DSC : Réseau WiFi de Dartmouth College

Antonin Barthélémy

6 octobre 2017

Plan

1. Raffinement

2. Transformation

3. Analyse

Raffinement ...

Présentation des données

1013576401 Feb 13 00:00:01 LibBldg2AP3 LibBldg2AP3 (Info):
Deauthenticating 00022d6f711f, reason "Must Authenticate Before Associating"

1013576401 Feb 13 00:00:01 LibBldg2AP3 LibBldg2AP3 (Info):
Station 00022d6f711f Authenticated

1013576401 Feb 13 00:00:01 LibBldg2AP3 LibBldg2AP3 (Info):
Station 00022d6f711f Reassociated

1013576401 Feb 13 00:00:01 LibBldg2AP14 LibBldg2AP14 (Info):
Station 00022d6f711f roamed

1013576403 Feb 13 00:00:03 ResBldg36AP1 ResBldg36AP1 (Info):
Deauthenticating 003065d0c649, reason "Inactivity"

Données fournit par le premier serveur

Présentation des données

1088568008 Jun 30 00:00:08 ResBldg44AP4 3171: ResBldg44AP4 Jun 30 04:00:07: %DOT11-6-DISASSOC:
Interface Dot11Radio0, Deauthenticating Station 00904b86f12a Reason:
Disassociated because sending station is leaving (or has left) BSS

1088568012 Jun 30 00:00:12 SocBldg3AP2 10190: SocBldg3AP2 Jun 30 04:00:12: %DOT11-4-MAXRETRIES:
Packet to client 00022dd9b5b2 reached max retries, removing the client

1088568013 Jun 30 00:00:13 SocBldg3AP2 10191: SocBldg3AP2 Jun 30 04:00:12: %DOT11-6-DISASSOC:
Interface Dot11Radio0, Deauthenticating Station 00022dd9b5b2 Reason:
Previous authentication no longer valid

1088568016 Jun 30 00:00:16 ResBldg97AP6 3928: ResBldg97AP6 Jun 30 04:00:15: %DOT11-6-ASSOC:
Interface Dot11Radio0, Station 0009b7f3ff1f Reassociated KEY_MGMT[NONE]

Données fournit par le second serveur

RegExp

^([\d]+) [\w]{3} +[\d]{1,2} [\d:]+ ([a-zA-Z]+[0-9]+)([\w]+) [\w]+ \(Info\): (.*)

1013576401 Feb 13 00:00:01 LibBldg2AP3 LibBldg2AP3 (Info):
Deauthenticating 00022d6f711f, reason "Must Authenticate Before Associating"

1013576401 Feb 13 00:00:01 LibBldg2AP3 LibBldg2AP3 (Info):
Station 00022d6f711f Authenticated

1013576401 Feb 13 00:00:01 LibBldg2AP3 LibBldg2AP3 (Info):
Station 00022d6f711f Reassociated

1013576401 Feb 13 00:00:01 LibBldg2AP14 LibBldg2AP14 (Info):
Station 00022d6f711f roamed

1013576403 Feb 13 00:00:03 ResBldg36AP1 ResBldg36AP1 (Info):
Deauthenticating 003065d0c649, reason "Inactivity"

Premier filtre d’expression régulière

RegExp

Le monstre !!!

```
^([\d]+) [\w]{3} +[\d]{1,2} [\d:]+ ([a-zA-Z]+[0-9]+)([\w]+) [\d]+:
      [\w\.]* ?[\.\*]?[\w]{3} +[\d]{1,2} [\d:\.]+ %[\w]+-[\d]+-[\w]+: (.*)
```

RegExp

```
^([\d]+) [\w]{3} +[\d]{1,2} [\d:]+ ([a-zA-Z]+[0-9]+)([\w]+) [\d]+:
        [\w\.]* ?[\.\*]?[\w]{3} +[\d]{1,2} [\d:\.]+ %[\w]+-[\d]+-[\w]+: (.*)
```

```
1088568008 Jun 30 00:00:08 ResBldg44AP4 3171: ResBldg44AP4 Jun 30 04:00:07: %DOT11-6-DISASSOC:
  Interface Dot11Radio0, Deauthenticating Station 00904b86f12a Reason:
    Disassociated because sending station is leaving (or has left) BSS

1088568012 Jun 30 00:00:12 SocBldg3AP2 10190: SocBldg3AP2 Jun 30 04:00:12: %DOT11-4-MAXRETRIES:
  Packet to client 00022dd9b5b2 reached max retries, removing the client
```

Deuxième filtre d’expression régulière

RegExp

```
regexp_g3_v1 = r'^Station (\w+) (\w+)'
regexp_g3_v2 = r'^(\w+) [from]* ?(\w+), reason'
regexp_g3_v3 = r'^Packet to client (\w+)'
regexp_g3_v4 = r'^Interface \w+, (\w+) Station (\w+)'
regexp_g3_v5 = r'^Interface \w+, Station (\w+) (\w+)'
```

Add_MAC

Action

Dernier filtre d’expression régulière

CSV

Timestamp,	Batiment,	AP,	Mac,	Action
0,	AcadBldg17,	AP3,	004096f9f2e4,	authenticated
145,	AcadBldg17,	AP3,	004096f9f2e4,	authenticated
146,	LibBldg1,	AP1,	003065beae9e,	deauthenticating

Résultat de l’extraction

Transformation !!!

Manipulation des données

Timestamp	Batiment	AP	Mac	Action
0	AcadBldg17	AP3	004096f9f2e4	authenticated
145	AcadBldg17	AP3	004096f9f2e4	authenticated
146	LibBldg1	AP1	003065beae9e	deauthenticating

```
Timestamp,Batiment,AP,Mac,Action
80937,ResBldg82,AP4,0002a55a0817,authenticated
80937,ResBldg82,AP4,0002a55a0817,associated
81063,ResBldg82,AP2,0002a55a0817,deauthenticating
81063,ResBldg82,AP2,0002a55a0817,authenticated
```

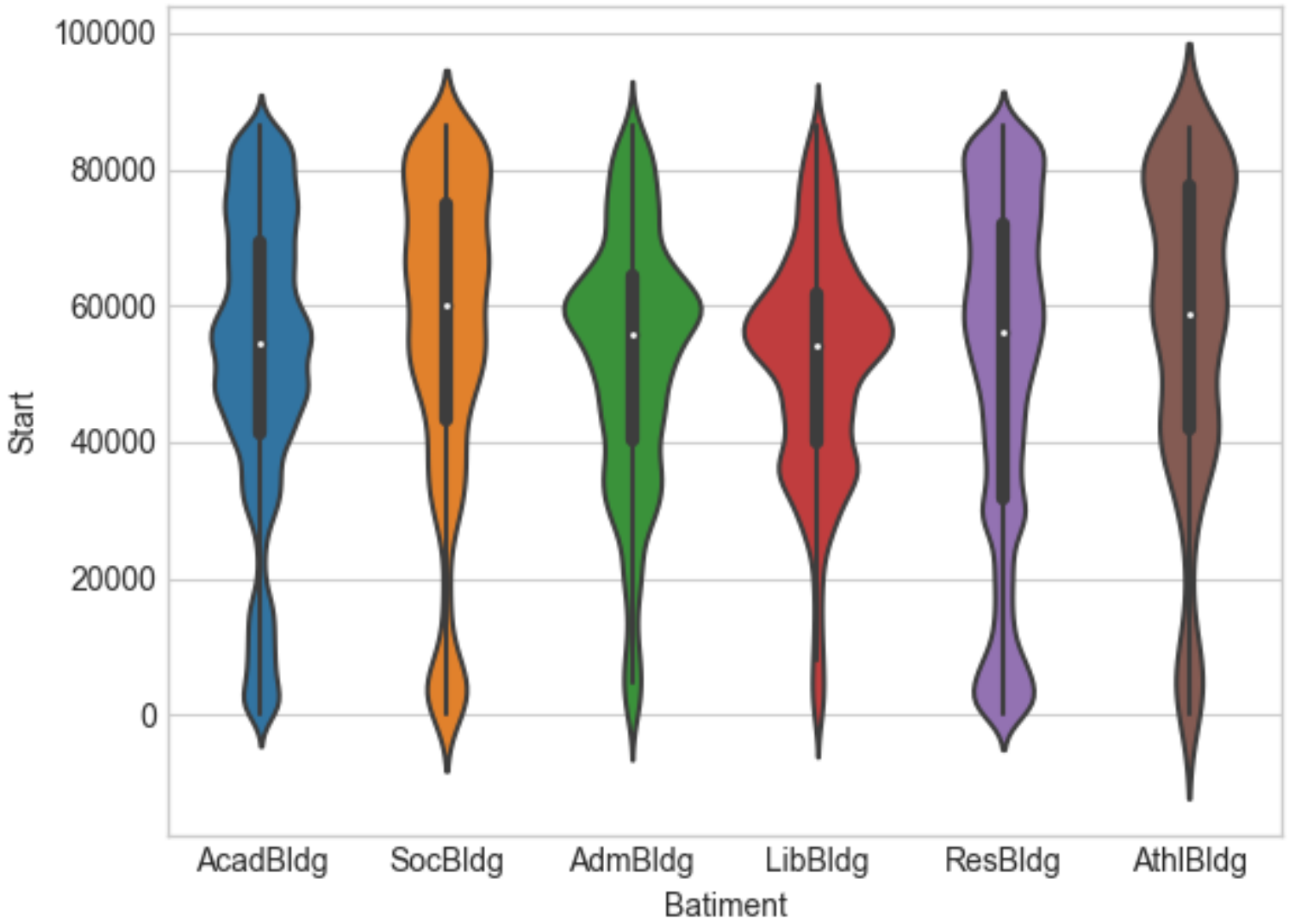
Isoler les utilisateurs

```
Mac_ID,AcadBldg,ResBldg,LibBldg,SocBldg,AdmBldg,OthBldg,AthBldg
00022db6cc78,8104,0,0,0,0,0,0
00306506fde2,0,0,0,0,0,0,0
0030650c0321,18022,0,0,0,0,0,0
```

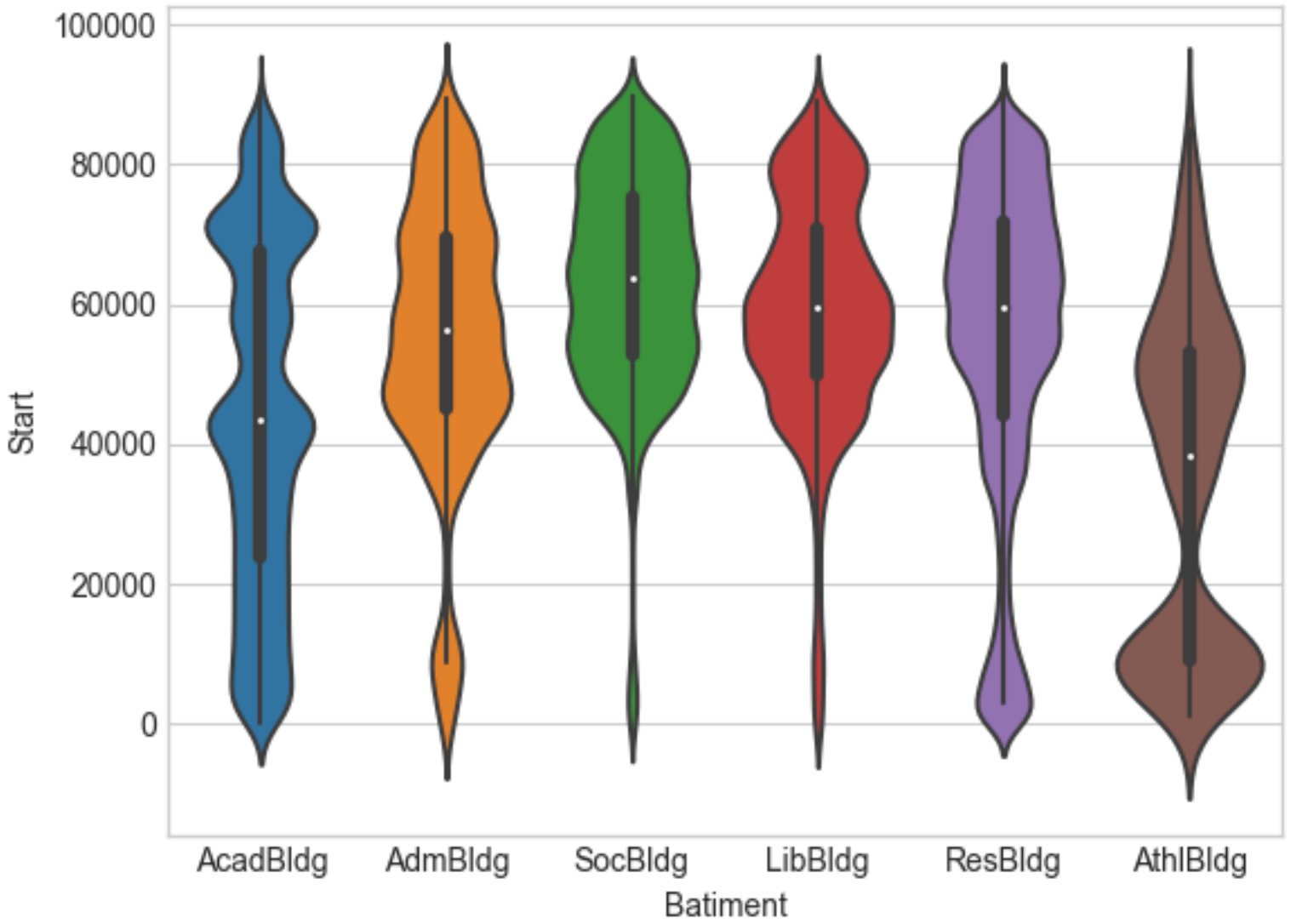
Agréger l'activitée par batiment

Analyse ???

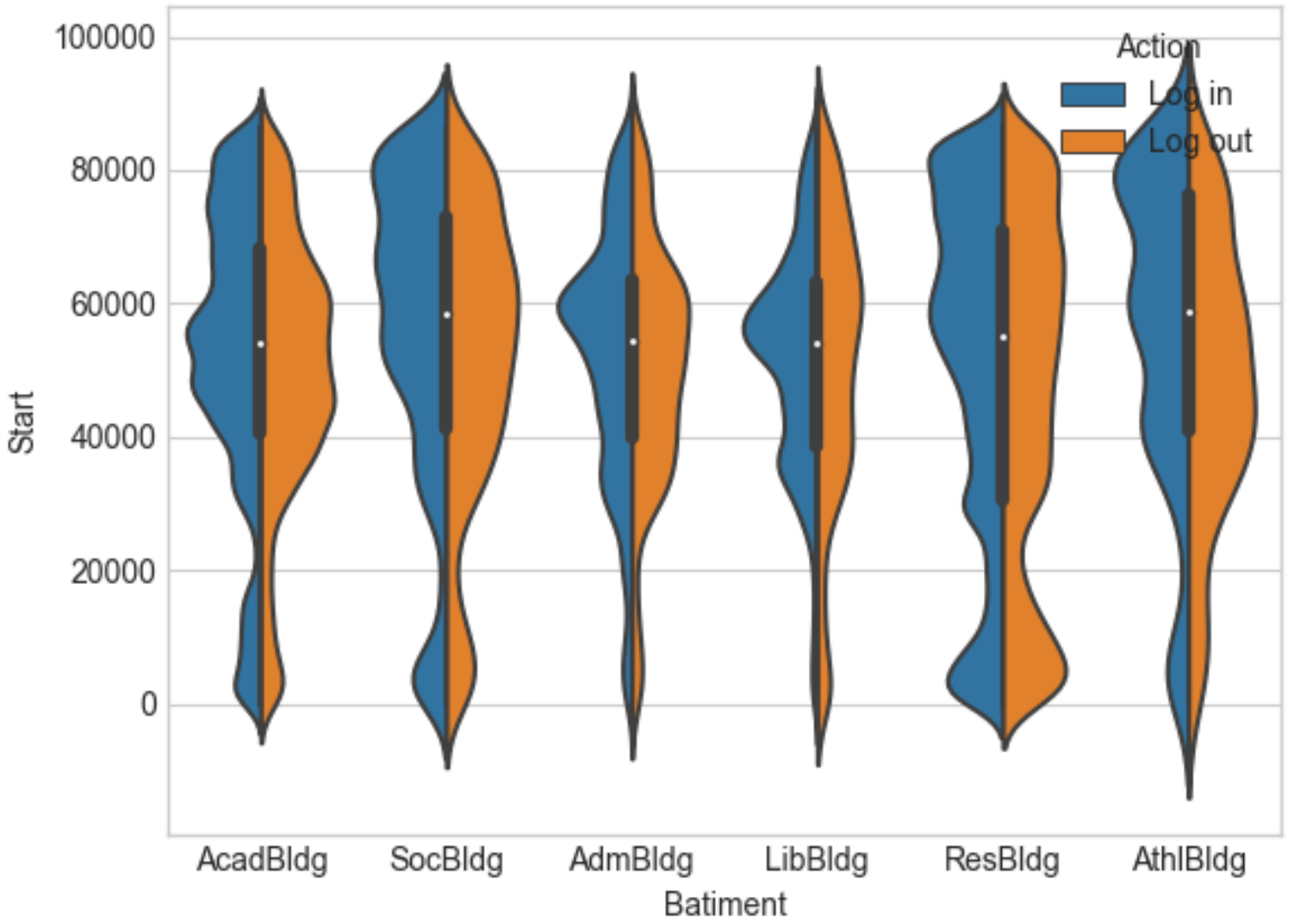
Violinplot : Lundi "Login"



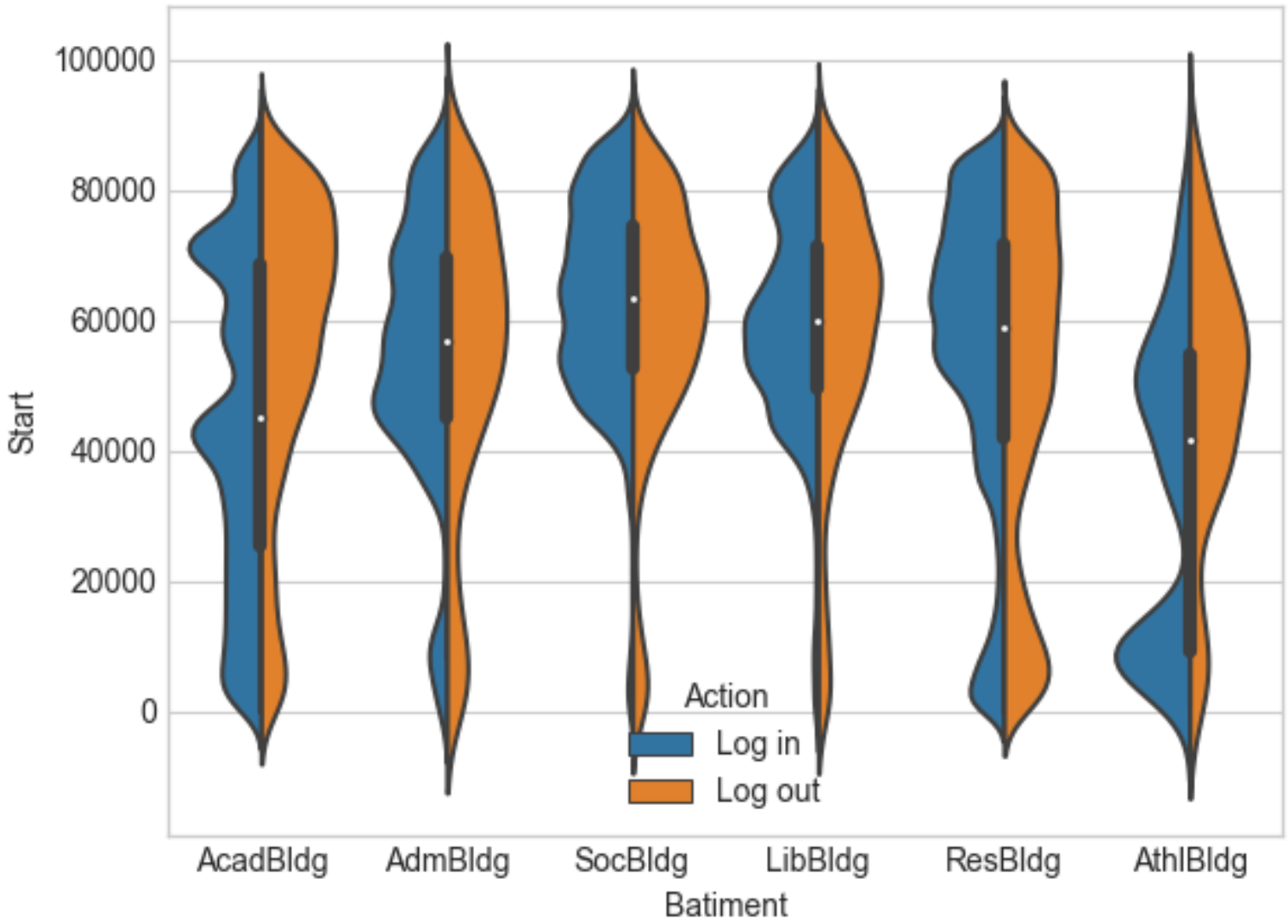
Violinplot : Dimanche "Login"



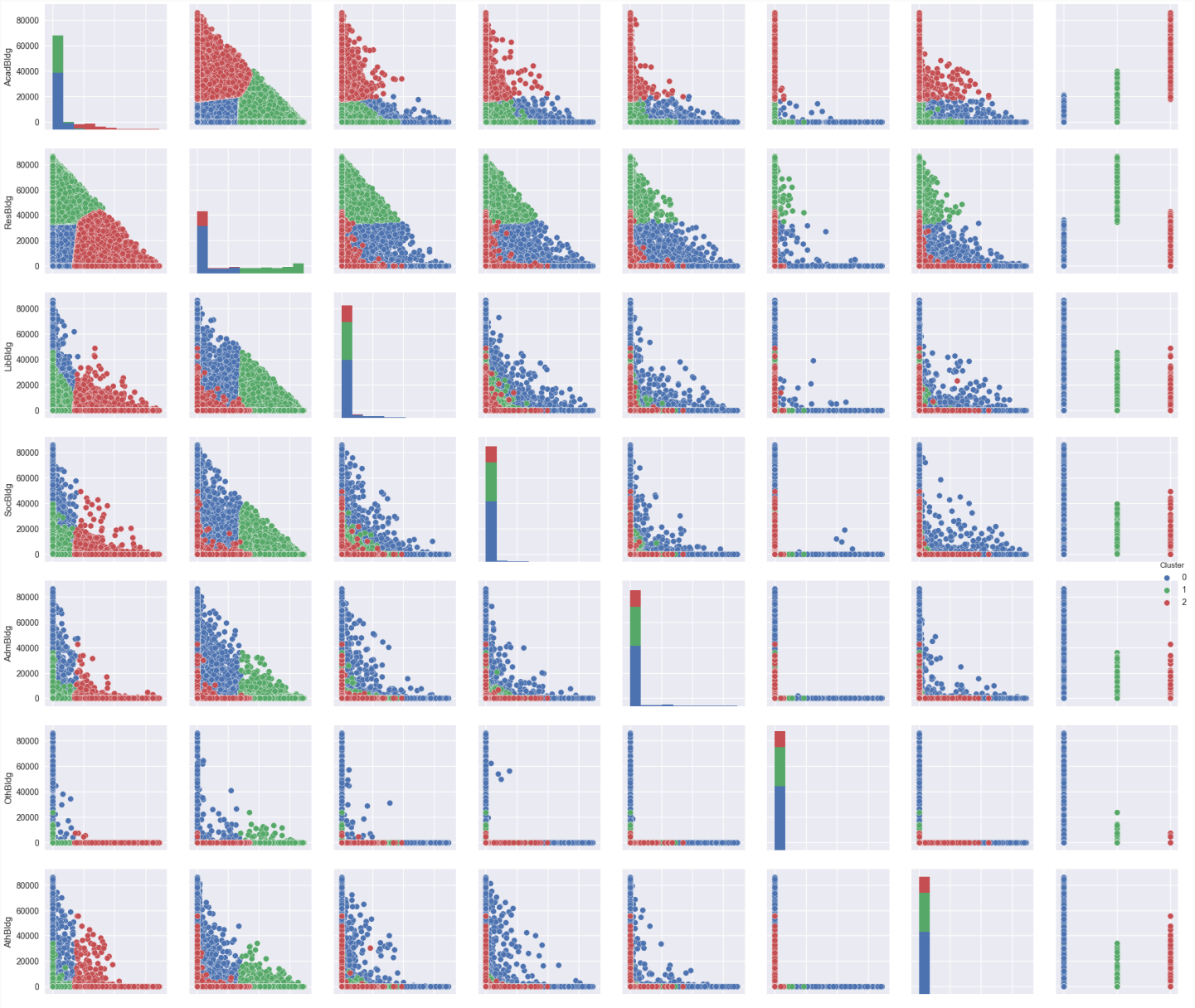
Violinplot : Lundi



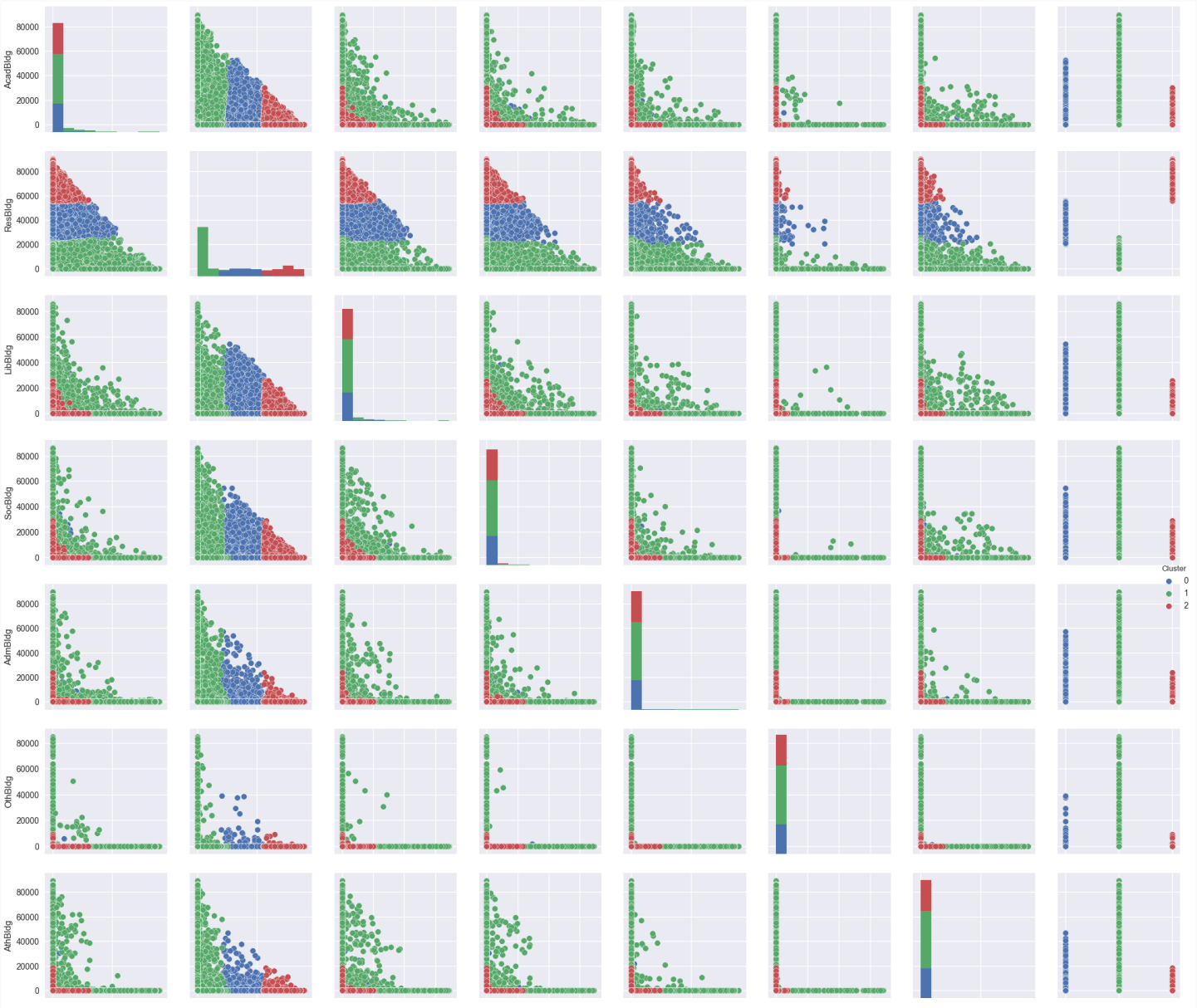
Violinplot : Dimanche



Clustering Lundi : Kmean



Clustering Dimanche : Kmean



La suite ...???