



ΔΙΟΙΚΗΣΗ ΨΗΦΙΑΚΗΣ ΕΠΙΧΕΙΡΗΣΗΣ

ΕΡΓΑΣΙΑ 5 ΑΣΚΗΣΗ ΣΤΟ DISCO

Κουστένης Χρίστος

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ | ΣΗΜΜΥ | Εργαστήριο Διοίκησης Πληροφοριακών
Συστημάτων | Ακ. Έτος 2023-2024



Εισαγωγή

Σκοπός της εργασίας είναι η αξιοποίηση τεχνικών Process Mining με τη χρήση του εργαλείου Disco για την ανακάλυψη και αξιολόγηση επιχειρηματικών διαδικασιών με την ανάλυση συνόλων δεδομένων της μορφής event log που εξάγονται από πληροφοριακά συστήματα. Επιλέξαμε το event log που αντιστοιχεί στη διαδικασία προμηθειών μιας επιχείρησης που και βρίσκεται στο [link](#), με τίτλο BPI Challenge 2017.

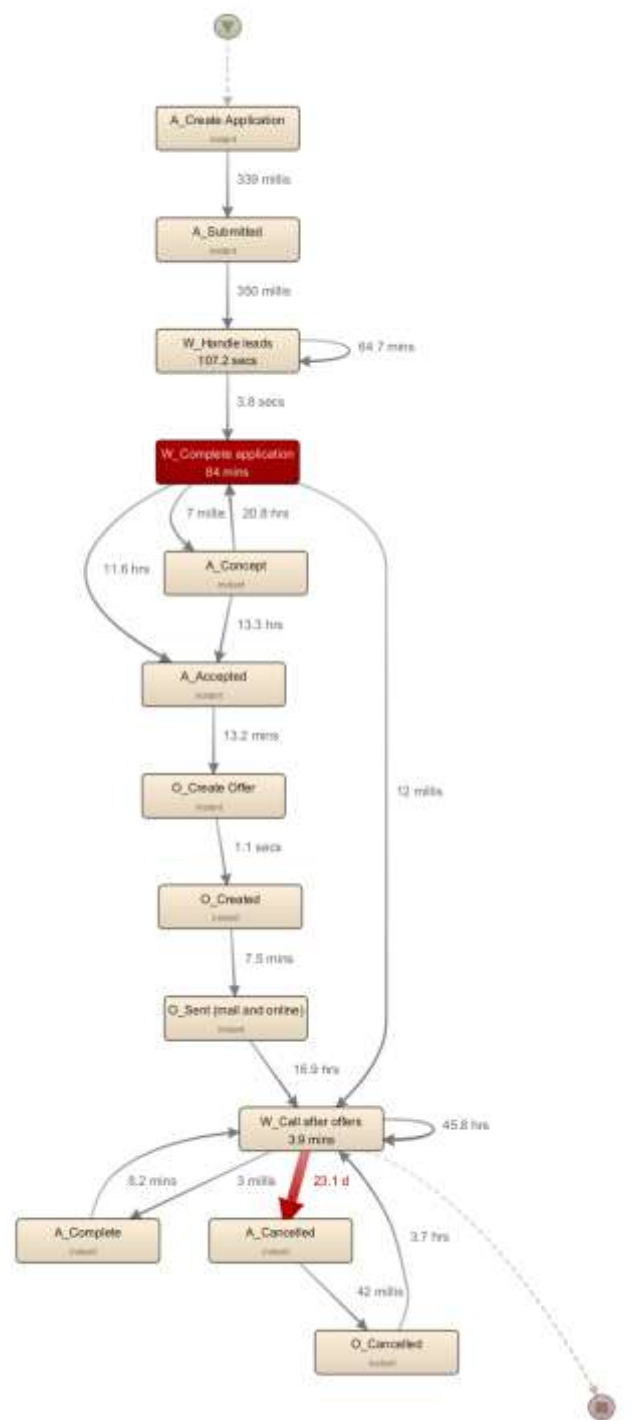
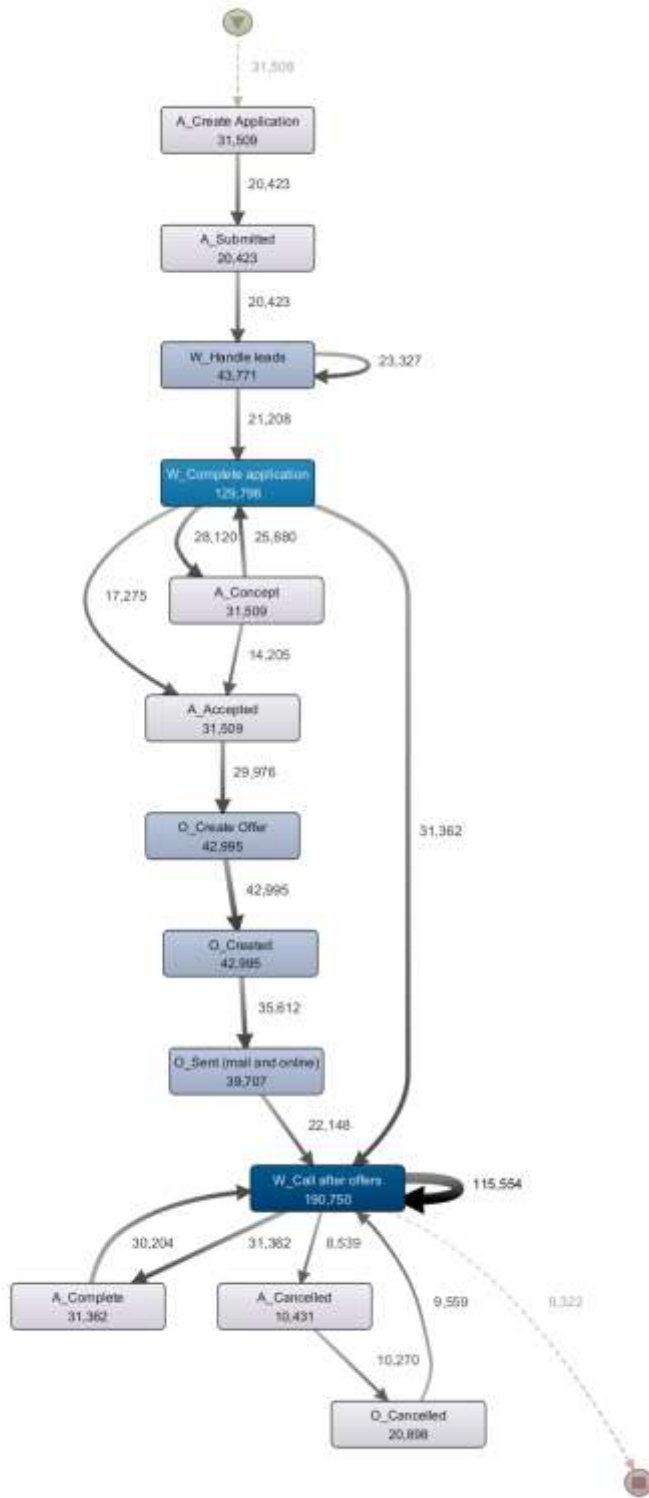
Για να αναλύσουμε τη διαδικασία, θα ακολουθήσουμε τα βήματα 3-9 του tutorial του Disco.

Dataset

Αυτό το αρχείο καταγραφής συμβάντων αφορά μια διαδικασία αίτησης δανείου ενός ολλανδικού χρηματοπιστωτικού ιδρύματος. Τα δεδομένα περιέχουν όλες τις αιτήσεις που κατατέθηκαν μέσω ενός ηλεκτρονικού συστήματος το 2016 και τις επακόλουθες εκδηλώσεις τους μέχρι την 1η Φεβρουαρίου 2017. Συγκεκριμένα, το σύστημα επιτρέπει πλέον πολλαπλές προσφορές ανά αίτηση δανείου. Αυτές οι προσφορές μπορούν να εντοπιστούν μέσω των αναγνωριστικών τους στο αρχείο καταγραφής.

Inspect Process

Έχοντας εισάγει το αρχείο .xes στο Disco, δημιουργείται αυτόματα ο χάρτης της διαδικασίας. Θέτοντας το Activity στο 0% (με το Paths στο 0% επίσης), παρουσιάζονται οι πιο βασικές και συνηθισμένες δραστηριότητες, με τις συχνότητες (absolute frequency) και με τη μέση διάρκεια τους (mean), όπως φαίνεται και παρακάτω.



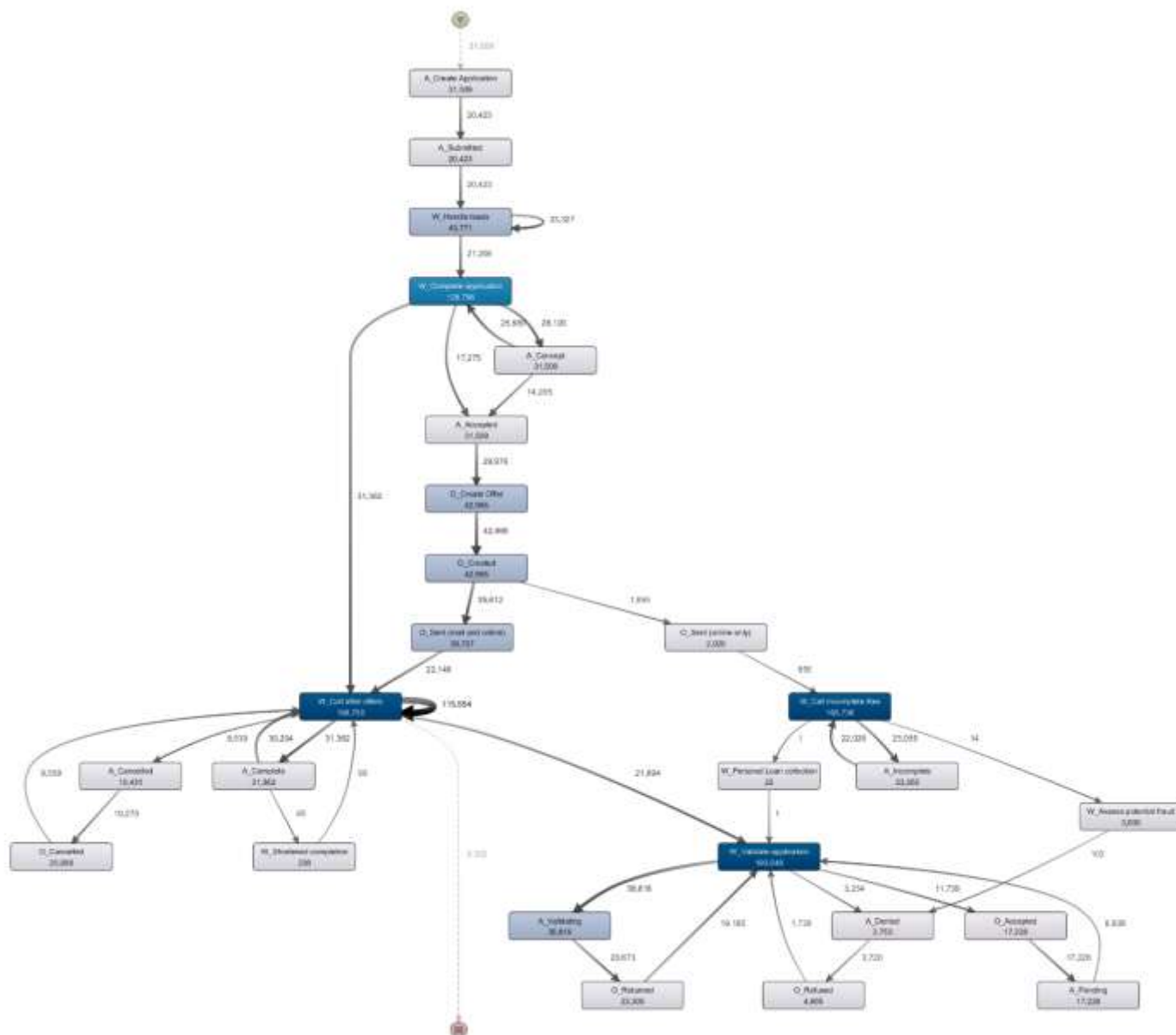


Οι βασικές δραστηριότητες(με τη μεγαλύτερη συχνότητα εμφάνισης σε events) που προκύπτουν είναι οι ακόλουθες:

Activity	Description
A_Create Application	Δημιουργία μιας νέας αίτησης από τον αιτούντα.
A_Submitted	Υποβολή της αίτησης για περαιτέρω επεξεργασία.
W_Handle leads	Διαχείριση των δυνητικών αιτούντων, συμπεριλαμβανομένων επικοινωνίας και αρχικής αξιολόγησης.
W_Complete application	Ολοκλήρωση της αίτησης με την παροχή όλων των απαραίτητων πληροφοριών και εγγράφων.
A_Concept	Εσωτερική αξιολόγηση και προετοιμασία της αίτησης για την επόμενη φάση.
A_Accepted (Αποδοχή)	Αποδοχή της αίτησης από την τράπεζα μετά την αξιολόγηση.
O_Create Offer (Δημιουργία Προσφοράς)	Δημιουργία προσφοράς για τον αιτούντα με βάση την αξιολογηθείσα αίτηση.
O_Created (Δημιουργία)	Εγγραφή της προσφοράς στο σύστημα της τράπεζας.
O_Sent (mail and online)	Αποστολή της προσφοράς στον αιτούντα μέσω ταχυδρομείου ή ηλεκτρονικά.
W_Call after offers	Αναμονή τηλεφωνήματος από τον πελάτη ως απόκριση της αποσταλείσας προσφοράς
A_Complete	Ολοκλήρωση της διαδικασίας αίτησης.
A_Cancelled	Ακύρωση της αίτησης από τον αιτούντα ή την τράπεζα.
O_Cancelled	Ακύρωση της προσφοράς που είχε δημιουργηθεί για τον αιτούντα.

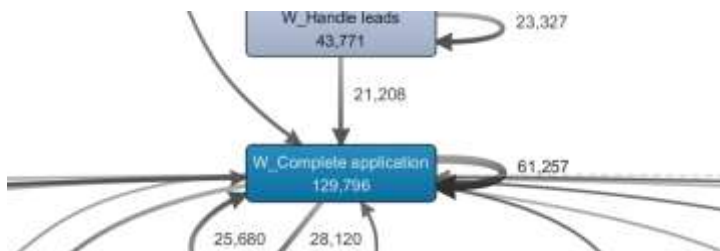
Όπως φαίνεται από το προηγούμενο διαγράμματα με Activities 0% και Paths 0% οι δραστηριότητες με τη μεγαλύτερη συχνότητα εμφάνισης είναι οι *W_Complete application* και *W_Call after offers*.

Ανεβάζοντας το Activity στο 100% για να εμφανιστούν όλες οι δραστηριότητες(26 συνολικά), παίρνουμε το παρακάτω αποτέλεσμα στο οποίο φαίνεται ότι εκτός από τις προαναφερθείσες σημαντική συχνότητα εμφανίζουν και τα activities *W_Call incomplete files* και *W_Validate application*:



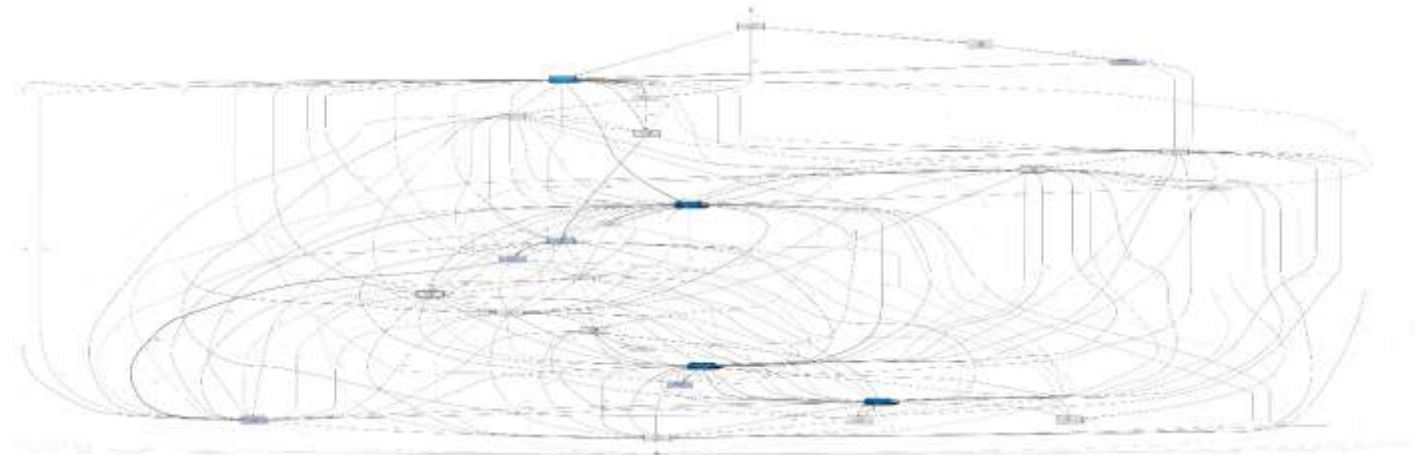
Περιγραφή κύριας διαδικασίας

Η διαδικασία ξεκινά με τη δημιουργία μιας νέας αίτησης από τον ενδιαφερόμενο. Σε αυτό το σημείο, δημιουργούνται 31.509 αιτήσεις. Από τις αιτήσεις που δημιουργήθηκαν 20.423 υποβάλλονται για περαιτέρω επεξεργασία. Αυτή η μείωση υποδηλώνει ότι δεν ολοκληρώνονται για υποβολή όλες οι αιτήσεις που δημιουργούνται. Οι υποβληθείσες αιτήσεις περνούν στη διαχείριση προοπτικών, όπου η τράπεζα επικοινωνεί/εξετάζει το προφίλ των αιτούντων ή επικοινωνεί μαζί τους για περαιτέρω πληροφορίες ή διευκρινίσεις. Υπάρχουν 43.771 επαφές σε αυτό το στάδιο, με 23.327 επαναλήψεις (loops), που υποδηλώνουν ότι πολλές αιτήσεις χρειάζονται επανεπεξεργασία. Στη συνέχεια προχωράμε στο στάδιο αναμονής της για ολοκλήρωση αίτησης στο οποίο έχουμε 129.796 εκτελέσεις που είναι σημαντικά περισσότερες από αυτές που προέρχονται από το activity W_handle_leads. Ρυθμίζοντας το path στο 2% παρατηρούμε ότι μεγάλο μέρος του αριθμού εκτέλεσης τους συγκεκριμένου activity οφείλεται στην επανατροφοδότηση από παρατεταμένη αναμονή για συμπλήρωση αίτησης (η οποία εισάγει καθυστέρηση 7.2 hrs κατά μέσο όρο όπως φαίνεται στην παρακάτω εικόνα.)



Στη συνέχεια 31.509 applications γίνονται αποδεκτές και έπειτα έχουμε ροή προς τα activities O_Create Offer (Δημιουργία Προσφοράς) --> O_Created (Δημιουργία με μηδανικές απώλειες καθώς φαίνεται πως τα περισσότερα αποδεκτά application οδηγούν σε αυτήν τη διαδρομή. Ακόμη, παρατηρούμε πως κάποια offers στέλνονται μόνο ηλεκτρονικά(2.026) ενώ τα περισσότερα στέλνονται και ταχυδρομικά και ηλεκτρονικά(39.707). Εν τέλει, όλες οι αιτήσεις θα περάσουν από τα 3 activities με το μεγαλύτερο frequency(W_Call after offers, W_Call incomplete files και W_Validate application) πριν γίνουν αποδεκτές ή απορριφθούν.

Για λόγους πληρότητας θα συμπεριλάβουμε και το πλήρες διάγραμμα του process με Activities 100% και Paths 100% το οποίο όπως ήταν αναμενόμενο ήταν αρκετά πολύπλοκο.



Αυξάνοντας το path percentage με σχετικά μικρό activity percentage παρατηρούμε διακεκκόμενες γραμμές με κόκκινο χρώμα οι οποίες δεν πηγάζουν μόνο από την W_Call after offers αλλά και από άλλα activities καθώς δεν τερματίζουν όλα τα variants με το ίδιο Activity.





Συμπεράσματα

- Το activity W_complete application αφορά τον πελάτη οπότε η τράπεζα για να επιταχύνει τη διαδικασία πρέπει να καταστήσει απλούστερη και ευκολότερη τη συμπλήρωση της αίτησης.
- Τα activities O_Create Offer και O_Created ίσως μπορούν να ενοποιηθούν σε ένα activity αφού ακόμα και το αναλυτικότερα διάγραμμα δείχνει ότι είναι στενά συνδεδεμένα και όσα applications περνούν από το πρώτο οδηγούνται όλα στο 2ο.

Inspect Statistics

Overview

Events	1,160,405
Cases	31,509
Activities	26
Median case duration	19.1 d
Mean case duration	21.9 d
Start	01/01/2016 11:51:15
End	02/01/2017 16:11:03

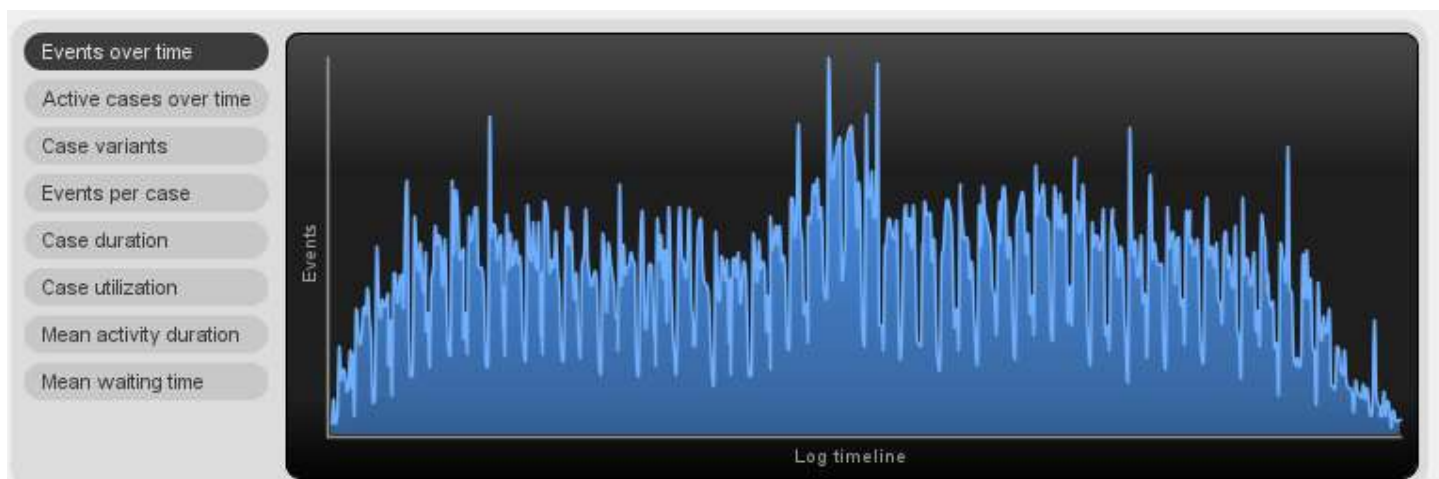
Παρατηρώντας τα στατιστικά, βλέπουμε ότι έχουμε συνολικά 1.160.405 events με 31.509 cases. Πράγματι επιβεβαιώνεται από το “Start” και το “End” ότι τα δεδομένα αυτά προέρχονται από την αρχή του 2016 έως τους πρώτους μήνες του 2017. Επίσης, ο μέσος χρόνος ολοκλήρωσης μίας χορήγησης δανείου είναι 21.9 μέρες ενώ η διάμεσος του συνόλου ανέρχεται στις 19.1 μέρες. Άρα χρειάζονται γύρω στις 3 εβδομάδες για την ολοκλήρωση της προναφερθείσας συναλλαγής.

Από το παρακάτω διάγραμμα (Events Over Time) παρατηρούμε μία άνοδο στη συχνότητα των events στο μέσο του έτους(Ιούλιο) γεγονός που μπορεί να οφείλεται σε 3 παράγοντες :

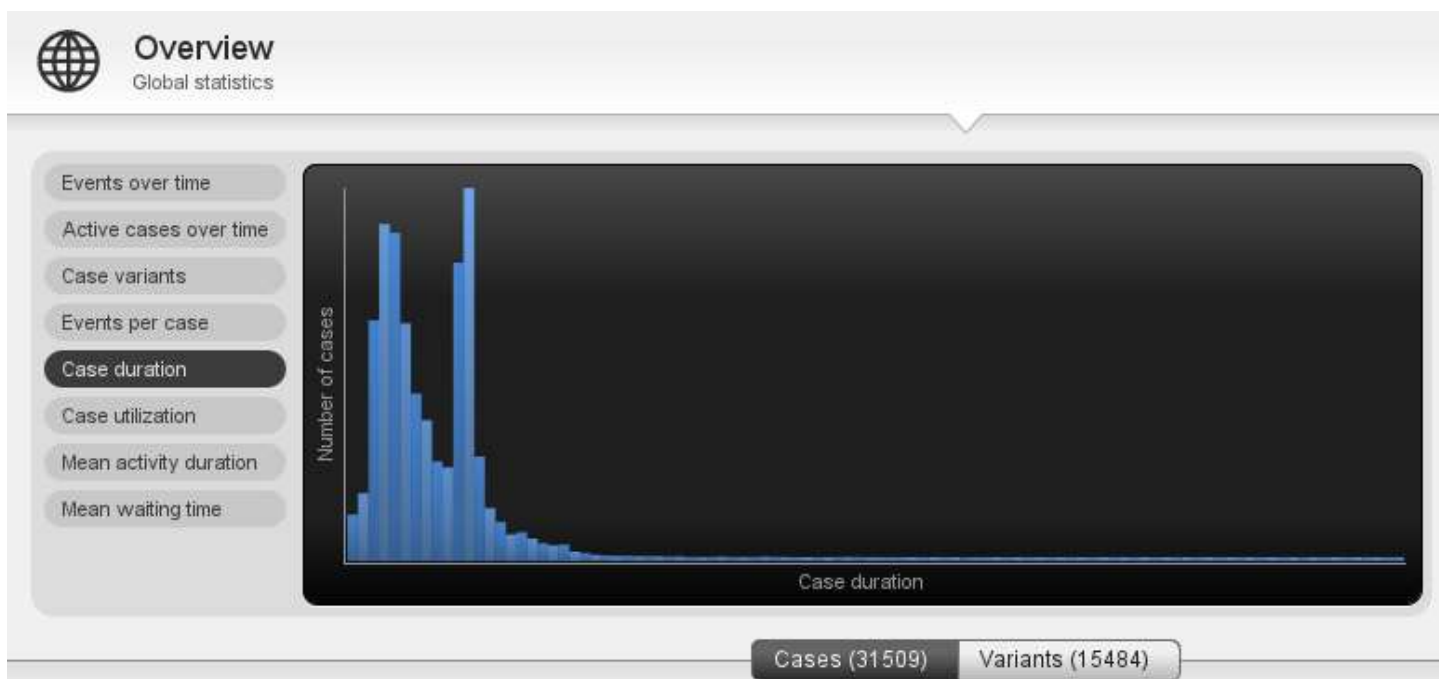
- Καλοκαιρινές διακοπές: Οι άνθρωποι μπορεί να υποβάλλουν αίτηση για δάνεια για τη χρηματοδότηση καλοκαιρινών διακοπών ή ταξιδιών.
- Εκπαίδευση παιδιών: Οι γονείς μπορεί να ζητούν δάνεια για την κάλυψη σχολικών εξόδων για τα παιδιά τους.



- Τέλος οικονομικού έτους: Ορισμένες επιχειρήσεις ενδέχεται να υποβάλουν αίτηση για δάνεια στο τέλος του οικονομικού έτους (το οποίο μπορεί να ποικίλλει ανάλογα με τη χώρα) για διάφορους σκοπούς, όπως χρηματοδότηση αγορών αποθεμάτων ή ανακαινίσεις στο τέλος της σεζόν.



Στην παρακάτω φωτογραφία φαίνεται η κατανομή του Case Duration:



Φαίνεται ότι οι περισσότερες περιπτώσεις είναι μέχρι τις 25 μέρες παρόλα αυτά υπάρχει αρκετά μεγάλο ποσοστό περιπτώσεων όπου φτάνουν τις 31-35 μέρες και πολύ λίγες δεν ξεπερνούν τις 5 μέρες.

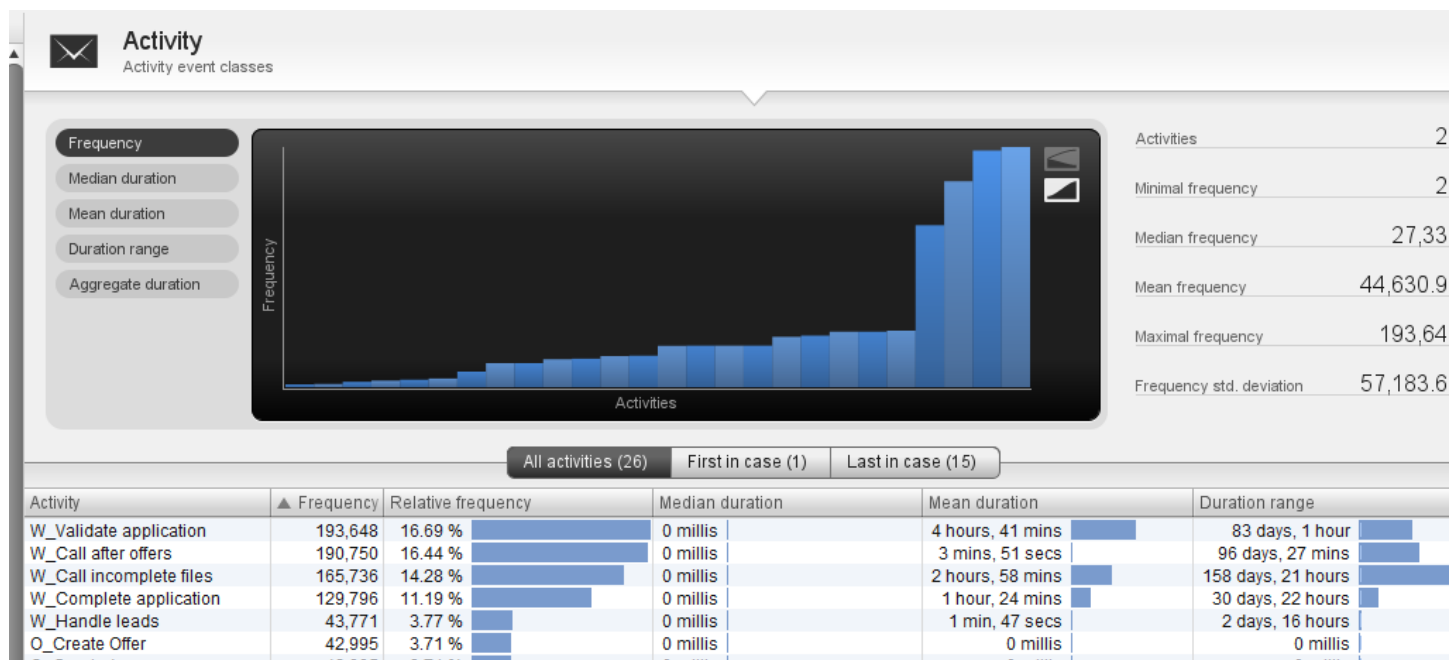


Activity

Βλέπουμε στην παρακάτω εικόνα ότι οι δραστηριότητες από τις οποίες περνάνε τα περισσότερα cases αφορούν αναμονή για συμβάν και είναι ακόλουθες:

- Αναμονή για επιβεβαίωση αίτησης.
- Αναμονή για τηλεφώνημα μετά την αποστολή προσφοράς.
- Αναμονή για αναζήτηση εγγράφων που καθιστούν ημιτελή την αίτηση μέσω κλήσης στον πελάτη.
- Αναμονή για ολοκλήρωση συμπλήρωσης αίτησης.

Επίσης παρατηρούμε ότι αυτά τα activities δαπανούν και τον περισσότερο χρόνο σε διάρκεια ολοκλήρωσης.



Παρακάτω φαίνονται τα στατιστικά για τις πρώτες και τις τελευταίες δραστηριότητες.

All activities (26)					
First in case (1)					
Last in case (15)					
Activity	Frequency	Relative frequency	Median duration	Mean duration	Duration range
A_Create Application	31,509	2.72 %	0 millis	0 millis	0 millis

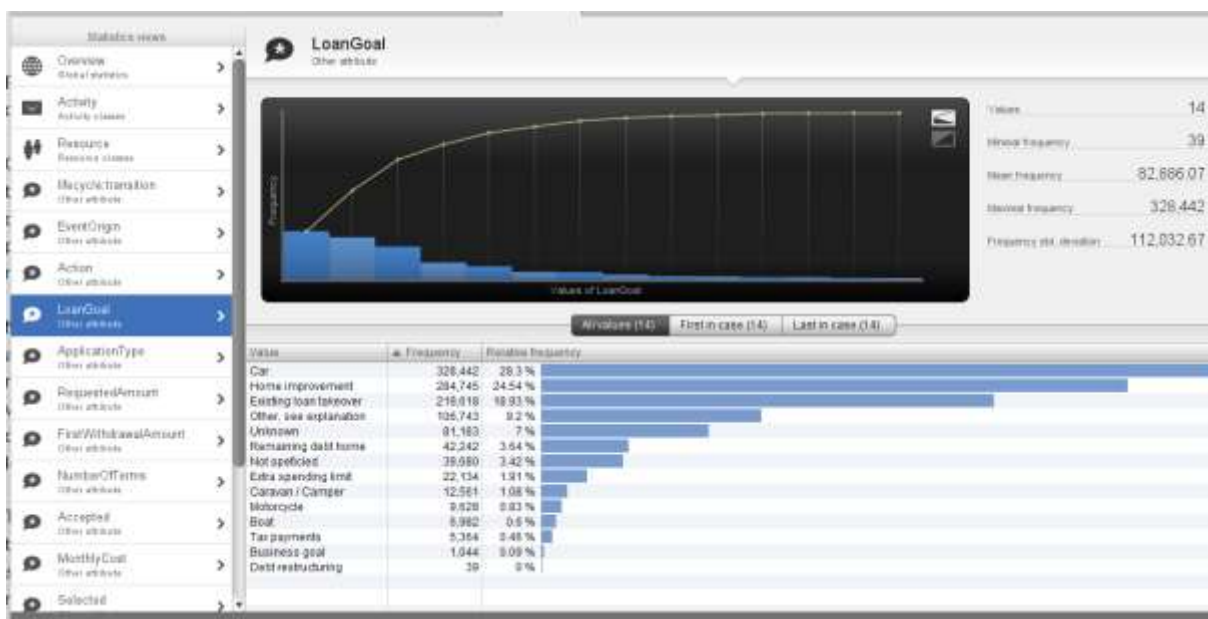
All activities (26)					
First in case (1)					
Last in case (15)					
Activity	Frequency	Relative frequency	Median duration	Mean duration	Duration range
W_Validate application	193,648	16.69 %	0 millis	4 hours, 41 mins	83 days, 1 hour
W_Call after offers	190,750	16.44 %	0 millis	3 mins, 51 secs	96 days, 27 mins
W_Call incomplete files	165,736	14.28 %	0 millis	2 hours, 58 mins	158 days, 21 hours
W_Complete application	129,796	11.19 %	0 millis	1 hour, 24 mins	30 days, 22 hours
O_Sent (mail and online)	39,707	3.42 %	0 millis	0 millis	0 millis
O_Returned	23,305	2.01 %	0 millis	0 millis	0 millis
O_Cancelled	20,898	1.8 %	0 millis	0 millis	0 millis
A_Pending	17,228	1.48 %	0 millis	0 millis	0 millis
A_Cancelled	10,431	0.9 %	0 millis	0 millis	0 millis
O_Refused	4,695	0.4 %	0 millis	0 millis	0 millis
A_Denied	3,753	0.32 %	0 millis	0 millis	0 millis
W_Assess potential fraud	3,000	0.26 %	0 millis	8 hours, 40 mins	88 days, 4 hours
O_Sent (online only)	2,026	0.17 %	0 millis	0 millis	0 millis
W_Shortened completion	238	0.02 %	0 millis	0 millis	0 millis
W_Personal Loan collection	22	0 %	0 millis	0 millis	0 millis



Επιβεβαιώνεται αυτό που είχε παρατηρηθεί από το Process Map, ότι όλες οι περιπτώσεις ξεκινάνε από ένα συγκεκριμένο σημείο και τερματίζουν σε διαφορετικά activities.

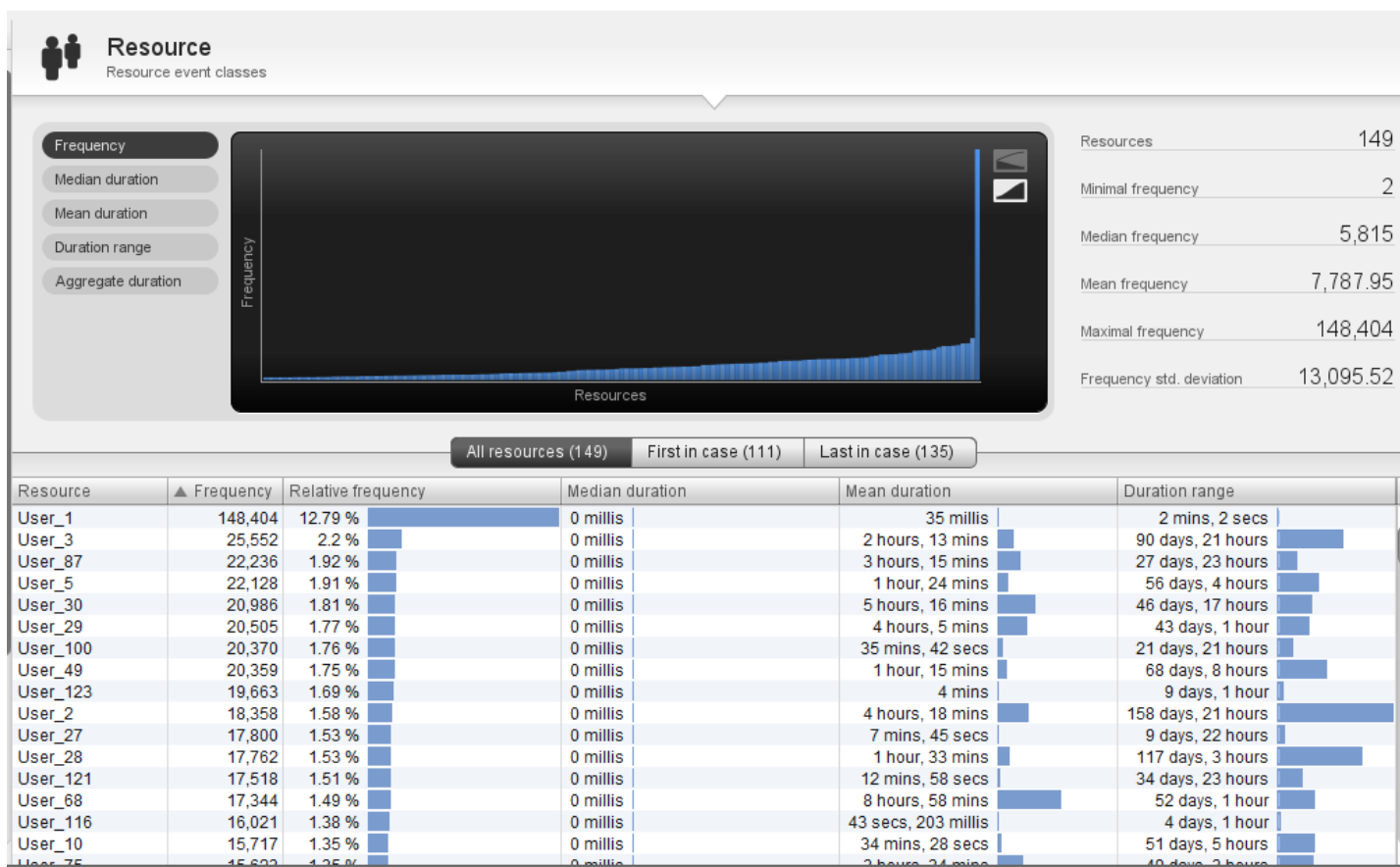
Loan Goal

Επιλέγοντας Loan Goal από τη στήλη των attributes διαπιστώνουμε ότι οι 3 συχνότεροι στόχοι δανεισμού είναι η αγορά αυτοκινήτου, οι ανακαινίσεις κατοικίας και η ανάληψη υπάρχοντος δανείου το οποίο πρέπει να αποσχολήσει τον process owner αφού υπονοεί την παροχή ανταγωνιστικών υπηρεσιών από άλλες τράπεζες.



Resource

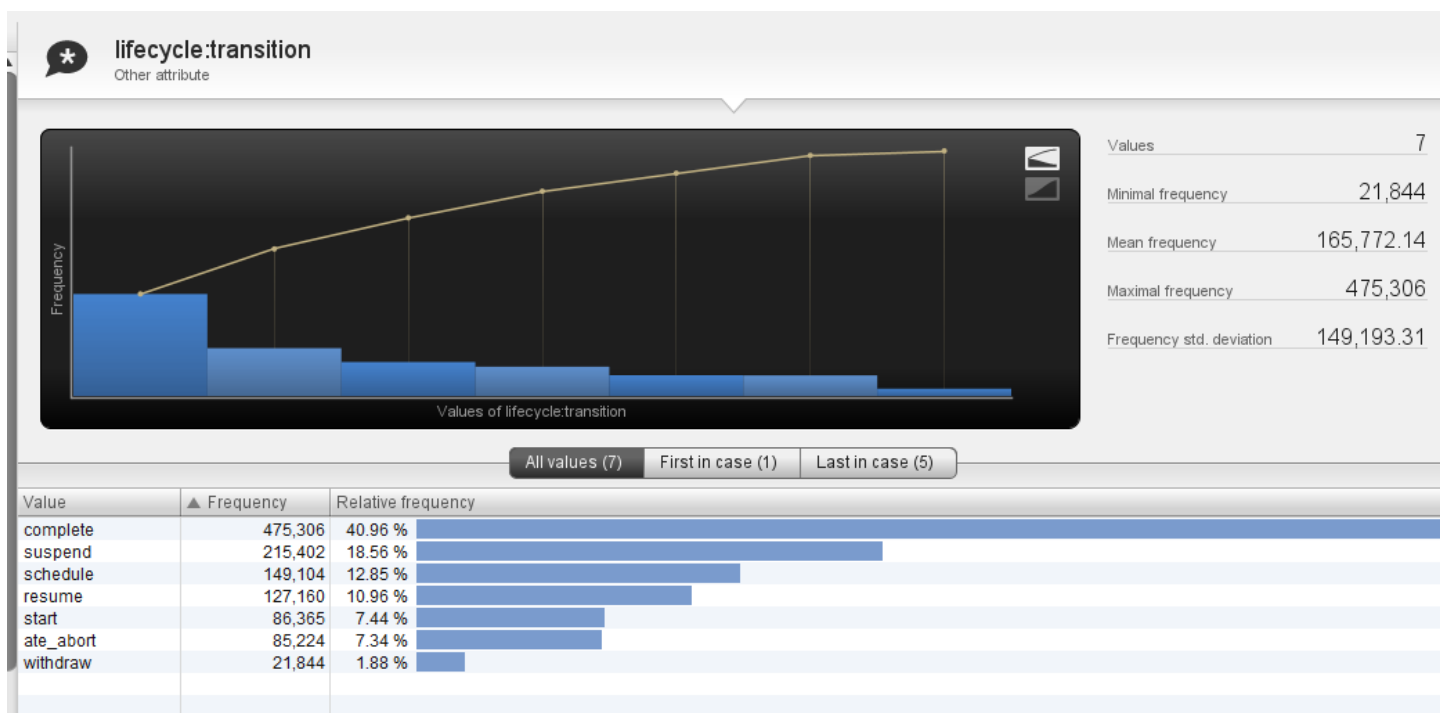
Παρατηρούμε 149 διαφορετικούς χρήστες. Ένας χρήστης (User_1) παρουσιάζει αρκετά υψηλό ποσοστό με λίγο λιγότερο από 13% και όλοι οι υπόλοιποι κυμαίνονται από το 2.2% και κάτω.



Παρόλα αυτά ο πρώτος user δεν εμφανίζει σημαντική διάρκεια, κατά μέσο όρο λιγότερο από δευτερόλεπτο σε αντίθεση με τους υπόλοιπους users που η συχνότητα κυμαίνεται από 2.2% και κάτω αλλά η διάρκεια είναι αρκετά μεγαλύτερη.

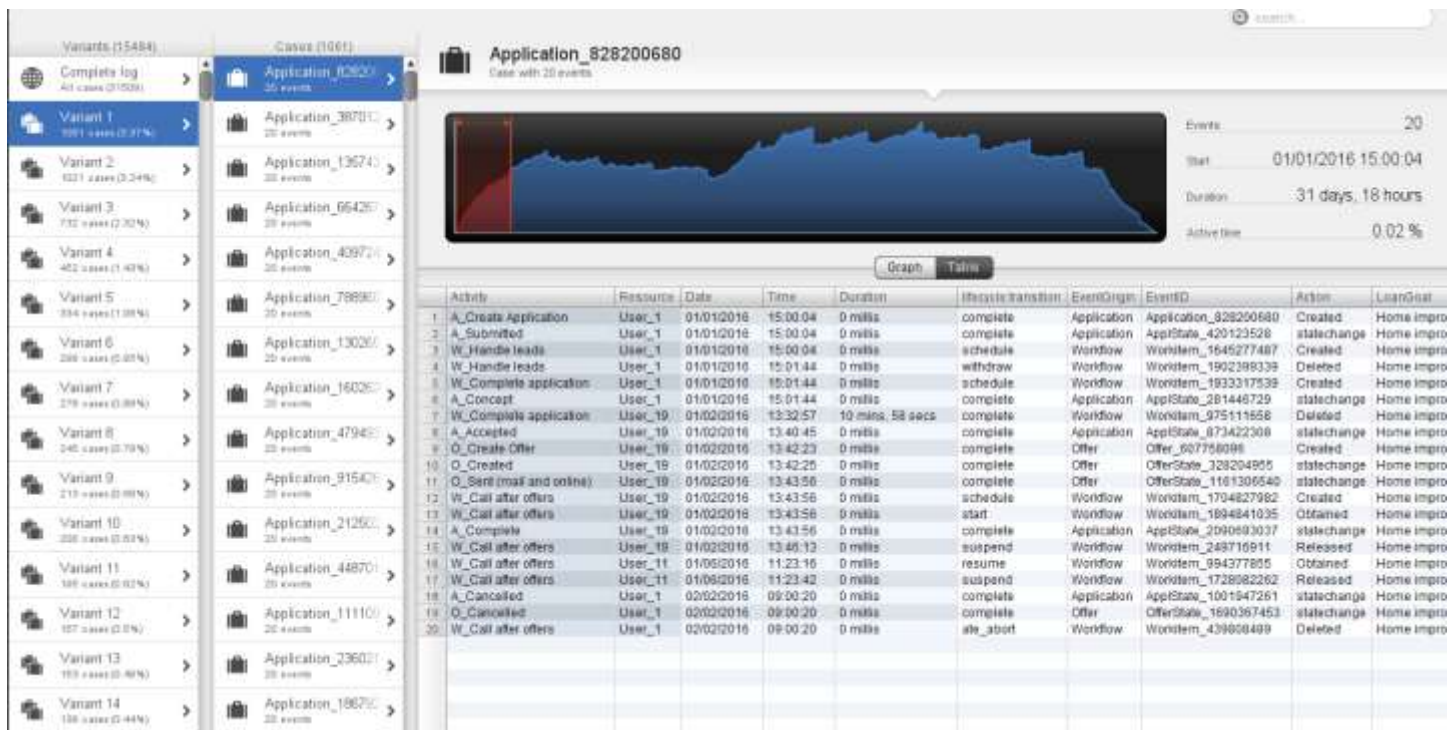
lifecycle:transition

Παρακάτω παρατηρείται ότι για τις 7 καταστάσεις του συστήματος οι περισσότερες καταστάσεις είναι αιτήσεις που έχουν ολοκληρωθεί με ποσοστό κοντά στο 41%. Στην συνέχεια είναι η Suspend με περίπου 18.6%, το Schedule κοντά στο 13%, το Resume στο 11%, το Start στο 7.5%, το Ate_abort στο 7.3% και τέλος το Withdraw κοντά στο 2%.



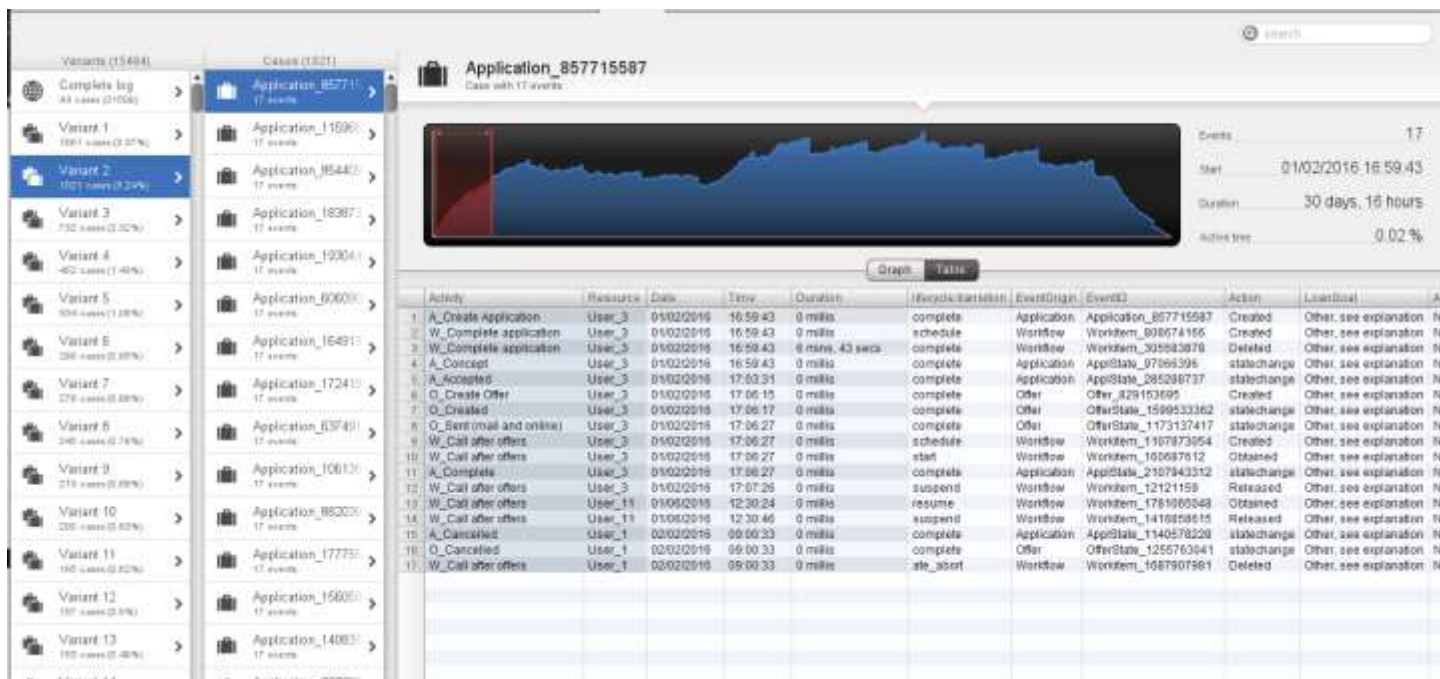
Inspect Cases

Τα συνολικά Cases είναι 31.509, όπως φαίνονται και παραπάνω. Βλέπουμε την παρουσία 15.484 Variants, με κυρίαρχο το Variant 1 να έχει 1.061 περιπτώσεις αντιπροσωπεύοντας βέβαια μόνο το 3.37%. Η κάθε περίπτωση του συγκεκριμένου Variant αποτελείται από 20 events με καταληκτικό το W_Call after offers.

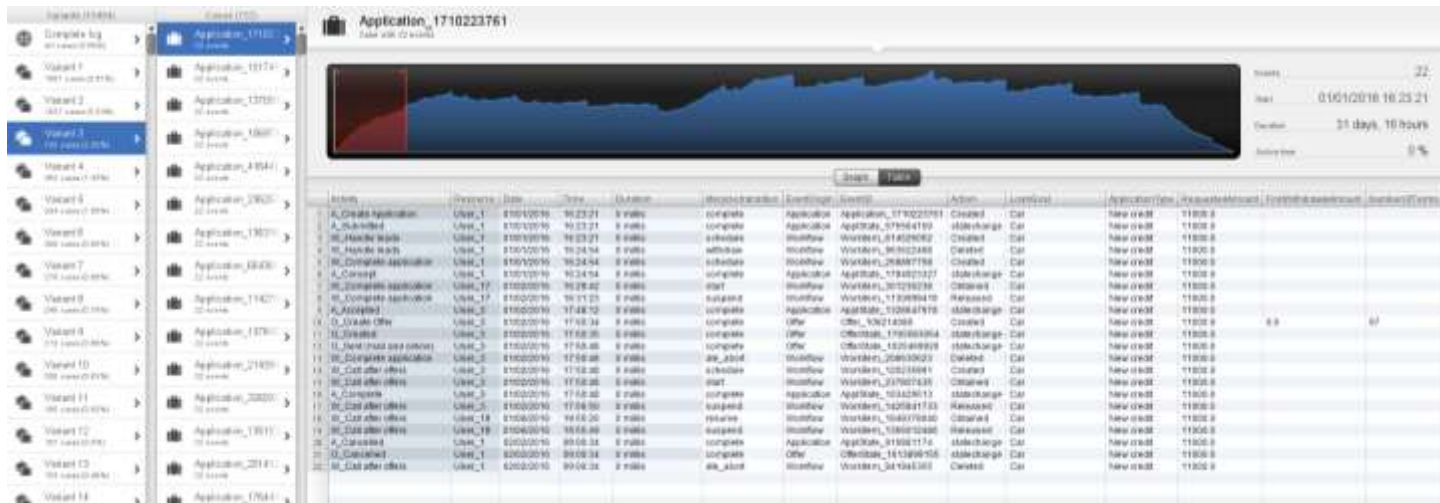


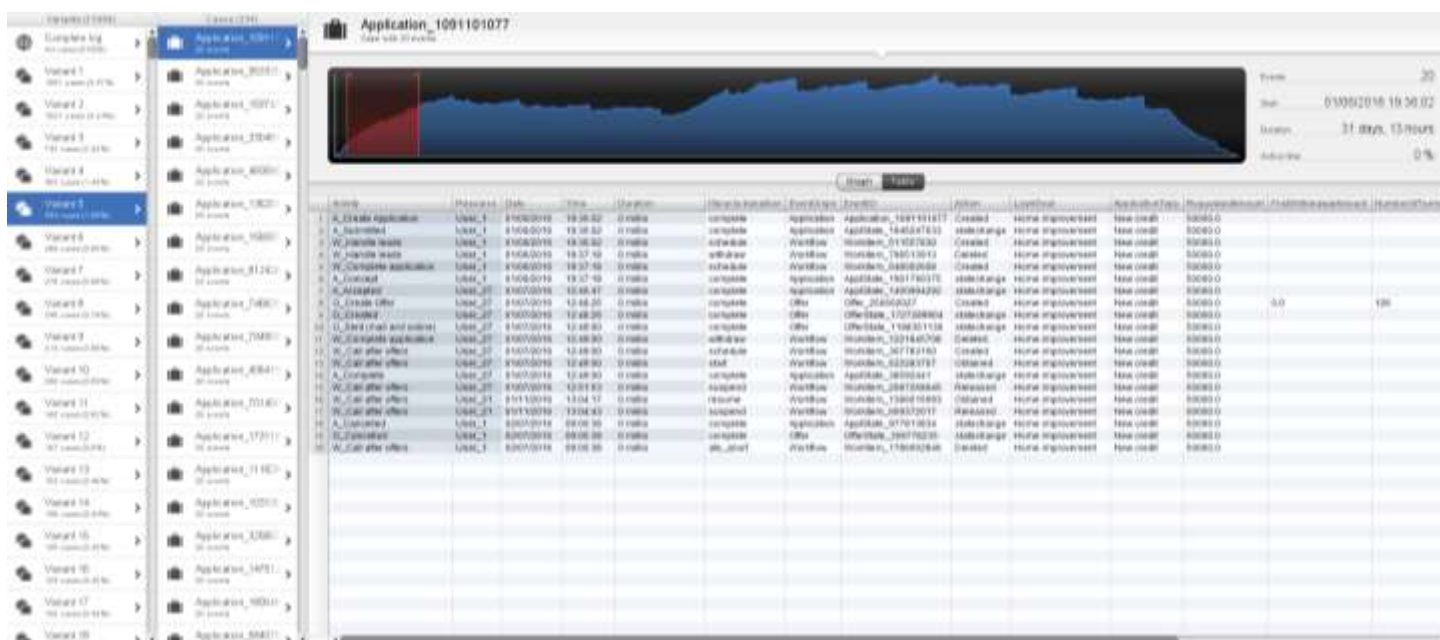
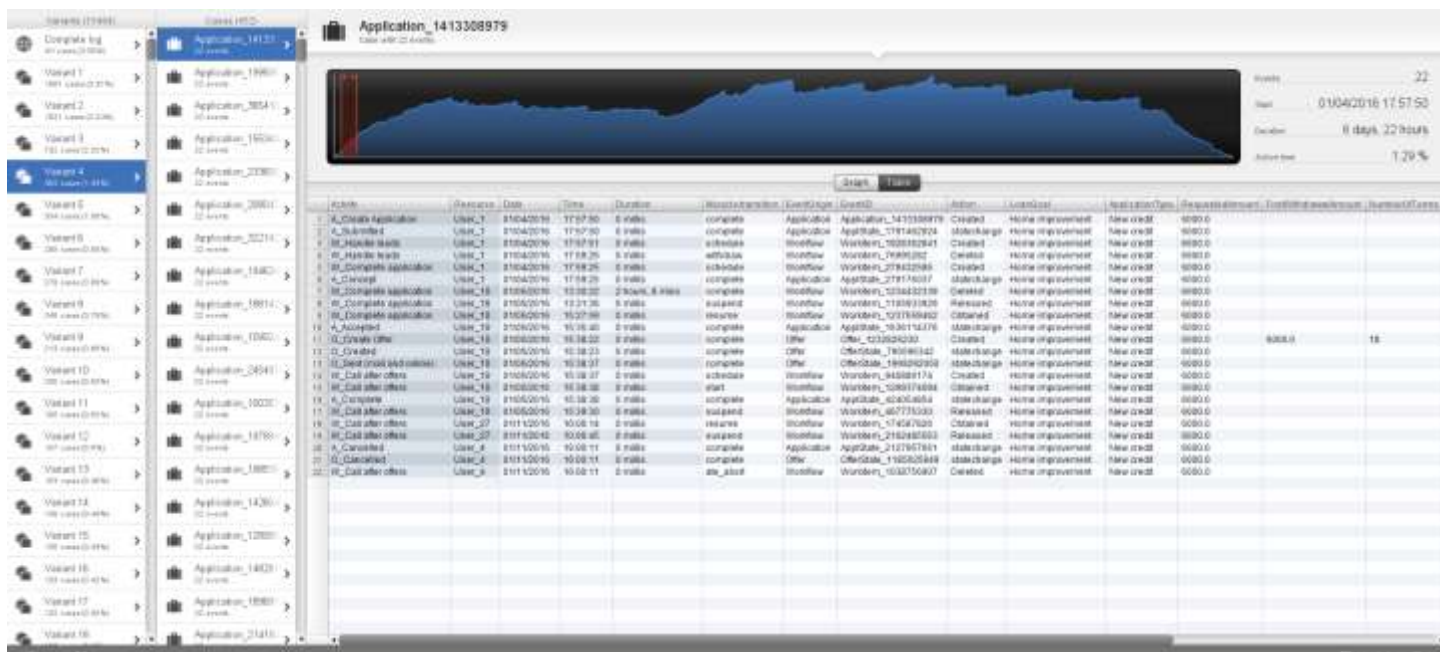


Υστερα ένα άλλο Variant είναι το Variant 2, με 1021 cases αντιπροσωπεύοντας το 3.24%. Η κάθε περίπτωση του συγκεκριμένου Variant αποτελείται από 17 events τα οποία φαίνονται παρακάτω.



Ακολουθούν τα Variant 3 με 22 events και 732 cases (2.32%), Variant 4 με 452 cases (1.43%) και Variant 5 με 334 cases (1.06%).







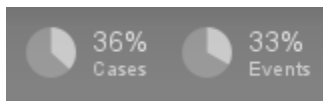
Θα χρησιμοποιηθεί ένα Performance Filter βασισμένο στο Case duration, με δεδομένα ότι η μέση διάρκεια είναι 19.1 μέρες και ο διάμεσος είναι 21.9 μέρες θα χωριστεί το dataset σε δύο μέρη. Σε αυτές που ολοκληρώνονται γρήγορα και σε αυτές που καθυστερούν παραπάνω, θα επιλέξω τα παρακάτω εύρη τιμών:

- ✓ 0 έως 28 μέρες (61%)
- ✓ 28 και παραπάνω (39%)

Επιλέγω αυτό το χρονικό διάστημα καθώς παρατηρώ μια αύξηση Cases στο διάστημα των 29 με 35 ημερών το οποίο χρήζει μελέτης.



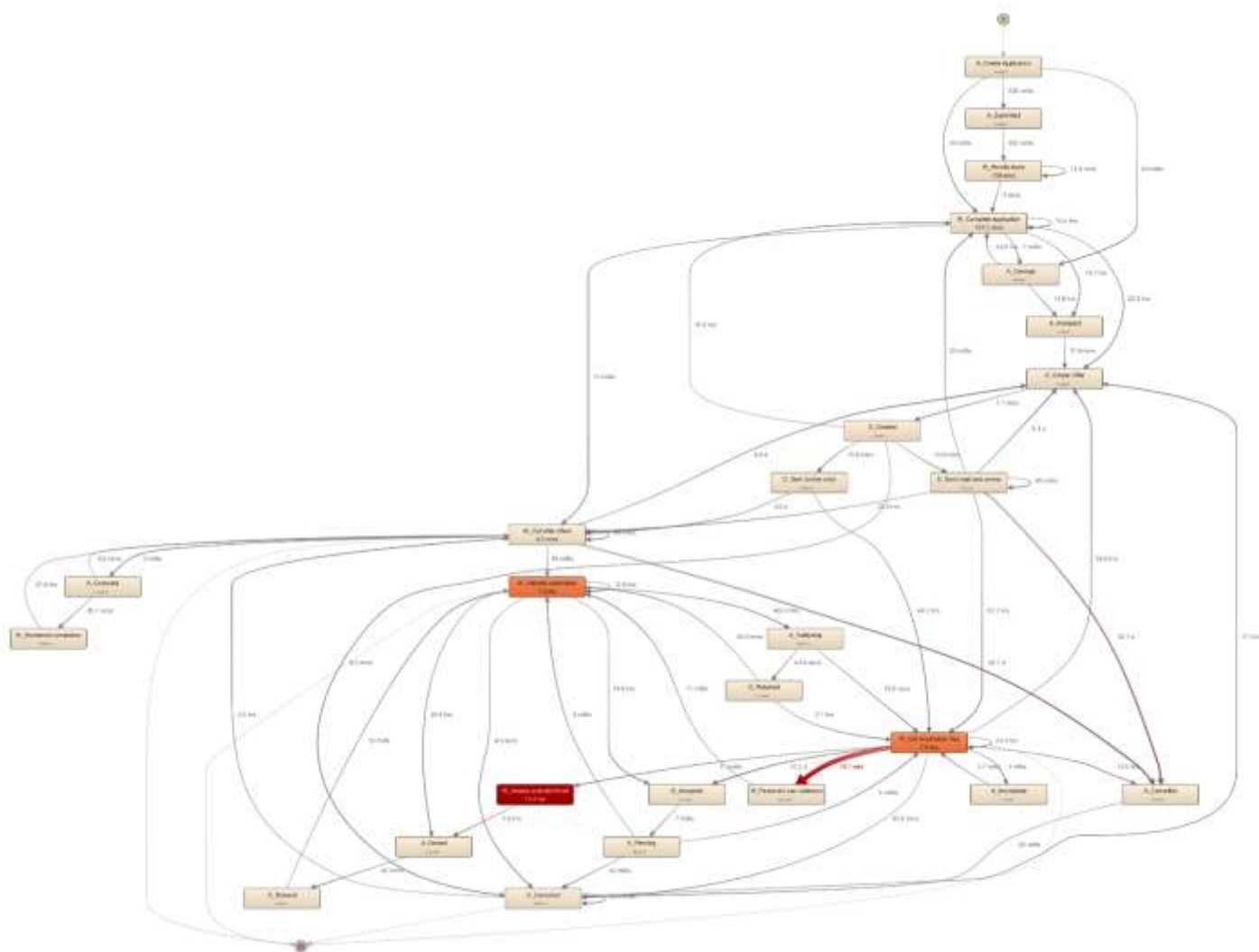
Για το επιλεγμένο διάστημα ισχύει :



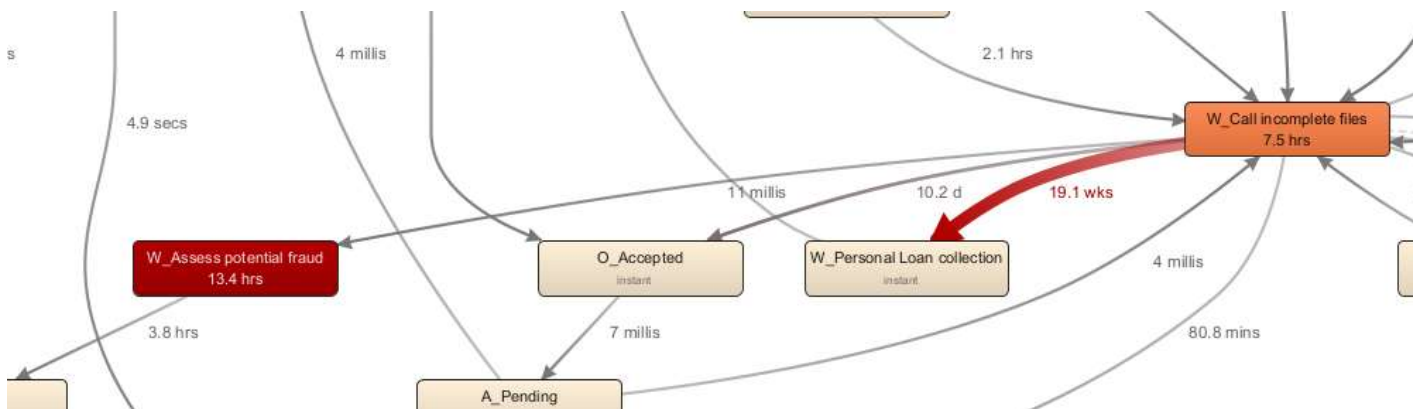


Visualize Bottlenecks

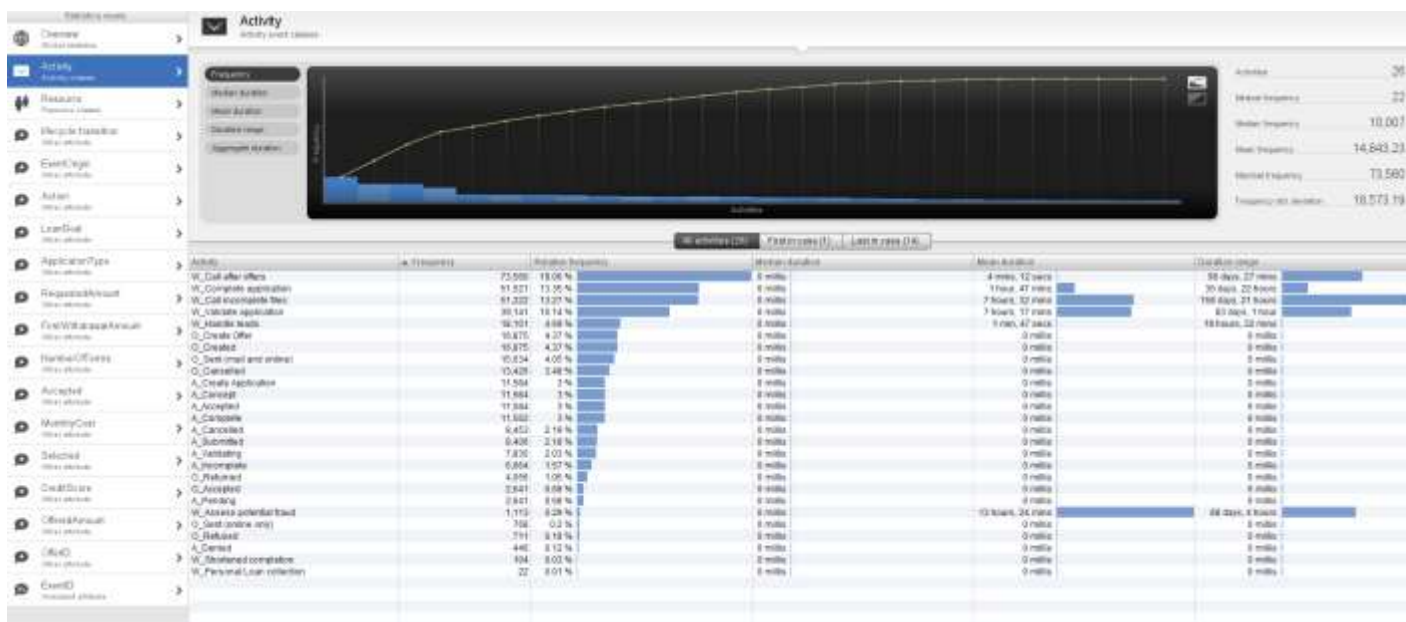
Παρακάτω παρατίθεται το Μαρ αυτών των περιπτώσεων προβάλλοντας το Mean duration από το Performance ώστε να δούμε την γενική εικόνα των καθυστερήσεων. Με το Activity στο 100% και τα Paths στο 30% για να αποκτηθεί ένα γενικό συμπέρασμα το αποτέλεσμα είναι το ακόλουθο:



Μεγεθύνουμε το κομμάτι που μας ενδιαφέρει περισσότερο.



Παρατηρούμε ότι παρά την καθυστέρηση κάποιων activity αυτή είναι μερικών ωρών σε σχέση με τα άλλα και δε δικαιολογεί την καθυστέρηση ημερών που έχουμε σε αυτά τα cases.

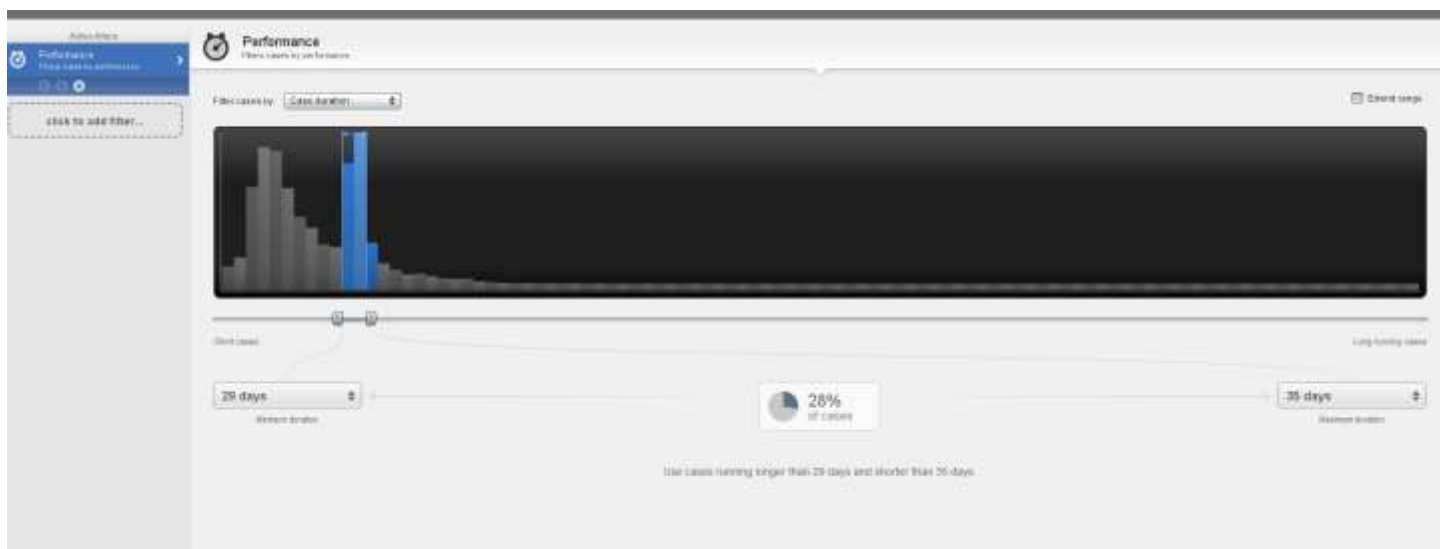


Το πρόβλημα φαίνεται να εντοπίζεται στο transition από W_Call incomplete files ---> W_Personal Loan collection. Ωστόσο, πραγματοποιείται μόνο μία φορά όπως φαίνεται από την παρακάτω εικόνα και άρα δεν είναι αντιπροσωπευτικό.

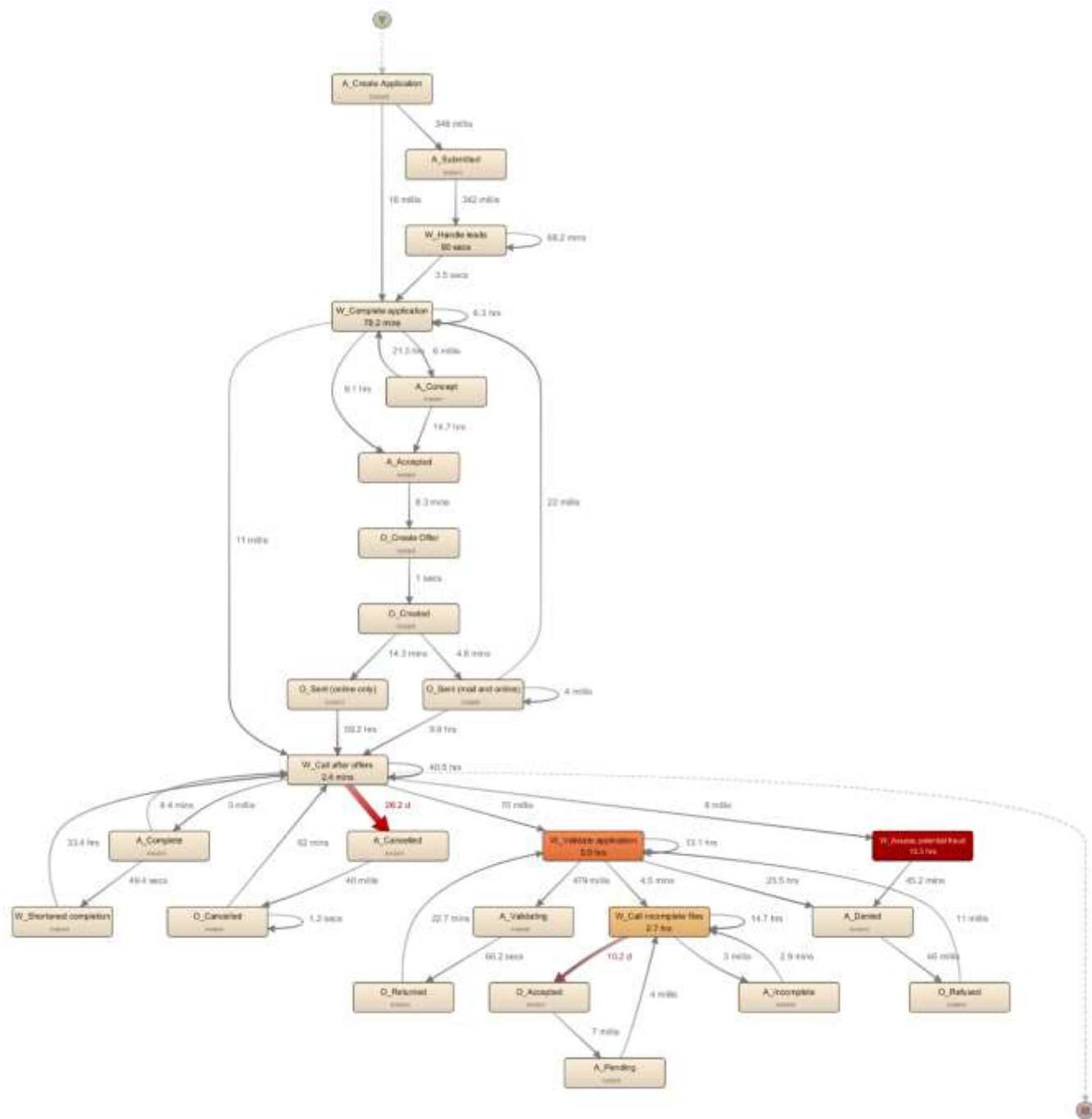




Προχωράμε επομένως σε μια πιο στοχευμένη μελέτη στο διάστημα 29 με 35 ημερών.



Για activity στο 100% και path στο 10% προκύπτει το παρακάτω Map.



Εδώ εντοπίζουμε το πρόβλημα στην περιοχή που φαίνεται στην παρακάτω εικόνα:



The diagram illustrates a process flow with several nodes and transitions. The nodes are labeled as follows:

- O_Sent (online only)**: instant
- W_Call after offers**: 2.4 mins
- A_Cancelled**
- Canceled**: instant

Transitions between nodes include durations such as 14.3 mins, 4.6 mins, 59.2 hrs, 3 mins, 62 mins, 40 millis, and 1.2 secs.

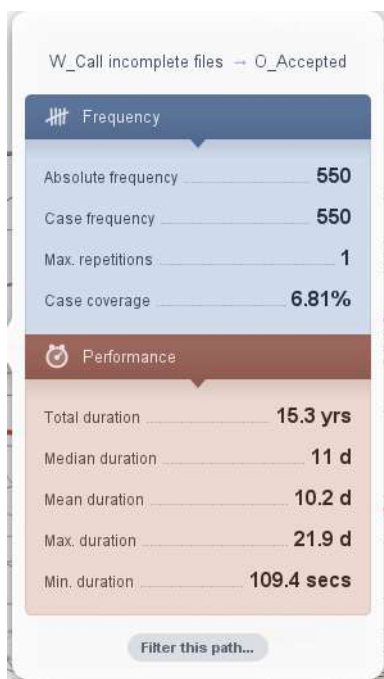
A pop-up window titled "W_Call after offers — A_Cancelled" provides detailed statistics for this specific path:

Frequency	
Absolute frequency	6,691
Case frequency	6,691
Max repetitions	1
Case coverage	82.8%

Performance	
Total duration	480.4 yrs
Median duration	26.8 d
Mean duration	26.2 d
Max duration	31 d
Min duration	3 mins

At the bottom of the pop-up window, there is a button labeled "Filter this path..."

{ 19 }



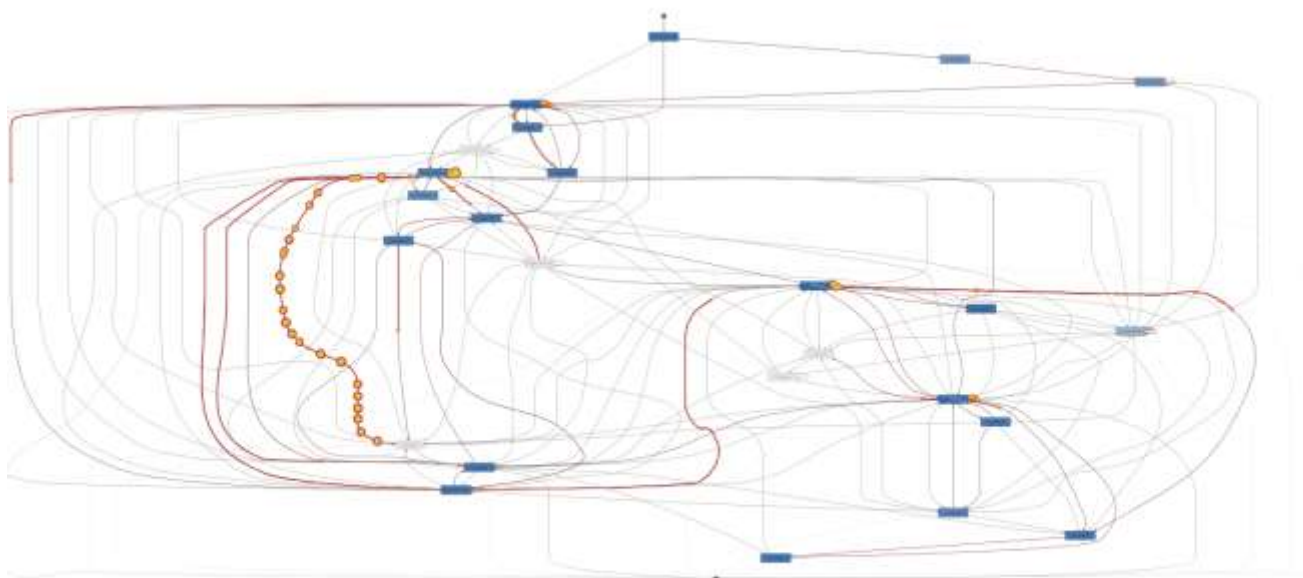
Συμπέρασμα

Θα προτείναμε στον Process Owner να περιορίσει την αναμονή στο activity *W_call after offer* αφού φαίνεται πως εκεί εντοπίζεται το bottleneck ενώ συμπεραίνουμε ακόμη πως η μεγάλη αναμονή οδηγεί σε ανεπιτυχείς συναλλαγές. Συγκεκριμένα,

- Απλοποίηση της *W_Call after offer* activity, αν είναι δυνατόν. Αυτό θα μπορούσε να περιλαμβάνει τη διάσπαση της δραστηριότητας σε μικρότερα βήματα ή τη δημιουργία μιας λίστας ελέγχου για να ακολουθήσει ο process owner.
- Αυξήστε τα επίπεδα στελέχωσης στο τηλεφωνικό κέντρο, αν είναι δυνατόν. Αυτό θα μπορούσε να συμβάλει στη μείωση του χρόνου αναμονής για τους πελάτες.

Animate Process

Με τη χρήση του animation, μπορούμε να δούμε σε ζωντανό χρόνο πως εξελίσσονται οι περιπτώσεις που μελετάμε και να παρατηρήσουμε το γεγονός ότι υπάρχουν καθυστερήσεις στην εξυπηρέτηση σε συγκεκριμένες δραστηριότητες, όπως παρατηρήθηκε παραπάνω, στην αναζήτηση των bottlenecks. Συνεπώς, έχοντας μια πιο ολοκληρωμένη εικόνα της αλληλουχίας και ταχύτητας πραγματοποίησης των μεταβάσεων και της συχνότητας και ταχύτητας διεκπεραίωσης των ενεργειών, ουσιαστικά επιβεβαιώνουμε τις προηγούμενες παρατηρήσεις μας.



Compliance Check

Στο τελευταίο βήμα, καλούμαστε να εντοπίσουμε περιπτώσεις που η διαδικασία παρεκκλίνει από την προβλεπόμενη ροή. Φυσικά, θα πρέπει να αντιμετωπίσουμε κάθε κατηγορία ροής με διαφορετικό τρόπο και να ελέγξουμε ότι τηρείται μια σειρά από διαδικασίες, χωρίς να υπάρχουν μεγάλες παρεκκλίσεις ή ενέργειες που να είναι ανησυχητικές ως προς την ασφάλεια και την ακρίβεια των δεδομένων.

Εστιάζουμε την προσοχή μας στο activity W_Assess potential fraud και φιλτράρουμε τα events που το έχουν end activity. Έτσι έχουμε το ακόλουθο νέο σύνολο.

Events	413
Cases	9
Activities	19
Median case duration	24 d
Mean case duration	23.1 d
Start	03/31/2016 11:36:37
End	01/24/2017 17:20:14

Στο αρχικό σύνολο των δεδομένων παρατηρούμε πως η χρήστες 138 και 144 έχουν πολύ μικρό ποσοστό συμμετοχής.



Με σχετικά υψηλό mean duration.

User_138	1,008	0.09 %	0 πηλίκ	19 hours, 36 mins	88 days, 4 hours
User_139	125	0.01 %	0 πηλίκ	0 mins	0 mins
User_140	295	0.02 %	0 πηλίκ	1 hour, 19 mins	8 days, 1 hour
User_141	400	0.03 %	0 πηλίκ	48 mins, 590 mins	4 hours, 25 mins
User_142	35	0 %	0 πηλίκ	1 hour, 20 mins	1 day, 19 hours
User_143	564	0.05 %	0 πηλίκ	9 hours, 39 mins	16 days, 18 hours
User_144	839	0.07 %	0 πηλίκ	5 hours, 56 mins	19 days, 2 hours

Αντιθέτως, στο φιλτραρισμένο σύνολο είναι στις θέσεις 2 και 3 με υψηλά ποσοστά συμμετοχής.

Resource	# Participants	Relative frequency	Median duration	Mean duration	Duration range
User_1	83	14.53 %	0 mins	0 mins	0 mins
User_128	37	6.56 %	0 mins	0 mins	0 mins
User_144	31	5.51 %	0 mins	0 mins	0 mins
Total: 48	151	26.60 %	0 mins	4 hours, 160 mins	4 days, 150 hours

Αυτό καθιστά αιτία ελέγχου της αξιοπιστίας τους.