810334	2
## 1474 acres82Q2	1474	0.09810334	2
## 1478 acres82Q2	1478	0.09810334	2
## 1485 acres82Q2	1485	0.09810334	2
## 1486 acres82Q2	1486	0.09810334	2
## 1487 acres82Q2	1487	0.09810334	2
## 1498 acres82Q2	1498	0.09810334	2
## 1500 acres82Q2	1500	0.09810334	2
## 1504 acres82Q2	1504	0.09810334	2
## 1526 acres82Q2	1526	0.09810334	2
## 1534 acres82Q2	1534	0.09810334	2
## 1553 acres82Q2	1553	0.09810334	2
## 1557 acres82Q2	1557	0.09810334	2
## 1559 acres82Q2	1559	0.09810334	2
## 1573 acres82Q2	1573	0.09810334	2
## 1581 acres82Q2	1581	0.09810334	2
## 1592 acres82Q2	1592	0.09810334	2
## 1607 acres82Q2	1607	0.09810334	2
## 1611 acres82Q2	1611	0.09810334	2
## 1617 acres82Q2	1617	0.09810334	2
## 1627 acres82Q2	1627	0.09810334	2
## 1634 acres82Q2	1634	0.09810334	2
## 1639 acres82Q2	1639	0.09810334	2
## 1668 acres82Q2	1668	0.09810334	2
## 1679 acres82Q2	1679	0.09810334	2
## 1682 acres82Q2	1682	0.09810334	2
## 1683 acres82Q2	1683	0.09810334	2
## 1688 acres82Q2	1688	0.09810334	2
## 1689 acres82Q2	1689	0.09810334	2
## 1696 acres82Q2	1696	0.09810334	2
## 1697 acres82Q2	1697	0.09810334	2
## 1701 acres82Q2	1701	0.09810334	2
## 1725 acres82Q2	1725	0.09810334	2
## 1735 acres82Q2	1735	0.09810334	2
## 1752 acres82Q2	1752	0.09810334	2
## 1771 acres82Q2	1771	0.09810334	2
## 1799 acres82Q2	1799	0.09810334	2
## 1802 acres82Q2	1802	0.09810334	2
## 1803 acres82Q2	1803	0.09810334	2
## 1808 acres82Q2	1808	0.09810334	2
## 1813 acres82Q2	1813	0.09810334	2
## 1818 acres82Q2	1818	0.09810334	2
## 1831 acres82Q2	1831	0.09810334	2
## 1842 acres82Q2	1842	0.09810334	2
## 1852 acres82Q2	1852	0.09810334	2
## 1854 acres82Q2	1854	0.09810334	2
## 1870 acres82Q2	1870	0.09810334	2
## 1877 acres82Q2	1877	0.09810334	2
## 1878 acres82Q2	1878	0.09810334	2
## 1915 acres82Q2	1915	0.09810334	2
## 1926 acres82Q2	1926	0.09810334	2

##	1939	acres82Q2	1939	0.09810334	2
##	1943	acres82Q2	1943	0.09810334	2
##	1951	acres82Q2	1951	0.09810334	2
##	1952	acres82Q2	1952	0.09810334	2
##	1953	acres82Q2	1953	0.09810334	2
##	1967	acres82Q2	1967	0.09810334	2
##	2000	acres82Q2	2000	0.09810334	2
##	2005	acres82Q2	2005	0.09810334	2
##	2011	acres82Q2	2011	0.09810334	2
##	2017	acres82Q2	2017	0.09810334	2
##	2046	acres82Q2	2046	0.09810334	2
##	2061	acres82Q2	2061	0.09810334	2
##	2066	acres82Q2	2066	0.09810334	2
##	2087	acres82Q2	2087	0.09810334	2
##	2097	acres82Q2	2097	0.09810334	2
##	2098	acres82Q2	2098	0.09810334	2
##	2099	acres82Q2	2099	0.09810334	2
##	2102	acres82Q2	2102	0.09810334	2
##	2104	acres82Q2	2104	0.09810334	2
##	2139	acres82Q2	2139	0.09810334	2
##	2149	acres82Q2	2149	0.09810334	2
##	2158	acres82Q2	2158	0.09810334	2
##	2160	acres82Q2	2160	0.09810334	2
##	2171	acres82Q2	2171	0.09810334	2
##	2186	acres82Q2	2186	0.09810334	2
##	2221	acres82Q2	2221	0.09810334	2
##	2223	acres82Q2	2223	0.09810334	2
##	2230	acres82Q2	2230	0.09810334	2
##	2252	acres82Q2	2252	0.09810334	2
##	2267	acres82Q2	2267	0.09810334	2
##	2280	acres82Q2	2280	0.09810334	2
##	2290	acres82Q2	2290	0.09810334	2
##	2294	acres82Q2	2294	0.09810334	2
##	2302	acres82Q3	2302	0.09803922	3
##	2311	acres82Q3	2311	0.09803922	3
##	2317	acres82Q3	2317	0.09803922	3
##	2328	acres82Q3	2328	0.09803922	3
##	2338	acres82Q3	2338	0.09803922	3
##	2339	acres82Q3	2339	0.09803922	3
##	2341	acres82Q3	2341	0.09803922	3
##	2361	acres82Q3	2361	0.09803922	3
##	2380	acres82Q3	2380	0.09803922	3
##	2384	acres82Q3	2384	0.09803922	3
##	2385	acres82Q3	2385	0.09803922	3
##	2421	acres82Q3	2421	0.09803922	3
##	2422	acres82Q3	2422	0.09803922	3
##	2425	acres82Q3	2425	0.09803922	3
##	2446	acres82Q3	2446	0.09803922	3
##	2462	acres82Q3	2462	0.09803922	3
##	2493	acres82Q3	2493	0.09803922	3
##	2511	acres82Q3	2511	0.09803922	3
##	2515	acres82Q3	2515	0.09803922	3
##	2516	acres82Q3	2516	0.09803922	3
##	2521	acres82Q3	2521	0.09803922	3

##	2539	acres82Q3	2539	0.09803922	3
##	2565	acres82Q3	2565	0.09803922	3
##	2568	acres82Q3	2568	0.09803922	3
##	2580	acres82Q3	2580	0.09803922	3
##	2624	acres82Q3	2624	0.09803922	3
##	2631	acres82Q3	2631	0.09803922	3
##	2639	acres82Q3	2639	0.09803922	3
##	2654	acres82Q3	2654	0.09803922	3
##	2658	acres82Q3	2658	0.09803922	3
##	2670	acres82Q3	2670	0.09803922	3
##	2672	acres82Q3	2672	0.09803922	3
##	2681	acres82Q3	2681	0.09803922	3
##	2682	acres82Q3	2682	0.09803922	3
##	2690	acres82Q3	2690	0.09803922	3
##	2705	acres82Q3	2705	0.09803922	3
##	2708	acres82Q3	2708	0.09803922	3
##	2711	acres82Q3	2711	0.09803922	3
##	2714	acres82Q3	2714	0.09803922	3
##	2722	acres82Q3	2722	0.09803922	3
##	2735	acres82Q3	2735	0.09803922	3
##	2749	acres82Q3	2749	0.09803922	3
##	2763	acres82Q3	2763	0.09803922	3
##	2767	acres82Q3	2767	0.09803922	3
##	2768	acres82Q3	2768	0.09803922	3
##	2782	acres82Q3	2782	0.09803922	3
##	2790	acres82Q3	2790	0.09803922	3
##	2797	acres82Q3	2797	0.09803922	3
##	2799	acres82Q3	2799	0.09803922	3
##	2819	acres82Q3	2819	0.09803922	3
##	2822	acres82Q3	2822	0.09803922	3
##	2824	acres82Q3	2824	0.09803922	3
##	2831	acres82Q3	2831	0.09803922	3
##	2844	acres82Q3	2844	0.09803922	3
##	2850	acres82Q3	2850	0.09803922	3
##	2851	acres82Q3	2851	0.09803922	3
##	2886	acres82Q3	2886	0.09803922	3
##	2896	acres82Q3	2896	0.09803922	3
##	2901	acres82Q3	2901	0.09803922	3
##	2905	acres82Q3	2905	0.09803922	3
##	2922	acres82Q4	2922	0.09803922	4
##	2929	acres82Q4	2929	0.09803922	4
##	2932	acres82Q4	2932	0.09803922	4
##	2934	acres82Q4	2934	0.09803922	4
##	2950	acres82Q4	2950	0.09803922	4
##	2957	acres82Q4	2957	0.09803922	4
##	2963	acres82Q4	2963	0.09803922	4
##	2975	acres82Q4	2975	0.09803922	4
##	2979	acres82Q4	2979	0.09803922	4
##	2993	acres82Q4	2993	0.09803922	4
##	2994	acres82Q4	2994	0.09803922	4
##	3008	acres82Q4	3008	0.09803922	4
##	3012	acres82Q4	3012	0.09803922	4
##	3019	acres82Q4	3019	0.09803922	4
##	3036	acres82Q4	3036	0.09803922	4


```
## checking sampling results
table (strsample [,1])

##
## acres82Q1 acres82Q2 acres82Q3 acres82Q4
##      75      150      60      15

# collecting data on sampled counties
agstrat <- agpop [strsample$ID_unit, ]
agstrat$weight <- 1/strsample$Prob

str_mean_estimate_data (agstrat, "acres92", "stratum", "weight")

##      Est.      S.E.      ci.low      ci.upp
## 306423.020  8086.189 290574.091 322271.950
```

4.2.4 Repeat stratified sampling with P2S allocation 2000 times

```
Nh <- tapply (1:nrow (agpop), agpop$stratum, length)
nh <- round(Nh/sum (Nh)*300) ## make sure the order matches Nh
data.frame(Nh, nh)

##      Nh  nh
## acres82Q1 765 75
## acres82Q2 1529 150
## acres82Q3 612 60
## acres82Q4 153 15

nres <- 2000
str_p2s_simulated <- matrix (0, nres, 4)
for (i in 1:nres)
{
  ## doing stratified sampling
  strsample <- strata (agpop, "stratum", size = nh, method = "srswor")
  agstrat <- agpop [strsample$ID_unit, ]
  agstrat$weight <- 1/strsample$Prob
  ## analyzing data
  str_p2s_simulated[i,] <- str_mean_estimate_data (agstrat, "acres92", "stratum", "weight")
}
```

4.2.5 Repeat stratified sampling with optimal allocation 2000 times

```
Nh <- tapply (1:nrow (agpop), agpop$stratum, length)
# find order of stratum
# unique (agpop$stratum)
Sh <- tapply (agpop$acres92, agpop$stratum, sd)
nh_opt <- round((Nh*Sh)/sum (Nh*Sh) * 300);nh_opt

## acres82Q1 acres82Q2 acres82Q3 acres82Q4
##      22      98      77      103

data.frame(Nh,Sh,nh_opt)

##      Nh      Sh  nh_opt
## acres82Q1 765 37409.71    22
```

```
## acres82Q2 1529 83714.96 98
## acres82Q3 612 162774.52 77
## acres82Q4 153 874495.84 103

nres <- 2000
str_neyman_simulated <- matrix(0, nres, 4)
for (i in 1:nres)
{
  ## doing stratified sampling
  strsample <- strata(agpop, "stratum", size = nh_opt, method = "srswor")
  # collecting data on sampled counties
  agstrat <- agpop[strsample$ID_unit, ]
  agstrat$weight <- 1/strsample$Prob

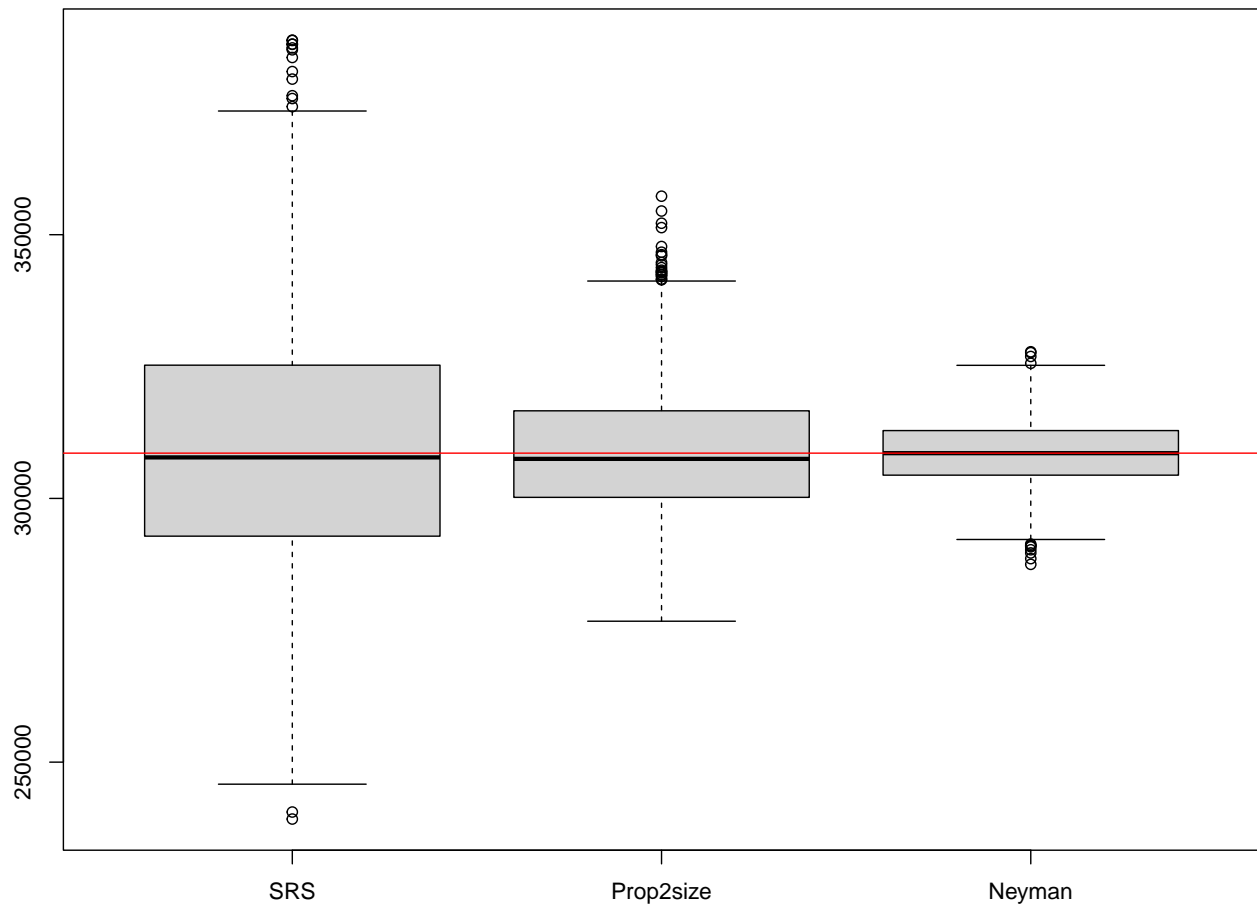
  str_neyman_simulated[i,] <- str_mean_estimate_data(agstrat, "acres92", "stratum", "weight")
}
```

4.2.6 Repeat simple random sampling 2000 times

```
nres <- 2000
srs_simulated <- matrix(0, nres, 4)
for (i in 1:nres)
{
  ## doing stratified sampling
  srs <- sample(sum(Nh), sum(nh))
  # collecting data on sampled counties
  agsrs <- agpop[srs, ]
  srs_simulated[i,] <- srs_mean_est(agsrs[, "acres92"], N = sum(Nh))
}
```

4.2.7 Compare the efficiency of different methods

```
sim_results_str_acres82 <- data.frame("SRS"=srs_simulated[,1],
                                     "Prop2size"=str_p2s_simulated[,1],
                                     "Neyman"=str_neyman_simulated[,1])
boxplot(sim_results_str_acres82)
abline(h = mean(agpop$acres92), col = "red")
```



```
sapply (sim_results_str_acres82, mean) -> sim_means
sapply (sim_results_str_acres82, var) -> sim_var
```

```
sim_var/sim_var[1] -> sim_var_relative
```

```
#Percentage of reduction of variances of estimates compared to SRS
```

```
1-sim_var/sim_var[1] -> sim_var_reduction
```

```
cbind("Mean"=sim_means, "Variance"=sim_var, "Relative Variance"= sim_var_relative, "Percentage of Variance Reduction"=sim_var_reduction)
```

##	Mean	Variance	Relative Variance	Percentage of Variance Reduction
## SRS	309395.3	563980266	1.00000000	0.0000000
## Prop2size	308903.1	150620233	0.26706650	0.7329335
## Neyman	308585.9	38522185	0.06830414	0.9316959