



The fourth project of the Data Science Bootcamp T5
Unsupervised Learning and Natural Learning Processing
(NLP)

Job Postings

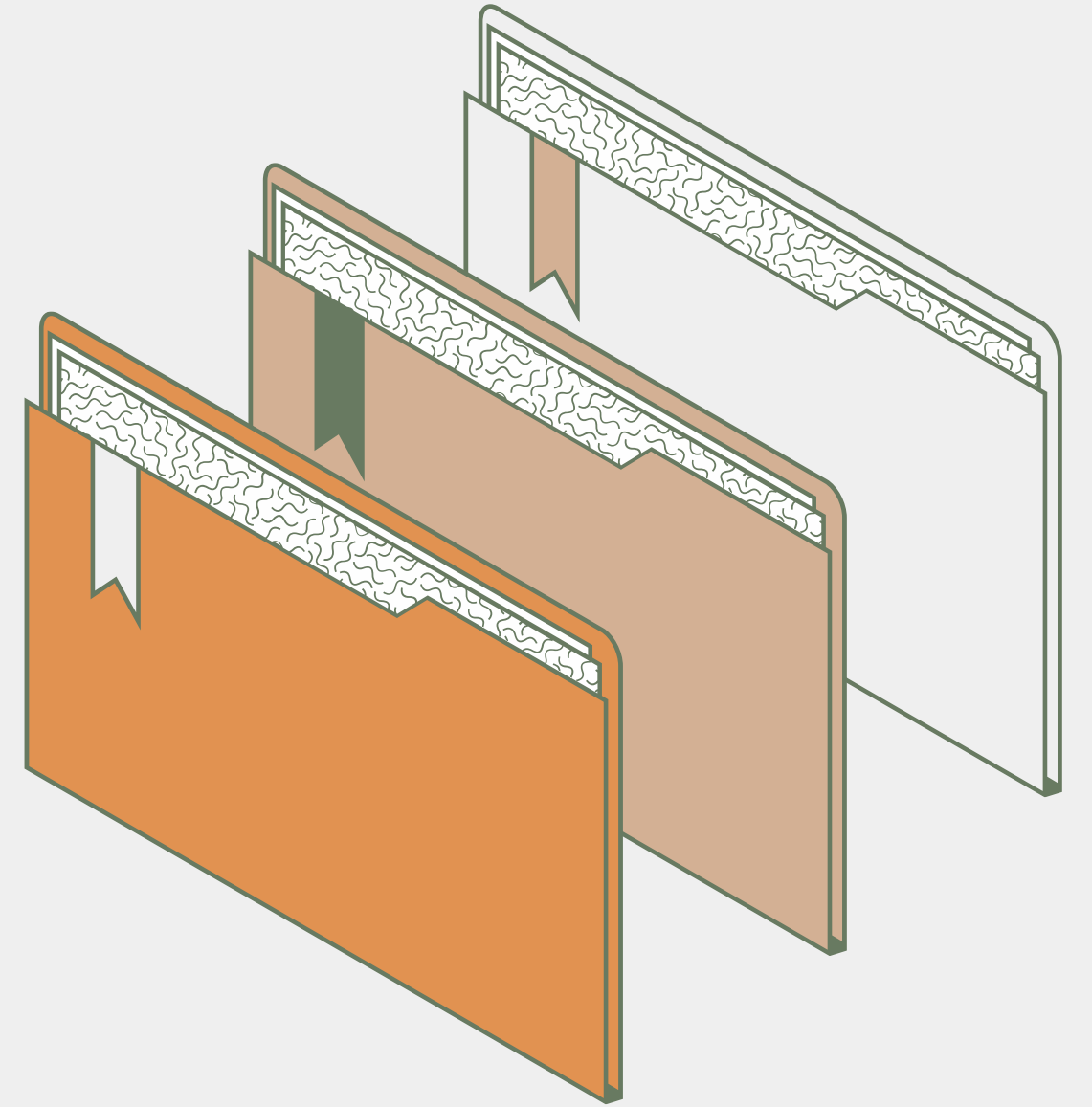
November 2021

By Modhi Alhbrdi and Rahaf Alyousef

Business Objective

Smart employers know their job posting plays an important role in their ability to attract top talent.

In the project, we will classify jobs based on their job description. That will give the requirement employee better performance to the advertiser for an available position, and the right position for the job seeker also, he can find similar jobs in his interest.



Tools

- Libraries:

- Pandas
- NumPy
- Matplotlib
- Seaborn
- WordCloud
- plotly
- bokeh
- Pandas_profiling
- re
- Sklearn
- nltk
- TSNE

- Technologies:

- Python.
- Jupyter Notebook.

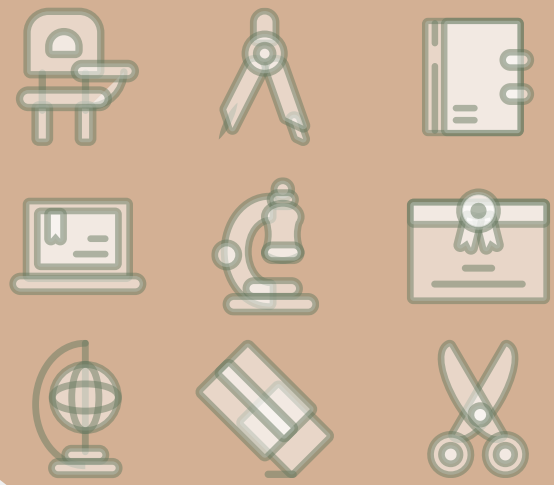


Data Set

- The Job Postings dataset is found on Kaggle.
- The data consists of 2,946 posts records and 9 features.

	Job ID	Agency	Posting Type	Business Title	Civil Service Title	Job Type	Division/Work Unit	Job Description	To Apply
0	87990	DEPARTMENT OF BUSINESS SERV.	Internal	Account Manager	CONTRACT REVIEWER (OFFICE OF L	NaN	Strategy & Analytics	Division of Economic & Financial Opportunity (...)	NaN
1	97899	DEPARTMENT OF BUSINESS SERV.	Internal	EXECUTIVE DIRECTOR, BUSINESS DEVELOPMENT	ADMINISTRATIVE BUSINESS PROMOT	NaN	Tech Talent Pipeline	The New York City Department of Small Business...	In addition to applying through this website, ...
2	132292	NYC HOUSING AUTHORITY	External	Maintenance Worker - Technical Services-Heatin...	MAINTENANCE WORKER	Maintenance & Operations	Management Services Department	Under direct supervision, assist in the routin...	Click the "Apply Now" button.
3	132292	NYC HOUSING AUTHORITY	Internal	Maintenance Worker - Technical Services-Heatin...	MAINTENANCE WORKER	Maintenance & Operations	Management Services Department	Under direct supervision, assist in the routin...	Click the "Apply Now" button.
4	133921	NYC HOUSING AUTHORITY	Internal	Temporary Painter	PAINTER	Maintenance & Operations	Dept of Management & Planning	Responsibilities of selected candidates will i...	Click the "Apply Now" button.

(2946, 9)



Data Preprocessing

1

NLP Preprocessing Pipeline:

- Remove Symbols.
- Remove Punctuations.
- Remove Non-alphabetic Characters.
- Convert to Lowercase.
- Convert text to tokens.
- Remove Stopwords.
- Lemmatizing tokens.

2

Check if there are any missing values then fill it by mode.

3

Check if there are any duplicate values then remove it.

Job Description Before NLP Preprocessing Pipeline

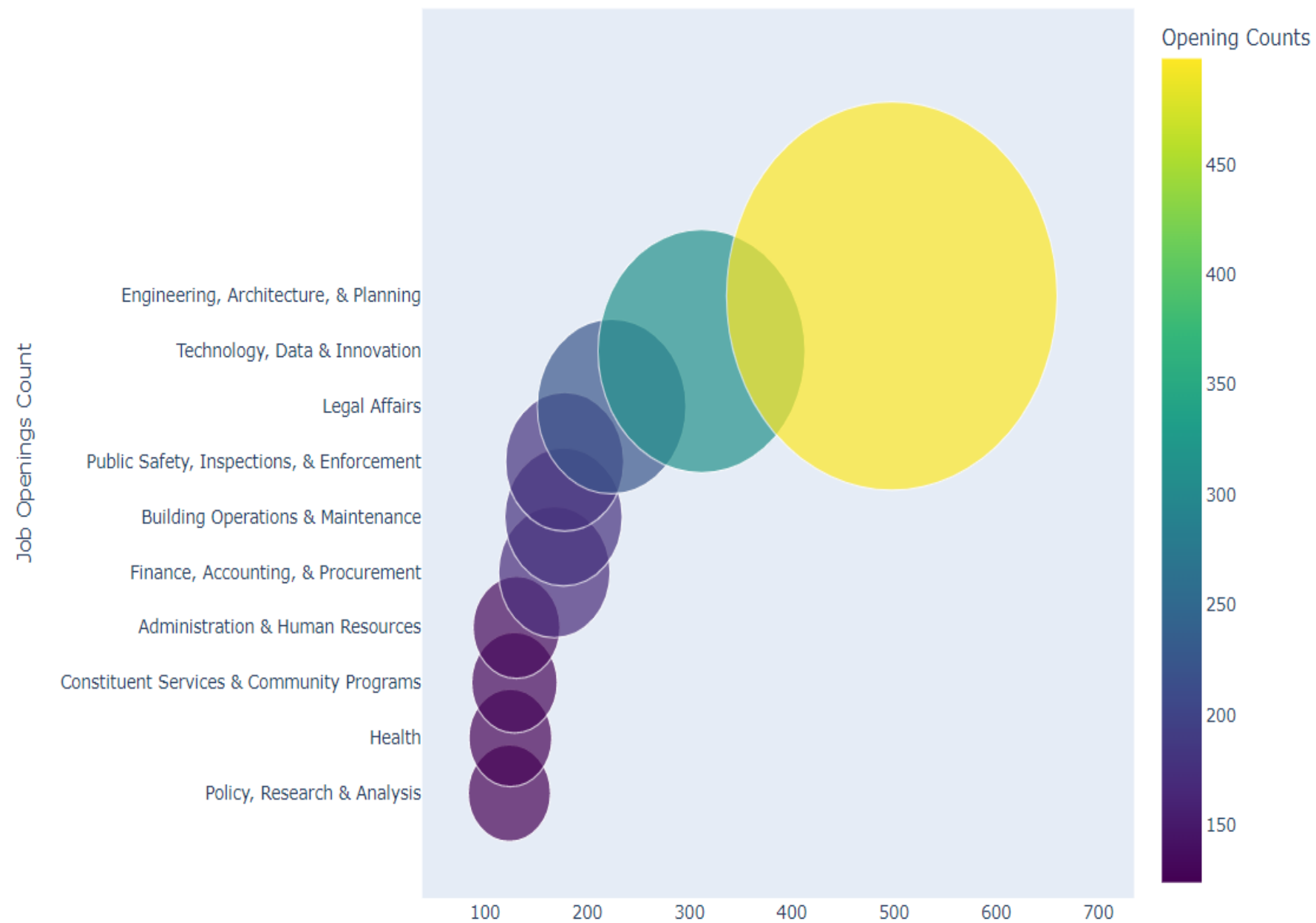
	Job Description
2	Under direct supervision, assist in the routine maintenance operation and repair of public buildings, structures, and the equipment they contain; perform related work. Responsibilities include, but are not limited to the following: 1. Perform minor and major repairs to boilers, burners, vacuum tank, pumps, motors and other various heating equipment citywide. 2. Survey and report on existing conditions of heating equipment. 3. Assist Heating Superintendent with periodic reports. 4. Assist skill trades staff. 5. Provide assistance during emergencies. 6. Respond to all heating/hot water service disruptions. Candidates selected must be available to work and travel throughout the five boroughs; and will be required to work rotating shifts, including holidays and weekends. 8:00 AM - 4:00 PM 4:00 PM - 12:00 AM 12:00 AM - 8:00 AM
3	Under direct supervision, assist in the routine maintenance operation and repair of public buildings, structures, and the equipment they contain; perform related work. Responsibilities include, but are not limited to the following: 1. Perform minor and major repairs to boilers, burners, vacuum tank, pumps, motors and other various heating equipment citywide. 2. Survey and report on existing conditions of heating equipment. 3. Assist Heating Superintendent with periodic reports. 4. Assist skill trades staff. 5. Provide assistance during emergencies. 6. Respond to all heating/hot water service disruptions. Candidates selected must be available to work and travel throughout the five boroughs; and will be required to work rotating shifts, including holidays and weekends. 8:00 AM - 4:00 PM 4:00 PM - 12:00 AM 12:00 AM - 8:00 AM
4	Responsibilities of selected candidates will include, but are not limited to the following: 1. Prepare, fill and prime surfaces for painting. 2. Mix paint components and match colors. 3. Apply paint with a brush, roller or spray gun. 4. Apply plaster to surfaces. 5. Work on and from ladders, platforms and scaffolds 6. Rig lines and scaffolds.
5	Responsibilities of selected candidates will include, but are not limited to the following: 1. Prepare, fill and prime surfaces for painting. 2. Mix paint components and match colors. 3. Apply paint with a brush, roller or spray gun. 4. Apply plaster to surfaces. 5. Work on and from ladders, platforms and scaffolds 6. Rig lines and scaffolds.

Job Description After NLP Preprocessing Pipeline

	Processed_Desc
2	direct supervision assist routine maintenance operation repair public building structure equipment contain perform related work responsibility include limited following perform minor major repair boiler burner vacuum tank pump motor various heating equipment citywide survey report existing condition heating equipment assist heating superintendent periodic report assist skill trade staff provide assistance emergency respond heatinghot water service disruption candidate selected must available work travel throughout five borough required work rotating shift including holiday weekend pm pm
3	direct supervision assist routine maintenance operation repair public building structure equipment contain perform related work responsibility include limited following perform minor major repair boiler burner vacuum tank pump motor various heating equipment citywide survey report existing condition heating equipment assist heating superintendent periodic report assist skill trade staff provide assistance emergency respond heatinghot water service disruption candidate selected must available work travel throughout five borough required work rotating shift including holiday weekend pm pm
4	responsibility selected candidate include limited following prepare fill prime surface painting mix paint component match color apply paint brush roller spray gun apply plaster surface work ladder platform scaffold rig line scaffold
5	responsibility selected candidate include limited following prepare fill prime surface painting mix paint component match color apply paint brush roller spray gun apply plaster surface work ladder platform scaffold rig line scaffold

EDA

Top 10 Job Openings Count



Job Description



Topic Modeling



LDA

```
from sklearn.decomposition import LatentDirichletAllocation
```



LSA

```
from sklearn.decomposition import TruncatedSVD
```



NMF

```
from sklearn.decomposition import NMF
```





Probabilistic Modeling: LDA

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Technology	develop	manag	program	agenc	support	data	work	plan	servic	provid
Environment	water	wastewat	manag	oper	protect	bureau	engin	project	environment	work
Engineering & Planning	project	manag	contract	program	agenc	review	report	construct	work	servic
Legal Affairs	case	assist	investig	perform	offic	agenc	work	staff	provid	servic

	Technology	Environment	Engineering & Planning	Legal Affairs	dominant_topic
Doc0	0.41	0.00	0.59	0.0	Engineering & Planning

	Processed_Desc	Business Title	Division/Work Unit	Job Type
108	supervision latitude independent action exercise independent judgment infrastructure project manager responsible scoping planning execution infrastructure project project focus delivery new capability application service infrastructure candidate specific duty include work project stakeholder clearly define manage expectation objective assumption measurement schedule milestone scope requirement risk issue escalation change metric success criterion communication change control successful deployment actively communicate project stakeholder ensure alignment visibility transparency aspect assigned project manage vendor engagement foster accountability deliverable timeline quality project delivery develop maintain project charter project planschedules implementation plan status report risk mitigation contingency plan qa plan project budget cost capital funding request rfp strategic plan business continuity plan account resource capacity organizational change competing initiative project constraint utilize project management process methodology tool ensure project delivered time within budget adhere high quality standard meet stakeholder expectation build lead mentor diverse project team throughout project lifecycles collaborate unit throughout project lifecycles successful implementation project deliverable	infrastructure project manager	mi adminnetwork sysms dvlpmt	information technology telecommunication

Matrix Factorization: LSA (SVD)

Topic 0

manag, project, develop, program, work, includ, agenc, water, respons, plan

???

Topic 1

water, wastewat, engin, bureau, protect, construct, project, sewer, environment, suppli

Topic 2

project, manag, energi, construct, design, capit, program, schedul, scope, effici

Topic 3

develop, program, build, plan, water, energi, hous, track, polici, sewer

Matrix Factorization: LSA (SVD)

Topic 0

manag, project, develop, program, work, includ, agenc, water, respons, plan

Topic 1

water, wastewat, engin, bureau, protect, construct, project, sewer, environment, suppli

Topic 2

project, manag, energi, construct, design, capit, program, schedul, scope, effici

Topic 3

develop, program, build, plan, water, energi, hous, track, polici, sewer

Environment ??

Matrix Factorization: LSA (SVD)

Topic 0

manag, project, develop, program, work, includ, agenc, water, respons, plan

Topic 1

water, wastewat, engin, bureau, protect, construct, project, sewer, environment, suppli

Topic 2

project, manag, energi, construct, design, capit, program, schedul, scope, effici

???

Topic 3

develop, program, build, plan, water, energi, hous, track, polici, sewer

Matrix Factorization: LSA (SVD)

Topic 0

manag, project, develop, program, work, includ, agenc, water, respons, plan

Topic 1

water, wastewat, engin, bureau, protect, construct, project, sewer, environment, suppli

Topic 2

project, manag, energi, construct, design, capit, program, schedul, scope, effici

Topic 3

develop, program, build, plan, water, energi, hous, track, polici, sewer

Engineering & Planning ??

Matrix Factorization: NMF

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9
Technology	develop	data	manag	support	technolog	busi	servic	agenc	team	plan
Environment	water	wastewat	protect	sewer	bureau	environment	oper	drink	drink water	gallon
Engineering & Planning	project	construct	contract	design	manag	engin	capit	budget	design construct	project manag
Legal Affairs	investig	case	attorney	assist	enforc	legal	nycha	perform	court	agenc

2787

nyc department environmental protection dep enriches environment protects public health new yorkers providing billion gallon high quality drinking water managing wastewater stormwater reducing air noise hazardous material pollution dep largest combined municipal water wastewater utility country nearly employee deps water supply system comprised reservoir controlled lake throughout system square mile watershed extends mile north west city dep office energy resource recovery program dep energy office housed within commissioner office committed implementing energy management best practice across deps portfolio building operating facility help city meet ambitious greenhouse gas emission energy reduction goal dep energy office work track overall energy usage identify energysaving opportunity building implement energyefficient operation maintenance practice deliver energy efficiency retrofit project provide energy resource recovery performance tracking program assessment scientific research emergency planning internal auditing service dep energy office seeking hire energy manager deps bureau wastewater treatment work direction director dep energy office duty include limited identify implement energy resource recovery research opportunity project apply external grant program enrollment including limited energy funding reliability program report project progress support advise bureau throughout procurement design policyssetting order influence decisionmaking process toward energy resource efficiency share energy resource recovery best practice identify employee candidate energy resource recovery training identify energy conservation opportunity potential energy efficiency project associated wastewater conveyance treatment process building system equipment contribute project scope work cost estimate support measurement verification effort manage bureau heat light power budget including monthly billing review annual budgetsetting process investigate correct potential billing error determine report root cause change energy usage demand coordinate energy utility meter account service request inspection replacement interruption bureau provide project management assistance including developing maintaining project plan document managing coordinating project element reviewing milestone ensuring efficient progress implementation consistent monitoring progress initiative commitment assist senior staff project management responsibility including adhering internal external timeline ensuring completion highquality deliverable managing budget may also responsible one discrete project expected manage project completion track provide energy resource recovery data needed reporting compliance purpose prepare edit proofread present highquality report presentation article meeting correspondence document report energy resource recovery communication material datasets participate ad hoc project arise provide overall support dep energy office mission

bureau
energy
manager

wtenergy

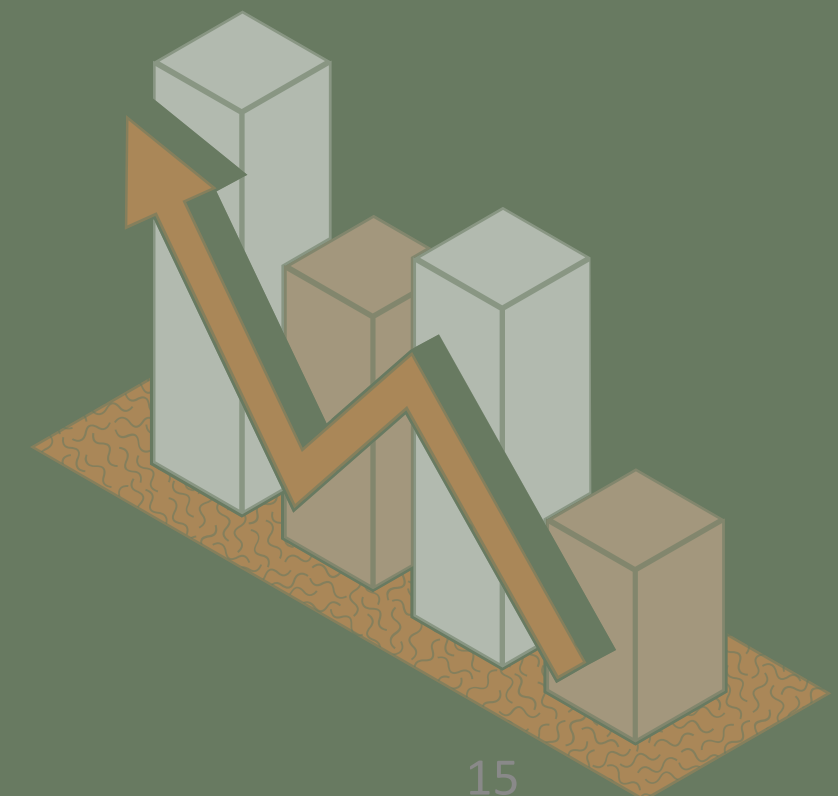
engineering
architecture
planning policy
research analysis

Matrix Factorization: NMF

	Technology	Environment	Engineering & Planning	Legal Affairs	dominant_topic
Doc0	0.050	0.000	0.162	0.000	Engineering & Planning
Doc1	0.138	0.005	0.014	0.000	Technology
Doc2	0.142	0.000	0.000	0.021	Technology
Doc3	0.031	0.051	0.080	0.000	Engineering & Planning
Doc4	0.019	0.000	0.084	0.059	Engineering & Planning

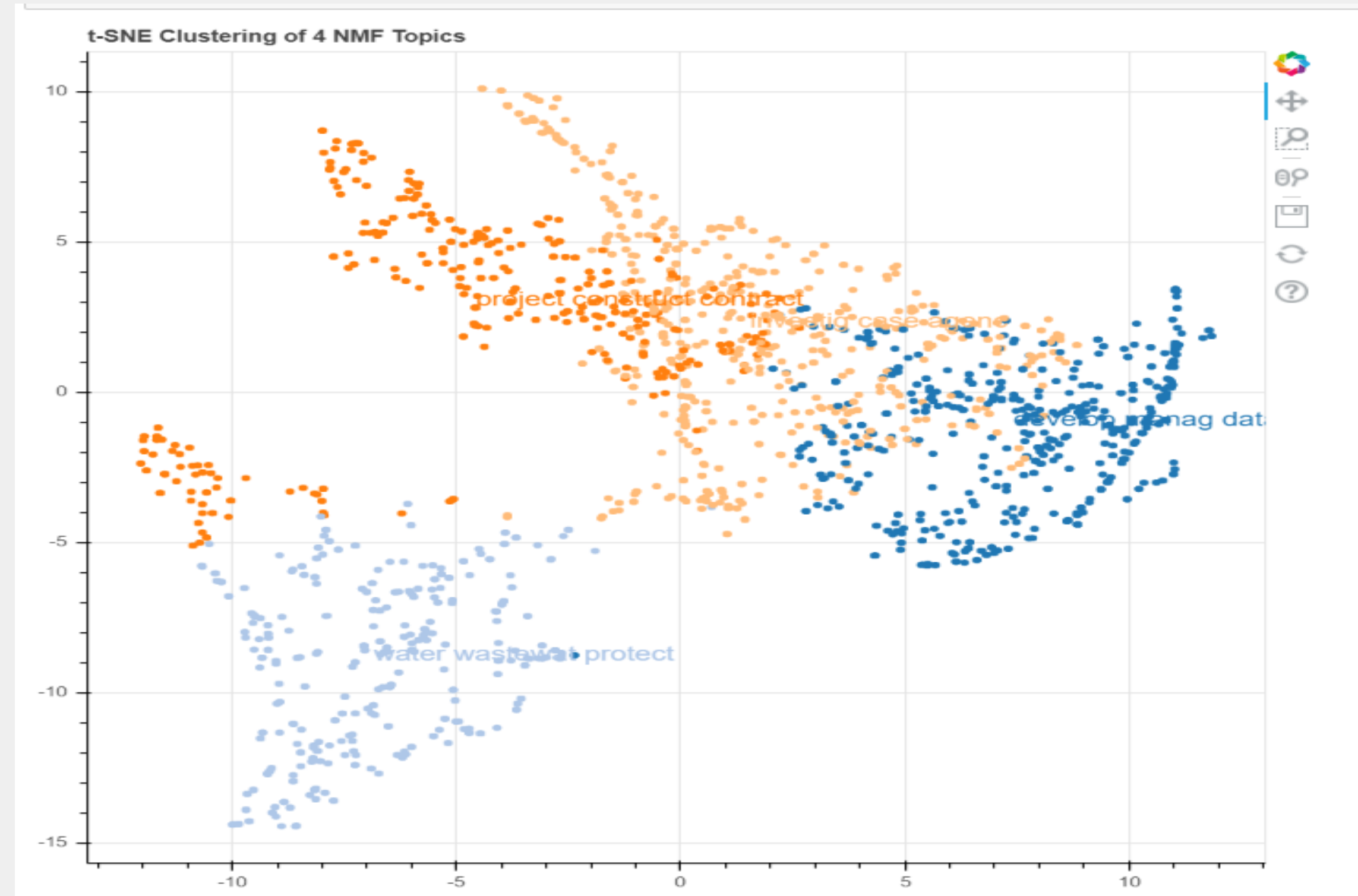
```
vec1 = [0.138,0.005,0.014,0.000]
vec2 = [0.142,0.000,0.000,0.021] #same
print(cosine_similarity([vec1, vec2]).round())

[[1. 1.]
 [1. 1.]]
```



Results

t-SNE Clustering of 4
Topics of NMF model



Conclusion

The job posting classifies into 4 topics basic are:

- ❖ Technology
- ❖ Engineering & Planning
- ❖ Environment
- ❖ Legal Affairs

Future Work:

- ❖ Resumes topic modeling.
- ❖ Job similarity.

Four paper airplanes are illustrated in a stylized, geometric manner. They are colored with a light green dotted pattern on the wings and a solid orange-brown on the fuselage. Each airplane has two dark green lines trailing behind it, suggesting motion. The airplanes are positioned in the upper left, center, lower left, and bottom center of the slide.

**Thank you
for listening**