Video Game Data Analysis

Introduction

The following sections provide detailed analysis of a massively multiplayer online role-playing game (MMORPG) called Glitch. An online visualization of the game's data can be found at

http://powerful-meadow-8588.herokuapp.com/

The dataset for this analysis tracks the user's migration through the final 14 months of its operation (Nov 2011 through Dec 2012). The dataset shows the movement of player populations from month to month. The dataset is in the JSON format and available from

http://powerful-meadow-8588.herokuapp.com/data/12months_departures_joiners.json

In this game analysis, the players are classified into 11 various levels based on their activity:

Hardcore

Forum

Moderate, Moderate Winners, Moderate Losers, Moderate Farmers, Moderate Miscellanea

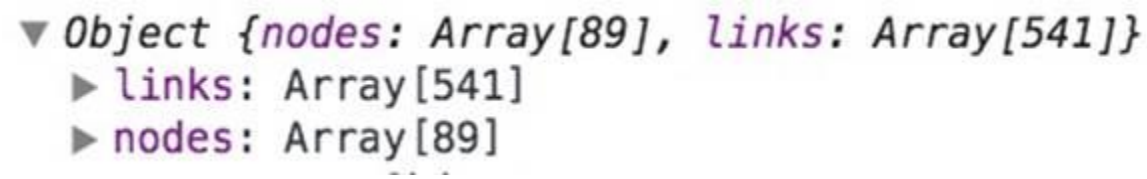
Casual, Casual Winners, Casual Losers, Casual Forum

A snapshot of the JSON data is shown below.


```
{
"nodes":[
{"name":"Casual Losers","node":0,"color":"#95d5af","month":"Nov-11","departing":451,"joining":0},
{"name":"Casual Winners","node":1,"color":"#53b67d","month":"Nov-11","departing":671,"joining":0},
{"name":"Moderate Farmers","node":2,"color":"#a898b6","month":"Nov-11","departing":270,"joining":0},
{"name":"Moderate","node":3,"color":"#332341","month":"Nov-11","departing":184,"joining":0},
{"name":"Forum","node":4,"color":"#f3bd4e","month":"Nov-11","departing":0,"joining":0},
{"name":"Hardcore","node":5,"color":"#5089a8","month":"Nov-11","departing":169,"joining":0},
{"name":"Casual Losers","node":6,"color":"#95d5af","month":"Dec-11","departing":912,"joining":859},
{"name":"Casual Winners","node":7,"color":"#53b67d","month":"Dec-11","departing":1247,"joining":1112},
{"name":"Moderate Farmers","node":8,"color":"#a898b6","month":"Dec-11","departing":535,"joining":465},
{"name":"Moderate","node":9,"color":"#332341","month":"Dec-11","departing":765,"joining":578},
{"name":"Forum","node":10,"color":"#f3bd4e","month":"Dec-11","departing":30,"joining":89},
{"name":"Hardcore","node":11,"color":"#5089a8","month":"Dec-11","departing":541,"joining":383},
{"name":"Casual","node":84,"color":"#398b5c","month":"Dec-12","departing":0,"joining":169},
{"name":"Moderate Miscellanea","node":85,"color":"#c7bfce","month":"Dec-12","departing":0,"joining":36},
{"name":"Moderate Farmers","node":86,"color":"#a898b6","month":"Dec-12","departing":0,"joining":60},
{"name":"Forum","node":87,"color":"#f3bd4e","month":"Dec-12","departing":0,"joining":12},
{"name":"Hardcore","node":88,"color":"#5089a8","month":"Dec-12","departing":0,"joining":28}
],

"links":[
{"source":3,"target":9,"value":112},
{"source":3,"target":8,"value":249},
{"source":3,"target":6,"value":93},
{"source":3,"target":10,"value":52},
{"source":3,"target":7,"value":92},
{"source":3,"target":11,"value":248},
```


The JSON data consists of two properties at the top level, nodes and links. The nodes property has 89 entries and the links property has 541 entries.



Each node entry is a JSON object capturing the data for the month for a particular level of the players. The node entry has the following properties:

node **– a unique node number (ranges from 0 to 88)**

name **– the level of the players for this node**

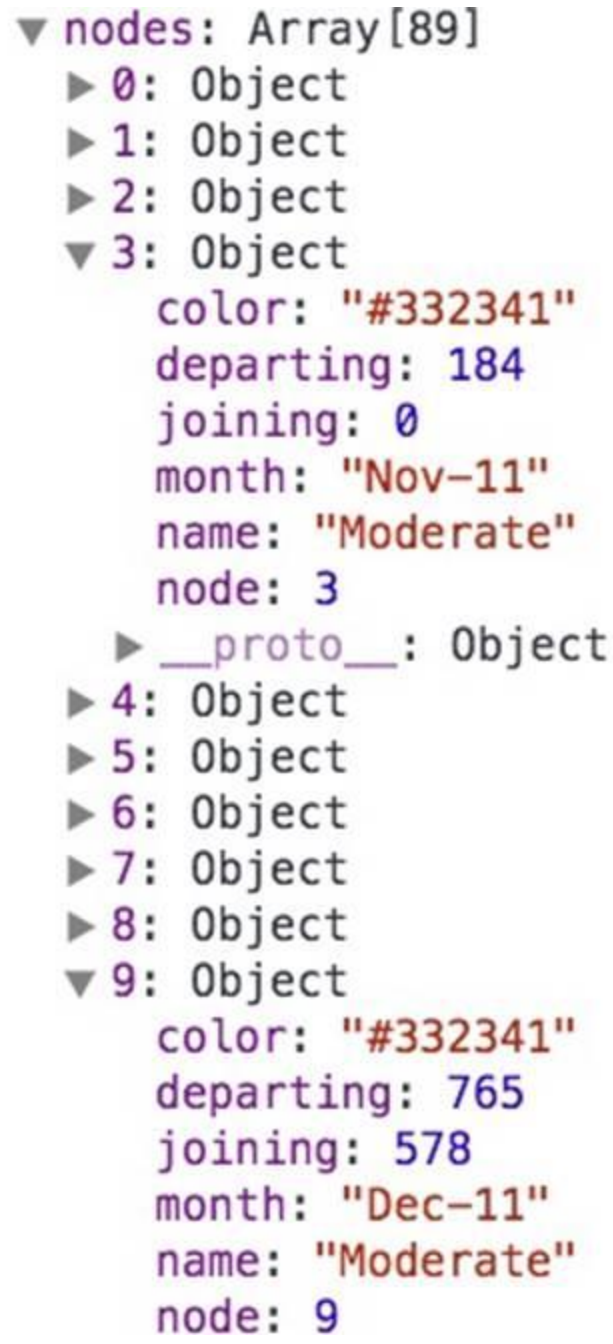
month **– the month associated with this node**

joining **– the number of players joining at this level for this month**

departing **– the number of players departing at this level for this month**

color **– the unique color for the level associated with this node**

Information about few of the nodes is shown below.



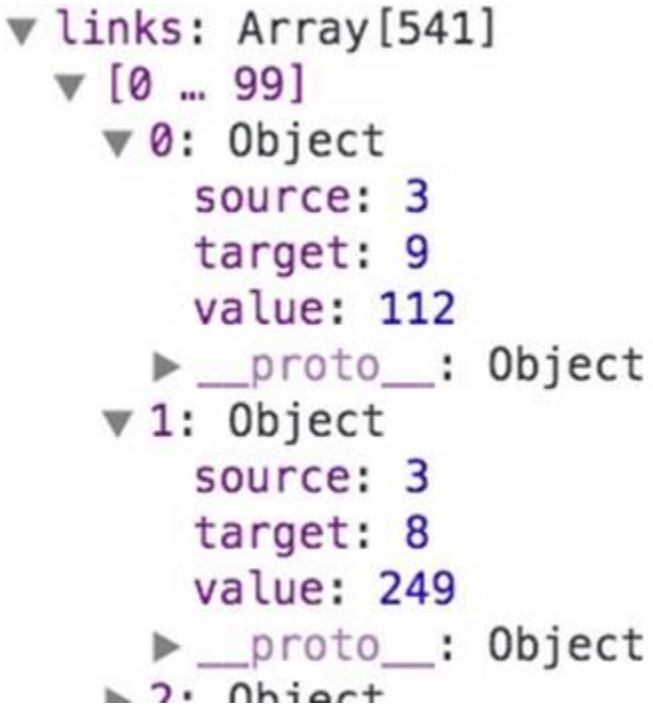
Each link entry captures the player migration data from the source node of one month to the target node of the subsequent month. Each entry has three properties:

source **– the source node of the link**

target **– the target node of the link (subsequent month)**

value **– the number of players going from the source node player level to the target node player level.**

Information about few of the links is shown below.



Importing Game Data into R

The RCurl and RJSONIO packages are used for importing the JSON data from the given URL and converting the data into an R data structure.

```
> library(RCurl)
> library(RJSONIO)
>
> webpage <-
+   paste0("http://powerful-meadow-8588.herokuapp.com/",
+          "data/12months_departures_joiners.json", sep="")
>
> data <- fromJSON(getURL(webpage))
```

The data variable is now a list of two components, nodes and links.

```
> typeof(data)
[1] "list"
> names(data)
[1] "nodes" "links"
```

Node Information

The nodes component of the list data structure imported in the previous section is again a list of 89 components, 1 component per node.

```
> typeof(data$nodes)
[1] "list"
> length(data$nodes)
[1] 89
```

The first node can be inspected as shown below.

```
> data$nodes[[1]]
$name
[1] "Casual Losers"

$node
[1] 0

$color
[1] "#95d5af"

$month
[1] "Nov-11"

$departing
[1] 451

$joining
[1] 0
```

The information regarding all the nodes can be converted to a data frame as shown below. Each node is first converted to a data frame with one row, and all these individual data frames are bound into one. The first six and the last six rows of the data frame are also shown here.

```
> nodes.info <-
+   do.call("rbind",
+           lapply(data$nodes, data.frame))
>
> head(nodes.info)
              name node   color  month departing joining
1    Casual Losers    0 #95d5af Nov-11       451       0
2   Casual Winners    1 #53b67d Nov-11       671       0
3 Moderate Farmers    2 #a898b6 Nov-11       270       0
4         Moderate    3 #332341 Nov-11       184       0
5            Forum    4 #f3bd4e Nov-11         0       0
6         Hardcore    5 #5089a8 Nov-11       169       0
> tail(nodes.info)
                  name node   color  month departing joining
84             Hardcore   83 #5089a8 Nov-12       800     306
85               Casual   84 #398b5c Dec-12         0     169
86 Moderate Miscellanea   85 #c7bfce Dec-12         0      36
87     Moderate Farmers   86 #a898b6 Dec-12         0      60
88                Forum   87 #f3bd4e Dec-12         0      12
89             Hardcore   88 #5089a8 Dec-12         0      28
```

The month column of the data frame shows the month name associated with each node. The 14 months for which the data is collected are as shown below.

```
> months <- unique(nodes.info$month)
> months
 [1] Nov-11 Dec-11 Jan-12 Feb-12 Mar-12 Apr-12 May-12
 [8] Jun-12 Jul-12 Aug-12 Sep-12 Oct-12 Nov-12 Dec-12
14 Levels: Nov-11 Dec-11 Jan-12 Feb-12 ... Dec-12
```

The number of nodes (player segments) for each month is computed as follows:

```
> table(nodes.info$month)

Nov-11 Dec-11 Jan-12 Feb-12 Mar-12 Apr-12 May-12
     6      6      7      7      7      7      7
Jun-12 Jul-12 Aug-12 Sep-12 Oct-12 Nov-12 Dec-12
     7      7      6      6      6      5      5
```

The name column of the data frame shows the player level associated with each node. The 11 levels into which the players are classified is as shown below.

```
> levels <- unique(nodes.info$name)
> levels
 [1] Casual Losers         Casual Winners
 [3] Moderate Farmers      Moderate
 [5] Forum                 Hardcore
 [7] Moderate Miscellanea Moderate Losers
 [9] Casual Forum          Moderate Winners
[11] Casual
11 Levels: Casual Losers Casual Winners ... Casual
```

The number of nodes associated with each player segment over the 14 months is computed as shown below.

```
> table(nodes.info$name)

        Casual Losers       Casual Winners
                   12                   12
     Moderate Farmers             Moderate
                   14                    2
                Forum             Hardcore
                   13                   14
Moderate Miscellanea       Moderate Losers
                    9                    6
         Casual Forum      Moderate Winners
                    4                    1
               Casual
                    2
```

Link Information

The links component of the list data structure shows the data about the player migration from a source node to a target node, from the given month to the next month. The links component is again a list of 541 components.

```
> typeof(data$links)
[1] "list"
> length(data$links)
[1] 541
```

The first link can be inspected as shown below.

```
> data$links[[1]]
source target  value
     3      9    112
```

Each component in the links is a named vector with three values, the source node number, the target node number, and the number of players who migrated to the target node level. The data about these links can be converted into a data frame as shown below. The first six and the last six rows of the resulting data frame are also shown.

```
> links.info <-
+   data.frame(do.call("rbind", data$links))
>
> head(links.info)
  source target value
1      3      9   112
2      3      8   249
3      3      6    93
4      3     10    52
5      3      7    92
6      3     11   248
```

```
> tail(links.info)
    source target value
536     80     85    42
537     83     84    50
538     83     86    55
539     83     88   145
540     83     87     5
541     83     85    80
```

In order to show the names of the player levels for each node in the graphs, the links.info data frame is modified my adding the SourceName column. The player level for the corresponding source node is obtained from the nodes.info data frame.

```
> links.info$SourceName <-
+   paste0(links.info$source, "-",
+          nodes.info[links.info$source+1, "name"])
>
> head(links.info)
  source target value SourceName
1      3      9   112 3-Moderate
2      3      8   249 3-Moderate
3      3      6    93 3-Moderate
4      3     10    52 3-Moderate
5      3      7    92 3-Moderate
6      3     11   248 3-Moderate
```

Similarly, the TargetName column is added to the links.info data frame. The player level for the corresponding target node is obtained from the nodes.info data frame. The first six rows of the modified data frame are also shown below.

```
> links.info$TargetName <-
+   paste0(links.info$target, "-",
+          nodes.info[links.info$target+1, "name"])
>
> head(links.info)
  source target value SourceName         TargetName
1      3      9   112 3-Moderate         9-Moderate
2      3      8   249 3-Moderate 8-Moderate Farmers
3      3      6    93 3-Moderate    6-Casual Losers
4      3     10    52 3-Moderate           10-Forum
5      3      7    92 3-Moderate   7-Casual Winners
6      3     11   248 3-Moderate        11-Hardcore
```

Visualizing Player Migration

The following example uses the data for the months of Nov 2011, Dec 2011, and Jan 2012 to show the migration of players across various levels from one month to the subsequent month. The first step in this process is to find the information about the source nodes for these months using the nodes.info data frame.

```
> source.data <-
+   nodes.info[nodes.info$month %in%
+                c("Nov-11", "Dec-11", "Jan-12"),
+              "node"]
>
> source.data
 [1]  0  1  2  3  4  5  6  7  8  9 10 11 12 13 14
[16] 15 16 17 18
```

The values associated with the above nodes show the players going from one level to other level in the subsequent month. The relevant information is obtained from the links.info data frame as shown below.

```
> links.data <-
+   links.info[links.info$source %in% source.data,
+              c("SourceName", "TargetName", "value")]
>
```

The first six and the last six rows of the player migration data for the first three months is shown below.

```
> head(links.data)
  SourceName          TargetName value
1 3-Moderate          9-Moderate   112
2 3-Moderate 8-Moderate Farmers   249
3 3-Moderate    6-Casual Losers    93
4 3-Moderate            10-Forum   52
5 3-Moderate   7-Casual Winners    92
6 3-Moderate        11-Hardcore   248
> tail(links.data)
     SourceName          TargetName value
122 18-Hardcore        25-Hardcore   444
123 18-Hardcore 22-Moderate Farmers   219
124 18-Hardcore            24-Forum   33
125 18-Hardcore  23-Moderate Losers    76
126 18-Hardcore    20-Casual Losers    24
127 18-Hardcore     19-Casual Forum    12
```

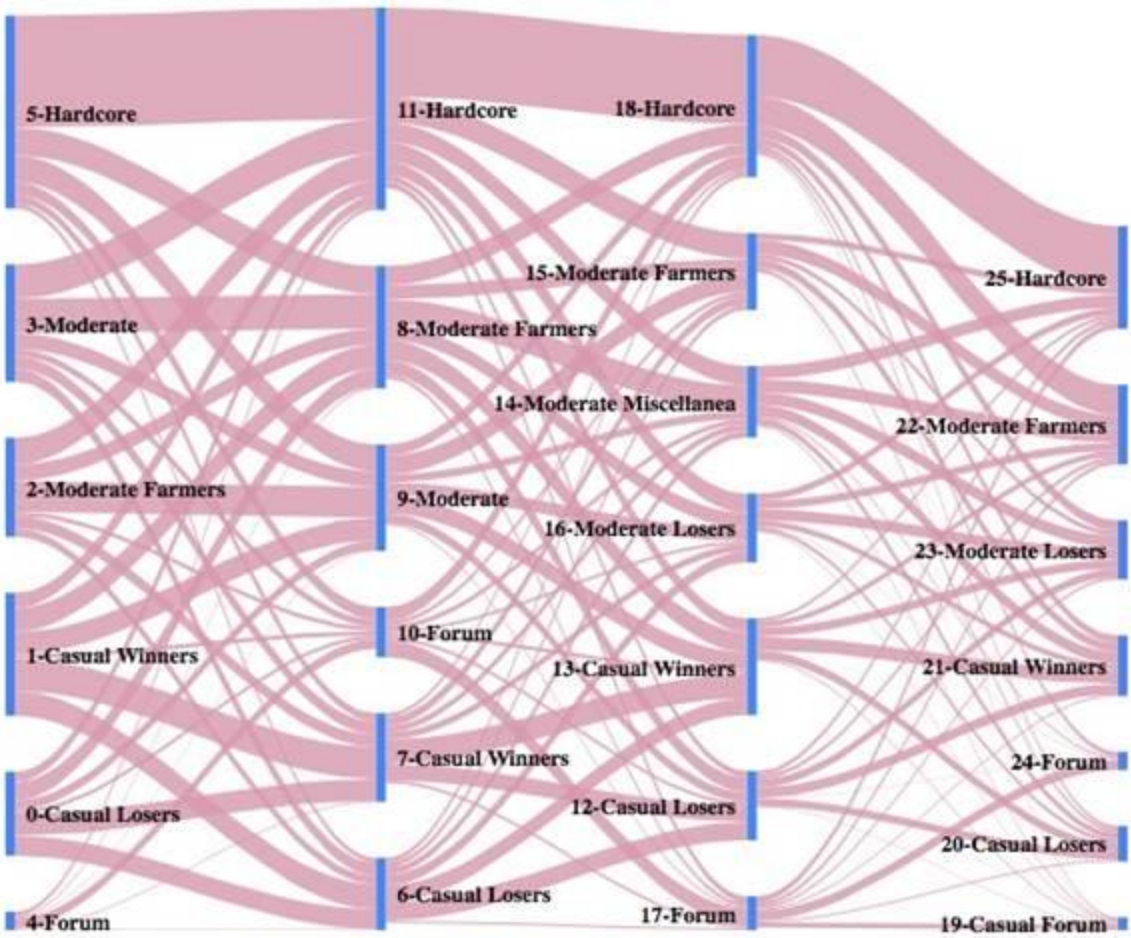
Using googleVis package for Google charts, the above data can be plotted using the gvisSankey function to create the Sankey diagram as shown below.

```
> chart.data <-
+   gvisSankey(links.data,
+              from="SourceName",
+              to="TargetName",
+              weight="value",
+              options=list(
+                width=600, height=500,
+                sankey=options
+              ))
> plot(chart.data)
```

The options used for the Sankey chart are:

```
options <- "{
  node: {
    label: {
      fontName: 'Times-Roman',
      fontSize: 12,
      color: '#000',
      bold: true,
      italic: false
    },
    labelPadding: 6,
    nodePadding: 30,
    width: 5
  },
  link: {
    color: { fill: '#d799ae' }
  }
}"
```

The plot shows the player migration data from Nov-2011 to Dec-2011, Dec-2011 to Jan-2012, and Jan-2012 to Feb-2012.



The data for the names of source months and target months can be built using a data frame as shown below.

```
> months.names <- data.frame(
+   rbind(c("Nov-11...",
+           "Dec-11...",
+           "Jan-12...")))
>
> months.names
         X1         X2         X3
1 Nov-11... Dec-11... Jan-12...
>
> names(months.names) <-
+   paste(c("Nov-11", "Dec-11", "Jan-12"),
+         c("Dec-11", "Jan-12", "Feb-12"),
+         sep="====>")
>
> months.names
  Nov-11====>Dec-11 Dec-11====>Jan-12 Jan-12====>Feb-12
1         Nov-11...          Dec-11...          Jan-12...
```

The above data frame can be plotted using the gvisTable function to create a Google Table chart.

```
> chart.names <- gvisTable(months.names,
+                          options=list(
+                            width=600))
>
> plot(chart.names)
```

Nov-11====>Dec-11	Dec-11====>Jan-12	Jan-12====>Feb-12
Nov-11...	Dec-11...	Jan-12...

Data: months.names • Chart ID: TableID8edb3c3fed • googleVis-0.5.8
R version 3.1.2 (2014-10-31) • Google Terms of Use • Documentation and Data Policy

The information about the nodes (node number, month, and player level) can be created using a data frame for the source and target nodes during this period as shown below. The data is plotted using the gvisTable function.

```
> months.info <-
+   nodes.info[nodes.info$month %in%
+                c("Nov-11", "Dec-11", "Jan-12", "Feb-12"),
+              c("month", "node", "name")]
>
> chart.table <- gvisTable(months.info)
>
> plot(chart.table)
```

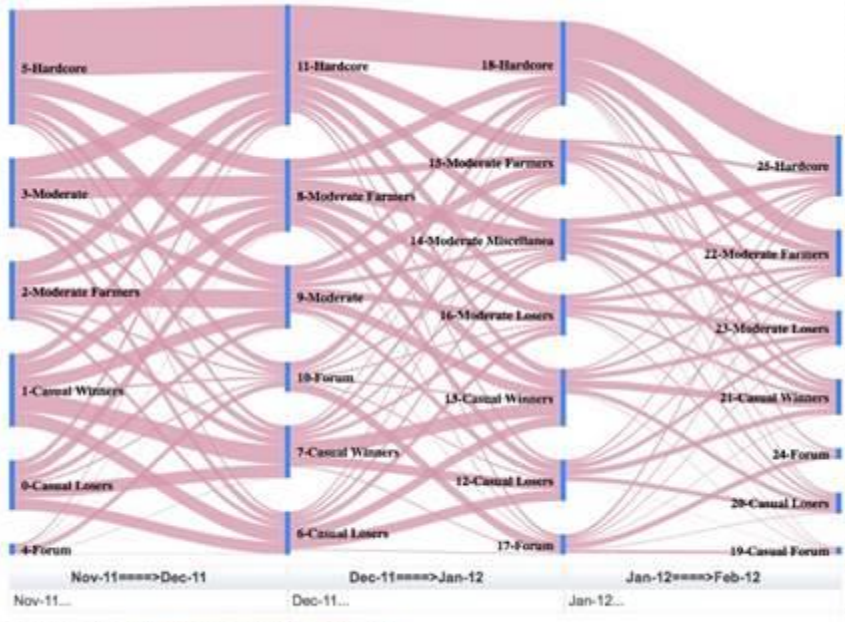
The Google Table chart for the above data is as shown below.

month	node	name
Nov-11	0	Casual Losers
Nov-11	1	Casual Winners
Nov-11	2	Moderate Farmers
Nov-11	3	Moderate
Nov-11	4	Forum
Nov-11	5	Hardcore
Dec-11	6	Casual Losers
Dec-11	7	Casual Winners
Dec-11	8	Moderate Farmers
Dec-11	9	Moderate
Dec-11	10	Forum
Dec-11	11	Hardcore
Jan-12	12	Casual Losers
Jan-12	13	Casual Winners
Jan-12	14	Moderate Miscellanea
Jan-12	15	Moderate Farmers
Jan-12	16	Moderate Losers
Jan-12	17	Forum
Jan-12	18	Hardcore
Feb-12	19	Casual Forum
Feb-12	20	Casual Losers
Feb-12	21	Casual Winners
Feb-12	22	Moderate Farmers
Feb-12	23	Moderate Losers
Feb-12	24	Forum
Feb-12	25	Hardcore

Data: months.info • Chart ID: TableID8ed2088e65e • googleVis-0.5.8
R version 3.1.2 (2014-10-31) • Google Terms of Use • Documentation and Data Policy

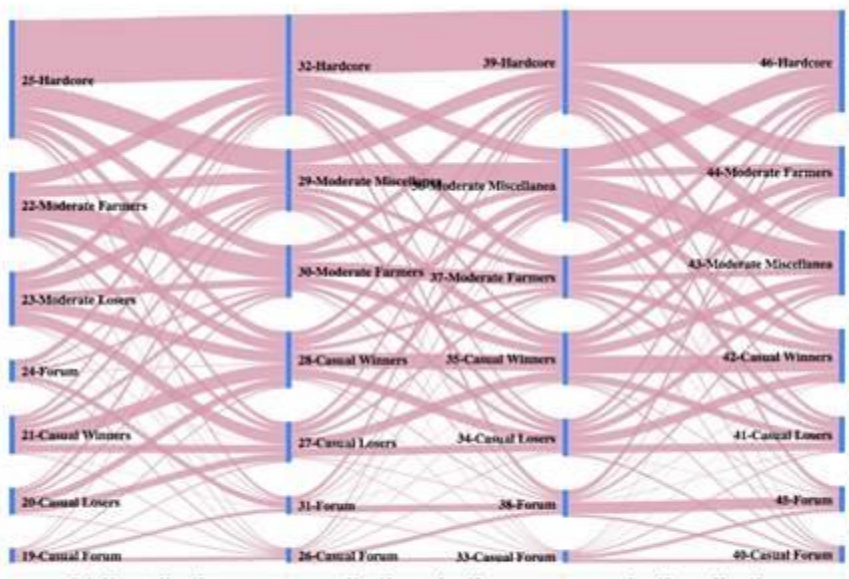
The three charts can be merged into one chart and plotted as shown below.

```
> chart1 <- gvisMerge(
+   gvisMerge(chart.data, chart.names),
+   chart.table, horizontal = TRUE)
>
> plot(chart1)
```



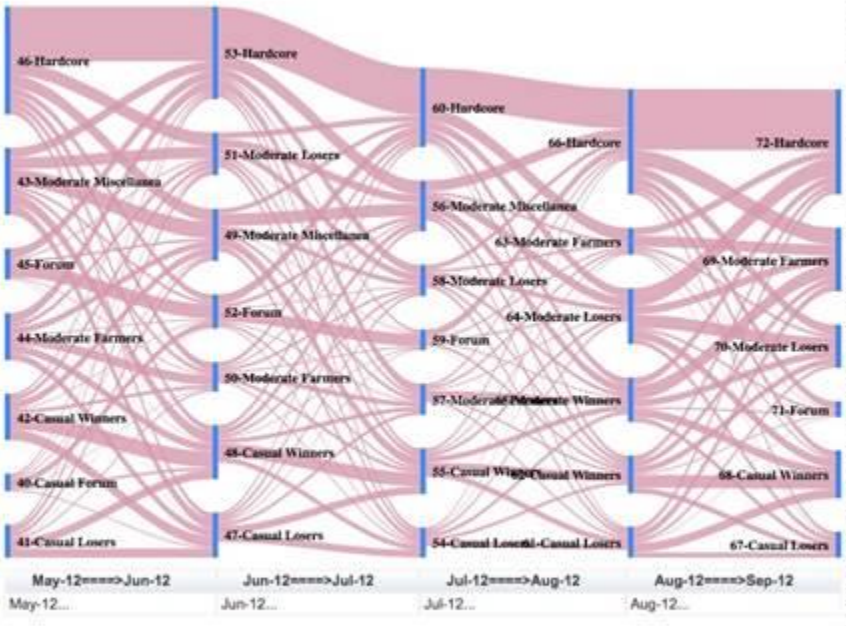
month	node	name
Nov-11	0	Casual Losers
Nov-11	1	Casual Winners
Nov-11	2	Moderate Farmers
Nov-11	3	Moderate
Nov-11	4	Forum
Nov-11	5	Hardcore
Dec-11	6	Casual Losers
Dec-11	7	Casual Winners
Dec-11	8	Moderate Farmers
Dec-11	9	Moderate
Dec-11	10	Forum
Dec-11	11	Hardcore
Jan-12	12	Casual Losers
Jan-12	13	Casual Winners
Jan-12	14	Moderate Miscellanea
Jan-12	15	Moderate Farmers
Jan-12	16	Moderate Losers
Jan-12	17	Forum
Jan-12	18	Hardcore
Feb-12	19	Casual Forum
Feb-12	20	Casual Losers
Feb-12	21	Casual Winners
Feb-12	22	Moderate Farmers
Feb-12	23	Moderate Losers
Feb-12	24	Forum
Feb-12	25	Hardcore

Similarly, the player migration information for the months Feb 2012, Mar 2012, and Apr 2012 is as shown below.



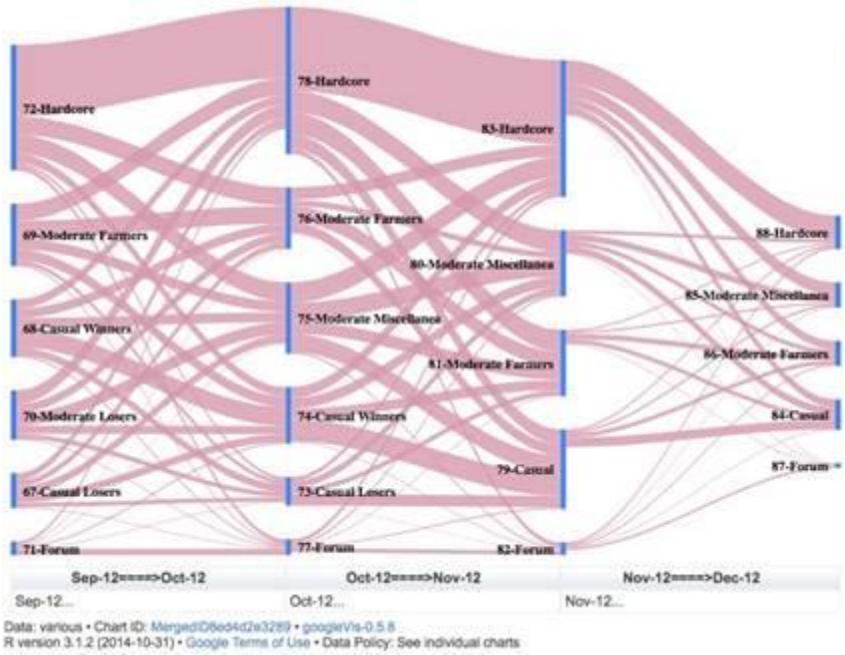
month	node	name
Feb-12	19	Casual Forum
Feb-12	20	Casual Losers
Feb-12	21	Casual Winners
Feb-12	22	Moderate Farmers
Feb-12	23	Moderate Losers
Feb-12	24	Forum
Feb-12	25	Hardcore
Mar-12	26	Casual Forum
Mar-12	27	Casual Losers
Mar-12	28	Casual Winners
Mar-12	29	Moderate Miscellanea
Mar-12	30	Moderate Farmers
Mar-12	31	Forum
Mar-12	32	Hardcore
Apr-12	33	Casual Forum
Apr-12	34	Casual Losers
Apr-12	35	Casual Winners
Apr-12	36	Moderate Miscellanea
Apr-12	37	Moderate Farmers
Apr-12	38	Forum
Apr-12	39	Hardcore
May-12	40	Casual Forum
May-12	41	Casual Losers
May-12	42	Casual Winners
May-12	43	Moderate Miscellanea
May-12	44	Moderate Farmers
May-12	45	Forum
May-12	46	Hardcore

Similarly, the player migration information for the months May 2012, Jun 2012, Jul 2012 and Aug 2012 is as shown below.



month	node	name
May-12	40	Casual Forum
May-12	41	Casual Losers
May-12	42	Casual Winners
May-12	43	Moderate Miscellanea
May-12	44	Moderate Farmers
May-12	45	Forum
May-12	46	Hardcore
Jun-12	47	Casual Losers
Jun-12	48	Casual Winners
Jun-12	49	Moderate Miscellanea
Jun-12	50	Moderate Farmers
Jun-12	51	Moderate Losers
Jun-12	52	Forum
Jun-12	53	Hardcore
Jul-12	54	Casual Losers
Jul-12	55	Casual Winners
Jul-12	56	Moderate Miscellanea
Jul-12	57	Moderate Farmers
Jul-12	58	Moderate Losers
Jul-12	59	Forum
Jul-12	60	Hardcore
Aug-12	61	Casual Losers
Aug-12	62	Casual Winners
Aug-12	63	Moderate Farmers
Aug-12	64	Moderate Losers
Aug-12	65	Moderate Winners
Aug-12	66	Hardcore
Sep-12	67	Casual Losers
Sep-12	68	Casual Winners
Sep-12	69	Moderate Farmers
Sep-12	70	Moderate Losers
Sep-12	71	Forum
Sep-12	72	Hardcore

Similarly, the player migration information for the months Sep 2012, Oct 2012, and Nov 2012 is as shown below.



month	node	name
Sep-12	67	Casual Losers
Sep-12	68	Casual Winners
Sep-12	69	Moderate Farmers
Sep-12	70	Moderate Losers
Sep-12	71	Forum
Sep-12	72	Hardcore
Oct-12	73	Casual Losers
Oct-12	74	Casual Winners
Oct-12	75	Moderate Miscellanea
Oct-12	76	Moderate Farmers
Oct-12	77	Forum
Oct-12	78	Hardcore
Nov-12	79	Casual
Nov-12	80	Moderate Miscellanea
Nov-12	81	Moderate Farmers
Nov-12	82	Forum
Nov-12	83	Hardcore
Dec-12	84	Casual
Dec-12	85	Moderate Miscellanea
Dec-12	86	Moderate Farmers
Dec-12	87	Forum
Dec-12	88	Hardcore

The above four plots can be merged into a single plot shown them in a vertical table using the nested gvisMerge functions.

```
> chart.final <- gvisMerge(
+   gvisMerge(chart1, chart2),
+   gvisMerge(chart3, chart4))
>
> plot(chart.final)
```

View HTML

Visualizing Player Categories

The node information for each month for each player level also includes information about the number of new players joining at that level, and also the players in that level who are completely quitting the game. The monthly information for the players classified as Hardcore is filtered from the nodes as shown below.

```
> hardcore <-
+   nodes.info[nodes.info$name == 'Hardcore', ]
>
> hardcore
       name node   color  month departing joining
6  Hardcore    5 #5089a8 Nov-11       169       0
12 Hardcore   11 #5089a8 Dec-11       541     383
19 Hardcore   18 #5089a8 Jan-12       278     118
26 Hardcore   25 #5089a8 Feb-12       184     132
33 Hardcore   32 #5089a8 Mar-12       138     136
40 Hardcore   39 #5089a8 Apr-12       173     194
47 Hardcore   46 #5089a8 May-12       206     232
54 Hardcore   53 #5089a8 Jun-12       141     132
61 Hardcore   60 #5089a8 Jul-12       128     107
67 Hardcore   66 #5089a8 Aug-12       218     416
73 Hardcore   72 #5089a8 Sep-12       341     498
79 Hardcore   78 #5089a8 Oct-12       363     512
84 Hardcore   83 #5089a8 Nov-12       800     306
89 Hardcore   88 #5089a8 Dec-12         0      28
```

The data about the joining and departing players for this level is plotted as shown below.

```
> chart5 <- gvisColumnChart(hardcore,
+                           xvar="month",
+                           yvar=c("joining","departing"),
+                           options=list(
+                             legend="top",
+                             title="Hardcore Players"))
> plot(chart5)
```

The player levels can be arranged in the natural hierarchical order as shown below.

```
order.levels <- c("Hardcore",
                  "Forum",
                  "Moderate", "Moderate Winners",
                  "Moderate Losers", "Moderate Farmers",
                  "Moderate Miscellanea",
                  "Casual", "Casual Winners",
                  "Casual Losers", "Casual Forum")
```

The months associated with the data are shown below.

```
> months <- as.character(months)
> months
 [1] "Nov-11" "Dec-11" "Jan-12" "Feb-12" "Mar-12" "Apr-12"
 [7] "May-12" "Jun-12" "Jul-12" "Aug-12" "Sep-12" "Oct-12"
[13] "Nov-12" "Dec-12"
```

From the node information, the total number players in each level for each month can now be calculated. The following matrix has a row for each month and a column for each player level. The column names are organized from the lowest level to the highest level.

```
> totals <- matrix(NA,
+                  nrow=length(months),
+                  ncol=length(order.levels))
>
> colnames(totals) <-
+   order.levels[length(order.levels):1]
> rownames(totals) <- months
```

A portion of the matrix initialized with NA values is shown below.

```
> totals[1:5,c(1:2, 10:11)]
       Casual Forum Casual Losers Forum Hardcore
Nov-11             NA             NA          NA
Dec-11             NA             NA          NA
Jan-12             NA             NA          NA
Feb-12             NA             NA          NA
Mar-12             NA             NA          NA
```

The player data for the first month is computed as follows. The nodes associated with the month are filtered as shown below.

```
> m <- months[1]
> m
[1] "Nov-11"
>
> m.nodes <- nodes.info[nodes.info$month == m, "node"]
> m.nodes
[1] 0 1 2 3 4 5
```

The links information for the first month's nodes showing the number of players migrating to other levels in the subsequent month is filtered as shown below. The first three and the last three rows of the data are also shown.

```
> m.links <- links.info[
+               links.info$source %in% m.nodes,
+               c("source", "value")]
```

```
> head(m.links, n = 3)
  source value
1      3   112
2      3   249
3      3    93
> tail(m.links, n = 3)
   source value
34      2    31
35      2    88
36      2   194
```

The data is aggregated for the source nodes for the first month as shown below.

```
> df <- aggregate(m.links$value,
+                 by=list(m.links$source),
+                 FUN = sum)
> df
  Group.1    x
1       0  605
2       1  886
3       2  714
4       3  846
5       4  132
6       5 1390
```

The column names for the data frame are set as node and Total showing the total players for those nodes.

```
> names(df) <- c("node", "Total")
> df
  node Total
1    0   605
2    1   886
3    2   714
4    3   846
5    4   132
6    5  1390
```

The player level names can be added to the above data frame as shown below.

```
> df$name <-
+   nodes.info[df$node+1, "name"]
> df
  node Total              name
1    0   605    Casual Losers
2    1   886   Casual Winners
3    2   714 Moderate Farmers
4    3   846         Moderate
5    4   132            Forum
6    5  1390         Hardcore
```

The total number of players for the first month is shown below. The totals matrix is also populated for the first month.

```
> print(sum(df$Total))
[1] 4573
>
> totals[m,as.character(df$name)] <- df$Total
>
> print(cbind(totals[m,]))
                        [,1]
Casual Forum              NA
Casual Losers            605
Casual Winners           886
Casual                    NA
Moderate Miscellanea      NA
Moderate Farmers         714
Moderate Losers           NA
Moderate Winners          NA
Moderate                 846
Forum                    132
Hardcore                1390
```

For the middle months, since new players are joining and existing players departing, the number of players is calculated look at both the source and target links. The following code shows the player calculations for the second month, Dec 2011.

```
> print(m)
[1] "Dec-11"
>   m.nodes <- nodes.info[nodes.info$month == m, "node"]
>   m.nodes
[1]  6  7  8  9 10 11
```

The number of players migrating to each of the nodes in this month from the previous month is calculated from the links information target column.

```
> m.targetlinks <-
+     links.info[links.info$target %in% m.nodes,
+                c("target", "value")]
>   head(m.targetlinks, n = 3)
  target value
1      9   112
2      8   249
3      6    93
```

Similarly, the source column in the links information show the migration of players of the current month nodes.

```
> m.sourcelinks <-
+     links.info[links.info$source %in% m.nodes,
+                c("source", "value")]
>   head(m.sourcelinks, n = 3)
   source value
37      9   136
38      9    13
39      9    71
```

The number of players coming to the nodes for this month is aggregated using the target values into a new data frame, df.

```
> df.targetlinks <-
+     aggregate(m.targetlinks$value,
+               by=list(m.targetlinks$target),
+               FUN = sum)
>
>   names(df.targetlinks) <-
+     c("node", "TargetTotal")
>
>   df.targetlinks
  node TargetTotal
1    6         520
2    7         640
3    8         883
4    9         767
5   10         303
6   11        1460
```

Similarly, the number of players migrating from the nodes for this month is aggregated using the source values using the same data frame, df.

```
> df.sourcelinks <-
+     aggregate(m.sourcelinks$value,
+               by=list(m.sourcelinks$source),
+               FUN = sum)
>
>   names(df.sourcelinks) <-
+     c("node", "SourceTotal")
>
>   df.sourcelinks
  node SourceTotal
1    6         467
2    7         505
3    8         813
4    9         580
5   10         362
6   11        1302
```

The data frame now captures the total number of players for each node in this month, having the nodes as target, and having the nodes as source.

```
>   df
  node TargetTotal SourceTotal
1    6         520         467
2    7         640         505
3    8         883         813
4    9         767         580
5   10         303         362
6   11        1460        1302
```

The nodes information for this month shows the departing and joining player counts.

```
> m.info <- nodes.info[nodes.info$month == m,]
>   m.info
                name node   color  month departing joining
7     Casual Losers    6 #95d5af Dec-11       912     859
8    Casual Winners    7 #53b67d Dec-11      1247    1112
9  Moderate Farmers    8 #a898b6 Dec-11       535     465
10         Moderate    9 #332341 Dec-11       765     578
11            Forum   10 #f3bd4e Dec-11        30      89
12         Hardcore   11 #5089a8 Dec-11       541     383
```

It can be verified that the source totals for the nodes is the same as target totals coming in from the previous month less the number of players departing each level and adding the number of players joining each level.

```
>   df$TargetTotal - m.info$departing + m.info$joining
[1]  467  505  813  580  362 1302
> df$SourceTotal
[1]  467  505  813  580  362 1302
```

Hence, the total player count for this month for each level is the maximum of the source and target totals for each level.

```
> df$Total <- apply(df[-1], 1, max)
>   df
  node TargetTotal SourceTotal Total
1    6         520         467   520
2    7         640         505   640
3    8         883         813   883
4    9         767         580   767
5   10         303         362   362
6   11        1460        1302  1460
```

The total number of players for the second month is shown below. The totals matrix is also populated for the second month.

```
> print(sum(df$Total))
[1] 4632
>
>   totals[m,as.character(df$name)] <- df$Total
>
>   print(cbind(totals[m,]))
                      [,1]
Casual Forum            NA
Casual Losers          520
Casual Winners         640
Casual                  NA
Moderate Miscellanea    NA
Moderate Farmers       883
Moderate Losers         NA
Moderate Winners        NA
Moderate               767
Forum                  362
Hardcore              1460
```

The other month totals can similarly be calculated. For the last month, only the target values exist.

```
> print(m)
[1] "Dec-12"
>
>   m.nodes <- nodes.info[nodes.info$month == m, "node"]
>   m.nodes
[1] 84 85 86 87 88
```

The links information for the last month nodes is filtered as shown below.

```
> m.links <- links.info[
+                links.info$target %in% m.nodes,
+                c("target", "value")]
>
>   head(m.links, n = 3)
    target value
517     84    27
518     86    33
519     88    10
```

The total player counts for each level for the last month are aggregated as shown below.

```
> df <- aggregate(m.links$value,
+                   by=list(m.links$target),
+                   FUN = sum)
>   names(df) <- c("node", "Total")
>   df
  node Total
1   84   186
2   85   153
3   86   152
4   87    29
5   88   203
```

The player levels for these nodes are also added as shown below.

```
> df$name <- nodes.info[df$node+1, "name"]
>   df
  node Total                    name
1   84   186                 Casual
2   85   153 Moderate Miscellanea
3   86   152     Moderate Farmers
4   87    29                 Forum
5   88   203              Hardcore
```

The total number of players for the last month is shown below. The totals matrix is also populated for the last month.

```
> print(sum(df$Total))
[1] 723
>
>   totals[m,as.character(df$name)] <- df$Total
>
>   print(cbind(totals[m,]))
                      [,1]
Casual Forum            NA
Casual Losers           NA
Casual Winners          NA
Casual                 186
Moderate Miscellanea   153
Moderate Farmers       152
Moderate Losers         NA
Moderate Winners        NA
Moderate                NA
Forum                   29
Hardcore               203
```

Now that the entire monthly totals for each player levels are calculated, a data frame is created with the month and the player totals. A portion of the data frame is displayed below.

```
> df <- data.frame(Month = months, totals)
>
> df[1:5, c(2:3, 11:12)]
       Casual.Forum Casual.Losers Forum Hardcore
Nov-11           NA           605   132     1390
Dec-11           NA           520   362     1460
Jan-12           NA           499   244     1024
Feb-12           91           255   131      740
Mar-12           86           237   105      587
```

The colors associated with the player levels are composed using a data frame as shown below.

```
> nodes.colors <-
+   data.frame(node=levels(nodes.info$name),
+              color=levels(nodes.info$color))
>
> rownames(nodes.colors) <- nodes.colors$node
>
> nodes.colors
                                        node   color
Casual Losers                 Casual Losers #95d5af
Casual Winners               Casual Winners #53b67d
Moderate Farmers         Moderate Farmers #a898b6
Moderate                             Moderate #332341
Forum                                   Forum #f3bd4e
Hardcore                             Hardcore #5089a8
Moderate Miscellanea Moderate Miscellanea #c7bfce
Moderate Losers           Moderate Losers #806b91
Casual Forum                 Casual Forum #c7f3d8
Moderate Winners         Moderate Winners #5d4a6c
Casual                                 Casual #398b5c
```

The colors are arranged as per the player levels using the indexing as shown below.

```
> color <- nodes.colors[
+   order.levels[length(order.levels):1],
+   "color"]
>
> color <- as.character(color)
> color
 [1] "#c7f3d8" "#95d5af" "#53b67d" "#398b5c" "#c7bfce"
 [6] "#a898b6" "#806b91" "#5d4a6c" "#332341" "#f3bd4e"
[11] "#5089a8"
```

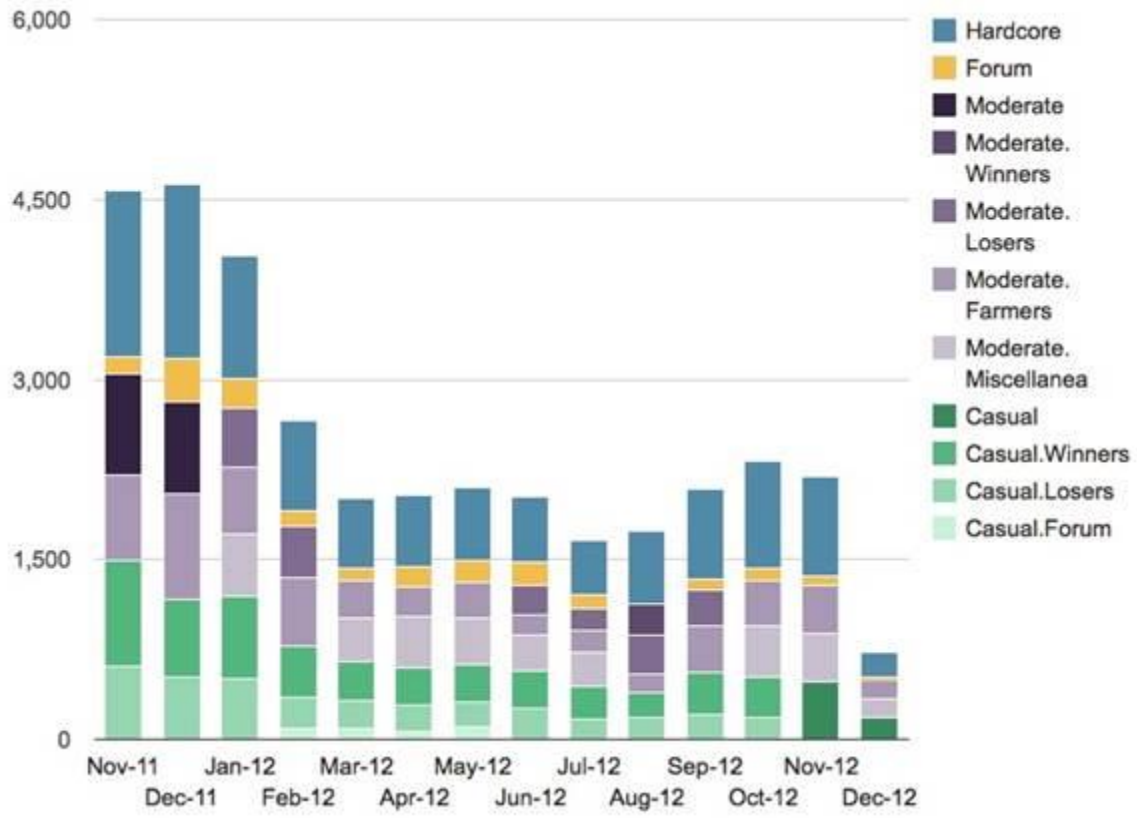
The information about the colors is pasted into a single string as required for the Google charts.

```
> color <- paste0('"', color, '"')
> color <- paste0(color, collapse=",")
> color <- paste0('[', color, ']')
>
> color
[1] "[\"#c7f3d8\",\"#95d5af\",\"#53b67d\",\"#398b5c\",\"#c
7bfce\",\"#a898b6\",\"#806b91\",\"#5d4a6c\",\"#332341\",\"
#f3bd4e\",\"#5089a8\"]"
```

The stacked column chart showing the information for the various months and the number of players in each level for those months is plotted using the gvisColumnChart function.

```
> chart6 <-
+   gvisColumnChart(
+     df,
+     options=list(
+       height=750, width=850,
+       isStacked = TRUE,
+       colors=color))
> plot(chart6)
```

The stacked column chart is shown below.



References

Glitch: Analyzing the MMO's Final Year of Activity http://powerful-meadow-8588.herokuapp.com/

Glitch, http://www.glitchthegame.com/

Using Google Charts https://developers.google.com/chart/interactive/docs/

Package 'googleVis': R Interface for Google Charts http://cran.r-project.org/web/packages/googleVis/googleVis.pdf