# GRTS Survey Designs for an Area Resource

#### Thomas Kincaid

#### October 23, 2015

#### Contents

1	Preliminaries	1
2	Create a shapefile	2
3	Shapefile attribute data	2
4	Unstratified, equal probability, GRTS survey design	6
5	Unstratified, unequal probability, GRTS survey design	8
6	Stratified, equal probability, GRTS survey design	10
7	Unstratified, unequal probability, GRTS survey design with an oversample and a panel structure for survey over time	13

### 1 Preliminaries

This document presents example GRTS survey designs for an area resource. The area resource used in the designs is Omernik level 3 ecoregions within Utah. Four survey designs will be presented: (1) an unstratified, equal probability design; (2) an unstratified, unequal probability design; (3) a stratified, equal probability design; and (4) an unstratified, unequal probability design with an oversample and a panel structure for survey over time. The sampling frame used for the survey designs is contained in either an ESRI shapefile or an sp package object. The frame contains the coordinates for a set of polygons that define the area resource in addition to attribute data associated with the polygons. The coordinate system for the set of points in the sampling frame is an equal area projection rather than latitude

and longitude. An equal area projection is used so that calculation of distance between points is valid.

The initial step is to use the library function to load the spsurvey package. After the package is loaded, a message is printed to the R console indicating that the spsurvey package was loaded successfully.

Load the spsurvey package

```
> # Load the spsurvey package
> library(spsurvey)
>
```

Version 3.1 of the spsurvey package was loaded successfully.

# 2 Create a shapefile

For creating a survey design using the spsurvey package, the standard form of input regarding the resource is a shapefile. In order to conserve storage space, shapefiles are not included with the package. Instead, a data set from which a shapefile can be created is included in the data directory of the package. The data function is used to load the data set stored in the data directory into an object named UT\_ecoregions. The sp2shape function is used to create a shapefile from the UT\_ecoregions object. Note that objects loaded from the data sets in the data directory are stored in formats that are defined in the sp package. See documentation for the sp package for additional information regarding format of the objects.

```
> # Load the sp object in the data directory
> data(UT_ecoregions)
> # Create a shapefile
> sp2shape(sp.obj=UT_ecoregions, shpfilename="UT_ecoregions")
>
```

# 3 Shapefile attribute data

The next step is to read the attribute data from the shapefile. The read.dbf function in the spsurvey package is used to read the attribute (dbf) file in the shapefile and assign it to a data frame named att. The att data frame is printed by entering the data frame name at the R prompt.

The ecoregion attribute will be used to define stratum codes and unequal selection probability (multidensity) categories for the survey designs. Ecoregion is contained in a variable named "level3\_nam" and includes seven unique values. Frame area is summarized for the ecoregion

attribute. Note that ecoregion area measured in hectares is contained in the variable named "area\_ha". The tapply function is used to calculate total area for each ecoregion. The addmargins function is applied to the output from tapply to calculate total area for all all ecoregions, and the round function is used to round value to whole numbers. Finally, the resulting table is displayed.

Read the attribute table from the shapefile

```
> # Read the attribute table from the shapefile
> att <- read.dbf("UT_ecoregions")
>
```

Display the attribute data frame

> # Display the attribute data frame
> att

	Level3	Level3_Nam	Area_ha	$area_mdm$
1	80	Northern Basin and Range	1.42202e+11	2639990439
2	18	Wyoming Basin	1.33312e+11	2910588302
3	13	Central Basin and Range	3.09949e+11	82064546355
4	19	Wasatch and Uinta Mountains	4.47240e+10	42569491524
5	20	Colorado Plateaus	1.26379e+11	85797163092
6	21	Southern Rockies	5.40909e+08	540909129
7	14	Mojave Basin and Range	1.29599e+11	1931860757
8	19	Wasatch and Uinta Mountains	3.43657e+08	343657270
9	19	${\tt Wasatch} \ {\tt and} \ {\tt Uinta} \ {\tt Mountains}$	6.26102e+08	626102339
10	21	Southern Rockies	4.05534e+08	405534288

>

Summarize frame area by ecoregion

```
> # Summarize frame area by ecoregion
> temp <- tapply(att$Area_ha, att$Level3_Nam, sum)
> temp <- round(addmargins(temp), 0)
> temp
```

```
Central Basin and Range Colorado Plateaus 309949000000 126379000000 Mojave Basin and Range Northern Basin and Range 129599000000 1422020000000 Southern Rockies Wasatch and Uinta Mountains
```

946443000 Wyoming Basin 133312000000

45693759000 Sum 888081202000

>

The seven ecoregions in Utah are displayed in Figure 1. Ecoregions are used during creation of several of the survey designs described in this vignette.

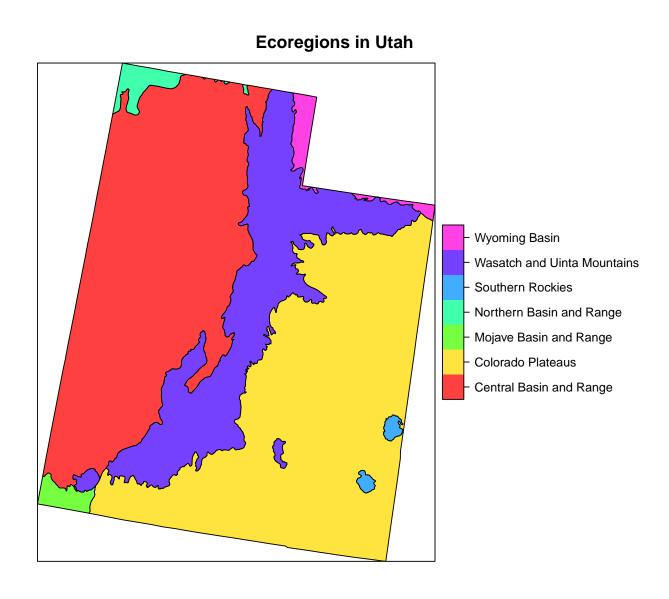


Figure 1: Location of the seven ecoregions in Utah.

# 4 Unstratified, equal probability, GRTS survey design

The first survey design is an unstratified, equal probability design. The set seed function is called so that, if necessary, the designs can be replicated.

The initial step is to create a list named Equaldsgn that contains information for specifying the survey design. Since the survey design is unstratified, the list contains a single item named "None" that also is a list. The "None" list includes two items: panel, which is used to specify the sample size for each panel, and seltype, which is used to input the type of random selection for the design. For this example, panel is assigned a single value named "PanelOne" that is set equal to 115, and seltype is assigned the value "Equal", which indicates equal probability selection.

The grts function in the spsurvey package is called to select the survey design. The following arguments are included in the call to grts: (1) design: the named list of stratum design specifications, which is assigned the Equaldsgn list; (2) DesignID: name for the design, which is used to create a site ID for each site and is assigned the value "EQUAL"; (3) type.frame: the type of frame, which is assigned the value "area" to indicate an area resource; (4) src.frame: source of the frame, which is assigned the value "shapefile" to indicate a shapefile frame; (5) in.shape: name of the input shapefile, which is assigned the value "UT\_ecoregions"; (6) att.frame: the data frame of attributes associated with elements in the frame, which is assigned the att data frame; and (7) shapefile: option to create a shapefile containing the survey design information, which is assigned FALSE.

During execution of the grts function, messages are printed that indicate the initial number of hierarchical levels used for the GRTS grid, the current number of levels, and the final number of levels. The set of messages is printed for each stratum, and is labeled with the stratum name. For this example, the set of messages is labeled "None", i.e., the name used in the Equaldsgn list. Upon completion of the call to grts, the initial six sites for the survey design and a design summary are printed. The output object created by the grts function is assigned class "SpatialDesign". The design summary is created using the summary method for that class. In addition to summary, a plot method is available for the SpatialDesign class. For assistance using the summary and plot methods, see documentation for "SpatialDesign-class" on the R help page for spsurvey.

Call the set.seed function so that the design can be replicated

```
> # Call the set.seed function so that the survey designs can be replicate
> set.seed(4447864)
>
Create the design list
> # Create the design list
> Equaldsgn <- list(None=list(panel=c(PanelOne=115), seltype="Equal"))</pre>
```

#### Select the sample

Stratum: None

Initial number of levels: 4 Current number of levels: 4 Final number of levels: 4

Print the initial six lines of the survey design

- > # Print the initial six lines of the survey design
- > head(Equalsites@data)

```
siteID
              xcoord ycoord mdcaty
                                           wgt stratum
                                                          panel EvalStatus
1 EQUAL-001 -1361882 1986078 Equal 1911563856
                                                  None PanelOne
                                                                   NotEval
2 EQUAL-002 -1304045 2195914 Equal 1911563856
                                                  None PanelOne
                                                                   NotEval
3 EQUAL-003 -1269657 1803336 Equal 1911563856
                                                  None PanelOne
                                                                   NotEval
4 EQUAL-004 -1311248 1689394 Equal 1911563856
                                                  None PanelOne
                                                                   NotEval
5 EQUAL-005 -1353690 2074387 Equal 1911563856
                                                  None PanelOne
                                                                   NotEval
6 EQUAL-006 -1389763 2190386 Equal 1911563856
                                                  None PanelOne
                                                                   NotEval
 EvalReason Level3
                                 Level3_Nam
                                                Area_ha
1
                 13 Central Basin and Range 3.09949e+11
2
                 13 Central Basin and Range 3.09949e+11
3
                 20
                          Colorado Plateaus 1.26379e+11
4
                 20
                          Colorado Plateaus 1.26379e+11
5
                 13 Central Basin and Range 3.09949e+11
6
                 13 Central Basin and Range 3.09949e+11
```

Print the survey design summary

- > # Print the survey design summary
- > summary(Equalsites)

```
Design Summary: Number of Sites
stratum
None Sum
115 115
```

# 5 Unstratified, unequal probability, GRTS survey design

The second survey design is an unstratified, unequal probability design. Ecoregions are used to identify multidensity categories. List Unequaldsgn is assigned design specifications. Since the survey design is unstratified, Unequaldsgn includes a single list named "None" that contains three items: panel, seltype, and caty.n. The value for panel is the same as for the equal probability design, and seltype is assigned "Unequal" to indicate unequal selection probabilities. The third item, caty.n, assigns sample sizes for each of seven multidensity categories, where ecoregion names are used as the categories. Note that the sum of sample sizes provided in caty.n must equal the value in panel.

For this survey design, a shapefile will be used as the sampling frame. The following arguments are included in the call to grts: (1) design: assigned the Unequaldsgn list; (2) DesignID: assigned the value "UNEQUAL"; (3) type.frame: assigned the value "area"; (4) src.frame: assigned the value "shapefile"; (5) in.shape: assigned the value "UT\_ecoregions"; (6) att.frame: assigned the att data frame; (7) mdcaty: name of the column in the attributes data frame that identifies the unequal probability category for each element in the frame, which is assigned the value "level3\_nam"; and (8) shapefile: assigned the value FALSE. Upon completion of the call to grts, the initial six sites for the survey design and a design summary are printed.

Create the design list

Select the sample

Print the initial six lines of the survey design

- > # Print the initial six lines of the survey design
- > head(Unequalsites@data)

```
siteID
                xcoord ycoord
                                                    mdcaty
                                                                   wgt stratum
1 UNEQUAL-001 -1471013 2106177
                                   Central Basin and Range 3282581854
                                                                          None
2 UNEQUAL-002 -1354503 1848229 Wasatch and Uinta Mountains 1741570045
                                                                          None
3 UNEQUAL-003 -1298067 1916720 Wasatch and Uinta Mountains 1741570045
                                                                          None
4 UNEQUAL-004 -1165284 1728905
                                          Southern Rockies
                                                              94644342
                                                                          None
5 UNEQUAL-005 -1395426 2161246
                                   Central Basin and Range 3282581854
                                                                          None
6 UNEQUAL-006 -1483098 1743104 Wasatch and Uinta Mountains 1741570045
                                                                          None
     panel EvalStatus EvalReason Level3
                                            Area_ha
1 PanelOne
              NotEval
                                     13 3.09949e+11
2 PanelOne
              NotEval
                                     19 4.47240e+10
3 PanelOne
                                     19 4.47240e+10
              NotEval
4 PanelOne
              NotEval
                                     21 4.05534e+08
5 PanelOne
              NotEval
                                     13 3.09949e+11
6 PanelOne
              NotEval
                                     19 4.47240e+10
```

>

Print the survey design summary

- > # Print the survey design summary
- > summary(Unequalsites)

```
Design Summary: Number of Sites Classified by mdcaty (Multidensity Category)
```

```
mdcaty
Central Basin and Range
27
25
Mojave Basin and Range
Northern Basin and Range
9
Southern Rockies Wasatch and Uinta Mountains
9
25
Wyoming Basin
11
115
```

>

# 6 Stratified, equal probability, GRTS survey design

The third survey design is a stratified, equal probability design. Ecoregions are used to identify strata. List Stratdsgn is assigned design specifications. The ecoregion attribute is used to identify strata. Stratdsgn includes seven lists, one for each stratum. The names for the lists match the levels of the stratum variable, i.e., the unique values of the ecoregion attribute. Each list in Stratdsgn contains two items: panel and seltype. The value for panel is the same as for the equal probability design, and seltype is assigned "Equal".

For this survey design, an sp package object will be used as the sampling frame. The read.shape function will be used to read the shapefile and assign its output to an sp object named shp. Note that the object created by the read.shape function is identical to the object that was loaded from the data directory at the beginning of this vignette. The following arguments are included in the call to grts: (1) design: assigned the Stratdsgn list; (2) DesignID: assigned the value "STRATIFIED"; (3) type.frame: assigned the value "area"; (4) src.frame: assigned the value "sp.object" to indicate that the sampling frame is provided by an sp object; (5) sp.object: name of the sp object, which is assigned the shp object; (6) att.frame: assigned the att data frame; 7) stratum: name of the column in the attributes data frame that identifies the stratum code for each element in the frame, which is assigned the value "level3\_nam"; and (8) shapefile: assigned the value FALSE. Upon completion of the call to grts, the initial six sites for the survey design and a design summary are printed.

```
> # Read the shapefile
> shp <- read.shape("UT_ecoregions")
>
Create the design list
> Stratdsgn <- list("Central Basin and Range"=list(panel=c(PanelOne=25),</pre>
```

```
seltype="Equal"),
                    "Colorado Plateaus"=list(panel=c(PanelOne=25),
                                              seltype="Equal"),
                    "Mojave Basin and Range"=list(panel=c(PanelOne=10),
                                                   seltype="Equal"),
                    "Northern Basin and Range"=list(panel=c(PanelOne=10),
                                                     seltype="Equal"),
                    "Southern Rockies"=list(panel=c(PanelOne=10),
                                             seltype="Equal"),
                    "Wasatch and Uinta Mountains"=list(panel=c(PanelOne=25),
                                                        seltype="Equal"),
                    "Wyoming Basin"=list(panel=c(PanelOne=10),
                                          seltype="Equal"))
Select the sample
> Stratsites <- grts(design=Stratdsgn,
                     DesignID="STRATIFIED",
+
                     type.frame="area",
                     src.frame="sp.object",
                     sp.object=shp,
                     att.frame=att,
                     stratum="Level3_Nam",
                     shapefile=FALSE)
Stratum: Central Basin and Range
Initial number of levels: 3
Current number of levels: 3
Current number of levels: 4
Final number of levels: 4
Stratum: Colorado Plateaus
Initial number of levels: 3
Current number of levels: 3
Current number of levels: 4
Final number of levels: 4
Stratum: Mojave Basin and Range
Initial number of levels: 2
Current number of levels: 2
Current number of levels: 3
Final number of levels: 3
```

Stratum: Northern Basin and Range

Initial number of levels: 2 Current number of levels: 2 Current number of levels: 4 Final number of levels: 4

Stratum: Southern Rockies
Initial number of levels: 2
Current number of levels: 2
Current number of levels: 3
Current number of levels: 4
Final number of levels: 4

Stratum: Wasatch and Uinta Mountains

Initial number of levels: 3 Current number of levels: 3 Current number of levels: 4 Final number of levels: 4

Stratum: Wyoming Basin
Initial number of levels: 2
Current number of levels: 2
Current number of levels: 3
Current number of levels: 4
Final number of levels: 4

Print the initial six lines of the survey design

- > # Print the initial six lines of the survey design
- > head(Stratsites@data)

```
xcoord ycoord mdcaty
          siteID
                                                wgt
                                                                     stratum
1 STRATIFIED-001 -1529774 1751713
                                   Equal 3282581854 Central Basin and Range
2 STRATIFIED-002 -1456864 2008542
                                   Equal 3282581854 Central Basin and Range
3 STRATIFIED-003 -1382436 1851915
                                   Equal 3282581854 Central Basin and Range
4 STRATIFIED-004 -1376583 2129989
                                   Equal 3282581854 Central Basin and Range
5 STRATIFIED-005 -1542866 1885916
                                   Equal 3282581854 Central Basin and Range
6 STRATIFIED-006 -1509108 1963177
                                   Equal 3282581854 Central Basin and Range
    panel EvalStatus EvalReason Level3
                                            Area_ha
1 PanelOne
                                     13 3.09949e+11
              NotEval
2 PanelOne
                                     13 3.09949e+11
              NotEval
3 PanelOne
              NotEval
                                     13 3.09949e+11
4 PanelOne
              NotEval
                                     13 3.09949e+11
5 PanelOne
              NotEval
                                     13 3.09949e+11
6 PanelOne
             NotEval
                                     13 3.09949e+11
```

>

Print the survey design summary

```
> # Print the survey design summary
```

> summary(Stratsites)

Design Summary: Number of Sites

stratum

Central Basin and Range	Colorado Plateaus
25	25
Mojave Basin and Range	Northern Basin and Range
10	10
Southern Rockies	Wasatch and Uinta Mountains
10	25
Wyoming Basir	Sum
10	115

>

# 7 Unstratified, unequal probability, GRTS survey design with an oversample and a panel structure for survey over time

The fourth survey design is an unstratified, unequal probability design with an oversample and a panel structure for survey over time. List Paneldsgn is assigned design specifications. Since the survey design is unstratified, Paneldsgn includes a single list named "None" that contains four items: panel, seltype, caty.n, and over. A vector identifying sample sizes for five panels is assigned to panel. The value "Unequal" is assigned to seltype, which indicates unequal selection probabilities. The third item, caty.n, assigns sample sizes for each of seven multidensity categories, where ecoregion names are used as the categories. Note that the sum of sample sizes provided in caty.n must equal the sum of sample sizes in panel. The value 100 is assigned to over, which specifies an oversample of 100 sites. An oversample is replacement sites for the survey design. The grts function attempts to distribute the oversample proportionately among sample sizes for the multidensity categories. If the oversample proportion for one or more categories is not a whole number, a warning message is printed and the proportionate to the category sample sizes, and the warning message is printed by calling the warnings function.

For this survey design, a shapefile will be used as the sampling frame. The following arguments are included in the call to grts: (1) design: assigned the Paneldsgn list; (2) DesignID: assigned the value "UNEQUAL"; (3) type.frame: assigned the value "area"; (4) src.frame: assigned the value "shapefile"; (5) in.shape: assigned the value "UT\_ecoregions"; (6) att.frame: assigned the att data frame; (7) mdcaty: assigned the value "level3\_nam"; and (8) shapefile: assigned the value FALSE. Upon completion of the call to grts, the initial six sites for the survey design and a design summary are printed.

Create the design list

#### Select the sample

Print the warning message

Current number of levels: 7 Final number of levels: 7

```
> # Print the warning message
> warnings()
>
```

#### Warning message:

In grts(design = Paneldsgn, DesignID = "UNEQUAL", type.frame = "area", :

Oversample size is not proportional to category sample sizes for stratum "None".

Print the initial six lines of the survey design

- > # Print the initial six lines of the survey design
- > head(Panelsites@data)

siteID	xcoord	ycoord	mdcaty	wgt	stratum pa	nel
UNEQUAL-001	-1157287	1661084	Colorado Plateaus	1361859732	None Ye	ar1
UNEQUAL-002	-1213951	1954093	Colorado Plateaus	1361859732	None Ye	ar1
UNEQUAL-003	-1542076	1770381	${\tt Central\ Basin\ and\ Range}$	1282258537	None Ye	ar1
UNEQUAL-004	-1261338	2135698	Wyoming Basin	194039220	None Ye	ar1
UNEQUAL-005	-1185851	1770761	Colorado Plateaus	1361859732	None Ye	ar1
UNEQUAL-006	-1138910	1801524	Southern Rockies	63096228	None Ye	ar1
EvalStatus E	EvalReasor	Level3	Area_ha			
${ t NotEval}$		20	1.26379e+11			
${ t NotEval}$		20	1.26379e+11			
${ t NotEval}$		13	3.09949e+11			
${ t NotEval}$		18	1.33312e+11			
${ t NotEval}$		20	1.26379e+11			
${ t NotEval}$		21	5.40909e+08			
	UNEQUAL-001 UNEQUAL-002 UNEQUAL-003 UNEQUAL-005 UNEQUAL-006 EvalStatus F NotEval NotEval NotEval NotEval NotEval	UNEQUAL-001 -1157287 UNEQUAL-002 -1213951 UNEQUAL-003 -1542076 UNEQUAL-004 -1261338 UNEQUAL-005 -1185851 UNEQUAL-006 -1138910 EvalStatus EvalReasor NotEval NotEval NotEval NotEval NotEval NotEval	UNEQUAL-004 -1261338 2135698 UNEQUAL-005 -1185851 1770761 UNEQUAL-006 -1138910 1801524 EvalStatus EvalReason Level3 NotEval 20 NotEval 20 NotEval 13 NotEval 18 NotEval 20	UNEQUAL-001 -1157287 1661084 Colorado Plateaus UNEQUAL-002 -1213951 1954093 Colorado Plateaus UNEQUAL-003 -1542076 1770381 Central Basin and Range UNEQUAL-004 -1261338 2135698 Wyoming Basin UNEQUAL-005 -1185851 1770761 Colorado Plateaus UNEQUAL-006 -1138910 1801524 Southern Rockies EvalStatus EvalReason Level3 Area_ha NotEval 20 1.26379e+11 NotEval 13 3.09949e+11 NotEval 18 1.33312e+11 NotEval 20 1.26379e+11	UNEQUAL-001 -1157287 1661084 Colorado Plateaus 1361859732 UNEQUAL-002 -1213951 1954093 Colorado Plateaus 1361859732 UNEQUAL-003 -1542076 1770381 Central Basin and Range 1282258537 UNEQUAL-004 -1261338 2135698 Wyoming Basin 194039220 UNEQUAL-005 -1185851 1770761 Colorado Plateaus 1361859732 UNEQUAL-006 -1138910 1801524 Southern Rockies 63096228 EvalStatus EvalReason Level3 Area_ha NotEval 20 1.26379e+11 NotEval 13 3.09949e+11 NotEval 18 1.33312e+11 NotEval 20 1.26379e+11	UNEQUAL-001 -1157287 1661084 Colorado Plateaus 1361859732 None Ye UNEQUAL-002 -1213951 1954093 Colorado Plateaus 1361859732 None Ye UNEQUAL-003 -1542076 1770381 Central Basin and Range 1282258537 None Ye UNEQUAL-004 -1261338 2135698 Wyoming Basin 194039220 None Ye UNEQUAL-005 -1185851 1770761 Colorado Plateaus 1361859732 None Ye UNEQUAL-006 -1138910 1801524 Southern Rockies 63096228 None Ye EvalStatus EvalReason Level3 Area_ha NotEval 20 1.26379e+11 NotEval 13 3.09949e+11 NotEval 13 1.33312e+11 NotEval 20 1.26379e+11

>

Print the survey design summary

- > # Print the survey design summary
- > summary(Panelsites)

Design Summary: Number of Sites Classified by mdcaty (Multidensity Category) and panel

	panel						
mdcaty	OverSamp	Year1	Year2	Year3	Year4	Year5	${\tt Sum}$
Central Basin and Range	28	11	13	11	12	14	89
Colorado Plateaus	23	18	14	10	14	11	90
Mojave Basin and Range	5	4	3	3	3	4	22
Northern Basin and Range	7	3	3	3	2	3	21

Southern Rockies	6	4	2	5	1	2	20
Wasatch and Uinta Mountains	26	8	10	15	16	15	90
Wyoming Basin	7	2	5	3	2	1	20
Sum	102	50	50	50	50	50	352

>