# Install and load required packages

install.packages(c("tm", "topicmodels"))

library(tm)

library(topicmodels)

# Sample documents

documents <- c(

"Understanding the fundamentals of topic modeling in R and its applications.",

"Exploring techniques to uncover hidden topics in a diverse collection of documents.",

"A deep dive into Latent Dirichlet Allocation (LDA) for advanced topic analysis.",

"R, a versatile and powerful language for comprehensive data analysis and visualization.",

"Enhancing topic modeling results by carefully curating and removing irrelevant stop words."

)

# Preprocess the data

corpus <- Corpus(VectorSource(documents))

corpus <- tm\_map(corpus, content\_transformer(tolower))

corpus <- tm\_map(corpus, removePunctuation)

corpus <- tm\_map(corpus, removeNumbers)

corpus <- tm\_map(corpus, removeWords, stopwords("en"))

corpus <- tm\_map(corpus, stripWhitespace)

# Create a document-term matrix

dtm <- DocumentTermMatrix(corpus)

# Build an LDA model with two topics

lda\_model <- LDA(dtm, k = 2)

# Extract topic proportions from the LDA model

topics <- as.data.frame(topics(lda\_model)$topics)

# Rename the columns

colnames(topics) <- c("Topic 1", "Topic 2")

# Visualize topics using a bar plot

barplot(t(topics), beside = TRUE, col = c("skyblue", "salmon"),

main = "Topic Distribution in Documents",

xlab = "Documents", ylab = "Topic Probability")

# Add legend

legend("topright", legend = colnames(topics), fill = c("skyblue", "salmon"),

title = "Topics")