Realtime Visualization & Analytics [Data Track] By Naveen Rajan & Vikas Manjunath

Contents

[Summary 2](#_Toc432872559)

[Assumptions 2](#_Toc432872560)

[Flow 3](#_Toc432872561)

[Sample Data Structures 3](#_Toc432872562)

[Retension 3](#_Toc432872563)

[Funnel 3](#_Toc432872564)

[Sample **Working** Kibana Dashboard for the above scenarios: 4](#_Toc432872565)

[Visualization Details 4](#_Toc432872566)

[Similarly, other Visualizations like ARPU, ARPPU, Avg Session Lenghts & Number of Sessions can be added to the dashboard. 5](#_Toc432872567)

[Clustering Users By Their Behaviour 5](#_Toc432872568)

[To Do 5](#_Toc432872569)

# Summary

Data including playtimes can give vital insights into player engagement. This can be used to for various other downstream activities like:

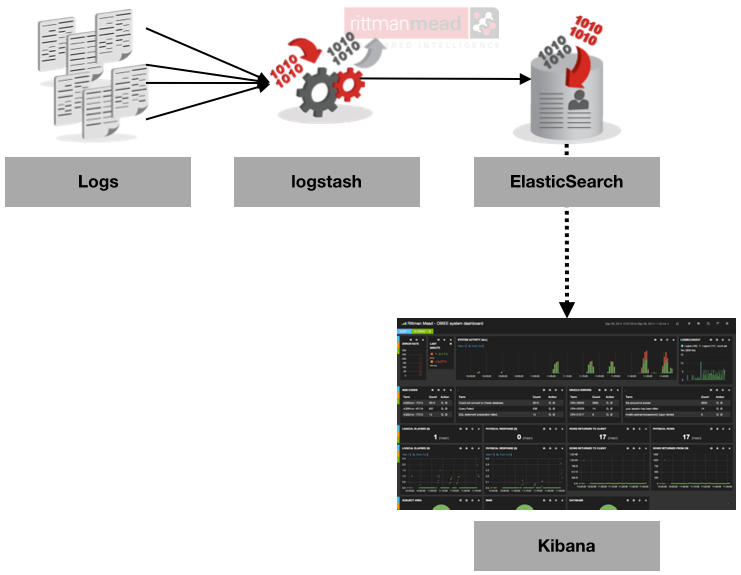
* Fixing issues in game/infrastructure
* Predicting workload
* A/B testing of new features etc

Further, clustering algorithms can be used to bucket players into various cluster & address their needs in a better manner.

# Assumptions

* Scale is very important for large MMP games.
* The ELK (Elasticsearch-Logstash-Kibana) framework is ideal.
* Data reaches Elastic after passing thru a data cleaning tier. This tier is missing in this implementation.

# Flow



# Sample Data Structures

The JSON data structures were idebtified for a few sample Visualization use cases:

## Retension

Shows the total Installed & Dropped user counts at any specific time interval. Sample JSON: *{"id":"1","datestamp":"2015-10-01","installed":100.0,"dropped":0.0}*

Sample data file:



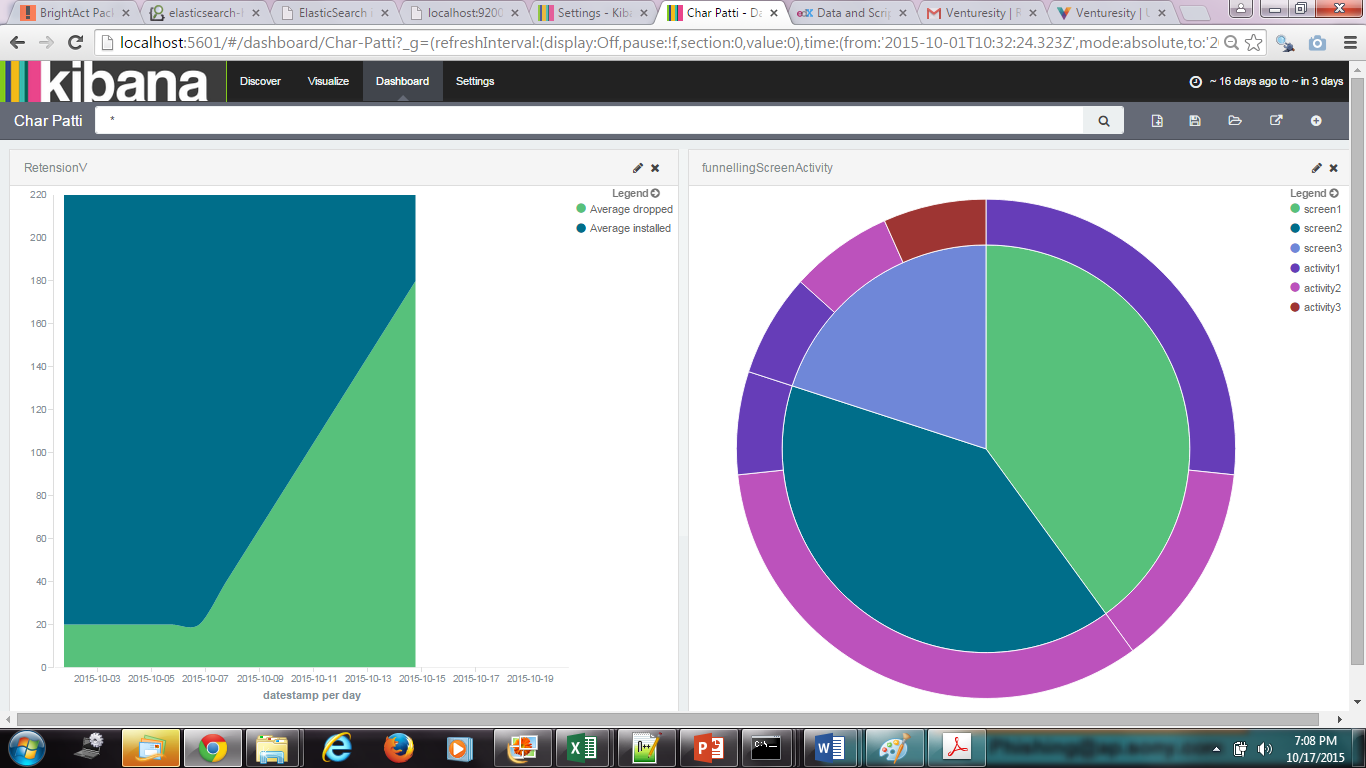
## Funnel

Shows the user drop out at various Screens & Activities inside the screens. This helps in tuning the app. Sample JSON: *{"index":{"\_index":"funnel","\_type":"funnel","\_id":"F3"}}*

*Sample data file:*

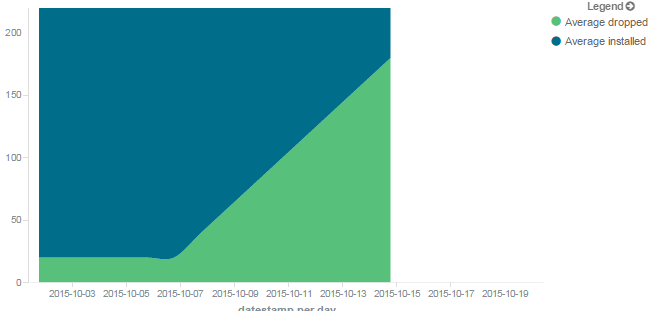


# Sample **Working** Kibana Dashboard for the above scenarios:

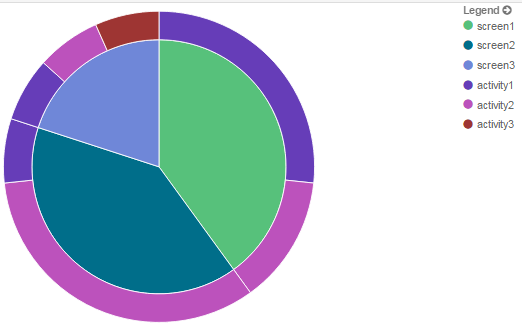


## Visualization Details

The Retension data is shown in an Area graph. The graph shows a linear increase users abandoning the game.



The Funnel showing abandonment by Screen & Activity uses a Layered Pie Chart to show the 2 dimensions.



## Similarly, other Visualizations like ARPU, ARPPU, Avg Session Lenghts & Number of Sessions can be added to the dashboard.

# Clustering Users By Their Behaviour

Hierarching Clustering can be used to classify the users based on their behaviour. Further, the same program can be used across problem statements as this is an unsupervised learning problem. Sample R code for Hierarchial Clustering is attached.



# To Do

*Watcher* is an interesting new sister product to Elasticsearch that is specifically targetted at alerts. Use this product to trigger notifications based on certain anamolies.