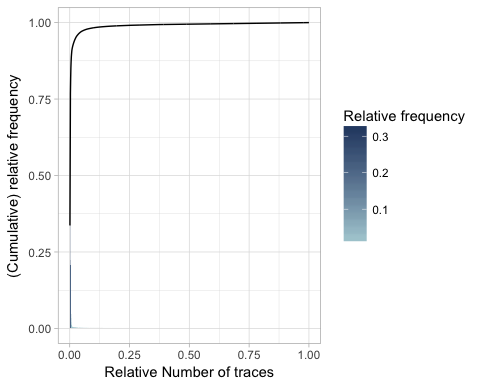
Project BUPRA

Wouter Goossens, Daan Roosen

22-11-2017

## Managementsamenvatting

Om een beter inzicht te krijgen in het facturatieproces van een ziekenhuis, namen wij de dataset 'Hospital Billing' onder de loop. We starten onze analyse met het samenvatten van enkele belangrijke elementen uit deze dataset. De dataset (of log) bevat informatie over 451.359 activiteiten die precies 100.000 gevallen (of cases) beschrijven. Deze activiteiten worden uitgevoerd door 1.151 verschillende medewerkers. Er zijn 1.022 verschillende manieren (traces) waarop een geval kan worden doorlopen. De 10 meest voorkomende traces nemen samen 91.3% van de totale log in.



De gemiddelde doorlooptijd bij deze cases is 30 dagen, terwijl de gemiddelde doorlooptijd over de hele log gemiddeld 127 dagen bedraagt.

De log bevat 18 soorten activiteiten waaronder de belangrijkste zijn: - 'NEW': het aanmaken van een factuur - 'FIN': het afwerken van een factuur - 'RELEASE': het uitgeven van een factuur - 'CODE OK': de status die wordt gegeven wanneer een factuur wordt goedgekeurd - 'BILLED': de status die wordt gegeven wanneer een factuur is betaald - 'CHANGE DIAGN': wanneer de diagnose waarvoor de factuur oorspronkelijk werd opgesteld wordt gewijzigd

## # A tibble: 18 x 2  
## Activity absolute\_frequency  
## <chr> <int>  
## 1 NEW 101289  
## 2 FIN 74738  
## 3 RELEASE 70926  
## 4 CODE OK 68006  
## 5 BILLED 67448  
## 6 CHANGE DIAGN 45451  
## 7 DELETE 8225  
## 8 REOPEN 4669  
## 9 CODE NOK 3620  
## 10 STORNO 2973  
## 11 REJECT 2016  
## 12 SET STATUS 705  
## 13 EMPTY 449  
## 14 MANUAL 372  
## 15 JOIN-PAT 358  
## 16 CODE ERROR 75  
## 17 CHANGE END 38  
## 18 ZDBC\_BEHAN 1

## Inhoudstafel

## Inleiding

De 'Hospital Billing' log bestaat uit 25 verschillende attributen. We vatten deze kort samen in de tabel #:

* Case ID: de unieke waarde die aan iedere case wordt toegewezen
* Activity: de naam van de activiteit die werd uitgevoerd
* Resource: de naam van de persoon die deze activiteit heeft uitgevoerd
* Complete Timestamp: het tijdstip waarop de activeit werd afgerond
* Variant: een waarde om aan te duiden wij welk soort trace de case hoort
* Variant index: enkel de numerieke waarde van variant
* actOrange
* actRed
* blocked
* caseType
* closeCode
* diagnosis
* flagA
* flagB
* flagC
* flagD
* isCancelled
* isClosed
* lifecycle:transition
* msgCode
* msgCount
* msgType
* speciality
* state
* version

## Resultaten

1. Hoeveel facturen worden uiteindelijk niet betaald?

## # A tibble: 36,502 x 2  
## case\_id last\_activity  
## <chr> <chr>  
## 1 AAA NEW  
## 2 AAAD NEW  
## 3 AAAE FIN  
## 4 AAB NEW  
## 5 AABB NEW  
## 6 AACE NEW  
## 7 AAD NEW  
## 8 AADA FIN  
## 9 AADE NEW  
## 10 AAEB SET STATUS  
## # ... with 36,492 more rows

1. Welke activiteiten moeten vaak opnieuw worden uitgevoerd?

## Activity Frequency  
## 1 CODE OK 3447  
## 2 RELEASE 3266  
## 3 FIN 3072  
## 4 BILLED 2150  
## 5 REOPEN 454  
## 6 STORNO 350  
## 7 CHANGE DIAGN 200  
## 8 CODE NOK 148  
## 9 REJECT 138  
## 10 MANUAL 46  
## 11 SET STATUS 24  
## 12 CODE ERROR 5  
## 13 CHANGE END 3  
## 14 DELETE 0  
## 15 EMPTY 0  
## 16 JOIN-PAT 0  
## 17 NEW 0  
## 18 ZDBC\_BEHAN 0

1. Hoe lang duren de verschillende activiteiten?
2. Wat zorgt ervoor dat cases langer duren? (bottlenecks)
3. Welke traces eindigen vaak geblokkeerd?
4. Welke diagnose wordt later het vaakst veranderd?

## Joining, by = "case\_id"

## # A tibble: 568 x 2  
## # Groups: diagnosis [568]  
## diagnosis frequency  
## <chr> <int>  
## 1 YC 163  
## 2 GKA 127  
## 3 NE 82  
## 4 JA 73  
## 5 Z 64  
## 6 RKA 54  
## 7 KM 51  
## 8 L 50  
## 9 XG 49  
## 10 VF 45  
## # ... with 558 more rows

1. Welke eerste diagnose was vaak toch de juiste?

## # A tibble: 71,885 x 2  
## case\_id n  
## <chr> <int>  
## 1 A 1  
## 2 AA 1  
## 3 AAAA 1  
## 4 AAAB 1  
## 5 AAAC 1  
## 6 AAAE 1  
## 7 AABA 1  
## 8 AABC 1  
## 9 AABD 1  
## 10 AABE 1  
## # ... with 71,875 more rows

1. Welke diagnoses volgen mekaar vaak op?

## Joining, by = "case\_id"

## # A tibble: 1,748 x 3  
## # Groups: previous\_diagnosis [568]  
## previous\_diagnosis diagnosis frequency  
## <chr> <chr> <int>  
## 1 GKA RKA 109  
## 2 YC XC 84  
## 3 YC PE 79  
## 4 RKA GKA 48  
## 5 L VB 41  
## 6 NE Z 32  
## 7 ZH XM 25  
## 8 XG LC 23  
## 9 NE DJ 22  
## 10 KM Q 19  
## # ... with 1,738 more rows

1. Hoe vaak zijn dokters het niet helemaal eens over de diagnose?

## Joining, by = "case\_id"

## # A tibble: 27,581 x 4  
## case\_id Activity Resource akkoord  
## <chr> <chr> <chr> <chr>  
## 1 AAAA NEW ResXG NEW  
## 2 AAAA CHANGE DIAGN ResJN niet akkoord  
## 3 AABD NEW ResSD NEW  
## 4 AABD CHANGE DIAGN ResYAA niet akkoord  
## 5 AACA NEW ResPP NEW  
## 6 AACA CHANGE DIAGN ResAI niet akkoord  
## 7 AADC NEW ResYW NEW  
## 8 AADC CHANGE DIAGN ResK niet akkoord  
## 9 AAEE NEW ResJU NEW  
## 10 AAEE CHANGE DIAGN ResJV niet akkoord  
## # ... with 27,571 more rows

1. Hoe vaak wordt een eerst afgewezen diagnose, toch opnieuw gesteld?

## # A tibble: 359 x 3  
## # Groups: case\_id [207]  
## case\_id diagnosis abs\_freq  
## <chr> <chr> <int>  
## 1 ABHB EM 2  
## 2 AGHA N 2  
## 3 AGTB QV 2  
## 4 AHLD DJ 2  
## 5 AHLD NE 2  
## 6 AHYC AL 2  
## 7 AOSD N 2  
## 8 APH AJ 15  
## 9 APH CM 5  
## 10 APH CN 8  
## # ... with 349 more rows

## # A tibble: 114 x 2  
## diagnosis abs\_freq  
## <chr> <int>  
## 1 Z 24  
## 2 AJ 20  
## 3 DJ 17  
## 4 MN 13  
## 5 NE 13  
## 6 CN 12  
## 7 RKA 12  
## 8 YC 12  
## 9 XC 11  
## 10 HB 10  
## # ... with 104 more rows

## # A tibble: 1,352 x 2  
## case\_id n  
## <chr> <int>  
## 1 AAAB 1  
## 2 AACA 1  
## 3 AADC 1  
## 4 AAE 1  
## 5 AALB 1  
## 6 ABG 1  
## 7 ABKD 1  
## 8 ABZA 1  
## 9 ACHE 1  
## 10 ACLD 1  
## # ... with 1,342 more rows

1. Hoeveel tijd zit er tussen de goedkeuring en betaling van een factuur?

## # A tibble: 1 x 1  
## avg  
## <time>  
## 1 48.94975 days

## Beperkingen

## Conclusie