GRTS Survey Designs for a Linear Resource

Thomas Kincaid

May 23, 2012

Contents

1 Preliminaries

This document presents example GRTS survey designs for a linear resource. The linear resource used in the designs is streams that comprise the Luckiamute watershed in Oregon. Four survey designs will be presented: (1) an unstratified, equal probability design; (2) a stratified, equal probability design with an oversample; (3) a stratified, unequal probability design with an oversample; and (4) a stratified, 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 line segments that define the linear resource in addition to attribute data associated with the line segments. 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 2.1 of the spsurvey package was loaded successfully.
```

2   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 initial six lines in the att data frame are printed using the head function

Next, two attributes, stream type and Strahler stream order, that will be used to define, respectively, stratum codes and unequal selection probability (multidensity) categories for the survey designs are examined. Stream type is contained in a variable named "perint", and Strahler stream order is contained in a variable named "strahcat". For stream type, streams are classified as either perennial or intermittent. For Strahler stream order, streams are classified as either first order ("1st"), second order ("2nd"), or third order and higher ("3rd+"). The table and addmargin functions are used to produce a table displaying number of stream segments for each combination of values for the strata and multidensity category variables.

Finally, frame stream length is summarized for the strata and multidensity category attributes. Note that stream length measured in kilometers is contained in the variable named "length_km". The tapply function is used to calculate total stream length for each combination of stream type and Strahler stream order. The addmargins function is applied to the output from tapply to calculate stream length for each category of stream type and Strahler stream order, and the round function is used to round value to two decimal places. Finally, the resulting cross-tabulation of sum of stream length in kilometers for Strahler stream order and stream type is displayed.

Read the attribute table from the shapefile

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

Display the initial six lines in the attribute data frame

```
> # Display the initial six lines in the attribute data frame
> head(att)


        perint strahcat length_km length_mdm
1    Perennial      2nd     2.326       2326
2 Intermittent      1st     0.579        579
3 Intermittent      1st     0.780        780
4    Perennial      1st     1.876       1876
5 Intermittent      1st     1.001       1001
6    Perennial      1st     1.646       1646


>
```

Display number of stream segments cross-classified by the strata and multidensity category variables

```
> # Display number of stream segments cross-classified by the  strata and
> # multidensity category variables
>
> addmargins(table("Stream Type"=att$perint, "Strahler Order"=att$strahcat))


              Strahler Order
Stream Type    1st 2nd 3rd+ Sum
  Intermittent 137  20    2 159
  Perennial    104  78   88 270
  Sum          241  98   90 429


>
```

Summarize frame stream length by stratum and multidensity category

```
> # Summarize frame stream length by stratum and multidensity category
> temp <- tapply(att$length_km, list(att$perint, att$strahcat), sum)
> temp <- round(addmargins(temp), 2)
> names(dimnames(temp)) <- list("Stream Type", "Strahler Order")
> temp


              Strahler Order
Stream Type    1st    2nd   3rd+ Sum
  Intermittent 306   20.5   3.03 329
  Perennial    201  133.1 159.79 493
  Sum          506  153.6 162.82 822
```

```
>
```

Streams in the Luckiamute watershed are displayed in Figure 1 classified by stream type and in Figure 2 classified by Strahler stream order category. To produce the figure, first the read.shape function in the spsurvey package is used to read the shapefile and assign it to an object named shp. The shp object takes the form of a spatial data object defined in the sp package. Specifically, shp belongs to class "SpatialLinesDataFrame". For further information about spatial data objects, see documentation for the sp package. The spplot function in the sp package is used to create the figures.

Read the shapefile

```
> shp <- read.shape("luck_ash")
```

Plot streams in the Luckiamute watershed classified by stream type

```
> spplot(shp, zcol="perint", col.regions=c("red", "blue"))
```

Plot streams in the Luckiamute watershed classified by Strahler stream order

```
> spplot(shp, zcol="strahcat", col.regions=c("red", "green", "blue"))
```

3 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 50, 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 "linear" to indicate a linear 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 "luck_ash"; (6) att.frame: the data frame of attributes associated with elements in the frame, which is
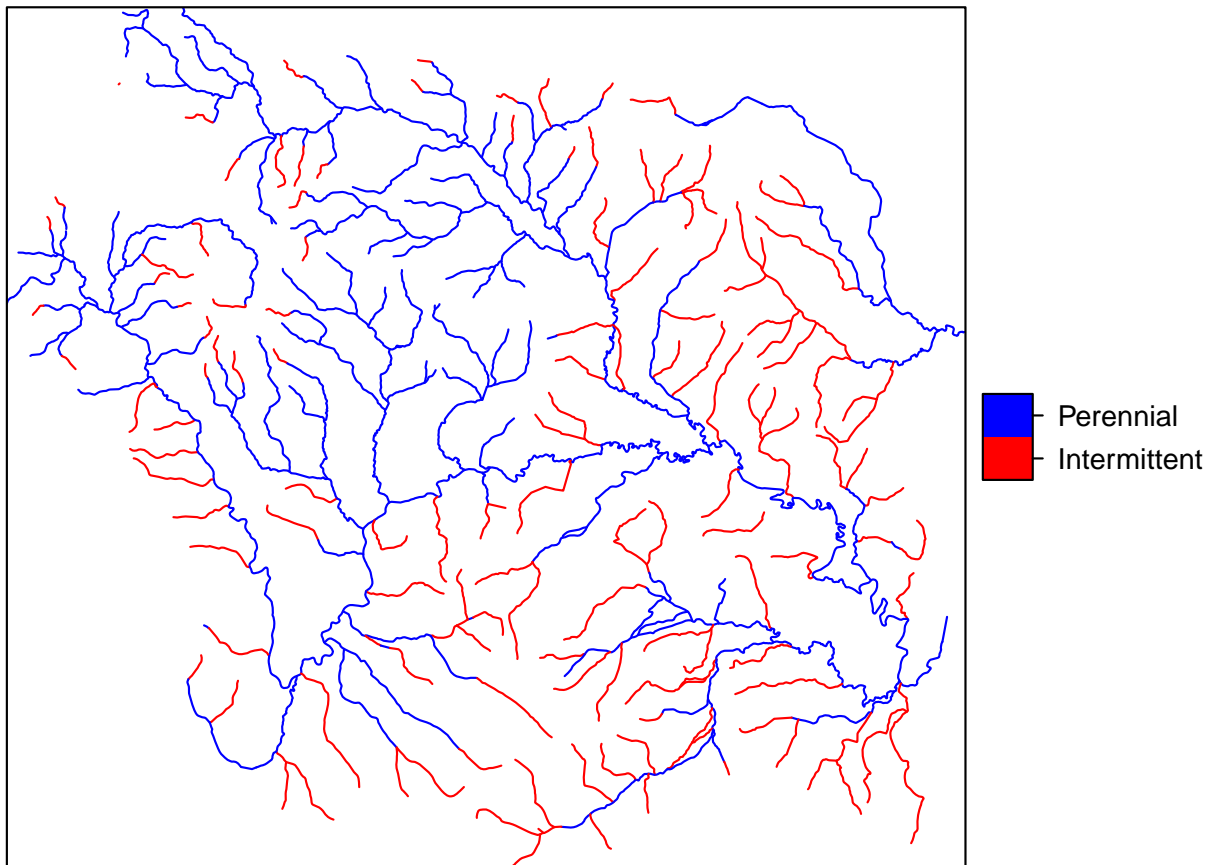
Figure 1: Streams in the Luckiamute Watershed Classified by Stream Type.
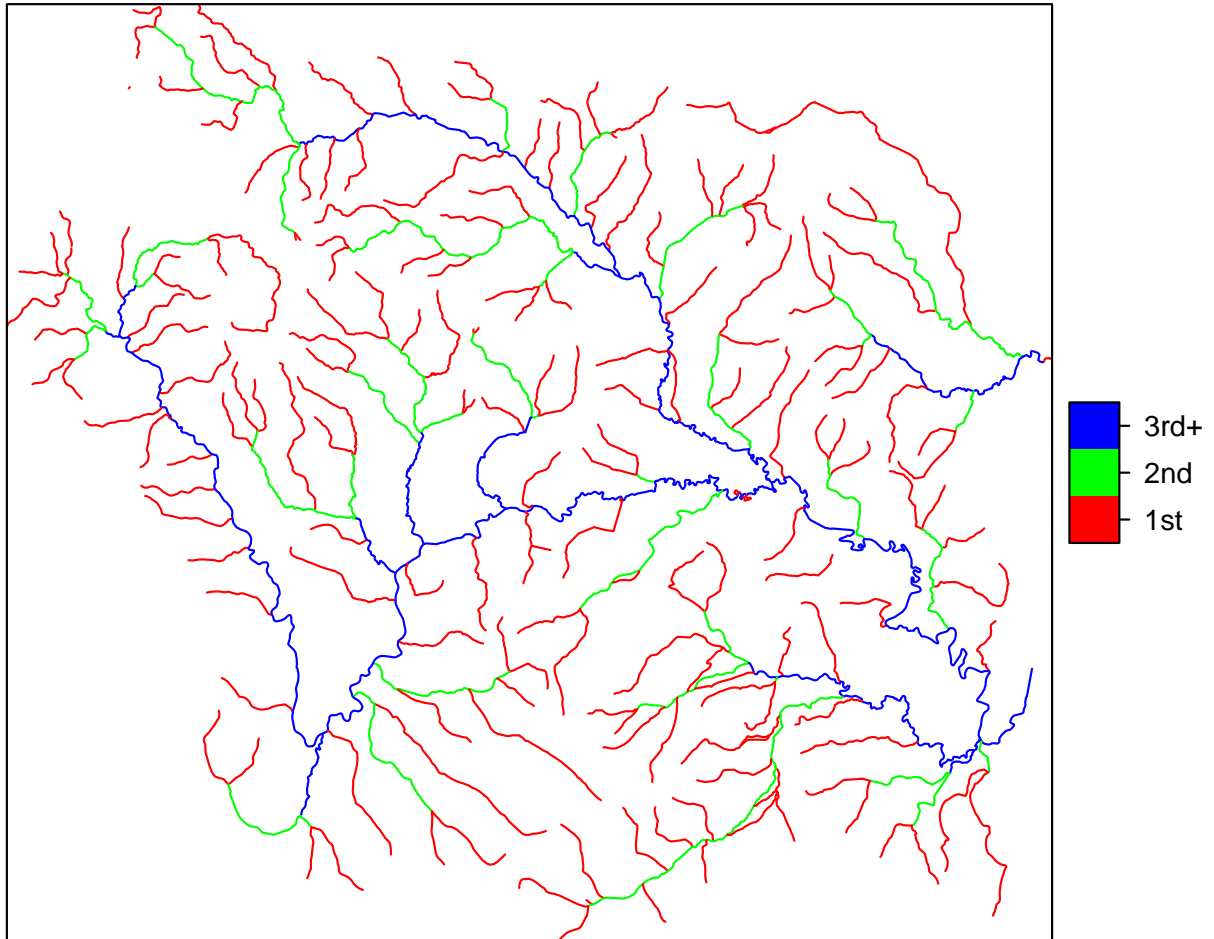
Figure 2: Streams in the Luckiamute Watershed Classified by Strahler Stream Order Category.

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.

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(19742003)
>
```

Create the design list

```
> # Create the design list
> Equaldsgn <- list(None=list(panel=c(PanelOne=50), seltype="Equal"))
>
```

Select the sample

```
> Equalsites <- grts(design=Equaldsgn,
+                    DesignID="EQUAL",
+                    type.frame="linear",
+                    src.frame="shapefile",
+                    in.shape="luck_ash",
+                    att.frame=att,
+                    shapefile=FALSE)


Stratum: None
Initial number of levels: 3
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(Equalsites@data)
```

```
      siteID   xcoord  ycoord mdcaty   wgt stratum       panel EvalStatus EvalReason
1 EQUAL-01 -2151443 2734150  Equal 16450    None PanelOne    NotEval
2 EQUAL-02 -2144834 2735275  Equal 16450    None PanelOne    NotEval
3 EQUAL-03 -2143643 2743030  Equal 16450    None PanelOne    NotEval
4 EQUAL-04 -2124733 2722643  Equal 16450    None PanelOne    NotEval
5 EQUAL-05 -2130690 2732801  Equal 16450    None PanelOne    NotEval
6 EQUAL-06 -2140629 2722765  Equal 16450    None PanelOne    NotEval
        perint strahcat length_km
1    Perennial      1st      1.91
2 Intermittent      1st      1.50
3    Perennial      2nd      2.51
4    Perennial     3rd+      2.49
5 Intermittent      1st      2.34
6    Perennial     3rd+      1.79


>
```

Print the survey design summary

```
> # Print the survey design summary
> dsgnsum(Equalsites)


Design Summary: Number of Sites

stratum
None   Sum
  50    50


>
```

4 Stratified, equal probability, GRTS survey design with an oversample

The second survey design is a stratified, equal probability design with an oversample. The stream type attribute is used to identify strata. List Stratdsgn is assigned design specifications. Since the survey design is stratified, Stratdsgn includes two lists named "Perennial" and "Intermittent" that contains three items: panel, seltype, and over. Note that the names for the two lists match the levels of the stratum variable. For both lists, the values for panel and seltype are the same as the ones used for the equal probability design. The third item, over, assigns the value 50 for size of the oversample. An oversample provides additional sample sites to replace sites that cannot be used, e.g., to replace sites in the sample that are not accessible.

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 Stratdsgn list; (2) DesignID: assigned the value "STRATIFIED"; (3) type.frame: assigned the value "linear"; (4) src.frame: assigned the value "shapefile"; (5) in.shape: assigned the value "luck_ash"; (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 "perint"; 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

```
> Stratdsgn <- list(Perennial=list(panel=c(PanelOne=50),
+                                  seltype="Equal",
+                                  over=50),
+                 Intermittent=list(panel=c(PanelOne=50),
+                                   seltype="Equal",
+                                   over=50))
```

Select the sample

```
> Stratsites <- grts(design=Stratdsgn,
+                DesignID="STRATIFIED",
+                type.frame="linear",
+                src.frame="shapefile",
+                in.shape="luck_ash",
+                att.frame=att,
+                stratum="perint",
+                shapefile=FALSE)


Stratum: Perennial
Initial number of levels: 4
Current number of levels: 4
Current number of levels: 5
Final number of levels: 5

Stratum: Intermittent
Initial number of levels: 4
Current number of levels: 4
Current number of levels: 5
Final number of levels: 5
```

Print the initial six lines of the survey design

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

```
          siteID   xcoord  ycoord mdcaty  wgt   stratum    panel EvalStatus
1 STRATIFIED-001 -2137017 2737685  Equal 9868 Perennial PanelOne    NotEval
2 STRATIFIED-002 -2147634 2745811  Equal 9868 Perennial PanelOne    NotEval
3 STRATIFIED-003 -2142407 2728762  Equal 9868 Perennial PanelOne    NotEval
4 STRATIFIED-004 -2147848 2732473  Equal 9868 Perennial PanelOne    NotEval
5 STRATIFIED-005 -2131391 2741066  Equal 9868 Perennial PanelOne    NotEval
6 STRATIFIED-006 -2139311 2724986  Equal 9868 Perennial PanelOne    NotEval
  EvalReason strahcat length_km
1              1st     1.781
2              2nd     3.100
3              1st     6.384
4              3rd+    0.611
5              1st     2.895
6              3rd+    1.653


>
```

Print the survey design summary

```
> # Print the survey design summary
> dsgnsum(Stratsites)


Design Summary: Number of Sites Classified by panel and stratum

          stratum
panel     Perennial Intermittent Sum
  OverSamp        50           50 100
  PanelOne        50           50 100
  Sum            100          100 200


>
```

5   Stratified, unequal probability, GRTS survey design with an oversample

The third survey design is a stratified, unequal probability design with an oversample. As for the second survey design, the stream type attribute is used to identify strata. Strahler order categories are used to identify multidensity categories. List Unequaldsgn is assigned design specifications. Unequaldsgn includes the same two lists with three items (panel, seltype, and over) as used for the stratified, equal probability design plus a value for caty.n. For both lists, panel specifies a single panel, and seltype is assigned "Unequal" to indicate

unequal probability sampling. Note that the value 0 is assigned to over for the "Intermittent" stratum, i.e., no oversample. The over item could have been omitted from the list for "Intermittent". The vector assigned to caty.n specifies sample sizes for each of the three multidensity categories. Note that the sum of values provided in caty.n must equal the value in panel.

For this survey design, an sp package object will be used as the sampling frame. Recall that the read.shape function was used to read the shapefile and assign its output to an sp object named shp. 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 "linear"; (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: assigned the value "perint"; (8) 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 "strahcat"; (9) 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

```
> Unequaldsgn <- list(Perennial=list(panel=c(PanelOne=75),
+                                     seltype="Unequal",
+                                     caty.n=c("1st"=25, "2nd"=25, "3rd+"=25),
+                                     over=36),
+                      Intermittent=list(panel=c(PanelOne=32),
+                                        seltype="Unequal",
+                                        caty.n=c("1st"=25, "2nd"=5, "3rd+"=2),
+                                        over=0))
```

Select the sample

```
> Unequalsites <- grts(design=Unequaldsgn,
+                  DesignID="UNEQUAL",
+                  type.frame="linear",
+                  src.frame="sp.object",
+                  sp.object=shp,
+                  att.frame=att,
+                  stratum="perint",
+                  mdcaty="strahcat",
+                  shapefile=FALSE)


Stratum: Perennial
Initial number of levels: 4
Current number of levels: 4
Current number of levels: 5
```

```
Final number of levels: 5


Stratum: Intermittent
Initial number of levels: 3
Current number of levels: 3
Current number of levels: 4
Current number of levels: 5
Final number of levels: 5
```

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      panel EvalStatus
1 UNEQUAL-001 -2127979 2739386    2nd 5324 Perennial PanelOne    NotEval
2 UNEQUAL-002 -2151221 2737138    2nd 5324 Perennial PanelOne    NotEval
3 UNEQUAL-003 -2143315 2724152   3rd+ 6392 Perennial PanelOne    NotEval
4 UNEQUAL-004 -2148760 2734240   3rd+ 6392 Perennial PanelOne    NotEval
5 UNEQUAL-005 -2133077 2740938    2nd 5324 Perennial PanelOne    NotEval
6 UNEQUAL-006 -2129966 2736424   3rd+ 6392 Perennial PanelOne    NotEval
  EvalReason length_km
1               0.279
2               2.958
3               4.612
4               2.141
5               2.640
6               2.294


>
```

Print the survey design summary

```
> # Print the survey design summary
> dsgnsum(Unequalsites)


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

      stratum
mdcaty Perennial Intermittent Sum
  1st         32           25  57
```

```
2nd              39              5  44
3rd+             40              2  42
Sum             111             32 143
```

Design Summary: Number of Sites Classified by panel and stratum

```
          stratum
panel        Perennial Intermittent Sum
  OverSamp          36            0  36
  PanelOne          75           32 107
  Sum              111           32 143
```

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

, , stratum = Perennial

```
      panel
mdcaty OverSamp PanelOne Sum
  1st        10       22  32
  2nd        11       28  39
  3rd+       15       25  40
  Sum        36       75 111
```

, , stratum = Intermittent

```
      panel
mdcaty OverSamp PanelOne Sum
  1st         0       25  25
  2nd         0        5   5
  3rd+        0        2   2
  Sum         0       32  32
```

, , stratum = Sum

```
      panel
mdcaty OverSamp PanelOne Sum
  1st        10       47  57
  2nd        11       33  44
  3rd+       15       27  42
  Sum        36      107 143
```

```
>
```

6 Stratified, unequal probability, GRTS survey design with an oversample and a panel structure for survey over time

The fourth survey design is a stratified, unequal probability design with an oversample and a panel structure for survey over time. List Paneldsgn is assigned design specifications. Analogous to the stratified, unequal probability design, Paneldsgn includes two lists named "Perennial" and "Intermittent". For the "Perennial" stratum, a vector identifying sample sizes for three panels is assigned to panel. For the "Intermittent" stratum, the sample size for a single panel named "YearOnce" is assigned to panel. The value "Unequal" is assigned to seltype for both lists, which indicates unequal selection probabilities. For both lists, the third item, caty.n, assigns sample sizes for each of the three multidensity categories. Again, note that the sum of sample sizes provided in caty.n must equal the sum of sample sizes in panel. For the "Perennial" stratum, the value 50 is assigned to over, which specifies an oversample of 50 sites. No oversample is specified for the "Intermittent" stratum, and so over is not included in the list. 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 proportion is rounded to the next higher integer. For the "Perennial" stratum, the oversample is not proportionate to the multidensity 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 "linear"; (4) src.frame: assigned the value "shapefile"; (5) in.shape: assigned the value "luck_ash"; (6) att.frame: assigned the att data frame; 7) stratum: assigned the value "perint"; (8) mdcaty: assigned the value "strahcat"; (9) 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

```
> Paneldsgn <- list(Perennial=list(panel=c(Year1=17, Year2=17, YearAll=16),
+                                  seltype="Unequal",
+                                  caty.n=c("1st"=15, "2nd"=15, "3rd+"=20),
+                                  over=50),
+                    Intermittent=list(panel=c(YearOnce=27),
+                                      seltype="Unequal",
+                                      caty.n=c("1st"=20, "2nd"=5, "3rd+"=2)))
```

Select the sample

```
> Panelsites <- grts(design=Paneldsgn,
+                    DesignID="UNEQUAL",
```

```
+                      type.frame="linear",
+                      src.frame="shapefile",
+                      in.shape="luck_ash",
+                      att.frame=att,
+                      stratum="perint",
+                      mdcaty="strahcat",
+                      shapefile=FALSE)


Stratum: Perennial
Initial number of levels: 4
Current number of levels: 4
Current number of levels: 5
Current number of levels: 6
Final number of levels: 6

Stratum: Intermittent
Initial number of levels: 3
Current number of levels: 3
Current number of levels: 4
Current number of levels: 5
Final number of levels: 5
```

Print the warning message

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


Warning message:
In grts(design = Paneldsgn, DesignID = "UNEQUAL", type.frame = "linear",  :

Oversample size is not proportional to category sample sizes for stratum
"Perennial".
```

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 panel EvalStatus
1 UNEQUAL-001 -2148784 2745360    1st 13369 Perennial Year1    NotEval
2 UNEQUAL-002 -2152267 2736268    1st 13369 Perennial Year1    NotEval
```

```
3 UNEQUAL-003 -2128882 2729846   3rd+  7990 Perennial Year1    NotEval
4 UNEQUAL-004 -2146814 2731447   3rd+  7990 Perennial Year1    NotEval
5 UNEQUAL-005 -2141802 2742260   3rd+  7990 Perennial Year1    NotEval
6 UNEQUAL-006 -2147981 2732448    1st 13369 Perennial Year1    NotEval
  EvalReason length_km
1               1.86
2               3.71
3               4.83
4               2.26
5               1.27
6               3.03


>
```

Print the survey design summary

```
> # Print the survey design summary
> dsgnsum(Panelsites)


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


      stratum
mdcaty Perennial Intermittent Sum
  1st         30           19  49
  2nd         28            5  33
  3rd+        42            3  45
  Sum        100           27 127



Design Summary: Number of Sites Classified by panel and stratum


         stratum
panel     Perennial Intermittent Sum
  OverSamp        50            0  50
  Year1           17            0  17
  Year2           17            0  17
  YearAll         16            0  16
  YearOnce         0           27  27
  Sum            100           27 127



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

```
, , stratum = Perennial

       panel
mdcaty OverSamp Year1 Year2 YearAll YearOnce Sum
  1st          16     5     5       4        0  30
  2nd          17     4     4       3        0  28
  3rd+         17     8     8       9        0  42
  Sum          50    17    17      16        0 100

, , stratum = Intermittent

       panel
mdcaty OverSamp Year1 Year2 YearAll YearOnce Sum
  1st           0     0     0       0       19  19
  2nd           0     0     0       0        5   5
  3rd+          0     0     0       0        3   3
  Sum           0     0     0       0       27  27

, , stratum = Sum

       panel
mdcaty OverSamp Year1 Year2 YearAll YearOnce Sum
  1st          16     5     5       4       19  49
  2nd          17     4     4       3        5  33
  3rd+         17     8     8       9        3  45
  Sum          50    17    17      16       27 127


>
```