Пакетна обрада података употребом *Apache Pig* алата

Владимир Инђић, Е2 82/2018

Скуп података

- Жалбе корисника на услуге различитих компанија
- Може се преузети са линка
- Увид у скуп доступан на линку.
- Скуп се одржава од стране СГРВ-а
- Текстуални структурирани документ (.csv формат)

Обележја торки скупа података

- dateReceived
- product
- subProduct
- issue
- subIssue
- consumerComplaintNarrative
- companyPublicResponse
- Company
- state

Обележја торки скупа података

- zipCode
- tags
- consumerConsentProvided
- submittedVia
- dateSentToCompany
- companyResponseToConsumer
- TimelyResponse
- ConsumerDisputed
- complaintId

Компаније (сотрапу)

(EQUIFAX, INC., 100580) (Experian Information Solutions Inc., 89465) (TRANSUNION INTERMEDIATE HOLDINGS, INC.,82438) (BANK OF AMERICA, NATIONAL ASSOCIATION, 79440) (WELLS FARGO & COMPANY, 67422) (JPMORGAN CHASE & CO.,56368) (CITIBANK, N.A., 45765) (CAPITAL ONE FINANCIAL CORPORATION, 31357) (OCWEN LOAN SERVICING LLC,27231) (Navient Solutions, LLC., 27080) (SYNCHRONY FINANCIAL, 19865) (NATIONSTAR MORTGAGE, 19587) (U.S. BANCORP, 16127) (Ditech Financial LLC,13535) (AMERICAN EXPRESS COMPANY, 12408) (PNC Bank N.A., 11132) (ENCORE CAPITAL GROUP INC., 10684) (DISCOVER BANK, 9476) (PORTFOLIO RECOVERY ASSOCIATES INC,9318)

Производи (product)

```
(Mortgage, 268183)
(Debt collection, 224672)
(Credit reporting, credit repair services, or other personal consumer reports, 172622)
(Credit reporting, 140432)
(Credit card, 89190)
(Bank account or service, 86206)
(Student loan, 47746)
(Credit card or prepaid card, 36813)
(Consumer Loan, 31605)
(Checking or savings account, 31282)
(Vehicle loan or lease, 8971)
(Money transfer, virtual currency, or money service, 8194)
(Payday loan, title loan, or personal loan, 6827)
(Payday loan, 5544)
(Money transfers, 5354)
(Prepaid card, 3819)
(Other financial service, 1059)
(Virtual currency, 18)
```

Потпроизводи (subProduct)

(,235166)(Credit reporting, 169175) (Other mortgage, 86636) (Checking account,82351) (Conventional fixed mortgage, 70615) (I do not know, 45748) (Other (i.e. phone, health club, etc.),44558) (FHA mortgage, 30147) (Credit card, 28700) (General-purpose credit card or charge card, 28624) (Conventional adjustable mortgage (ARM),25381) (Non-federal student loan, 25165) (Other debt, 24738) (Conventional home mortgage, 22824) (Medical, 21192) (Federal student loan servicing, 19748) (Other bank product/service, 17825) (Vehicle loan, 17782)

Проблеми (issue)

```
(Loan modification, collection, foreclosure, 112313)
(Incorrect information on credit report, 102686)
(Incorrect information on your report, 100430)
(Loan servicing, payments, escrow account, 77337)
(Cont'd attempts collect debt not owed, 60695)
(Problem with a credit reporting company's investigation into an existing
problem, 38619)
(Account opening, closing, or management, 37961)
(Attempts to collect debt not owed, 33340)
(Communication tactics, 33151)
(Disclosure verification of debt, 30802)
(Improper use of your report, 28176)
(Deposits and withdrawals, 22851)
(Written notification about debt, 19572)
(Managing an account, 19463)
(Managing the loan or lease, 18573)
(False statements or representation, 18526)
(Trouble during payment process, 17972)
```

Потпроблеми (subIssue)

```
(,516235)
(Information belongs to someone else, 40100)
(Account status, 37057)
(Debt is not mine, 36734)
(Information is not mine, 32384)
(Their investigation did not fix an error on your report, 27522)
(Debt was paid, 24077)
(Account status incorrect, 22058)
(Not given enough info to verify debt, 21816)
(Frequent or repeated calls, 19522)
(Account information incorrect, 18786)
(Debt is not yours, 18426)
(Credit inquiries on your report that you don't recognize, 17422)
(Attempted to collect wrong amount, 14570)
(Didn't receive enough information to verify debt, 14498)
(Account terms, 10995)
(Reporting company used your report improperly, 10183)
(Public record, 8876)
(Personal information, 7529)
```

Америчке државе (state)

(CA,161712)

(FL,112753)

(TX,97002)

(NY,78505)

(GA,60380)

(IL,45266)

(NJ,44174)

(PA,40376)

(NC,35969)

(OH,35068)

(VA,34583)

(MD,34131)

(MI, 27873)

(AZ,25292)

(WA,22737)

(MA,21147)

(TN,19090)

(CO, 18976)

(SC,17231)

(MO,16121)

(,15963)

Тагови (tags) (ради лакше претраге)

(,1008695) (Servicemember,74548) (Older American,72651) (Older American, Servicemember,12643)

Зип код (zipCode)

(,91996)(300XX,5930) (770XX,5000) (606XX,4153) (750XX,4083) (331XX,3976) (330XX,3929) (334XX,3738) (303XX,3594) (900XX,3581) (945XX,3480) (070XX,3000)(112XX,2966) (302XX,2916) (891XX,2826) (207XX,2753) (191XX,2720)

Детаљан текст жалбе (consumerComplaintNarrative)

(,824277)

(There are many mistakes appear in my report without my understanding.,611) (Equifax mishandled my information which has led to a breach that puts myself and millions of others at potential risk. I am extremely disappointed with how equifax has handled reporting this breach. Very little was done to notify the public for nearly a month after the breach was detected. I received no email, letter, or phone call and instead had to discover it via social media.,118)

(I am filing this complaint because Experian has ignored my request to provide me with the documents that their company has on file that was used to verify the accounts I disputed. Being that they have gone past the 30 day mark and can not verify these accounts, under Section 611 (5) (A) of the FCRA - they are required to ... promptly delete all information which can not be verified that I have disputed. Please resolve this manner as soon as possible. Thank you.,103)



Одговор који компанија дозвољава да *CFPB* прикаже (companyPublicResponse)

(,776830)

(Company has responded to the consumer and the CFPB and chooses not to provide a public response, 262124)

(Company believes it acted appropriately as authorized by contract or law,55375)

(Company chooses not to provide a public response,52473)

(Company believes the complaint is the result of a misunderstanding, 5206)

(Company disputes the facts presented in the complaint, 4841)

(Company believes complaint caused principally by actions of third party outside the control or direction of the company,3730)

(Company believes complaint is the result of an isolated error, 3422)

(Company believes complaint represents an opportunity for improvement to better serve consumers, 2318)

(Company can't verify or dispute the facts in the complaint, 2157)

(Company believes complaint relates to a discontinued policy or procedure, 61)



Да ли је корисник дао сагласност да му се текст жалбе објави (consumerConsentProvided)

(N/A,543948) (Consent provided,344570) (Consent not provided,248511) (Other,15779) (,14628) (Consent withdrawn,1101)

Комуникационо средство обавештења (submittedVia)

(Web,849739) (Referral,165410) (Phone,70358) (Postal mail,64572) (Fax,18082) (Email,376)

Компанија одговорила кориснику (companyResponseToConsumer)

(Closed with explanation,903316)
(Closed with non-monetary relief,143604)
(Closed with monetary relief,69233)
(Closed without relief,17868)
(Closed,17611)
(Untimely response,5825)
(In progress,5770)
(Closed with relief,5304)
(,6)

Компанија одговорила на време (timelyResponse)

(Yes,1138129) (No,30408)

Корисник се споречкао (consumerDisputed)

(No,62<mark>0149)</mark> (N/A,400010) (Yes,148378)

Датум када је *CFPB* послао жалбу компанији (dateSentToCompany)

(09/08/2017,3387) (09/09/2017,2656) (01/19/2017,1613) (09/13/2017,1535) (01/20/2017,1467) (09/14/2017,1264) (01/24/2017,1235) (04/10/2018,1192) (01/25/2017,1148) (09/12/2017,1138) (09/26/2017,1137) (04/24/2018,1132) (04/11/2018,1129)

Датум када је *CFPB* примио жал**бу** (dateReceived)

(09/08/2017,3553) (09/09/2017,2709) (01/19/2017,2070) (01/20/2017,1633) (09/13/2017,1600) (04/05/2018,1262) (09/12/2017,1230) (04/10/2018,1178) (09/11/2017,1177) (09/14/2017,1153) (04/24/2018,1142) (01/24/2018,1103)

Могућности за анализу

- Колико времена треба CFPB-у да проследи жалбу и да ли то можда зависи од начина на који је жалба достављена (просечно време по средству јављања)?
- Који вид комуникације најчешће користе старији Американци?
- Да ли су можда број жалби и неодговарање на време директно пропорционални?
- Како изгледа број жалби по фирми и по производу (да ли се корисници најчешће жале на банке и кредите)?
- Какви проблеми постоје са кредитним картицама?



Хвала на пажњи. Сви Ваши предлози су више него добро дошли! :)