

PS5_AndrewYu

July 29, 2023

```
[1]: from google.colab import drive

drive.mount('/content/drive')
FOLDERNAME = "Stanford Summer Session/SOC 128D"

#Packages for working with data
# !pip install "modin[dask]"
# import modin.pandas as pd
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sb
import plotnine as pn
import numpy as np

plt.style.use("ggplot")
```

Mounted at /content/drive

1. Run a topic model with 4 topics on the “guardian_co_river.csv” dataset.

```
[15]: import json
import nltk
nltk.download('stopwords')
from nltk.corpus import stopwords
from nltk import word_tokenize
nltk.download('punkt')
from string import punctuation
from sklearn.feature_extraction.text import CountVectorizer
import regex as re
from nltk import pos_tag
nltk.download('averaged_perceptron_tagger')
from nltk.stem import WordNetLemmatizer
nltk.download('wordnet')

with open("stopwords_extra.json", "r") as file: stopwords_extra = set(json.
    ↪load(file))
stopwords_nltk = set(stopwords.words('english'))
punctuation += "-- ' , " " „ † ‡ • … % < > / "
```

```

stopwords_punct = set(punctuation)
combined_stoplist = list(set.union(stopwords_extra, stopwords_nltk,
    stopwords_punct))

# I used the dataset with 18548 articles, for practice working with bigger
    datasets
df = pd.read_csv(f"drive/My Drive/{FOLDERNAME}/data/guardian_co_river.csv").
    dropna(subset=["body_text"])
# Unicode Categories C (Control), M (Mark), P (Punctuation), S (Symbol), Z
    (Separator) + emojis
df["body_text"] = df["body_text"].apply(lambda x: re.
    compile(r'[\p{C}|\p{M}|\p{P}|\p{S}|\p{Z}]+', re.UNICODE).sub(" ", x).strip())

# Lemmatization
def get_POS_tags(pos_tag):
    POS_tag = {'NN': 'n', 'JJ': 'a', 'VB': 'v', 'RB': 'r'}
    try:
        # Getting first 2 letters of pos_tag
        return POS_tag[pos_tag[:2]]
    except:
        # Fallback to noun (Default)
        return 'n'

# In doing this, token_pattern parameter in CountVectorizer won't work, so just
    add any patterns here
class LemmaTokenizer(object):
    def __init__(self):
        self.wnl = WordNetLemmatizer()
    def __call__(self, docs):
        return [self.wnl.lemmatize(word, pos=get_POS_tags(tag)) for word, tag
    in pos_tag(word_tokenize(docs)) if not re.search("\d+", word)]

# This python library automatically integrates anti-joining stopwords and takes
    in corpus as list of texts, converting into list of words via word_tokenize,
    to output DTM. Thus unnest_tokens not used
tf_vectorizer = CountVectorizer(stop_words=combined_stoplist,
    tokenizer=LemmaTokenizer(), strip_accents=None)
# tf is in the form (doc, word): freq
tf = tf_vectorizer.fit_transform(df["body_text"])

print(f"No. of words per topic: {len(tf_vectorizer.get_feature_names_out())}")
pd.DataFrame({"word": tf_vectorizer.get_feature_names_out()})

```

```

[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Package punkt is already up-to-date!

```



```

else:
    for topic_i, topic_freq in enumerate(model.components_):
        top_freq_words_i = topic_freq.argsort()[::-n_top_words - 1:-1]
        # {topic: [word, ...]}
        word_dict[topics[topic_i]] = [words[i] for i in top_freq_words_i]
    return pd.DataFrame(word_dict)

from sklearn.decomposition import LatentDirichletAllocation as LDA
lda4 = LDA(n_components=4, random_state=1)
# probability of each (4) topic per (17982) document
topic_per_document = lda4.fit_transform(tf)
# Frequency of each (157463) word-term per (4) topic
word_per_topic = lda4.components_

# Log Likelihood: Less negative
print("Log Likelihood: ", lda4.score(tf))

# Perplexity: exp(-1 * log likelihood), Lower
print("Perplexity: ", lda4.perplexity(tf))

topics_words_df4 = get_model_topics(lda4, tf_vectorizer, ["Topic 1", "Topic 2",
↳ "Topic 3", "Topic 4"])
topics_words_df4

```

Log Likelihood: -86926244.37087119

Perplexity: 5463.206371294159

```

[17]:
      Topic 1      Topic 2 Topic 3 Topic 4
0      people      trump   year   year
1      water      state   make   day
2      year republican people  city
3 government president   time  river
4      state      win    work   time
5      report      vote   life  world
6      country      make  river  park
7      river      year   find   make
8      case      time  child   run
9      make      election film   back

```

2. Plot the top words for each topic, similar to what is done in section 5.2. What would you label these topics?

```

[104]: topics_words_df4 = get_model_topics(lda4, tf_vectorizer, ["Topic 1", "Topic 2",
↳ "Topic 3", "Topic 4"], detailed=True)
topics_words_df4

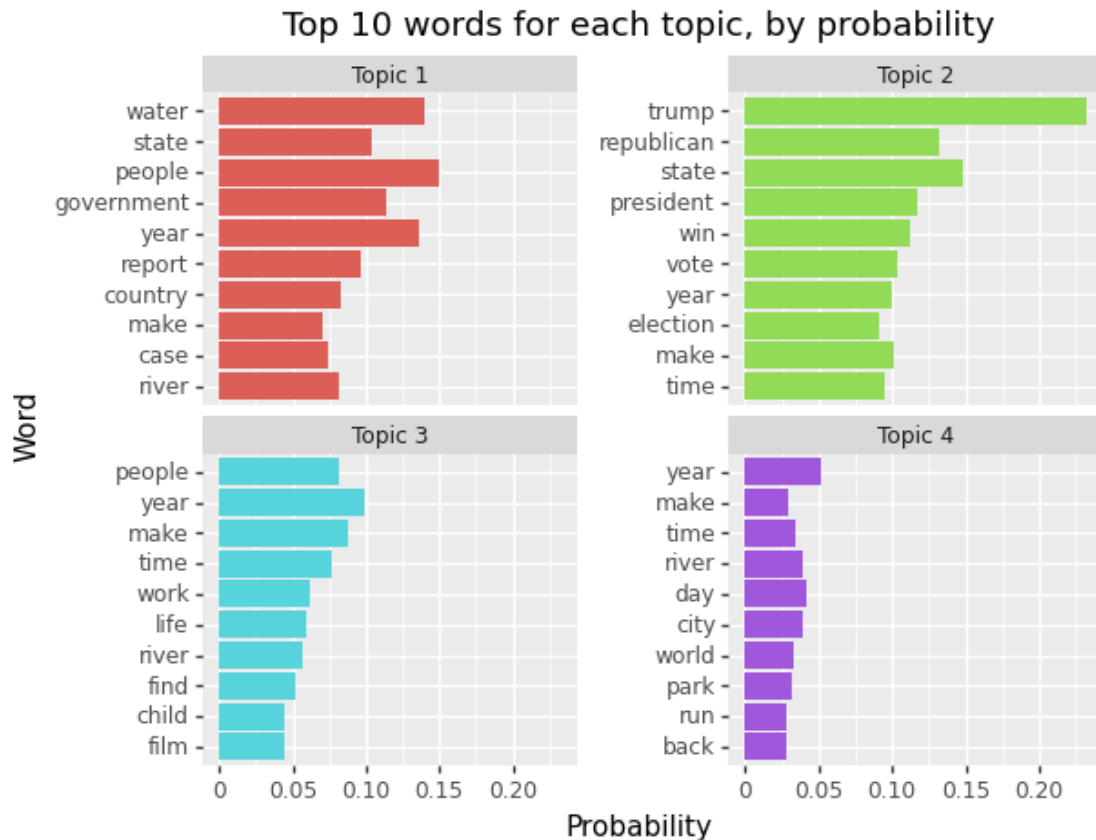
```

```
[104]:
```

	Topic	Word	Probability
0	Topic 1	people	0.150188
1	Topic 1	water	0.139437
2	Topic 1	year	0.135564
3	Topic 1	government	0.113529
4	Topic 1	state	0.104199
5	Topic 1	report	0.096412
6	Topic 1	country	0.082658
7	Topic 1	river	0.081916
8	Topic 1	case	0.074222
9	Topic 1	make	0.070491
10	Topic 2	trump	0.232615
11	Topic 2	state	0.147654
12	Topic 2	republican	0.132483
13	Topic 2	president	0.117492
14	Topic 2	win	0.112478
15	Topic 2	vote	0.102859
16	Topic 2	make	0.101298
17	Topic 2	year	0.099786
18	Topic 2	time	0.095326
19	Topic 2	election	0.091645
20	Topic 3	year	0.098372
21	Topic 3	make	0.087154
22	Topic 3	people	0.081191
23	Topic 3	time	0.076403
24	Topic 3	work	0.061427
25	Topic 3	life	0.059303
26	Topic 3	river	0.057298
27	Topic 3	find	0.051575
28	Topic 3	child	0.044530
29	Topic 3	film	0.044246
30	Topic 4	year	0.051727
31	Topic 4	day	0.041634
32	Topic 4	city	0.039051
33	Topic 4	river	0.038604
34	Topic 4	time	0.033675
35	Topic 4	world	0.033473
36	Topic 4	park	0.031982
37	Topic 4	make	0.028890
38	Topic 4	run	0.028120
39	Topic 4	back	0.027374

```
[105]: # I've tried searching but I'm not sure why the sorting is not working, df is
        ↪sorted, and x is also reordered within
plot = (pn.ggplot(topics_words_df4, pn.aes(x='reorder(Word, Probability)',
        ↪y='Probability', fill='Topic')) +
        pn.geom_col(show_legend=False) +
```

```
pn.coord_flip() +
pn.facet_wrap('Topic', scales = "free_y") +
pn.theme(subplots_adjust={'wspace': 0.4}) +
pn.labs(title="Top 10 words for each topic, by probability", x="Word", y="Probability")
print(plot)
```



They are words are not very distinct, due to choice of number of topics, but I would label them

Topic 1: Government

Topic 2: Politics

Topic 3: Work

Topic 4: Time

3. Try running the LDA again, this time with a different number of topics. Does the new topic model make more or less sense to you?

```
[ ]: from sklearn.model_selection import GridSearchCV

lda = LDA()
model = GridSearchCV(lda, param_grid={'n_components': [2, 4, 6, 8, 10]})
model.fit(tf)

print("Best Model's Params: ", model.best_params_)
print("Best Log Likelihood Score: ", model.best_score_)
print("Model Perplexity: ", model.best_estimator_.perplexity(tf))
```

```
Best Model's Params: {'n_components': 2}
Best Log Likelihood Score: -20246738.786500275
Model Perplexity: 9943.179337936572
```

```
[19]: lda2 = LDA(n_components=2, random_state=1)
_ = lda2.fit_transform(tf)

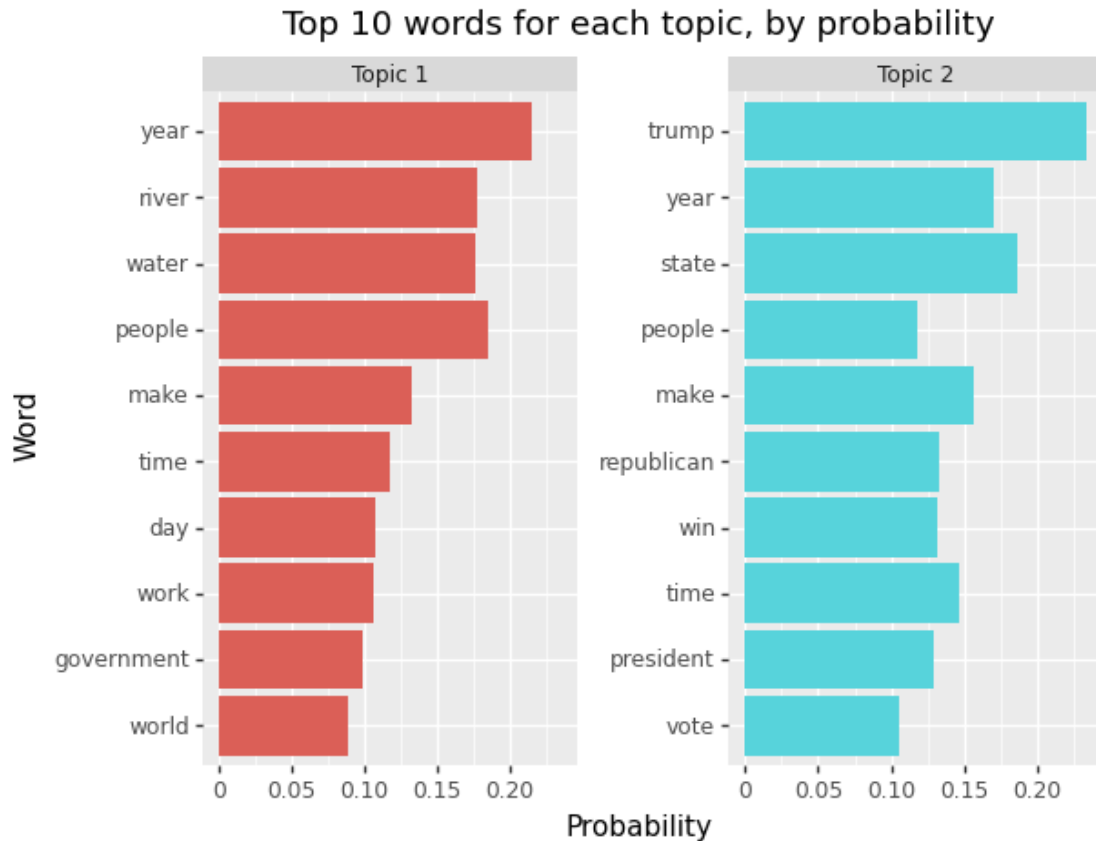
topics_words_df2 = get_model_topics(lda2, tf_vectorizer, ["Topic 1", "Topic 2"])
topics_words_df2
```

```
[19]:
```

	Topic 1	Topic 2
0	year	trump
1	people	state
2	river	year
3	water	make
4	make	time
5	time	republican
6	day	win
7	work	president
8	government	people
9	world	vote

```
[107]: topics_words_df2 = get_model_topics(lda2, tf_vectorizer, ["Topic 1", "Topic_
↳2"], detailed=True)
plot = (pn.ggplot(topics_words_df2, pn.aes(x='reorder(Word, Probability)',
↳y='Probability', fill='Topic')) +
        pn.geom_col(show_legend=False) +
        pn.coord_flip() +
        pn.facet_wrap('Topic', scales = "free_y") +
        pn.theme(subplots_adjust={'wspace': 0.4}) +
        pn.labs(title="Top 10 words for each topic, by probability", x="Word",
↳y="Probability")
        )

print(plot)
```



My intuition was 2 topics (based on the in-class materials), but I also wanted to test the optimal number using the GridSearch, which confirmed my intuition.

As evident, 2 topics summarises all the documents more clearly and sensibly, with Topic 1 being about rivers and Topic 2 being more political

- Using whichever topic model made the most sense to you, create a dataset similar to `co_documents` in section 5.3. Choose two articles to examine in more detail. Discuss the topic probabilities.

```
[123]: def get_model_docs(model, dtm, detailed=False):
    topic_dict = {}
    doc_topic = model.transform(dtm)
    if detailed:
        for doc_i, topic_probs_of_doc in enumerate(doc_topic):
            num_topics = doc_topic.shape[1]
            # Sorting topics by top probabilities
            top_probable_topics = topic_probs_of_doc.argsort()[::-num_topics - 1:-1]
            topic_dict[f"Doc {doc_i}"] = [(topic+1, topic_probs_of_doc[topic]) for
            ↪topic in top_probable_topics]
```



```

df = pd.DataFrame([(doc, topic, prob) for doc, topics_probs in topic_dict.
↳ items() for topic, prob in topics_probs], columns=["Document", "Topic",
↳ "Probability"])

# Chaining doesn't work for some reason
df['Dominant_topic'] = df[df.index % num_topics == 0]["Topic"]
df['Dominant_topic'] = df['Dominant_topic'].fillna(method='ffill')
df['Dominant_topic'] = df['Dominant_topic'].astype(int)
return df

else:
    for doc_i, topic_probs_of_doc in enumerate(doc_topic):
        num_topics = doc_topic.shape[1]
        top_probable_topics = topic_probs_of_doc.argsort()[: -num_topics - 1: -1]
        topic_dict[f"Doc {doc_i}"] = [topic for topic in top_probable_topics]
    return pd.DataFrame(topic_dict)

docs_topics_df = get_model_docs(lda2, tf, detailed=True)
docs_topics_df
# Probabilities of each topic add up to 1, which confirms this is similar to
↳ co_documents (gamma)

```

```

[123]:
      Document  Topic  Probability  Dominant_topic
0         Doc 0      1      0.645597              1
1         Doc 0      2      0.354403              1
2         Doc 1      1      0.999263              1
3         Doc 1      2      0.000737              1
4         Doc 2      1      0.921166              1
...         ...    ...          ...             ...
35959  Doc 17979      2      0.004749              1
35960  Doc 17980      2      0.987856              2
35961  Doc 17980      1      0.012144              2
35962  Doc 17981      1      0.952474              1
35963  Doc 17981      2      0.047526              1

```

[35964 rows x 4 columns]

```

[125]: # To find a topic that has a very fair probability of belonging to either Topic
docs_topics_df[docs_topics_df["Probability"].between(0.45, 0.65)]
# I chose      Doc 17975

```

```

[125]:
      Document  Topic  Probability  Dominant_topic
0         Doc 0      1      0.645597              1
18        Doc 9      1      0.593177              1
48        Doc 24     1      0.539227              1
49        Doc 24     2      0.460773              1
56        Doc 28     2      0.571738              2
...         ...    ...          ...             ...

```

35902	Doc 17951	2	0.586099	2
35938	Doc 17969	2	0.555969	2
35946	Doc 17973	1	0.635390	1
35950	Doc 17975	2	0.506625	2
35951	Doc 17975	1	0.493375	2

[2900 rows x 4 columns]

```
[124]: chosen_docs_df = docs_topics_df[docs_topics_df["Document"].isin(["Doc 1", "Doc_
↪17975", "Doc 17980"])]
chosen_docs_df
```

```
[124]:
```

	Document	Topic	Probability	Dominant_topic
2	Doc 1	1	0.999263	1
3	Doc 1	2	0.000737	1
35950	Doc 17975	2	0.506625	2
35951	Doc 17975	1	0.493375	2
35960	Doc 17980	2	0.987856	2
35961	Doc 17980	1	0.012144	2

The probabilities show that

- Doc 1 is extremely likely to be Topic 1 (Rivers)
 - Doc 17975 is quite on the fence, but ultimately leans towards Topic 2 (Politics)
 - Doc 17980 is extremely likely to be Topic 2 (Politics)
5. Considering what we know about the Colorado River and water rights in the U.S. (from readings and class), are we surprised by the topics generated by these news articles? Describe how your results compared to expectations.

```
[131]: for index, row in df.iloc[[1, 17975, 17980], df.columns.
↪get_indexer(['web_title', 'web_url'])].iterrows():
print(f"Title: {row['web_title']}, URL:{row['web_url']}\n")
```

Title: The farmers dealing with water shortages even before historic Colorado River deal, URL:<https://www.theguardian.com/global/2023/may/31/arizona-farmers-water-colorado-river-cuts>

Title: Paraguayan dictator toppled in family coup,
URL:<https://www.theguardian.com/theguardian/1989/feb/04/fromthearchive>

Title: The American Presidency,
URL:<https://www.theguardian.com/world/1912/nov/07/usa.fromthearchive>

Upon reading the articles, the most probable topics mostly makes sense, as also suggested by the title of the article.

Doc 1 discusses the lives of the farmers near the Colorado River, which checks out with it being assigned Topic 1 (Rivers)

Doc 17975 discusses the the events surrounding the political party in Colorado, and has little to do with the Colorado River, yet is assigned such a fair probability, which is confusing. This could be because even Topic 1 itself isn't such a clearly defined "Rivers" topic - there are also quite a lot of political terms. What is the true differentiating factor between Topic 1 and 2 is that 2 is more specific ("trump", "republican", "president"). However, the model did predict correctly that this article is more likely Topic 2 than 1 anyway.

Doc 17980 discusses mostly the American Presidency which checks out with it being assigned Topic 2 (Politics)

Doc 1, which talks the most regarding the Colorado River situation, aligns with what we learnt from the readings - Arizona is fighting for its rights to usage of the water from the river, with California and Nevada, aligning with the [reading](#) - The farmers were already struggling with water, in growing the most profitable and marketable commodity crops - alfalfa and cotton. This aligns with the [reading](#) that agriculture sucks up 79% of all water usage from the river.

```
[ ]: # If it contains spaces, use '\ ' to represent each space E.g. 'Summer\ PSet\ 1.
      ↳ipynb'
FILENAME = "PS5_AndrewYu.ipynb"

%cd drive/My\ Drive
%cd $FOLDERNAME
!sudo apt-get install texlive-xetex texlive-fonts-recommended
      ↳texlive-plain-generic
!pip install PyPDF2

!jupyter nbconvert --log-level CRITICAL --to pdf $FILENAME
```

```
/content/drive/My Drive
/content/drive/My Drive/Stanford Summer Session/SOC 128D
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
```

The following additional packages will be installed:

```
dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono
fonts-texgyre fonts-urw-base35 libapache-pom-java libcommons-logging-java
libcommons-parent-java libfontbox-java libfontenc1 libgs9 libgs9-common
libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsynchronet2 libteckit0 libtexlua53 libtexluajit2 libwoff1
libzip-0-13 lmodern poppler-data preview-latex-style rake ruby
ruby-net-telnet ruby-rubygems ruby-webrick ruby-xmlrpc ruby3.0
rubygems-integration tlutils teckit tex-common tex-gyre texlive-base
texlive-binaries texlive-latex-base texlive-latex-extra
texlive-latex-recommended texlive-pictures tipa xfonts-encodings
xfonts-utils
```

Suggested packages:

```
fonts-noto fonts-freefont-otf | fonts-freefont-ttf libavalon-framework-java
libcommons-logging-java-doc libexcalibur-logkit-java liblog4j1.2-java
poppler-utils ghostscript fonts-japanese-mincho | fonts-ipafont-mincho
```

```

fonts-japanese-gothic | fonts-ipafont-gothic fonts-arphic-ukai
fonts-arphic-uming fonts-nanum ri ruby-dev bundler debhelper gv
| postscript-viewer perl-tk xpdf | pdf-viewer xzdec
texlive-fonts-recommended-doc texlive-latex-base-doc python3-pygments
icc-profiles libfile-which-perl libspreadsheet-parseexcel-perl
texlive-latex-extra-doc texlive-latex-recommended-doc texlive-luatex
texlive-pstricks dot2tex prerex texlive-pictures-doc vprerex
default-jre-headless tipa-doc

```

The following NEW packages will be installed:

```

dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono
fonts-texgyre fonts-urw-base35 libapache-pom-java libcommons-logging-java
libcommons-parent-java libfontbox-java libfontenc1 libgs9 libgs9-common
libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsynchronet2 libteckit0 libtexlua53 libtexluajit2 libwoff1
libzip-0-13 lmodern poppler-data preview-latex-style rake ruby
ruby-net-telnet ruby-rubygems ruby-webrick ruby-xmlrpc ruby3.0
rubygems-integration tlutils teckit tex-common tex-gyre texlive-base
texlive-binaries texlive-fonts-recommended texlive-latex-base
texlive-latex-extra texlive-latex-recommended texlive-pictures
texlive-plain-generic texlive-xetex tipa xfonts-encodings xfonts-utils

```

0 upgraded, 54 newly installed, 0 to remove and 9 not upgraded.

Need to get 182 MB of archives.

After this operation, 571 MB of additional disk space will be used.

Get:1 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-droid-fallback all 1:6.0.1r16-1.1build1 [1,805 kB]

Get:2 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-lato all 2.0-2.1 [2,696 kB]

Get:3 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 poppler-data all 0.4.11-1 [2,171 kB]

Get:4 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 tex-common all 6.17 [33.7 kB]

Get:5 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-urw-base35 all 20200910-1 [6,367 kB]

Get:6 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libgs9-common all 9.55.0~dfsg1-0ubuntu5.3 [751 kB]

Get:7 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libidn12 amd64 1.38-4ubuntu1 [60.0 kB]

Get:8 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libijs-0.35 amd64 0.35-15build2 [16.5 kB]

Get:9 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libjbig2dec0 amd64 0.19-3build2 [64.7 kB]

Get:10 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libgs9 amd64 9.55.0~dfsg1-0ubuntu5.3 [5,039 kB]

Get:11 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libkpathsea6 amd64 2021.20210626.59705-1ubuntu0.1 [60.3 kB]

Get:12 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libwoff1 amd64 1.0.2-1build4 [45.2 kB]

Get:13 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 dvisvgm amd64

2.13.1-1 [1,221 kB]
Get:14 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 fonts-lmodern all 2.004.5-6.1 [4,532 kB]
Get:15 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-noto-mono all 20201225-1build1 [397 kB]
Get:16 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 fonts-texgyre all 20180621-3.1 [10.2 MB]
Get:17 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libapache-pom-java all 18-1 [4,720 B]
Get:18 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcommons-parent-java all 43-1 [10.8 kB]
Get:19 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcommons-logging-java all 1.2-2 [60.3 kB]
Get:20 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libfontenc1 amd64 1:1.1.4-1build3 [14.7 kB]
Get:21 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libptexenc1 amd64 2021.20210626.59705-1ubuntu0.1 [39.1 kB]
Get:22 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 rubygems-integration all 1.18 [5,336 B]
Get:23 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 ruby3.0 amd64 3.0.2-7ubuntu2.4 [50.1 kB]
Get:24 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby-rubygems all 3.3.5-2 [228 kB]
Get:25 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby amd64 1:3.0~exp1 [5,100 B]
Get:26 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 rake all 13.0.6-2 [61.7 kB]
Get:27 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby-net-telnet all 0.1.1-2 [12.6 kB]
Get:28 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 ruby-webrick all 1.7.0-3 [51.8 kB]
Get:29 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 ruby-xmlrpc all 0.3.2-1ubuntu0.1 [24.9 kB]
Get:30 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libruby3.0 amd64 3.0.2-7ubuntu2.4 [5,113 kB]
Get:31 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libsynchronet2 amd64 2021.20210626.59705-1ubuntu0.1 [55.5 kB]
Get:32 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libteckit0 amd64 2.5.11+ds1-1 [421 kB]
Get:33 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libtexlua53 amd64 2021.20210626.59705-1ubuntu0.1 [120 kB]
Get:34 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libtexluajit2 amd64 2021.20210626.59705-1ubuntu0.1 [267 kB]
Get:35 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libzip-0-13 amd64 0.13.72+dfsg.1-1.1 [27.0 kB]
Get:36 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 xfonts-encodings all 1:1.0.5-0ubuntu2 [578 kB]
Get:37 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 xfonts-utils amd64

```

1:7.7+6build2 [94.6 kB]
Get:38 http://archive.ubuntu.com/ubuntu jammy/universe amd64 lmodern all
2.004.5-6.1 [9,471 kB]
Get:39 http://archive.ubuntu.com/ubuntu jammy/universe amd64 preview-latex-style
all 12.2-1ubuntu1 [185 kB]
Get:40 http://archive.ubuntu.com/ubuntu jammy/main amd64 t1utils amd64
1.41-4build2 [61.3 kB]
Get:41 http://archive.ubuntu.com/ubuntu jammy/universe amd64 teckit amd64
2.5.11+ds1-1 [699 kB]
Get:42 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tex-gyre all
20180621-3.1 [6,209 kB]
Get:43 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 texlive-
binaries amd64 2021.20210626.59705-1ubuntu0.1 [9,848 kB]
Get:44 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-base all
2021.20220204-1 [21.0 MB]
Get:45 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-fonts-
recommended all 2021.20220204-1 [4,972 kB]
Get:46 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-base
all 2021.20220204-1 [1,128 kB]
Get:47 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libfontbox-java all
1:1.8.16-2 [207 kB]
Get:48 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libpdfbox-java all
1:1.8.16-2 [5,199 kB]
Get:49 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-
recommended all 2021.20220204-1 [14.4 MB]
Get:50 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-pictures
all 2021.20220204-1 [8,720 kB]
Get:51 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-extra
all 2021.20220204-1 [13.9 MB]
Get:52 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-plain-
generic all 2021.20220204-1 [27.5 MB]
Get:53 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tipa all 2:1.3-21
[2,967 kB]
Get:54 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-xetex all
2021.20220204-1 [12.4 MB]
Fetched 182 MB in 9s (19.2 MB/s)
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based
frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line 78,
<> line 54.)
debconf: falling back to frontend: Readline
debconf: unable to initialize frontend: Readline
debconf: (This frontend requires a controlling tty.)
debconf: falling back to frontend: Teletype
dpkg-preconfigure: unable to re-open stdin:
Selecting previously unselected package fonts-droid-fallback.
(Reading database ... 120493 files and directories currently installed.)
Preparing to unpack .../00-fonts-droid-fallback_1%3a6.0.1r16-1.1build1_all.deb

```

```

...
Unpacking fonts-droid-fallback (1:6.0.1r16-1.1build1) ...
Selecting previously unselected package fonts-lato.
Preparing to unpack .../01-fonts-lato_2.0-2.1_all.deb ...
Unpacking fonts-lato (2.0-2.1) ...
Selecting previously unselected package poppler-data.
Preparing to unpack .../02-poppler-data_0.4.11-1_all.deb ...
Unpacking poppler-data (0.4.11-1) ...
Selecting previously unselected package tex-common.
Preparing to unpack .../03-tex-common_6.17_all.deb ...
Unpacking tex-common (6.17) ...
Selecting previously unselected package fonts-urw-base35.
Preparing to unpack .../04-fonts-urw-base35_20200910-1_all.deb ...
Unpacking fonts-urw-base35 (20200910-1) ...
Selecting previously unselected package libgs9-common.
Preparing to unpack .../05-libgs9-common_9.55.0~dfsg1-0ubuntu5.3_all.deb ...
Unpacking libgs9-common (9.55.0~dfsg1-0ubuntu5.3) ...
Selecting previously unselected package libidn12:amd64.
Preparing to unpack .../06-libidn12_1.38-4ubuntu1_amd64.deb ...
Unpacking libidn12:amd64 (1.38-4ubuntu1) ...
Selecting previously unselected package libijs-0.35:amd64.
Preparing to unpack .../07-libijs-0.35_0.35-15build2_amd64.deb ...
Unpacking libijs-0.35:amd64 (0.35-15build2) ...
Selecting previously unselected package libjbig2dec0:amd64.
Preparing to unpack .../08-libjbig2dec0_0.19-3build2_amd64.deb ...
Unpacking libjbig2dec0:amd64 (0.19-3build2) ...
Selecting previously unselected package libgs9:amd64.
Preparing to unpack .../09-libgs9_9.55.0~dfsg1-0ubuntu5.3_amd64.deb ...
Unpacking libgs9:amd64 (9.55.0~dfsg1-0ubuntu5.3) ...
Selecting previously unselected package libkpathsea6:amd64.
Preparing to unpack .../10-libkpathsea6_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libkpathsea6:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package libwoff1:amd64.
Preparing to unpack .../11-libwoff1_1.0.2-1build4_amd64.deb ...
Unpacking libwoff1:amd64 (1.0.2-1build4) ...
Selecting previously unselected package dvisvgm.
Preparing to unpack .../12-dvisvgm_2.13.1-1_amd64.deb ...
Unpacking dvisvgm (2.13.1-1) ...
Selecting previously unselected package fonts-lmodern.
Preparing to unpack .../13-fonts-lmodern_2.004.5-6.1_all.deb ...
Unpacking fonts-lmodern (2.004.5-6.1) ...
Selecting previously unselected package fonts-noto-mono.
Preparing to unpack .../14-fonts-noto-mono_20201225-1build1_all.deb ...
Unpacking fonts-noto-mono (20201225-1build1) ...
Selecting previously unselected package fonts-texgyre.
Preparing to unpack .../15-fonts-texgyre_20180621-3.1_all.deb ...
Unpacking fonts-texgyre (20180621-3.1) ...

```

```

Selecting previously unselected package libapache-pom-java.
Preparing to unpack .../16-libapache-pom-java_18-1_all.deb ...
Unpacking libapache-pom-java (18-1) ...
Selecting previously unselected package libcommons-parent-java.
Preparing to unpack .../17-libcommons-parent-java_43-1_all.deb ...
Unpacking libcommons-parent-java (43-1) ...
Selecting previously unselected package libcommons-logging-java.
Preparing to unpack .../18-libcommons-logging-java_1.2-2_all.deb ...
Unpacking libcommons-logging-java (1.2-2) ...
Selecting previously unselected package libfontenc1:amd64.
Preparing to unpack .../19-libfontenc1_1%3a1.1.4-1build3_amd64.deb ...
Unpacking libfontenc1:amd64 (1:1.1.4-1build3) ...
Selecting previously unselected package libptexenc1:amd64.
Preparing to unpack .../20-libptexenc1_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libptexenc1:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package rubygems-integration.
Preparing to unpack .../21-rubygems-integration_1.18_all.deb ...
Unpacking rubygems-integration (1.18) ...
Selecting previously unselected package ruby3.0.
Preparing to unpack .../22-ruby3.0_3.0.2-7ubuntu2.4_amd64.deb ...
Unpacking ruby3.0 (3.0.2-7ubuntu2.4) ...
Selecting previously unselected package ruby-rubygems.
Preparing to unpack .../23-ruby-rubygems_3.3.5-2_all.deb ...
Unpacking ruby-rubygems (3.3.5-2) ...
Selecting previously unselected package ruby.
Preparing to unpack .../24-ruby_1%3a3.0~exp1_amd64.deb ...
Unpacking ruby (1:3.0~exp1) ...
Selecting previously unselected package rake.
Preparing to unpack .../25-rake_13.0.6-2_all.deb ...
Unpacking rake (13.0.6-2) ...
Selecting previously unselected package ruby-net-telnet.
Preparing to unpack .../26-ruby-net-telnet_0.1.1-2_all.deb ...
Unpacking ruby-net-telnet (0.1.1-2) ...
Selecting previously unselected package ruby-webrick.
Preparing to unpack .../27-ruby-webrick_1.7.0-3_all.deb ...
Unpacking ruby-webrick (1.7.0-3) ...
Selecting previously unselected package ruby-xmlrpc.
Preparing to unpack .../28-ruby-xmlrpc_0.3.2-1ubuntu0.1_all.deb ...
Unpacking ruby-xmlrpc (0.3.2-1ubuntu0.1) ...
Selecting previously unselected package libruby3.0:amd64.
Preparing to unpack .../29-libruby3.0_3.0.2-7ubuntu2.4_amd64.deb ...
Unpacking libruby3.0:amd64 (3.0.2-7ubuntu2.4) ...
Selecting previously unselected package libsyntax2:amd64.
Preparing to unpack .../30-libsyntax2_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libsyntax2:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package libteckit0:amd64.

```



```

Preparing to unpack .../31-libteckit0_2.5.11+ds1-1_amd64.deb ...
Unpacking libteckit0:amd64 (2.5.11+ds1-1) ...
Selecting previously unselected package libtexlua53:amd64.
Preparing to unpack .../32-libtexlua53_2021.20210626.59705-1ubuntu0.1_amd64.deb
...
Unpacking libtexlua53:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package libtexluajit2:amd64.
Preparing to unpack
.../33-libtexluajit2_2021.20210626.59705-1ubuntu0.1_amd64.deb ...
Unpacking libtexluajit2:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package libzip-0-13:amd64.
Preparing to unpack .../34-libzip-0-13_0.13.72+dfsg.1-1.1_amd64.deb ...
Unpacking libzip-0-13:amd64 (0.13.72+dfsg.1-1.1) ...
Selecting previously unselected package xfonts-encodings.
Preparing to unpack .../35-xfonts-encodings_1%3a1.0.5-0ubuntu2_all.deb ...
Unpacking xfonts-encodings (1:1.0.5-0ubuntu2) ...
Selecting previously unselected package xfonts-utils.
Preparing to unpack .../36-xfonts-utils_1%3a7.7+6build2_amd64.deb ...
Unpacking xfonts-utils (1:7.7+6build2) ...
Selecting previously unselected package lmodern.
Preparing to unpack .../37-lmodern_2.004.5-6.1_all.deb ...
Unpacking lmodern (2.004.5-6.1) ...
Selecting previously unselected package preview-latex-style.
Preparing to unpack .../38-preview-latex-style_12.2-1ubuntu1_all.deb ...
Unpacking preview-latex-style (12.2-1ubuntu1) ...
Selecting previously unselected package t1utils.
Preparing to unpack .../39-t1utils_1.41-4build2_amd64.deb ...
Unpacking t1utils (1.41-4build2) ...
Selecting previously unselected package teckit.
Preparing to unpack .../40-teckit_2.5.11+ds1-1_amd64.deb ...
Unpacking teckit (2.5.11+ds1-1) ...
Selecting previously unselected package tex-gyre.
Preparing to unpack .../41-tex-gyre_20180621-3.1_all.deb ...
Unpacking tex-gyre (20180621-3.1) ...
Selecting previously unselected package texlive-binaries.
Preparing to unpack .../42-texlive-
binaries_2021.20210626.59705-1ubuntu0.1_amd64.deb ...
Unpacking texlive-binaries (2021.20210626.59705-1ubuntu0.1) ...
Selecting previously unselected package texlive-base.
Preparing to unpack .../43-texlive-base_2021.20220204-1_all.deb ...
Unpacking texlive-base (2021.20220204-1) ...
Selecting previously unselected package texlive-fonts-recommended.
Preparing to unpack .../44-texlive-fonts-recommended_2021.20220204-1_all.deb ...
Unpacking texlive-fonts-recommended (2021.20220204-1) ...
Selecting previously unselected package texlive-latex-base.
Preparing to unpack .../45-texlive-latex-base_2021.20220204-1_all.deb ...
Unpacking texlive-latex-base (2021.20220204-1) ...
Selecting previously unselected package libfontbox-java.

```

```

Preparing to unpack .../46-libfontbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libfontbox-java (1:1.8.16-2) ...
Selecting previously unselected package libpdfbox-java.
Preparing to unpack .../47-libpdfbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libpdfbox-java (1:1.8.16-2) ...
Selecting previously unselected package texlive-latex-recommended.
Preparing to unpack .../48-texlive-latex-recommended_2021.20220204-1_all.deb ...
Unpacking texlive-latex-recommended (2021.20220204-1) ...
Selecting previously unselected package texlive-pictures.
Preparing to unpack .../49-texlive-pictures_2021.20220204-1_all.deb ...
Unpacking texlive-pictures (2021.20220204-1) ...
Selecting previously unselected package texlive-latex-extra.
Preparing to unpack .../50-texlive-latex-extra_2021.20220204-1_all.deb ...
Unpacking texlive-latex-extra (2021.20220204-1) ...
Selecting previously unselected package texlive-plain-generic.
Preparing to unpack .../51-texlive-plain-generic_2021.20220204-1_all.deb ...
Unpacking texlive-plain-generic (2021.20220204-1) ...
Selecting previously unselected package tipa.
Preparing to unpack .../52-tipa_2%3a1.3-21_all.deb ...
Unpacking tipa (2:1.3-21) ...
Selecting previously unselected package texlive-xetex.
Preparing to unpack .../53-texlive-xetex_2021.20220204-1_all.deb ...
Unpacking texlive-xetex (2021.20220204-1) ...
Setting up fonts-lato (2.0-2.1) ...
Setting up fonts-noto-mono (20201225-1build1) ...
Setting up libwoff1:amd64 (1.0.2-1build4) ...
Setting up libtexlua53:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up libijs-0.35:amd64 (0.35-15build2) ...
Setting up libtexluajit2:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up libfontbox-java (1:1.8.16-2) ...
Setting up rubygems-integration (1.18) ...
Setting up libzip-0-13:amd64 (0.13.72+dfsg.1-1.1) ...
Setting up fonts-urw-base35 (20200910-1) ...
Setting up poppler-data (0.4.11-1) ...
Setting up tex-common (6.17) ...
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based
frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line
78.)
debconf: falling back to frontend: Readline
update-language: texlive-base not installed and configured, doing nothing!
Setting up libfontenc1:amd64 (1:1.1.4-1build3) ...
Setting up libjbig2dec0:amd64 (0.19-3build2) ...
Setting up libteckit0:amd64 (2.5.11+ds1-1) ...
Setting up libapache-pom-java (18-1) ...
Setting up ruby-net-telnet (0.1.1-2) ...
Setting up xfonts-encodings (1:1.0.5-0ubuntu2) ...
Setting up t1utils (1.41-4build2) ...

```

```

Setting up libidn12:amd64 (1.38-4ubuntu1) ...
Setting up fonts-texgyre (20180621-3.1) ...
Setting up libkpathsea6:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up ruby-webrick (1.7.0-3) ...
Setting up fonts-lmodern (2.004.5-6.1) ...
Setting up fonts-droid-fallback (1:6.0.1r16-1.1build1) ...
Setting up ruby-xmlrpc (0.3.2-1ubuntu0.1) ...
Setting up libsyntax2:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up libgs9-common (9.55.0~dfsg1-0ubuntu5.3) ...
Setting up teckit (2.5.11+ds1-1) ...
Setting up libpdfbox-java (1:1.8.16-2) ...
Setting up libgs9:amd64 (9.55.0~dfsg1-0ubuntu5.3) ...
Setting up preview-latex-style (12.2-1ubuntu1) ...
Setting up libcommons-parent-java (43-1) ...
Setting up dvisvgm (2.13.1-1) ...
Setting up libcommons-logging-java (1.2-2) ...
Setting up xfonts-utils (1:7.7+6build2) ...
Setting up libptexenc1:amd64 (2021.20210626.59705-1ubuntu0.1) ...
Setting up texlive-binaries (2021.20210626.59705-1ubuntu0.1) ...
update-alternatives: using /usr/bin/xdvi-xaw to provide /usr/bin/xdvi.bin
(xdvi.bin) in auto mode
update-alternatives: using /usr/bin/bibtex.original to provide /usr/bin/bibtex
(bibtex) in auto mode
Setting up lmodern (2.004.5-6.1) ...
Setting up texlive-base (2021.20220204-1) ...
/usr/bin/ucfr
/usr/bin/ucfr
/usr/bin/ucfr
/usr/bin/ucfr
mktexlsr: Updating /var/lib/texmf/ls-R-TEXLIVEDIST...
mktexlsr: Updating /var/lib/texmf/ls-R-TEXMFMAIN...
mktexlsr: Updating /var/lib/texmf/ls-R...
mktexlsr: Done.

```