

Metode 3, Porteføljeopgave 2

Udleveres tirsdag d. 6-11 kl. 14 – skal afleveres i BB senest tirsdag d. 20-11 kl. 16

Sideomfang: 4-8 normalsider (eksklusiv figurer og henvisninger).

NB. Alle spørgsmål skal forsøges besvaret! – der skal selvfølgelig anføres kildehenvisninger, gerne i MLA-format.

Der skal afleveres individuelle besvarelser. Der må samarbejdes i grupper, men i så fald anføres, hvem man har arbejdet sammen med.

Opgaver kan besvares med følgende tools: *Ngram-viewer* (Google Ngram eller Smurf), *Voyant Tools* eller i *Python* med en Jupyter-Notebook eller et selvstændigt script.

1. **UNDERSØGELSESPØRGSMÅL:** Opstil et undersøgelsesspørgsmål, som du vil besvare vha. en seedlist (se punkt 2), et datasæt (se punkt 3) og et tool (se punkt 4).
2. **SEEDLIST:** Opstil en *seedlist* (dvs. en liste bestående af mindst 5 nøgleord), som skal danne grundlag for din historiske undersøgelse. Listen kan være selvstændigt formuleret, hentet fra undervisning, eller andet relevant materiale, eller oprinde i statistiske egenskaber i et udvalgt datasæt (fx. vores [US Inaugural Address datasæt](#)). Der skal redegøres for hvordan denne *seedlist* kan bidrage med at besvare undersøgelsesspørgsmål, samt hvorfor og hvordan disse nøgleord er valgt.
3. **DATASÆT:** Redegør for det datasæt, som din *seedlist* anvendes på. Hvilke data udgør det empiriske fundament for undersøgelsen, hvilke fordele og hvilke ulemper er der forbundet med datasættet?
4. **TOOL:** Redegør for det anvendte tool. Karakteriser dit tool (brug eventuel dokumentation) og redegør for fordele og ulemper ved at anvende netop dette tool.
5. **FORPROCESSERING:** Redegør for hvilke skridt der er taget for at normalisere data (fx. *casefolding*, stopords filtrering, inkludering af flere former af samme ord i seedlist) og hvorfor disse er anvendt (eller undladt).
6. **RESULTATER:** Beskriv resultater af søgninger med din *seedlist* gennem visualisering og forklaring. I skal ikke fortolke resultaterne i denne sektion, blot beskrive resultaterne (fx. hvordan udvikler jeres nøgleord sig over tid). Vigtigt, visualisering skal inkluderes i opgaven.

7. DISKUSSION: Diskuter resultaterne i konteksten af jeres undersøgelsesspørgsmål og fagviden (*domain knowledge*). Reflekter over metodens anvendelse (fordele og ulemper) samt forslå udkast til en videre undersøgelse.
8. KONKLUSION: Kort sammenfatning af besvarelsen af punkt 1-7.