

Text Mining Exercise week 2

Guido Zuidhof, s4160703

My approach uses regular expressions to recognize and replace certain parts of the text. All regular expressions are Python regex. The text is processed as follows:

1. Subsequent newlines are flagged, matched using `'\n\n'`
2. Hyphens at the end of sentences are removed, the two parts of the words get merged. Achieved using `re.sub(r'\n([^\A-Z])',r"\1", text)`
3. All newlines that do not have a capital letter as next character get removed: `re.sub(r'\n([^\A-Z])',r"\1", unhyphenated)`
4. Newlines are inserted where the following pattern is present: dot, space, capital letter, (NOT (whitespace AND dot) OR dot) . The operation for this is `re.sub(r'[.][]([A-Z])([^\s.])', r'.\n\1\2', text)` .
5. Newlines are re-inserted where the flags were placed in step 1.

This simple approach is quite succesful, I worked from simple rules to more advanced rules, back to simpler rules. It is not perfect, but it gets very close.

Encountered problems

I encountered many problems, the first was the encoding of the file. The windows commandline seemed to be unable to print certain characters, and Python wanted to convert everything to unicode. After a lot of googling I found out I could simply decode and encode the file.

The second was the use of hyphens mid-sentence. The solution was to only match hyphens at the end of sentences and remove these. The third problem had to do with capitals and dots used throughout the text, especially in names of people (with initials). By setting a minimal length of two for a sentence to be placed on a new line in the regex many of these problem areas seem to have been solved.

Finally, there was the page numbering/foot note problem. It turned out to be trivial to solve, as there were already newlines present in the original text, which I could simply restore.

Output

```
Distributie van levensmiddelen
Met de distributie van levensmiddelen was al tijdens de neutraliteitsperiodes begin gemaakt.
In de eerste wereldoorlog was die distributie aanvankelijk een gemeentelijke aangelegenheid geweest - nu
Er was een Centraal
Distributiekantoor, het C D K , in Den Haag gekomen, het land was in ruim 500 'distributiekringen' ingedeeld.
De winkelier plakte de ontvangen bonnen op vellen, leverde die vellen bij de distributiedienst in (waar
Uit de levensmiddelensector waren in de neutraliteitsperiode slechts twee artikelen 'op de bon' gekomen:
De rantsoenen waren evenwel zo ruim dat niet eens alle bonnen ingewisseld werden.
Waarom dan toch die distributie? 'De inwerkstelling (was)' , aldus minister Steenberglic na de oorlog, '
Welnu, het werkte.
Natuurlijk greep die distributie van het begin van de bezetting af steeds verder om zich heen en niet af.
Wij hebben er in dit hoofdstuk al voorbeelden van gegeven: voor de aanschaffing van fietsbanden, fietse
Kreeg men zulk een bon, bijvoorbeeld voor een fietsband of voor een paar schoenen, dan werd een vakje op
In '41 werd aan de stamkaart een z.g. inlegvel toegevoegd.
Op dat inlegvel werd in een vakje de datum gestempeld waarop men een bonkaart voor levensmiddelen uitgaaf.

II Getuige M. P. L.
Steenburch, Eng. , dl.
II c, p. 308.
```

ALLES ' OP DE B O n '

Van tijd tot tijd kreeg men een nieuw inlegvel.

Met dat al raakte toch ook destamkaart van velen in de loop van '42 aardig vol.

Kruidenierswaren, peulvruchten, kaas en zeep ontving men tegen inlevering van bonnen van de z.g. algeme De bonnen waren als regelgenummerd en welke bonnen in een bepaalde periode geldig waren, werd elkeweek Praktisch alles kwam'op de bon'1 - wat de levensmiddelen betrof, bleven alleen groente, fruit en vis lan Eind '40 had in gezinnen met een gemiddeld inkomen van ruim f 2 000 per jaar (gezinnen van geschoolde a Anders gezegd: de distributie werd de grondslag van de gehele levensmiddelenvoorziening.

Dat die distributie voor de burgerij een enorme rompslomp met zichbracht, spreekt vanzelf.

In zekere zin waren de distributiestamkaart (plusinlegvel) en de bonkaarten het belangrijkste wat men i Het verlies van een bonkaart, laat staan van alle distributiebescheiden tezamen, leek steeds een ramp.

Aanvankelijk kon men dan na aangifte bij de distributiedienst nieuwe bescheiden krijgen, maar in de zomer Er kwamen toen per week niet minder dan 20 000 opgaven van vermissingen binnen en men was bij de departe

1 De distributie van de volgende, voor alle consumenten belangrijke artikelen (onze opgave is niet volle

157

VERA RM EN D NEDERLANDgekomen dat een groot aantal daarvan gefingeerd was.

Er werd derhalve besloten, vermiste bonkaarten niet meer te vervangen, 'bijzondere gevallen daargelaten'1 Nadien nam het aantal vermissingen af.

Met dat al werd bijvoorbeeld in Den Haag in '43 nog altijd in bijna XI 000 gevallen aangifte gedaan; in Als regel werden de bonnen door de vrouw des huizes beheerd; zij was het immers die er de dagelijkse ink In menig z.g. a-sociaal gezin kwam van dat beheer niet veel terecht.

Was men daar bonkaarten kwijt, dan moest men ze 'zwart' kopen bij handelaren die het geld voorschoten ma De gemeente Amsterdam begon daar in '42 een stokje voor te steken: zij richtte een apart bureau op waa Gelijk reeds gezegd: het land was in distributiekringen ingedeeld.

Op het platteland behoorden als regel verscheidene gemeenten tot één kring.

In de geografische groepering deden zich nogal eens wijzigingen voor, als regel terwille van de efficienc Eind '42 waren er 382 distributiediensten die er tezamen 685 agentschappen op nahielden.

In een gemeente die geen eigen distributiedienst bezat, was in elk geval een agentschap van een distribu Uiteraard hadden die distributiediensten in de grootste gemeenten een aanzienlijke omvang.

Daar was nieuw personeel voor aangenomen hetwelk, als er extra veel werkzaamheden waren (bij de uitreik De

Amsterdamse distributiedienst had bijvoorbeeld eind '43 een vaste staf van meer dan elfhonderd krachten, In het gehele land waren eind '43 ca. twaalfduizend personen in vaste dienst (wij bedoelen dit niet in j

1 H N S, Research-bureau: 'Rapport over de toestand van het Nederlandse bedrijfsleven over de maand okt

158

DIST K IB U T IE -A PP A RAAT

Naast die twaalfduizend waren in de lente van '43 bijna zeventienhonderd personen werkzaam in het kader Uiteraard was het C D K een enorme papierasswinkel.

Het stond in vast contact, enerzijds met de Duitse en

Nederlandse centrale overheid, anderzijds met de meer dan duizend distributiediensten en agentschappen Van begin mei '40 tot eind april '45 gingen van het C D K bijna 10 000 circulaire uit, gemiddeld dus ve In de eerste jaren heerste bij het C D K een vrij strakke discipline.

Getieralkotmmissar Schmidt vond de gehele Nederlandse distributie zo bewonderenswaardig geregeld dat hi In april '43 kwam een van de Duitse hoofdbanen van de Hauptabteilung Ernährung und Landwirtschaft O f deze reis doorgegaan is, weten wij niet. Wel menen wij dat de Duitse banen minder enthousiast Om te beginnen hadden de frequente overvallen die vooral door de z.g.

Knokploegen op de distributiebureaus uitgevoerd waren, de ordelijke gang van zaken toen lelijk verstoort De evacuatie naar Zwolle en Hattem die in '43 voltrokken was², had de dienst geen goed gedaan.

Het kwam toen vaak voor dat functionarissen, ook leidende, afwezig waren omdat zij hun gezinnen, voorzo Gevallen van fraude hadden zich van meet af aan

1 Er bevonden zich onder hen slechts weinig N S B 'ers.

Volgens de directie waren heter in '42 vijf; onder pressie van de N S B kwamen er toen twee bij, 'doch d

159

VE RAR ME ND NEDERLANDbij verscheidene distributiediensten voorgedaan (als regel slaagde Hirschfelds d Naast de op het eigenbelanggerichte fraude was er echter ook een fraude tot ontwikkeling gekomen die alt

Hulp aan Onderduikers, ondervond ook van functionarissen van het C D K zoveel medewerking dat er toen en De zaak was dus deze dat van de zomer van '43 af een deel van de ambtenaren van het C D K het min of meer C D K gebruikte voor illegale doeleinden - de directeur, S. de Hoo, had deze zaak niet meer in de hand.

★

Sybrecht de Hoo, in 1900 in Assen geboren, had in '31 de Koninklijke Marine verlaten en was ambtenaar bij Hij werd capabel geacht, bereikte begin '40 de rang van referendaris en was toen belast met de dagelijkse C D K.

Een andere hoofdambtenaar, dr. W. L.

Groeneveld Meyer, directeur-generaal voor de middenstand, was daar directeur van.

Deze laatste werd, verdacht van tegenwerking, eind '40 op last van de Duitsers ontslagen.

De

Hoo werd de nieuwe directeur.

Dat was van de zijde van Hirschfeld een gelukkige keuze.

Spoedig bleek namelijk dat de Hoo de sterke neiging had

1 In februari '42 zei Hirschfelds perschef op de dagelijkse persconferentie: 'De tijd is zijn al zo verwa Ik herinner nogmaals aan het verzoek om fraude bij de distributiediensten, voorzover ze bij de ambtena

100

CENTRAAL DISTRIBUTIEKANTOOR met N S B 'ers en Duitsers aan te pakken.

Hij voelde zich min of meer de koning van de Nederlandse distributie en vond het kennelijk aangenaam dat Woltersom en diens naaste medewerkers, alsmede van tijd tot tijd ook diverse particulieren van extra-bo Hij protesteerde nimmer tegen dergelijke instructies en oefende bovendien weinig controle uit op de uitvoer N S B in de zomer van '44 zelfs de beschikking kon krijgen over door het

C D K afgetekende blanco-toewijzingen waarop die distributiedienst dus zelfde hoeveelheden kon invullen De correspondentie die de Hoo in zijn functie met die N S B - en Duitse instanties voerde, was van zijn In 'foute' kringen werd het dan ook vrij algemeen bekend dat een persoonlijk briefje aan de directeur van Kriminalkommissar Oskar Wenzky, aan de Rijksresearchcentrale toegevoegd was: 'En last not least, verzo Waar nu K.K.

Wenzky kort geleden toch banden voor dienstpersoneel heeft aangevraagd aan uw bureau, moge ik u verzoeken Is voor u een klein kunstje.

Van dit laatste weet K. K.

Wenzky niets afhoort Directeur.

Dit is slechts een onderhands verzoekje van de heer Gerhards . . . U doet mij er tevens een groot genoeg

1 Nog een voorbeeld dat buiten het C D K ligt: de N S B 'er die directeur was van het rijksinkoopbureau

161

VEERMEER NEDERLAND allerlei soort 'gevalletjes' die zo dagelijks aan de orde zijn. U begrijpt mij hoe Ik hoor u al mopperen, daar komt die 'rotvent' ook weer aandraven met denodige verzoekjes.

Ja Hooggeachte Directeur, het is waar, maar wij moeten toch trachten met elkaar door de moeilijkheden heen Meer dan ooit zijn wij thans op elkaar aangewezen.

Nach dem Krieg werden wir auch leiden müssen.'1

Niet alleen echter legde de Hoo jegens Duitsers en N S B 'ers een stuitende gedienstigheid aan de dag, h In augustus '43 verzocht hij Harster, de Beehlshaber der Sicherheitspolizei und des SD, om een document dat hem het recht zou geven, onmiddellijk de hulp van al Duitse, in te roepen.

In februari '44 ging hij er prat op dat hij het C D K perfect tegen overvallen beschermd had.

Begin mei '44 klaagde hij zich persoonlijk bij Rauter o.m. over het feit dat Harsters opvolger hem, te Nederlanders die bij de 'SD' in dienst waren; de Hoo vond dat blijkbaar onvoorzichtig.

Kortom, tot laat in de bezetting trachtte hij bij de bezetterin het gevlug te komen.

Het paste in dat kader dat hij, wanneer hij schriftelijk bezwaren uitte tegen het beleid van Hirschfelds Louwes om een afschrift zond van zijn brieven.

Daar kwam dan nog bij dat hij zichzelf en ook wel organisaties waar hij lid van was (de Jagersvereniging Het personeel vraagt zich af, waarom sommige firma's . . . zo goed worden bedacht.

Indertijd werden door de vele leidende figuren royaal geleefd . . .

Alles tezamen genomen is het

CD K nu niet bepaald een modelkantoor '2 en de Hoo (hij was in die tijd, van een ziekte herstellend, ken Dat laatste was onmiddellijk na het bezoek van die ambtenaar gebleken, ja op de dag zelf waarop deze zijn Die dag was de Hoo namelijk in Zwolle

1 Brief, 10 jan. 1944, van K.L. aan S. de Hoo (Doc I-739, a-4). 2 H. L.

162

CENTRAAL DISTRIBUTIEKANTOORgaan lunchen met drie directieleden van enkele van de grootste levensmiddelen Enschedé &

Zonen te Haarlem (waar alle distributiebescheiden gedrukt werden).

Deheren dronken een kruik jenever leeg plus enkele flessen wijn - toen de

Hoo in Hattem terugkwam, was hij dronken.

Drie dagen later kreeg hij van Hirschfeld ziekteverlof.

Met de leiding van het C D K werd toen een van de adjunct-directeuren belast, J. G.

Japikse.

De Hoo deelde nadien

Rauters adjudant schriftelijk mee dat hij ontslag wilde nemen. 'Kommt nicht in Frage!' schreef Seyss-Inq

Sie für den gesamten Apparat des Zentral Distributiekantors in vollem Umfang verantwortlich.¹ De beze

Dat had óók te maken met de diensten die de Hoo bij de invoering en uitreiking van de z.g. tweede distri

Bij de levensmiddelendistributie werd er met kracht naar gestreefd, medische normen die voor een voldoende
Reeds enkele weken na de capitulatie had Hirschfeld uit de bestaande Gezondheidsraad een aparte commissie
Voedingsraad, die de samenstelling van het distributiepakket regelmatig analyseerde en op grond daarvan
Als dat maar enigszins mogelijk was, werden die adviezen door Hirschfeld en Louwes overgenomen en door h
Toen Engeland in '40 de strijd bleek voort te zetten, stond voor Hirschfeld en Louwes vast dat men op de
Het getuigde hunnerzijds dan ook van vooruitziend beleid dat zij in de loop van '40 de stoot gaven tot de
Centrale Keukens.

Een enquête bij de gemeentebesturen toonde aan dat er in het land reeds 98 bestonden die elk een capacite

1 Brief, 12 juni 1944, van Rauter aan S. de Hoo (HSSuPF, 7 A, b).

163

VERA RM EN D NEDERLAND per dag.

Dat aantal werd te gering geacht.

Om te beginnen werden in de grote steden 50 keukens ingericht die eind '40 al een gemiddelde capaciteit
Die 50 keukens werden door de gemeenten geëxploiteerd.

Daarnaast richtte het rijksbureau voor de voedselvoorziening in oorlogstijd ook een aantal eigen keukens
Eind '43 was de capaciteit van alle keukens tezamen 2 miljoen porties per dag. Ook dat was

Hirschfeld en Louwes nog te weinig; de Geallieerde invasie kon, zo meenden zij, heel wel een toestand
Van alle hotels, restaurants en inrichtingen als ziekenhuizen werden de kookinstallaties geïnventariseerd
Voor alle zekerheid werd tevens door het rijksbureau een hoeveelheid van ca. 20 000 kilo stampot ingev
De centrale keukens (die in de hongerwinter een volstrekt onmisbare functie zouden vervullen) wierpen i
Daarbij trad bij de afnemers wel een zekere 'stampotmoetheid' op - het aantal personen die aldus hun war
Daar kwam bij dat men, teneinde tot deze voorziening toegelaten te worden, een kaart moest bezitten die
Nederlandse Volksdienst afgegeven werd.

In Amsterdam waren er begin '41 zesduizend personen die van de centrale keukens gebruik maakten; zomer '41
Gelijk gezegd: voor dat voedsel moest men een deel van zijn bonnen afstaan.

De bijvoeding die door een aantal grote bedrijven verstrekt werd, was boneloos en daarvan profiteerden in

164

CENTRALE KEUKENSenige gegevens die wij bezitten) meer dan driehonderdzeventigduizend arbeiders.¹ Van he
Riistungsinspektion van speciaal belang geacht werden.

Kortweg kan men stellen dat in de zomer van '42 ruim één derde van de in de industrie werkzame arbeiders
Een niet onaanzienlijk deel van de werkende bevolking had recht op die verhoogde rantsoenen: in '43 bij
Om te beginnen, maar dan louter naar verhouding, de kinderen; zij ontvingen rantsoenen die hun voeding be
De rantsoenen voor kinderen van 0 tot 3 jaar waren (dit alles uiteraard tot de hongerwinter begon) ongev
Maar kwamen die relatief verhoogde rantsoenen inderdaad steeds bij de kinderen terecht? In menig gezin w
Ook zieken en zwangere vrouwen hadden recht op extra-rantsoenen.

Die moesten door de behandelende huisarts aangevraagd worden; de aanvraag werd beoordeeld, in gemeenten
Geschillen werden aan provinciale medische commissies voorgelegd. Wij hebben geen gegevens over de aant
Banning, opgedrongen had.

Hij had vooral geconstateerd dat die ziekenrantsoenen, eenmaal toegekend, min of meer automatisch verle
De

1 Bijna 70 000 porties kwamen daarbij uit de centrale keukens van de overheid.

VE RA RM EN D NEDERLAND

Hauptabteilung Ernährung und Landivirtschaft gaf toen medio juni het C D Kde nogal botte instructie om :
In de paragraaf over de visserij vermeldde wij dat opvarenden van vissersschepen uit de vangst officieel
Die bepaling was onderdeel van een algemene regeling: allen die agrarische bedrijven bezaten dan wel als
Landarbeiders mochten van de oogst 25 kilo graan per gezinslid mee naar huis nemen.

Wij achten het aannemelijk dat men zich van al die normen van den beginne af niet veel aangetrokken heeft.
Boeren en landarbeiders hadden in elk geval voldoende te eten en van de daling van de brood- en vleesran-
Zij konden zich allen over de gehele linie of ten aanzien van bepaalde levensmiddelen redden.

Het waren de 'normale verbruikers' die in moeilijkheden kwamen.

Wanneer wij dan ook in een vorige paragraaf ten naaste bij becijferd hebben, welk percentage van de in-
Van de aardappel oogst '43-'44 bijvoorbeeld, ging vermoedelijk 15 % naar de Duitse toe, maar van het
In verschillende publikaties zijn overzichten verschenen van het verloop van de rantsoenen en is ook nau-

1 Een overzicht van de totale hoeveelheden der gerantsoeneerde levensmiddelen per persoon en per jaar, n-
Dezelfde uitgave bevat op pag. 257 een tabel van het totale jaarlijkse verbruik op de consumentenbonnen.