

VAJE_03

March 16, 2023

1 Uvozim knjižice

```
[1]: import spacy
import nltk
from tqdm.notebook import tqdm
import networkx as nx
import re #regular expression
```

2 Priprava besedila

```
[2]: #1 Naložimo besedilo
book=open("GOT.txt",encoding="utf8").read()

#2 Čiščenje besedila
book=re.sub(r"\"", "", book)
book=re.sub(r"\"", "", book)

#3 Diskretizacija besedila
stavki=nltk.sent_tokenize(book)
print(f"Število stavkov v besedilu je {len(stavki)}")
```

Število stavkov v besedilu je 89040

3 NER (Named Entities Recognition): Spacy

```
[3]: nlp=spacy.load("en_core_web_md")
```

```
[5]: liki={}

N=int(0.2*len(stavki))

for s in tqdm(range(N),total=N):
    doc=nlp(stavki[s])
    for ent in doc.ents:
        if ent.label_=="PERSON":
            oseba=ent.text
```

```

if oseba not in liki:
    liki[oseba]=[]
    #print(s,oseba, ent.start_char, ent.end_char, ent.label_)
    liki[oseba].append(s)

```

0%| | 0/17808 [00:00<?, ?it/s]

```

[6]: for oseba in list(liki.keys())[:10]:
    print(oseba,liki[oseba])

```

```

Gared [0, 11, 20, 25, 26, 49, 54, 58, 111, 114, 126, 134, 138, 183, 187, 195,
196, 212, 215, 217, 218]
Ser Waymar Royce [3, 23, 47, 160, 173, 236, 301, 345, 373, 5980]
Will [16, 43, 56, 75, 86, 139, 144, 170, 201, 215, 243, 316]
Royce [16, 12986]
Wall [32, 39, 45, 143, 147, 752, 810, 1377, 1614, 1635, 1639, 1846, 2142, 2485,
2970, 3073, 3853, 3992, 5231, 5232, 5706, 5822, 5836, 5848, 5902, 5911, 5930,
5932, 5969, 5969, 6510, 6639, 6656, 6715, 6720, 6735, 6749, 6752, 6756, 6834,
6836, 6838, 6838, 7621, 7636, 7897, 7906, 7922, 8235, 8397, 8399, 8438, 8522,
8921, 9167, 10049, 10833, 10836, 10841, 10848, 10875, 11473, 12663, 13801,
13802, 13805, 13806, 13807, 13851, 13866, 14650, 14653, 15977, 16112, 16154,
16167, 16169, 17131, 17150, 17155, 17165, 17194, 17209, 17277, 17311, 17322,
17345, 17500, 17511]
Ser Waymar [51, 112, 126, 129, 133, 177, 213, 249, 253, 285, 312, 322, 328, 335,
342]
Mallisters [73, 73, 8992]
ridge [75, 179]
m'lord [108, 4349, 7894, 9128, 10847, 11923, 12645, 14361, 14411, 14417, 14418,
14970, 15796, 15810, 16222]
Maester Aemon [135, 369, 5710, 5794, 6577, 6581, 6583, 6649, 8396, 8571, 13851,
13872, 13879, 13900, 13906, 13912, 13921, 13929, 13938, 13951, 13966, 13980,
13993, 15969, 15982, 16080, 16084, 16103, 17139, 17276, 17334, 17356, 17588]

```

4 Ustvarjanje sociograma

4.1 Mrežo poselimo z vozlišli, ki ponazarjajo posamezne like

```

[8]: G=nx.Graph()
    for oseba in liki:
        G.add_node(oseba)

    print(f"Število oseb = {len(G)}")

    nodes=G.nodes()

```

Število oseb = 610

```
[15]: for p1 in tqdm(G,total=len(G)):
        x=np.array(like[p1])
        for p2 in G:
            y=np.array(like[p2])
            if p1!=p2:
                xy=np.array([np.min(np.abs(y-xi)) for xi in x])
                wij=len(np.where(xy<5)[0])
                if wij>3:
                    #print(p1,p2,wij)
                    G.add_edge(p1,p2,weight=wij)
```

0%| | 0/158 [00:00<?, ?it/s]

```
[16]: remove_nodes=[]
for node in G:
    if G.degree(node)<1:
        remove_nodes.append(node)

for node in remove_nodes:
    G.remove_node(node)

nx.write_graphml(G,"sociogram.graphml")
```

5 Kateri hiši pripadajo liki?

```
[34]: appendix=open("appendix.txt",encoding="utf8").read().split("Other Houses Great_
↪And Small")[1]
appendix=re.sub("\n\n+","\n",appendix)

house=None
pattern = r"\b[A-Z]+\b"
for p in appendix.split("\n")[1::]:
    if "HOUSE" in p:
        house=p.split("HOUSE")[1]
    elif p[0]=="-":
        for pp in p.split(","):
            matches = ' '.join(re.findall(pattern, pp))
            if matches!='':
                print(house,matches)
```

```
ARRAYN MARILLION
ARRAYN MAESTER COLEMON
ARRAYN SER MARWYN BELMORE
ARRAYN MORD
ARRAYN LORD NESTOR ROYCE
ARRAYN SER ALBAR
ARRAYN MYRANDA
```

ARRYN MYA STONE
ARRYN I
ARRYN LORD YOHNN ROYCE
ARRYN BRONZE YOHNN
ARRYN SER ANDAR
ARRYN SER ROBAR
ARRYN SER WAYMAR
ARRYN SER LYN CORBRAY
ARRYN MYCHEL REDFORT
ARRYN LADY ANYA WAYNWOOD
ARRYN SER MORTON
ARRYN SER DONNEL
ARRYN EON HUNTER
ARRYN HORTON REDFORT
FLORENT ALEKYNE
FLORENT MELESSA
FLORENT RHEA
FLORENT SER AXELL
FLORENT SER RYAM
FLORENT QUEEN SELYSE
FLORENT SER IMRY
FLORENT SER ERREN
FLORENT SER COLIN
FLORENT DELENA
FLORENT SER HOSMAN NORCROSS
FLORENT EDRIC STORM
FLORENT I
FLORENT ALESTIR NORCROSS
FLORENT RENLY NORCROSS
FLORENT MAESTER OMER
FLORENT MERRELL
FLORENT RYLENE
FREY SER RYMAN
FREY EDWYN
FREY WALDA
FREY WALDER
FREY BLACK WALDER
FREY PETYR
FREY PETYR PIMPLE
FREY PERRA
FREY AEGON
FREY JINGLEBELL
FREY MAEGELLE
FREY MARIANNE
FREY WALDER VANCE
FREY PATREK VANCE
FREY WALTON
FREY STEFFON

FREY THE SWEET
FREY WALDA
FREY FAIR WALDA
FREY BRYAN
FREY SER EMMON
FREY SER CLEOS
FREY TYWIN
FREY WILLEM
FREY SER LYONEL
FREY TION
FREY WALDER
FREY RED WALDER
FREY SER AENYS
FREY AEGON BLOODBORN
FREY RHAEGAR
FREY ROBERT
FREY WALDA
FREY WHITE WALDA
FREY JONOS
FREY PERRIANE
FREY SER HARYS HAIGH
FREY WALDER HAIGH
FREY SER DONNEL HAIGH
FREY ALYN HAIGH
FREY LADY CYRENNIA
FREY SER JARED
FREY SER TYTOS
FREY ZIA
FREY ZACHERY
FREY KYRA
FREY WALDER GOODBROOK
FREY JEYNE GOODBROOK
FREY SEPTON LUCEON
FREY LADY AMAREI SER HOSTEEN
FREY SER ARWOOD
FREY RYELLA
FREY ANDROW ALYN
FREY LADY LYTHENE
FREY ELYANA
FREY RICKARD WYLDE
FREY SER DAMON VYPREN
FREY SYMOND
FREY ALESANDER
FREY ALYX
FREY BRADAMAR
FREY SER DANWELL
FREY MERRETT
FREY AMEREI

FREY AMI
FREY WALDA
FREY FAT WALDA
FREY MARISSA
FREY WALDER
FREY LITTLE WALDER
FREY SER GEREYMY
FREY SANDOR
FREY CYNTHIA
FREY SER RAYMUND
FREY ROBERT
FREY MALWYN
FREY SERRA SARRA
FREY CERSEI
FREY LITTLE BEE
FREY LADY ALYSSA
FREY LOTHAR
FREY LAME LOTHAR
FREY TYSANE
FREY WALDA
FREY EMERLEI
FREY SER JAMMOS
FREY WALDER
FREY BIG WALDER
FREY DICKON MATHIS
FREY SER WHALEN
FREY HOSTER
FREY MERIANNE
FREY MERRY
FREY LADY MORYA
FREY ROBERT BRAX
FREY WALDER BRAX
FREY JON BRAX
FREY TYTA
FREY TYTA THE MAID
FREY LADY SARYA
FREY LADY BETHANY SER PERWYN
FREY SER BENFREY
FREY DELLA
FREY DEAF DELLA
FREY OSMUND
FREY MAESTER WILLAMEN
FREY OLYVAR
FREY ROSLIN
FREY LADY ANNARA ARWYN
FREY WENDEL
FREY COLMAR
FREY WALTIR

FREY TYR
FREY ELMAR
FREY SHIREI
FREY LADY JOYEUSE
FREY WALDER RIVERS
FREY BASTARD WALDER
FREY SER AEMON RIVERS
FREY WALDA RIVERS
FREY MAESTER MELWYS
FREY JEYNERIVERS
FREY MARTYNRIVERS
FREY RYGERRIVERS
FREY RONELRIVERS
FREY MELLARARIVERS
LANNISTER PRINCESS MYRCELLA BARATHEON
LANNISTER PRINCE TOMMEN BARATHEON
LANNISTER TYRION
LANNISTER THE IMP
LANNISTER HALFMAN
LANNISTER SER KEVAN
LANNISTER DORNA
LANNISTER SER LANCEL
LANNISTER WILLEM
LANNISTER MARTYN
LANNISTER JANEI
LANNISTER GENNA
LANNISTER SER CLEOS FREY
LANNISTER SER LYONEL
LANNISTER TION FREY
LANNISTER WALDER
LANNISTER RED WALDER
LANNISTER SER TYGETT
LANNISTER DARLESSA
LANNISTER TYREK
LANNISTER GERION
LANNISTER JOY
LANNISTER SER STAFFORD LANNISTER
LANNISTER CERENNA MYRIELLE
LANNISTER SER DAVEN
LANNISTER SER DAMION LANNISTER
LANNISTER SER LUCION
LANNISTER LANNA
LANNISTER MARGOT
LANNISTER MAESTER CREYLEN
LANNISTER VYLARR
LANNISTER LUM RED LESTER
LANNISTER WHITESMILE WAT
LANNISTER SER BENEDICT BROOM

LANNISTER DAMON MARBRAND
LANNISTER SER ADDAM MARBRAND
LANNISTER ROLAND CRAKEHALL
LANNISTER SER BURTON CRAKEHALL
LANNISTER SER TYBOLT CRAKEHALL
LANNISTER SER LYLE CRAKEHALL
LANNISTER STRONGBOAR
LANNISTER SER MERLON CRAKEHALL
LANNISTER ANDROS BRAXI
LANNISTER SER RUPERT BRAX
LANNISTER SER TYTOS BRAX
LANNISTER SER ROBERT BRAX
LANNISTER SER FLEMENT BRAX
LANNISTER LORD LEO LEFFORD
LANNISTER REGENARD ESTREN
LANNISTER GAWEN WESTERLING
LANNISTER LADY SYBELL
LANNISTER SER ROLPH SPICER
LANNISTER SER SAMWELL SPICER
LANNISTER SER RAYNALD WESTERLING
LANNISTER JEYNE
LANNISTER ELEYNA
LANNISTER ROLLAM
LANNISTER LEWYS LYDDEN
LANNISTER LORD ANTARIO JAST
LANNISTER LORD PHILIP PLUMM
LANNISTER SER DENNIS PLUMM
LANNISTER SER PETER PLUMM
LANNISTER SER HARWYN PLUMM
LANNISTER HARDSTONE
LANNISTER QUENTEN BANEFORT
LANNISTER SER HARYS SWYFT
LANNISTER SER STEFFON SWYFT
LANNISTER JOANNA
LANNISTER SHIERLE
LANNISTER SER FORLEY PRESTER
LANNISTER SER GARTH GREENFIELD
LANNISTER SER LYMOND VIKARY
LANNISTER LORD SELMOND STACKSPEAR
LANNISTER SER STEFFON STACKSPEAR
LANNISTER SER ALYN STACKSPEAR
LANNISTER TERRENCE KENNING
LANNISTER SER KENNOS OF KAYCE
LANNISTER SER GREGOR CLEGANE
LANNISTER POLLIVER
LANNISTER CHISWYCK
LANNISTER RAFF THE SWEETLING
LANNISTER DUNSEN

LANNISTER THE TICKLER
LANNISTER SER AMORY LORCH
MARTELL PRINCESS ARIANNE
MARTELL PRINCE QUENTYN
MARTELL PRINCE TRYSTANE
MARTELL PRINCESS ELIA
MARTELL PRINCESS RHAENYS
MARTELL PRINCE AEGON
MARTELL PRINCE OBERYN
MARTELL THE RED VIPER
MARTELL ELLARIA SAND
MARTELL OBARA
MARTELL NYMERIA
MARTELL TYENE
MARTELL SARELLA
MARTELL ELIA
MARTELL OBELLA
MARTELL DOREA
MARTELL LOREZA
MARTELL THE SAND SNAKES
MARTELL HARMEN ULLER
MARTELL SER ULVVYCK ULLER
MARTELL SER RYON ALLYRION
MARTELL SER DAEMON SAND
MARTELL DAGOS MANWOODY
MARTELL MORS DICKON
MARTELL SER MYLES MANWOODY
MARTELL SER ARRON QORGYLE
MARTELL SER DEZIEL DALT
MARTELL MYRIA JORDAYNE
MARTELL LARRA BLACKMONT
MARTELL JYNESSA BLACKMONT
MARTELL PERROS BLACKMONT
MARTELL AREO HOTAH
MARTELL MAESTER CALEOTTE
MARTELL EDRIC DAYNE
MARTELL DELONNE ALLYRION
MARTELL DAGOS MANWOODY
MARTELL LARRA BLACKMONT
MARTELL TREMOND GARGALEN
MARTELL ANDERS YRONWOOD
MARTELL NYMELLA TOLAND
TULLY CATELYN
TULLY ROBB STARK
TULLY SANSA STARK
TULLY ARYA STARK
TULLY BRANDON STARK
TULLY RICKON STARK

TULLY LYSA
TULLY ROBERT
TULLY SER EDMURE
TULLY SER MARQ PIPER
TULLY LORD LYMOND GOODBROOK
TULLY SER RONALD VANCE
TULLY THE BAD
TULLY SER HUGO
TULLY SER ELLERY
TULLY KIRTH
TULLY PATREK MALLISTER
TULLY LUCAS BLACKWOOD
TULLY SER PERWYN FREY
TULLY TRISTAN RYGER
TULLY SER ROBERT PAEGE
TULLY SER BRYNDEN
TULLY MAESTER VYMAN
TULLY SER DESMOND GRELL
TULLY SER ROBIN RYGER
TULLY LONG LEW
TULLY ELWOOD
TULLY DELP
TULLY UTHERYDES WAYN
TULLY RYMUND THE RHYMER
TULLY JONOS BRACKEN
TULLY JASON MALLISTER
TULLY WALDER FREY
TULLY CLEMENT PIPER
TULLY KARYL VANCE
TULLY NORBERT VANCE
TULLY THEOMAR SMALLWOOD
TULLY LADY RAVELLA
TULLY CARELLEN
TULLY WILLIAM MOOTON
TULLY SHELLA WHENT
TULLY SER HALMON PAEGE
TULLY TYTOS BLACKWOOD
TYRELL WILLAS
TYRELL SER GARLAN
TYRELL THE GALLANT
TYRELL LADY LEONETTE
TYRELL SER LORAS
TYRELL MARGAERY
TYRELL I
TYRELL MEGGA
TYRELL ALLA
TYRELL ELINOR TYRELL
TYRELL ALYN AMBROSE

TYRELL LADY ALYSANNE BULWER
TYRELL MEREDYTH CRANE
TYRELL MERRY
TYRELL TAENA OF MYR
TYRELL LORD ORTON MERRYWEATHER
TYRELL LADY ALYCE GRACEFORD
TYRELL SEPTA NYSTERICA
TYRELL LADY OLENNA
TYRELL ARRYK ERRYK
TYRELL LEFT RIGHT
TYRELL LADY MINA
TYRELL SER HORAS REDWYNE
TYRELL HORROR
TYRELL SER HOBBER REDWYNE
TYRELL SLOBBER
TYRELL DESMERA REDWYNE
TYRELL LADY JANNA
TYRELL GARTH
TYRELL THE GROSS
TYRELL GARSE GARRETT FLOWERS
TYRELL SER MORYN
TYRELL SER LUTHOR
TYRELL SER THEODORE
TYRELL ELINOR
TYRELL LUTHOR
TYRELL MAESTER MEDWICK
TYRELL OLENE
TYRELL LEO
TYRELL LEO THE LAZY
TYRELL MAESTER GORMON
TYRELL SER QUENTIN
TYRELL SER OLYMER
TYRELL RAYMUND RICKARD
TYRELL MEGGA
TYRELL MAESTER NORMUND
TYRELL SER VICTOR
TYRELL VICTARIA
TYRELL LADY ALYSANNE BULWER
TYRELL SER LEO
TYRELL ALLA LEONA
TYRELL LYONEL
TYRELL LUCAS
TYRELL LORENT
TYRELL MAESTER LOMYS
TYRELL IGON VYRWEL
TYRELL SER VORTIMER CRANE
TYRELL BUTTERBUMPS
TYRELL RANDYLL TARLY

TYRELL PAXTER REDWYNE
TYRELL ARWYN OAKHEART
TYRELL MATHIS ROWAN
TYRELL ALESTER FLORENT
TYRELL LEYTON HIGHTOWER
TYRELL ORTON MERRYWEATHER
TYRELL LORD ARTHUR AMBROSE
TYRELL SER MARK MULLENDORE
TYRELL SER JON FOSSOWAY
TYRELL SER TANTON FOSSOWAY
TYRELL EDDISON TOLLETT
TYRELL DOLOROUS EDD
TYRELL THOREN SMALLWOOD
TYRELL DYWEN
TYRELL DIRK
TYRELL SOFTFOOT
TYRELL GRENN
TYRELL BEDWYCK GIANT
TYRELL OLLO LOPHAND
TYRELL GRUBBS
TYRELL BERNARR BROWN BERNARR
TYRELL BERNARR BLACK BERNARR
TYRELL TIM STONE
TYRELL ULMER OF KINGSWOOD
TYRELL GARTH GREYFEATHER
TYRELL GARTH OF GREENAWAY
TYRELL GARTH OF OLDTOWN
TYRELL ALAN OF ROSBY
TYRELL RONNEL HARCLAY
TYRELL AETHAN
TYRELL RYLES
TYRELL MAWNEY
TYRELL JARMEN BUCKWELL
TYRELL BANNEN
TYRELL KEDGE WHITEYE
TYRELL TUMBERJON
TYRELL FORNIO
TYRELL GOADY
TYRELL SER OTTYN WYTHERS
TYRELL SER MALADOR LOCKE
TYRELL DONNEL HILL
TYRELL SWEET DONNEL
TYRELL HAKE
TYRELL CHETT
TYRELL SAMWELL TARLY
TYRELL SER PIGGY
TYRELL LARK THE SISTERMAN
TYRELL ROLLEY OF SISTERTON

TYRELL CLUBFOOT KARL
TYRELL MASLYN
TYRELL SMALL PAUL
TYRELL SAWWOOD
TYRELL LEFT HAND LEW
TYRELL ORPHAN OSS
TYRELL MUTTERING BILL
TYRELL QHORIN HALFHAND
TYRELL SQUIRE DALBRIDGE
TYRELL EGGEN
TYRELL STONESNAKE
TYRELL BLANE
TYRELL SER BYAM FLINT
TYRELL BENJEN STARK
TYRELL SER WYNTON STOUT
TYRELL SER ALADALE WYNCH
TYRELL PYPAR
TYRELL DEAF DICK FOLLARD
TYRELL HAIRY HAL
TYRELL BLACK JACK BULWER
TYRELL ELRON
TYRELL MATTHAR
TYRELL OTHELL YARWYCK
TYRELL SPARE BOOT
TYRELL YOUNG HENLY
TYRELL HAIDER
TYRELL ALBETT
TYRELL KEGS
TYRELL SPOTTED PATE OF MAIDENPOOL
TYRELL DONAL NOYE
TYRELL THREE FINGER HOBBS
TYRELL TIM TANGLETONGUE
TYRELL EASY
TYRELL MULLY
TYRELL OLD HENLY
TYRELL CUGEN
TYRELL RED ALYN OF THE ROSEWOOD
TYRELL JEREN
TYRELL SEPTON CELLADOR
TYRELL SER ENDREW TARTH
TYRELL RAST
TYRELL ARRON
TYRELL EMRICK
TYRELL SATIN
TYRELL HOP ROBIN
TYRELL CONWY
TYRELL GUEREN
TYRELL SER ALLISER THORNE

TYRELL JANOS SLYNT
TYRELL SER GLENDON HEWETT
TYRELL DAREON
TYRELL IRON EMMETT
TYRELL MAESTER MULLIN
TYRELL BERIC DONDARRION
TYRELL THE LIGHTNING LORD
TYRELL THOROS OF MYR
TYRELL EDRIC DAYNE
TYRELL LEM
TYRELL LEM LEMONCLOAK
TYRELL HARWIN
TYRELL GREENBEARD
TYRELL TOM OF SEVENSTREAMS
TYRELL TOM SEVENSTRINGS TOM SEVENS
TYRELL ANGUY THE ARCHER
TYRELL JACK BE LUCKY
TYRELL THE MAD HUNTSMAN
TYRELL KYLE
TYRELL NOTCH
TYRELL DENNETT
TYRELL MERRIT O MOONTOWN
TYRELL WATTY THE MILLER
TYRELL LIKELY LUKE
TYRELL MUDGE
TYRELL BEARDLESS DICK
TYRELL SHARNA
TYRELL HUSBAND
TYRELL BOY
TYRELL TANSY
TYRELL ALYCE
TYRELL CASS
TYRELL LANNA
TYRELL JYZENE
TYRELL HELLY
TYRELL BELLA
TYRELL LADY RAVELLA
TYRELL LORD LYMOND LYCHESTER
TYRELL MAESTER ROONE
TYRELL HARMA
TYRELL DOGSHEAD
TYRELL THE LORD OF BONES
TYRELL RATTLESHIRT
TYRELL YGRITTE
TYRELL RYK
TYRELL LONGSPEAR
TYRELL RAGWYLE
TYRELL LENYL

TYRELL JON SNOW
TYRELL GHOST
TYRELL STYR
TYRELL JARL
TYRELL GRIGG THE GOAT
TYRELL ERROK
TYRELL QUORT
TYRELL BODGER
TYRELL DEL
TYRELL BIG BOIL
TYRELL HEMPEN DAN
TYRELL HENK THE HELM
TYRELL LENN
TYRELL TOEFINGER
TYRELL STONE THUMBS
TYRELL TORMUND
TYRELL GIANTSBANE
TYRELL TALL TALKER
TYRELL HORN BLOWER
TYRELL BREAKER OF ICE
TYRELL THUNDERFIST
TYRELL HUSBAND TO BEARS
TYRELL SPEAKER TO GODS
TYRELL FATHER OF HOSTS
TYRELL TOREGG THE TALL
TYRELL TORWYRD THE TAME
TYRELL DORMUND
TYRELL DRYN
TYRELL MUNDA
TYRELL ORELL
TYRELL ORELL THE EAGLE
TYRELL MAG MAR TUN DOH WEG
TYRELL MAG THE MIGHTY
TYRELL VARAMYR SIXSKINS
TYRELL THE WEEPER
TYRELL ALFYN CROWKILLER
TYRELL DYAH
TYRELL FERNY
TYRELL NELLA

[]: