

MDS503

8016013

2024-06-27

6

```
##-----
```

```
library(igraph)
```

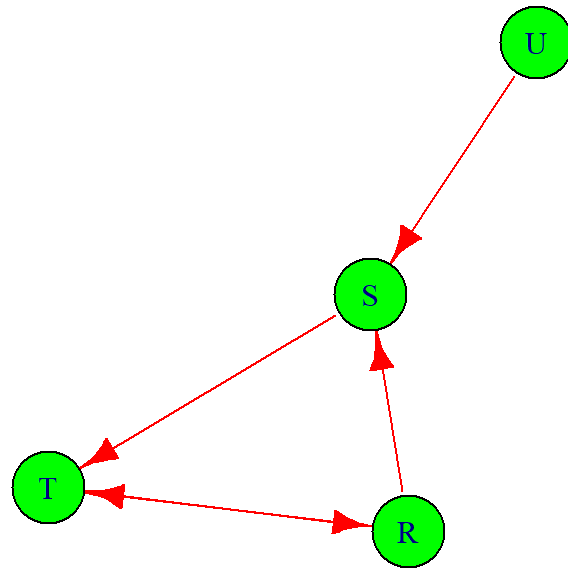
```
## Warning: package 'igraph' was built under R version 4.3.3
```

```
##  
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:stats':  
##  
##      decompose, spectrum
```

```
## The following object is masked from 'package:base':  
##  
##      union
```

```
#a  
#Defining a graph object  
g1<-graph(c("R","S","S","T","T","R","R","T","U","S"))  
  
#b  
plot(g1,vertex.color="green",vertex.size=30,edge.color="red",edge.size=5)
```



#this graph represent the relation between various nodes

```
#c
degree(g1)
```

```
## R S T U
## 3 3 3 1
```

*#degree states that how many relations a single node is holding then
#from the result we see that R,S,T have 3 relations but U has only one relation*

```
closeness(g1)
```

```
##      R      S      T      U
## 0.5000000 0.3333333 0.3333333 0.1666667
```

#closeness of the nodes is how close the node is with the other nodes

```
betweenness(g1) #it gives how many nodes have the relation with that node
```

```
## R S T U
## 1 2 2 0
```

#here S has 2 betweenness that means 2 nodes have relation with S

```
#d
# Identify hubs in the graph
hubs <- which(degree(g1) == max(degree(g1)))
hubs
```

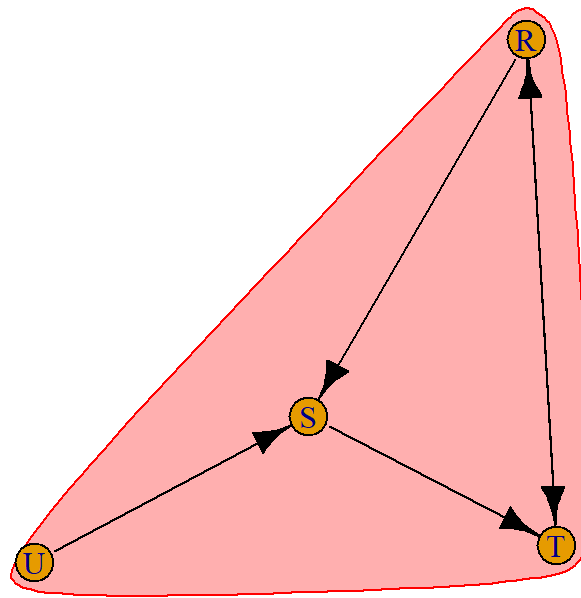
```
## R S T
## 1 2 3
```

*#Hubs in a graph refer to nodes with high connectivity or degree that
serve as central points of the network.*

```
# Find communities in the graph
communities <- cluster_walktrap(g1)
cat("Number of communities: ", length(communities), "\n")
```

```
## Number of communities: 1
```

```
# Visualize the graph with communities highlighted
plot(communities, g1)
```



*#Communities in a graph represent groups of nodes that are more densely connected
#within the group compared to connections between groups*

7

```
# a  
getwd()
```

```
## [1] "C:/Users/kaush/Desktop/8016013"
```

```
library(pdftools)
```

```
## Warning: package 'pdftools' was built under R version 4.3.3
```

```
## Using poppler version 23.08.0
```

```
library(tm)
```

```
## Warning: package 'tm' was built under R version 4.3.3
```

```
## Loading required package: NLP
```

```
library(wordcloud)
```

```
## Warning: package 'wordcloud' was built under R version 4.3.3
```

```
## Loading required package: RColorBrewer
```

```
library(RColorBrewer)  
library(topicmodels)
```

```
## Warning: package 'topicmodels' was built under R version 4.3.3
```

```
# a  
getwd()
```

```
## [1] "C:/Users/kaush/Desktop/8016013"
```

```
# Specify the path to your PDF file  
#Question b  
pdf_file <- "R-intro.pdf"
```

```
library(tm)  
#Question c  
# Extract text from PDF using pdftools  
pdf_text <- pdf_text(pdf_file)
```

```
# Convert the text into a corpus using tm package  
corpus <- Corpus(VectorSource(pdf_text))
```

```
# Preprocessing: Convert to lower case, remove numbers and punctuation  
corpus <- tm_map(corpus, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(corpus, content_transformer(tolower)):  
## transformation drops documents
```

```
corpus <- tm_map(corpus, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(corpus, removeNumbers): transformation drops  
## documents
```

```
corpus <- tm_map(corpus, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(corpus, removePunctuation): transformation drops
## documents
```

```
corpus <- tm_map(corpus, removeWords, stopwords("en")) # Remove English stopwords
```

```
## Warning in tm_map.SimpleCorpus(corpus, removeWords, stopwords("en")):
## transformation drops documents
```

```
# Strip whitespace
corpus <- tm_map(corpus, stripWhitespace)
```

```
## Warning in tm_map.SimpleCorpus(corpus, stripWhitespace): transformation drops
## documents
```

```
corpus<-tm_map(corpus,stemDocument)
```

```
## Warning in tm_map.SimpleCorpus(corpus, stemDocument): transformation drops
## documents
```

```
#Question d
# Create Document Term Matrix
myTdm<- TermDocumentMatrix(corpus,control = list(wordLengths=c(1,Inf)))

# Get term frequencies
(freq.terms<-findFreqTerms(myTdm,lowfreq = 150))
```

```
## [1] "data"      "r"         "function"  "use"       "valu"      "vector"
## [7] "model"     "plot"      "can"       "x"         "y"
```

```
# Subset term frequencies for most frequent terms
m<-as.matrix(myTdm)
# Sort frequencies in descending order
(freq<-sort(rowSums(m),decreasing = T))
```

##	function	use
##	420	397
##	r	vector
##	363	254
##	x	can
##	239	200
##	data	plot
##	194	175
##	model	y
##	165	164
##	valu	file
##	155	146
##	exampl	list
##	144	135
##	graphic	name
##	134	134
##	argument	will
##	126	120
##	object	command
##	116	114
##	array	may
##	112	109
##	variabl	matrix
##	103	101
##	packag	b
##	93	90
##	c	charact
##	86	86
##	number	see
##	82	82
##	line	one
##	81	80
##	form	way
##	80	80
##	factor	also
##	79	78
##	paramet	chapter
##	77	77
##	fit	set
##	76	75
##	frame	two
##	74	73
##	devic	distribut
##	68	67
##	first	oper
##	66	65
##	case	express
##	64	64
##	point	result
##	63	62
##	given	compon
##	62	61

##	window	default
##	60	60
##	avail	•
##	59	58
##	assign	mean
##	56	56
##	m	defin
##	55	55
##	specifi	index
##	55	54
##	column	statist
##	54	53
##	differ	give
##	53	52
##	page	current
##	52	52
##	make	length
##	51	51
##	director	call
##	51	51
##	need	allow
##	49	49
##	type	element
##	48	48
##	environ	n
##	47	47
##	note	numer
##	47	47
##	linear	follow
##	46	46
##	posit	label
##	45	45
##	help	includ
##	44	44
##	produc	general
##	44	43
##	mode	order
##	43	43
##	work	facil
##	42	42
##	matric	user
##	42	42
##	regress	provid
##	42	41
##	class	system
##	41	40
##	print	normal
##	40	40
##	attribut	session
##	38	38
##	creat	text
##	38	38

##	simpl	run
##	37	37
##	z	d
##	37	36
##	term	chang
##	36	36
##	search	inform
##	36	36
##	f	sampl
##	36	36
##	s	interact
##	36	35
##	read	figur
##	35	35
##	section	second
##	35	35
##	multipl	sequenc
##	34	34
##	either	new
##	34	33
##	mani	howev
##	33	33
##	size	squar
##	33	32
##	within	possibl
##	32	32
##	path	method
##	31	31
##	return	option
##	31	31
##	row	level
##	31	31
##	analysi	''
##	30	30
##	appendix	particular
##	30	30
##	singl	control
##	30	29
##	invok	alway
##	29	29
##	start	contain
##	29	29
##	special	usual
##	29	29
##	version	select
##	28	28
##	write	standard
##	28	28
##	unix	k
##	28	28
##	program	generat
##	27	27

##	test	famili
##	27	27
##	calcul	t
##	27	27
##	execut	tabl
##	26	26
##	determin	refer
##	26	26
##	must	often
##	26	26
##	respect	true
##	26	26
##	respons	axi
##	26	26
##	product	edit
##	25	25
##	like	time
##	25	25
##	remov	generic
##	24	24
##	display	formula
##	24	24
##	e	find
##	24	24
##	error	estim
##	24	24
##	output	attach
##	23	23
##	procedur	detail
##	23	23
##	similar	v
##	23	23
##	logic	varianc
##	22	22
##	input	larg
##	22	22
##	rather	thus
##	22	22
##	altern	evalu
##	22	22
##	addit	three
##	22	22
##	group	structur
##	21	21
##	place	part
##	21	21
##	-	add
##	21	21
##	document	least
##	20	20
##	commandlin	describ
##	20	20

##	suppli	consid
##	20	20
##	end	condit
##	20	19
##	axe	filepath
##	19	19
##	postscript	correspond
##	19	19
##	main	separ
##	19	19
##	string	sinc
##	19	19
##	subscript	arg
##	19	19
##	lst	df
##	19	19
##	w	anoth
##	18	18
##	except	outer
##	18	18
##	load	margin
##	18	18
##	nonlinear	comput
##	18	18
##	save	want
##	18	18
##	code	known
##	18	18
##	process	even
##	18	18
##	na	suppos
##	18	18
##	expr	startup
##	18	18
##	driver	extract
##	17	17
##	font	mark
##	17	17
##	par	effect
##	17	17
##	well	discuss
##	17	17
##	quit	consist
##	17	17
##	unless	just
##	17	17
##	complex	eg
##	17	17
##	left	nsw
##	17	17
##	cmd	manipul
##	17	16

##	specif	access
##	16	16
##	detach	loop
##	16	16
##	statement	maco
##	16	16
##	assum	local
##	16	16
##	problem	termin
##	16	16
##	automat	link
##	16	16
##	symbol	complet
##	16	16
##	entri	indic
##	16	16
##	explicit	total
##	16	16
##	look	dimens
##	16	16
##	languag	concaten
##	15	15
##	construct	probabl
##	15	15
##	highlevel	script
##	15	15
##	import	written
##	15	15
##	collect	direct
##	15	15
##	exist	later
##	15	15
##	much	shell
##	15	15
##	take	pass
##	15	15
##	quot	next
##	15	15
##	restor	common
##	15	15
##	log	experi
##	15	15
##	right	qld
##	15	15
##	vic	appropri
##	15	15
##	initi	state
##	15	14
##	arithmet	block
##	14	14
##	dataset	tick
##	14	14

##	style	support
##	14	14
##	appli	done
##	14	14
##	doubl	wish
##	14	14
##	locat	omit
##	14	14
##	workspac	sever
##	14	14
##	compar	everi
##	14	14
##	custom	design
##	13	13
##	repeat	compress
##	13	13
##	contrast	os
##	13	13
##	now	via
##	13	13
##	conveni	store
##	13	13
##	depend	instal
##	13	13
##	key	togeth
##	13	13
##	whose	equal
##	13	13
##	empti	datafram
##	13	13
##	dim	sa
##	13	13
##	wa	show
##	13	13
##	last	u
##	13	13
##	p	fm
##	13	13
##	featur	get
##	12	12
##	previous	frequenc
##	12	12
##	invers	editor
##	12	12
##	summari	origin
##	12	12
##	formal	tool
##	12	12
##	henc	prompt
##	12	12
##	without	various
##	12	12

##	rang	connect
##	12	12
##	requir	varieti
##	12	12
##	densiti	interv
##	12	12
##	fa	home
##	12	12
##	intercept	introduct
##	12	11
##	copi	entir
##	11	11
##	manual	relat
##	11	11
##	rule	cbind
##	11	11
##	lowlevel	base
##	11	11
##	john	ad
##	11	11
##	limit	readlin
##	11	11
##	occur	still
##	11	11
##	taken	seq
##	11	11
##	becom	act
##	11	11
##	affect	h
##	11	11
##	space	legend
##	11	11
##	ident	permiss
##	10	10
##	etc	recycl
##	10	10
##	transpos	arbitrari
##	10	10
##	effici	integr
##	10	10
##	scan	maximum
##	10	10
##	namespac	actual
##	10	10
##	made	reader
##	10	10
##	chamber	appear
##	10	10
##	might	q
##	10	10
##	continu	cursor
##	10	10

##	five	perform
##	10	10
##	simplest	zero
##	10	10
##	fals	short
##	10	10
##	sometim	identifi
##	10	10
##	r's	scale
##	10	10
##	pvalu	region
##	10	10
##	fun	go
##	10	10
##	user'	gaussian
##	10	10
##	plotx	batch
##	10	10
##	width	posix
##	10	10
##	modifi	notic
##	9	9
##	perman	recal
##	9	9
##	decomposit	eigenvalu
##	9	9
##	rbind	scope
##	9	9
##	extend	splus
##	9	9
##	precis	thing
##	9	9
##	full	necessari
##	9	9
##	say	definit
##	9	9
##	enter	insid
##	9	9
##	move	associ
##	9	9
##	context	cours
##	9	9
##	equival	integ
##	9	9
##	purpos	exact
##	9	9
##	negat	treat
##	9	9
##	account	nt
##	9	9
##	caus	obtain
##	9	9

##	achiev	poisson
##	9	9
##	profil	impli
##	9	9
##	residu	β
##	9	9
##	mous	titl
##	9	9
##	bottom	filesystem
##	9	9
##	stdin	preliminari
##	9	8
##	tappli	unord
##	8	8
##	binari	anova
##	8	8
##	hershey	likelihood
##	8	8
##	temporari	updat
##	8	8
##	almost	small
##	8	8
##	basic	develop
##	8	8
##	four	frequent
##	8	8
##	handl	intend
##	8	8
##	hold	minim
##	8	8
##	other	format
##	8	8
##	mechan	put
##	8	8
##	vertic	recommend
##	8	8
##	match	simpli
##	8	8
##	begin	real
##	8	8
##	quantiti	less
##	8	8
##	unit	statef
##	8	8
##	tas	implicit
##	8	8
##	repres	clean
##	8	8
##	onlin	binomi
##	8	8
##	global	imag
##	8	8

##	color	accept
##	8	8
##	meta	copyright
##	8	7
##	miss	builtin
##	7	7
##	orient	qr
##	7	7
##	readtabl	recurs
##	7	7
##	glm	action
##	7	7
##	concept	implement
##	7	7
##	prefer	site
##	7	7
##	although	expect
##	7	7
##	somewhat	step
##	7	7
##	convent	tri
##	7	7
##	word	consol
##	7	7
##	histori	sourc
##	7	7
##	rdata	arrang
##	7	7
##	complic	ordinari
##	7	7
##	sens	observ
##	7	7
##	replac	alreadi
##	7	7
##	individu	side
##	7	7
##	applic	slight
##	7	7
##	ttest	desir
##	7	7
##	ab	found
##	7	7
##	long	chosen
##	7	7
##	o	interpret
##	7	7
##	lm	coeffici
##	7	7
##	project	break
##	7	7
##	extens	top
##	7	7

##	visibl	final
##	7	7
##	age	dummi
##	7	7
##	lti	hypothesi
##	7	7
##	valid	sum
##	7	7
##	vari	cube
##	7	7
##	fb	balanc
##	7	7
##	measur	rprofil
##	7	7
##	μ	build
##	7	7
##	open	contour
##	7	7
##	fill	press
##	7	7
##	pch	correct
##	7	6
##	mix	regular
##	6	6
##	subset	advanc
##	6	6
##	drop	eigenvector
##	6	6
##	examin	singular
##	6	6
##	cran	dynam
##	6	6
##	back	book
##	6	6
##	httpscranrprojectorg	littl
##	6	6
##	“	futur
##	6	6
##	issu	whether
##	6	6
##	digit	restrict
##	6	6
##	solv	technic
##	6	6
##	topic	extern
##	6	6
##	horizont	librari
##	6	6
##	gui	temp
##	6	6
##	intern	nan
##	6	6

##	sqrt	warn
##	6	6
##	coerc	item
##	6	6
##	best	classif
##	6	6
##	incom	suitabl
##	6	6
##	confid	functionx
##	6	6
##	split	simpler
##	6	6
##	stand	longer
##	6	6
##	operand	present
##	6	6
##	grid	better
##	6	6
##	instead	an
##	6	6
##	choic	orthogon
##	6	6
##	latter	notat
##	6	6
##	inp	nn
##	6	6
##	boxplot	messag
##	6	6
##	ep	lim
##	6	6
##	panel	shown
##	6	6
##	fd	free
##	6	6
##	unset	inclus
##	6	6
##	strata	encod
##	6	6
##	lmi	weight
##	6	6
##	vm	xfit
##	6	6
##	mar	tree
##	6	6
##	scatterplot	draw
##	6	6
##	mai	-----
##	6	6
##	pdf	slash
##	6	6
##	addon	noinitfil
##	6	6

##	foor	rscrip
##	6	6
##	venabl	grant
##	5	5
##	-	subsect
##	5	5
##	equat	twosampl
##	5	5
##	annot	familiar
##	5	5
##	happen	suggest
##	5	5
##	among	classic
##	5	5
##	j	keep
##	5	5
##	typic	letter
##	5	5
##	subsequ	wherea
##	5	5
##	helpstart	sent
##	5	5
##	capabl	entiti
##	5	5
##	ess	high
##	5	5
##	lead	constant
##	5	5
##	got	longest
##	5	5
##	perhap	shorter
##	5	5
##	max	min
##	5	5
##	smallest	rep
##	5	5
##	bracket	immedi
##	5	5
##	sensibl	easier
##	5	5
##	delet	hand
##	5	5
##	combin	otherwis
##	5	5
##	major	clear
##	5	5
##	amount	histogram
##	5	5
##	diagon	avoid
##	5	5
##	close	sign
##	5	5

##	bind	onto
##	5	5
##	seen	pair
##	5	5
##	interfac	ncp
##	5	5
##	quantil	kolmogorovsmirnov
##	5	5
##	coplot	extra
##	5	5
##	enough	1
##	5	5
##	earlier	rest
##	5	5
##	approxim	ld
##	5	5
##	0	coordin
##	5	5
##	drawn	subtitl
##	5	5
##	height	button
##	5	5
##	-----	unixalik
##	5	5
##	noenviron	norestor
##	5	5
##	nosav	rguiex
##	5	5
##	vanilla	insert
##	5	5
##	smith	divert
##	4	4
##	introductori	sensit
##	4	4
##	softwar	partit
##	4	4
##	contribut	mathemat
##	4	4
##	multivari	nonstandard
##	4	4
##	iv	faq
##	4	4
##	hardcopi	hasti
##	4	4
##	mention	regard
##	4	4
##	techniqu	think
##	4	4
##	ask	easi
##	4	4
##	expert	guid
##	4	4

##	occasion	seek
##	4	4
##	shall	escap
##	4	4
##	html	man
##	4	4
##	arrow	byte
##	4	4
##	forward	keyboard
##	4	4
##	ls	menu
##	4	4
##	record	hard
##	4	4
##	cx	”
##	4	4
##	exp	lengthx
##	4	4
##	neither	power
##	4	4
##	sort	var
##	4	4
##	ind	kind
##	4	4
##	append	evid
##	4	4
##	past	sub
##	4	4
##	fact	coercion
##	4	4
##	care	certain
##	4	4
##	former	old
##	4	4
##	outsid	rare
##	4	4
##	australia	territori
##	4	4
##	sd	someth
##	4	4
##	confus	dimz
##	4	4
##	ignor	doubli
##	4	4
##	xi	cosi
##	4	4
##	random	rank
##	4	4
##	uniform	diag
##	4	4
##	smaller	xt
##	4	4

##	absolut	diagnost
##	4	4
##	eigensm	lsfit
##	4	4
##	leav	autoload
##	4	4
##	globalenv	util
##	4	4
##	usercontribut	gamma
##	4	4
##	logi	loglikelihood
##	4	4
##	shape	wilcoxon
##	4	4
##	≤	rd
##	4	4
##	smooth	dopointsfals
##	4	4
##	qq	verticalstru
##	4	4
##	aspect	stat
##	4	4
##	higher	distanc
##	4	4
##	compil	backslash
##	4	4
##	bodi	fund
##	4	4
##	graph	graphtru
##	4	4
##	exit	divid
##	4	4
##	sq	behavior
##	4	4
##	coef	transform
##	4	4
##	op	slope
##	4	4
##	finish	mm
##	4	4
##	predict	influenc
##	4	4
##	quasilikelihood	φ
##	4	4
##	blind	kalytho
##	4	4
##	logit	probit
##	4	4
##	success	optim
##	4	4
##	yfit	highlight
##	4	4

##	treebas	quartz
##	4	4
##	contourx	pointer
##	4	4
##	layout	convert
##	4	4
##	ensur	jpeg
##	4	4
##	png	usag
##	4	4
##	ose	windows'
##	4	4
##	check	circl
##	4	4
##	rtermex	nositefil
##	4	4
##	debugg	g
##	4	4
##	mb	latex
##	4	4
##	sweav	hall
##	4	4
##	york	core
##	4	3
##	team	preserv
##	3	3
##	content	intrins
##	3	3
##	rag	archiv
##	3	3
##	comparison	come
##	3	3
##	comment	allan
##	3	3
##	becker	built
##	3	3
##	coher	inde
##	3	3
##	methodolog	suit
##	3	3
##	wilk	abbrevi
##	3	3
##	advantag	box
##	3	3
##	cd	receiv
##	3	3
##	yes	elementari
##	3	3
##	enclos	helpsearch
##	3	3
##	inbuilt	invis
##	3	3

##	lost	proceed
##	3	3
##	strong	syntact
##	3	3
##	understand	anywher
##	3	3
##	backward	discard
##	3	3
##	emac	sink
##	3	3
##	speak	analys
##	3	3
##	bar	rmx
##	3	3
##	thought	colon
##	3	3
##	covari	increas
##	3	3
##	independ	largest
##	3	3
##	parallel	permut
##	3	3
##	incomplet	realli
##	3	3
##	distinct	join
##	3	3
##	xisnax	categor
##	3	3
##	unlik	yy
##	3	3
##	treatment	along
##	3	3
##	alpha	convers
##	3	3
##	exercis	temporarili
##	3	3
##	tax	irregular
##	3	3
##	ca	dimc
##	3	3
##	xv	extrem
##	3	3
##	artifici	ycoordin
##	3	3
##	crossprod	matlab
##	3	3
##	seem	xxt
##	3	3
##	eval	qrx
##	3	3
##	dimnam	status
##	3	3

##	incomef	statefr
##	3	3
##	especi	lstname
##	3	3
##	alter	detachlentil
##	3	3
##	lentil	lentilsu
##	3	3
##	packagebas	area
##	3	3
##	centheat	export
##	3	3
##	forc	presum
##	3	3
##	xnew	xold
##	3	3
##	cumul	deviat
##	3	3
##	dxxx	logist
##	3	3
##	lowertail	prob
##	3	3
##	simul	tail
##	3	3
##	bw	good
##	3	3
##	stem	empir
##	3	3
##	erupt	linesx
##	3	3
##	parpti	reason
##	3	3
##	us	qqlinex
##	3	3
##	shapirowilk	percent
##	3	3
##	signific	tend
##	3	3
##	addtru	null
##	3	3
##	round	tie
##	3	3
##	els	ifels
##	3	3
##	lattic	parenthes
##	3	3
##	versus	yb
##	3	3
##	choos	functiondata
##	3	3
##	job	turn
##	3	3

##	attr	difficult
##	3	3
##	r-	safeti
##	3	3
##	answer	integrand
##	3	3
##	interest	activ
##	3	3
##	lexic	deposit
##	3	3
##	openaccount	person
##	3	3
##	accommod	aov
##	3	3
##	summar	degre
##	3	3
##	polynomi	xp
##	3	3
##	yx	abc
##	3	3
##	whole	descript
##	3	3
##	fittedmodel	updatefm
##	3	3
##	dobson	infer
##	3	3
##	practic	predictor
##	3	3
##	stimulus	.
##	3	3
##	glmi	male
##	3	3
##	quasi	superimpos
##	3	3
##	se	ssmicmenx
##	3	3
##	loess	nlme
##	3	3
##	outlier	imaginari
##	3	3
##	maintain	histx
##	3	3
##	score	third
##	3	3
##	colour	dotchart
##	3	3
##	imagex	surfac
##	3	3
##	typen	polygon
##	3	3
##	textx	near
##	3	3

##	oldpar	paroldpar
##	3	3
##	unfortun	tck
##	3	3
##	inch	mfrowc
##	3	3
##	omi	reduc
##	3	3
##	pointsiz	printer
##	3	3
##	bitmap	horizontalfals
##	3	3
##	memori	packages”
##	3	3
##	prevent	retriev
##	3	3
##	cp	map
##	3	3
##	port	rapp
##	3	3
##	rgui	drive
##	3	3
##	recent	summaryfm
##	3	3
##	expt	light
##	3	3
##	speed	renviron
##	3	3
##	frontend	metacharact
##	3	3
##	assert	commandargstru
##	3	3
##	enabl	quiet
##	3	3
##	redirect	configur
##	3	3
##	emacsstyl	pl
##	3	3
##	runfoo	esc
##	3	3
##	gnu	resubmit
##	3	3
##	ret	cf
##	3	3
##	chapman	london
##	3	3
##	deriv	translat
##	2	2
##	verbatim	prefac
##	2	2
##	manag	repetit
##	2	2

##	diagram	typeset
##	2	2
##	bill	david
##	2	2
##	materi	preced
##	2	2
##	univers	wait
##	2	2
##	welcom	basi
##	2	2
##	character	cover
##	2	2
##	elsewher	essenti
##	2	2
##	increment	inflex
##	2	2
##	intermedi	internet
##	2	2
##	piec	releas
##	2	2
##	richard	storag
##	2	2
##	trevor	yet
##	2	2
##	's'	"standard"
##	2	2
##	adjust	aim
##	2	2
##	bring	cancel
##	2	2
##	click	customiz
##	2	2
##	folder	launch
##	2	2
##	machin	mkdir
##	2	2
##	noth	occas
##	2	2
##	seri	shortcut
##	2	2
##	straightforward	subdirectori
##	2	2
##	workstat	alphanumeric
##	2	2
##	browser	countri
##	2	2
##	offer	plus
##	2	2
##	request	though
##	2	2
##	unlimit	web
##	2	2

##	brace	del
##	2	2
##	everyth	newlin
##	2	2
##	silent	“r
##	2	2
##	conduct	reload
##	2	2
##	rhistori	rm
##	2	2
##	fraction	middl
##	2	2
##	obvious	reciproc
##	2	2
##	strict	unchang
##	2	2
##	concern	cos
##	2	2
##	flexibl	minimum
##	2	2
##	pmax	pmin
##	2	2
##	sin	tan
##	2	2
##	available”	inf
##	2	2
##	isnax	overwritten
##	2	2
##	replic	repx
##	2	2
##	reserv	situat
##	2	2
##	socal	“”
##	2	2
##	denot	emerg
##	2	2
##	iter	lab
##	2	2
##	nonmiss	ss
##	2	2
##	substr	admit
##	2	2
##	compact	lie
##	2	2
##	lunch	sixth
##	2	2
##	attributesobject	cater
##	2	2
##	consult	lengthobject
##	2	2
##	properti	raw
##	2	2

##	represent	unambigu
##	2	2
##	agre	retain
##	2	2
##	unclass	winter
##	2	2
##	briefli	idea
##	2	2
##	learn	alphabet
##	2	2
##	discret	south
##	2	2
##	tapplyincom	incster
##	2	2
##	natur	percentag
##	2	2
##	stderror	subclass
##	2	2
##	fortran	multipli
##	2	2
##	nonneg	onedimension
##	2	2
##	referenc	cbindn
##	2	2
##	incid	trivial
##	2	2
##	encount	fewer
##	2	2
##	tableblock	aperma
##	2	2
##	fr	ineffici
##	2	2
##	outerx	perm
##	2	2
##	surpris	bb
##	2	2
##	cbindx	crossprodx
##	2	2
##	eigen	potenti
##	2	2
##	quadrat	rbindx
##	2	2
##	root	solut
##	2	2
##	solvea	solveab
##	2	2
##	tx	absdet
##	2	2
##	alli	det
##	2	2
##	ev	orthonorm
##	2	2

##	prodsvdmd	sm
##	2	2
##	svdm	asvector
##	2	2
##	complement	cyclic
##	2	2
##	discov	res
##	2	2
##	ultim	vec
##	2	2
##	cross	cut
##	2	2
##	besid	sublist
##	2	2
##	asdatafram	conform
##	2	2
##	anyth	comfort
##	2	2
##	mask	safer
##	2	2
##	vw	\b
##	2	2
##	abl	floor
##	2	2
##	held	housepric
##	2	2
##	perl	presumpt
##	2	2
##	price	readtablehousesdata
##	2	2
##	room	around
##	2	2
##	headertru	ncol
##	2	2
##	remain	scaninputdat
##	2	2
##	editxold	packagedataset
##	2	2
##	beta	cauchi
##	2	2
##	cdf	comprehens
##	2	2
##	logp	multinomi
##	2	2
##	pxxx	qxxx
##	2	2
##	rate	student'
##	2	2
##	upper	weibul
##	2	2
##	'r'	bandwidth
##	2	2

##	bin	decim
##	2	2
##	eleg	fivenum
##	2	2
##	hist	histerupt
##	2	2
##	qu	ecdf
##	2	2
##	far	let
##	2	2
##	quantilequantil	righthand
##	2	2
##	qqnormx	rice
##	2	2
##	twosid	ratio
##	2	2
##	ttesta	cdfs
##	2	2
##	kstesta	maxim
##	2	2
##	qqplot	wilcoxtesta
##	2	2
##	accord	https://cran.r-project.org/package=lattice
##	2	2
##	larger	yci
##	2	2
##	clearer	skip
##	2	2
##	view	bslash
##	2	2
##	emul	lengthi
##	2	2
##	meani	ordinarili
##	2	2
##	tst	tstat
##	2	2
##	xy	x-
##	2	2
##	altogeth	furthermor
##	2	2
##	involv	minor
##	2	2
##	parpch	adapt
##	2	2
##	algorithm	canon
##	2	2
##	functionf	heavi
##	2	2
##	illustr	nodimnam
##	2	2
##	puzzl	routin
##	2	2

##	svd	funf
##	2	2
##	bank	catamount
##	2	2
##	don't	functionamount
##	2	2
##	ifamount	mimic
##	2	2
##	rosswithdraw	toplevel
##	2	2
##	withdraw	graphicsoff
##	2	2
##	hidden	rossbal
##	2	2
##	rprofilesit	rprofileus
##	2	2
##	objectcoef	variant
##	2	2
##	ziszaz	ei
##	2	2
##	polyx	suffici
##	2	2
##	yi	~
##	2	2
##	period	yax
##	2	2
##	contrpoli	introduc
##	2	2
##	optionscontrast	paper
##	2	2
##	textbook	brief
##	2	2
##	concis	criterion
##	2	2
##	devianc	submodel
##	2	2
##	vcov	blocks"
##	2	2
##	datadatafram	multistratum
##	2	2
##	newformula	npk
##	2	2
##	strataformula	"within
##	2	2
##	assumpt	explor
##	2	2
##	mccullagh	nelder
##	2	2
##	prior	regressor
##	2	2
##	subject	wide
##	2	2

##	η	cloglog
##	2	2
##	inhabit	island
##	2	2
##	mu	silvey
##	2	2
##	analyz	ez
##	2	2
##	failur	glmymat
##	2	2
##	ldl	twocolumn
##	2	2
##	year	bate
##	2	2
##	converg	critic
##	2	2
##	curv	fn
##	2	2
##	functionp	guess
##	2	2
##	nlm	nlminb
##	2	2
##	qualiti	sse
##	2	2
##	watt	dataframexx
##	2	2
##	hessian	linessplinesfit
##	2	2
##	nlmfn	outestim
##	2	2
##	outminimum	\pm
##	2	2
##	ace	ava
##	2	2
##	bruto	insight
##	2	2
##	lme	lqs
##	2	2
##	mass	nonparametr
##	2	2
##	pick	reduct
##	2	2
##	resist	rlm
##	2	2
##	robust	know
##	2	2
##	plotf	versatil
##	2	2
##	coplota	straight
##	2	2
##	axesfals	count
##	2	2

##	persp	suppress
##	2	2
##	typel	xaxi
##	2	2
##	across	clockwis
##	2	2
##	mtext	rotat
##	2	2
##	adj	disabl
##	2	2
##	identifyx	justif
##	2	2
##	nearbi	parcol
##	2	2
##	bold	center
##	2	2
##	col	expans
##	2	2
##	greek	ital
##	2	2
##	justifi	palett
##	2	2
##	plain	edg
##	2	2
##	mgpc	surround
##	2	2
##	mfgc	oma
##	2	2
##	share	devoff
##	2	2
##	onefilefals	devlist
##	2	2
##	devnext	devprev
##	2	2
##	whichk	clash
##	2	2
##	davison	hinkley
##	2	2
##	protect	writer
##	2	2
##	doublecolon	getanywher
##	2	2
##	analogu	casesensit
##	2	2
##	compliant	filetest
##	2	2
##	foopng	lower
##	2	2
##	platform	servic
##	2	2
##	crossplatform	header
##	2	2

##	normalizepath	pathexpand
##	2	2
##	recogn	slower
##	2	2
##	url	xz
##	2	2
##	zip	ablin
##	2	2
##	ablinecoeffm	datadummi
##	2	2
##	lrf	rnorm
##	2	2
##	len	mainresidu
##	2	2
##	michelson	pi
##	2	2
##	rmfm	seqpi
##	2	2
##	'runs'	linesz
##	2	2
##	nlevel	plotw
##	2	2
##	xlimc	ylabi
##	2	2
##	ylimc	norestoredata
##	2	2
##	nosit	osspecif
##	2	2
##	resid	sitewid
##	2	2
##	tmpdir	'namevalue'
##	2	2
##	"directly"	noninteract
##	2	2
##	norconsol	rconsol
##	2	2
##	reencod	rhome
##	2	2
##	cell	flag
##	2	2
##	noecho	noreadlin
##	2	2
##	stdout	trigger
##	2	2
##	verbos	'tk'
##	2	2
##	'x'	rdconv
##	2	2
##	rdtxt	shlib
##	2	2
##	stderr	texifi
##	2	2

##	vignett	consolebas
##	2	2
##	directory'	fail
##	2	2
##	latexex	mdi
##	2	2
##	mydoc	parent
##	2	2
##	sdi	sh
##	2	2
##	texinput	'home
##	2	2
##	background	eof
##	2	2
##	usrbinenv	cn
##	2	2
##	motion	resubmiss
##	2	2
##	vi	cb
##	2	2
##	"save"	approv
##	2	1
##	gentleman	ihaka
##	1	1
##	ion	maechler
##	1	1
##	ii	iii
##	1	1
##	address	adelaid
##	1	1
##	digest	email
##	1	1
##	expand	feedback
##	1	1
##	instant	novic
##	1	1
##	pleas	reflect
##	1	1
##	rhelprprojectorg	thank
##	1	1
##	warm	accret
##	1	1
##	bell	coauthor
##	1	1
##	ephemer	evolut
##	1	1
##	fulli	laboratori
##	1	1
##	latest	mind
##	1	1
##	modern	newli
##	1	1

##	peopl	plan
##	1	1
##	prepar	rapid
##	1	1
##	rick	r”
##	1	1
##	ssplus	vehicl
##	1	1
##	“environment”	“recommended”
##	1	1
##	bulk	copious
##	1	1
##	dialog	difficulti
##	1	1
##	discrimin	field
##	1	1
##	icon	interrog
##	1	1
##	philosophi	respond
##	1	1
##	said	sas
##	1	1
##	spss	tedious
##	1	1
##	vast	whenev
##	1	1
##	accent	advis
##	1	1
##	bear	breadth
##	1	1
##	brows	engin
##	1	1
##	exampletop	feel
##	1	1
##	great	helpsolv
##	1	1
##	htmlbase	hyperlink
##	1	1
##	keywords’	portabl
##	1	1
##	quick	string”
##	1	1
##	syntax	-za-z-
##	1	1
##	‘search	’s
##	1	1
##	“charact	amongst
##	1	1
##	commandsr	compound
##	1	1
##	emacs”	excess
##	1	1

##	hashmark	indent
##	1	1
##	reexecut	scroll
##	1	1
##	semicolon	sinkrecord
##	1	1
##	sourcecommandsr	decid
##	1	1
##	foo	ink
##	1	1
##	junk	meaning
##	1	1
##	opportun	“dot”
##	1	1
##	assignx	lastvalu
##	1	1
##	sidebysid	‘points’
##	1	1
##	“less	“minus”
##	1	1
##	byvalu	cminx
##	1	1
##	fromvalu	irrelev
##	1	1
##	lengthvalu	maxx
##	1	1
##	meanx	nby
##	1	1
##	pby	prioriti
##	1	1
##	prodx	pvariat
##	1	1
##	rais	seqfrom
##	1	1
##	seqlength	seqto
##	1	1
##	sortlist	sortx
##	1	1
##	sqrtd	sumx
##	1	1
##	sumxlengthx	sumxmeanxlengthx
##	1	1
##	tovalu	varx
##	1	1
##	“numbers”	alongvector
##	1	1
##	cna	counterpart
##	1	1
##	endtoend	fifth
##	1	1
##	inequ	intersect
##	1	1

##	isnaz	lengthvector
##	1	1
##	marker	met
##	1	1
##	motiv	undecid
##	1	1
##	union	value"
##	1	1
##	"miss	"missing"
##	1	1
##	backspace-se	blank
##	1	1
##	collapsess	cstyle
##	1	1
##	delimit	differenti
##	1	1
##	isnaxx	isnaxx
##	1	1
##	pastecxi	recreat
##	1	1
##	sep	sepstr
##	1	1
##	tab	xvalu
##	1	1
##	,	absi
##	1	1
##	appl	banana
##	1	1
##	corang	cxyrepc
##	1	1
##	exclud	fruit
##	1	1
##	fruitcappleorang	matrices'
##	1	1
##	matrixlik	meet
##	1	1
##	multidimension	namesfruit
##	1	1
##	peach	per
##	1	1
##	rememb	subvector
##	1	1
##	vectorindexvector	'data
##	1	1
##	project'	amalgam
##	1	1
##	appar	ascharacterz
##	1	1
##	asintegerdigit	assometh
##	1	1
##	atom	attributes"
##	1	1

##	constitu	explain
##	1	1
##	fundament	indirect
##	1	1
##	invest	lengthz
##	1	1
##	modeobject	modez
##	1	1
##	possess	reconstruct
##	1	1
##	revers	roundoff
##	1	1
##	"atomic"	"intrins
##	1	1
##	"property"	"recursive"
##	1	1
##	"value"	attrobject
##	1	1
##	attrz	circumst
##	1	1
##	creation	date
##	1	1
##	lengthalpha	nonintrins
##	1	1
##	objectori	react
##	1	1
##	report	truncat
##	1	1
##	'formal'	"empty"
##	1	1
##	unclasswint	asfactor
##	1	1
##	australian	capit
##	1	1
##	ctas	eight
##	1	1
##	factorst	incmean
##	1	1
##	levelsstatef	mnemon
##	1	1
##	money	northern
##	1	1
##	queensland	state'
##	1	1
##	tasmania	victoria
##	1	1
##	wale	western
##	1	1
##	'tapplyincom	"real"
##	1	1
##	"sorted"	categori
##	1	1

##	instanc	investig
##	1	1
##	liner	qt
##	1	1
##	sex	sqrtvarxlengthx
##	1	1
##	tdistribut	alon
##	1	1
##	comma	dimension
##	1	1
##	fastest	kdimension
##	1	1
##	order”	slowest
##	1	1
##	x	xx
##	1	1
##	“column	arrayc
##	1	1
##	cbindxb	crossprodx
##	1	1
##	ib	unreduc
##	1	1
##	xb	xbib
##	1	1
##	xviv	arraydatavector
##	1	1
##	arrayh	byelement
##	1	1
##	dimh	dimvector
##	1	1
##	elementbyel	mismatch
##	1	1
##	precipit	quirki
##	1	1
##	reliabl	signal
##	1	1
##	bc	cute
##	1	1
##	easiest	impract
##	1	1
##	jth	neat
##	1	1
##	noncommut	outera
##	1	1
##	permj	plotfr
##	1	1
##	singledigit	ta
##	1	1
##	tableouterd	transposit
##	1	1
##	twice	xlabdetermin
##	1	1

##	ylabfrequ	“obvious”
##	1	1
##	“sees”	accuraci
##	1	1
##	adopt	algebra
##	1	1
##	ambigu	choleski
##	1	1
##	diagk	diagm
##	1	1
##	diagv	loss
##	1	1
##	ncola	nrowa
##	1	1
##	nvector	promot
##	1	1
##	scalar	solveax
##	1	1
##	unstabl	whichev
##	1	1
##	“crossproducts”	absdetm
##	1	1
##	eigensmvalu	evval
##	1	1
##	evvec	followup
##	1	1
##	functionm	grand
##	1	1
##	hint	lsdiag
##	1	1
##	lsfitx	modulus
##	1	1
##	onlyvalu	symmetr
##	1	1
##	tr	trace
##	1	1
##	truevalu	tv
##	1	1
##	xplus	asvectorx
##	1	1
##	cbindarg	columnwis
##	1	1
##	matter	offici
##	1	1
##	older	qrcoefxplus
##	1	1
##	qrfittedxplus	qrresidxplus
##	1	1
##	redund	rough
##	1	1
##	rowwis	sideeffect
##	1	1

##	'backslash'	class"
##	1	1
##	factorcutincom	higherway
##	1	1
##	kway	tableincomefstatef
##	1	1
##	tablestatef	tapplystatef
##	1	1
##	twoway	"incom
##	1	1
##	childagesc	distinguish
##	1	1
##	forget	fred
##	1	1
##	ie	lengthlst
##	1	1
##	listnamefr	lstchildag
##	1	1
##	lstcoe	lstcoeffici
##	1	1
##	lstcov	lstcovari
##	1	1
##	lstwife	lstx
##	1	1
##	mari	namecomponentnam
##	1	1
##	nochildren	transfer
##	1	1
##	uniqu	wifemari
##	1	1
##	accountantshom	clista
##	1	1
##	dataframehomestatef	freeli
##	1	1
##	listabc	listb
##	1	1
##	listc	listmatrixmat
##	1	1
##	listnameobject	lootincom
##	1	1
##	namemobjectm	satisfi
##	1	1
##	scratch	shotincomef
##	1	1
##	somehow	attachanyoldlist
##	1	1
##	attachlentil	gather
##	1	1
##	greater	leftov
##	1	1
##	lentilsv	lentilsw
##	1	1

##	resort	unwant
##	1	1
##	'database'	confirm
##	1	1
##	track	awk
##	1	1
##	fair	imp
##	1	1
##	nonnumer	ortexport
##	1	1
##	primit	sed
##	1	1
##	byrowtru	datainfert
##	1	1
##	elabor	establish
##	1	1
##	head	implic
##	1	1
##	inpi	inpid
##	1	1
##	inputdat	inpx
##	1	1
##	listid	matrixscanlightdat
##	1	1
##	obsolet	reassign
##	1	1
##	datapackagerpart	datapuruomycin
##	1	1
##	editdatafram	fixxold
##	1	1
##	rich	spreadsheet
##	1	1
##	spreadsheetlik	accur
##	1	1
##	andom	binom
##	1	1
##	chisq	chisquar
##	1	1
##	dmultinom	exponenti
##	1	1
##	geom	geometr
##	1	1
##	hazard	ht
##	1	1
##	httpscranrprojectorgpackagesuppdist	hyper
##	1	1
##	hypergeometr	lambda
##	1	1
##	lnorm	lognorm
##	1	1
##	meanlog	nbinom
##	1	1

##	noncentr	norm
##	1	1
##	notabl	poi
##	1	1
##	prefix	pt
##	1	1
##	ptukey	pxxt
##	1	1
##	qf	qtukey
##	1	1
##	randomnumb	rhyper
##	1	1
##	rmultinom	rng
##	1	1
##	rsignrank	rwilcox
##	1	1
##	rxxx	sdlog
##	1	1
##	signrank	student
##	1	1
##	suppdist	unif
##	1	1
##	wilcox	'd'
##	1	1
##	'p'	'q'
##	1	1
##	"integrated"	attachfaith
##	1	1
##	autom	fivenumerupt
##	1	1
##	leaf"	linesdensityerupt
##	1	1
##	median	probtru
##	1	1
##	rugerupt	sj
##	1	1
##	st	stemandleaf
##	1	1
##	stemerupt	summaryerupt
##	1	1
##	trialanderror	univari
##	1	1
##	"interesting"	"stem
##	1	1
##	ecdflong	eruptionserupt
##	1	1
##	fnx	meanmeanlong
##	1	1
##	minut	overlay
##	1	1
##	plotecdfserupt	plotecdflong
##	1	1

##	pnormx	qqlinelong
##	1	1
##	qqnormlong	sdsqrtvarlong
##	1	1
##	agreement	calgm
##	1	1
##	dsn	fusion
##	1	1
##	heat	ice
##	1	1
##	kstestlong	latent
##	1	1
##	meanlong	onesampl
##	1	1
##	pnorm	qqplotqtppoint
##	1	1
##	q-q	rt
##	1	1
##	shapirotestlong	sqrtvarlong
##	1	1
##	theoret	theori
##	1	1
##	xlab	"classical"
##	1	1
##	boxplota	denom
##	1	1
##	num	popul
##	1	1
##	unpair	vartesta
##	1	1
##	welch	due
##	1	1
##	ecdf'	mann
##	1	1
##	plotecdfa	plotecdfb
##	1	1
##	shift	varequaltru
##	1	1
##	whitney	xlimrangea
##	1	1
##	ablinelsfitxci	ai
##	1	1
##	bi	conditioni
##	1	1
##	elementwis	exprm
##	1	1
##	ifelsecondit	lengthyc
##	1	1
##	most	plotxci
##	1	1
##	spliti	splitx
##	1	1

##	subexpress	wherev
##	1	1
##	xc	xyplot
##	1	1
##	yc	“shortcircuit”
##	1	1
##	abnorm	cycl
##	1	1
##	discontin	faster
##	1	1
##	object’	‘whole
##	1	1
##	“next”	bit
##	1	1
##	bslashxmat	datafemal
##	1	1
##	emphas	enorm
##	1	1
##	functionarg	functioni
##	1	1
##	gain	legitim
##	1	1
##	nameexpr	ns
##	1	1
##	nsnn	pay
##	1	1
##	qrcoefx	regcoeff
##	1	1
##	safe	steps”
##	1	1
##	tricki	tstatist
##	1	1
##	twosam	twosamdatamal
##	1	1
##	ybsqrtsn	yvar
##	1	1
##	counterintuit	dataframedf
##	1	1
##	fundatad	isol
##	1	1
##	qrcoef	unnam
##	1	1
##	“nameobject”	asfactorblock
##	1	1
##	asfactorvarieti	assignment”
##	1	1
##	asvectortableblock	asvectortablevarieti
##	1	1
##	bdeff	functionblock
##	1	1
##	govern	lengthlevelsblock
##	1	1

##	lengthlevelsvarieti	liter
##	1	1
##	notion	n'th
##	1	1
##	omiss	outlin
##	1	1
##	pedestrian	unmatch
##	1	1
##	'n'	"super
##	1	1
##	around"	blockcvsvu
##	1	1
##	competit	concentr
##	1	1
##	dima	dimnamesa
##	1	1
##	dimnamestemp	dl
##	1	1
##	farthest	fori
##	1	1
##	functiona	inherit
##	1	1
##	listeff	listrep
##	1	1
##	naiv	ncolx
##	1	1
##	nodimnamesx	nrowx
##	1	1
##	onepanel	overhead
##	1	1
##	pattern	qualit
##	1	1
##	repb	repsqrtr
##	1	1
##	rmtemp	sqrtdk
##	1	1
##	sv	svda
##	1	1
##	trapezium	varietycvsvv
##	1	1
##	"wrap	ascertain
##	1	1
##	functionn	ifabsa
##	1	1
##	nsq	printi
##	1	1
##	printx	printz
##	1	1
##	resolv	returna
##	1	1
##	returnfunf	static
##	1	1

##	therefor	'area'
##	1	1
##	'fun'	catyour
##	1	1
##	dump	functiontot
##	1	1
##	moneyn	mutabl
##	1	1
##	positiven	reach
##	1	1
##	robert	ross
##	1	1
##	stopdeposit	stopyou
##	1	1
##	withdrawn	catpastedatenadiosn
##	1	1
##	continuet	lack
##	1	1
##	librarymass	mystuffr
##	1	1
##	optionsdigit	optionsprompt
##	1	1
##	printout	robertbal
##	1	1
##	rossdeposit	sourcefilepathsysgetenvhom
##	1	1
##	task	asmatrix
##	1	1
##	asterisk	coefaov
##	1	1
##	coefarima	coefdefault
##	1	1
##	coeflistof	coefnl
##	1	1
##	coefsummarynl	fashion
##	1	1
##	getanywherecoefaov	getsmethodcoef
##	1	1
##	methodsclassdatafram	methodscoef
##	1	1
##	methodsplot	namespacestat
##	1	1
##	none	nonvis
##	1	1
##	presenc	question
##	1	1
##	regist	six
##	1	1
##	usemethod	usemethodcoef
##	1	1
##	'coefaov'	cl
##	1	1

##	gen	gencl
##	1	1
##	ambiti	broad
##	1	1
##	extractor	homoscedast
##	1	1
##	ial	interlock
##	1	1
##	ix	nid
##	1	1
##	pictur	polynom
##	1	1
##	spectrum	templat
##	1	1
##	xij	x β
##	1	1
##	β_j	σ
##	1	1
##	abstract	ax
##	1	1
##	errorc	exclus
##	1	1
##	genstat	glim
##	1	1
##	inevit	nest
##	1	1
##	nonaddit	roger
##	1	1
##	subplot	subspac
##	1	1
##	wilkinson	ya
##	1	1
##	ym	ccontrhelmert
##	1	1
##	ccontrtreat	deliber
##	1	1
##	helmert	im
##	1	1
##	insul	klevel
##	1	1
##	kth	mn
##	1	1
##	newcom	parameter
##	1	1
##	scheme	statistician
##	1	1
##	stori	tensor
##	1	1
##	“interactions”	“subclasses”
##	1	1
##	aic	akaike’
##	1	1

##	alia	anovaobject
##	1	1
##	coefficientsobject	coefobject
##	1	1
##	devianceobject	formulaobject
##	1	1
##	hierarchi	kappa
##	1	1
##	lmformula	newdatadatafram
##	1	1
##	plotobject	predictobject
##	1	1
##	printobject	proj
##	1	1
##	regardless	residobject
##	1	1
##	residualsobject	stepobject
##	1	1
##	stepwis	streamlin
##	1	1
##	anovafittedmodel	aovformula
##	1	1
##	aovyield	comprehend
##	1	1
##	datafarmdata	decreas
##	1	1
##	errorfarmsblock	errorstrataformula
##	1	1
##	farm	farms"
##	1	1
##	formula"	hierarch
##	1	1
##	inconsequenti	interblock
##	1	1
##	meanformula	newmodel
##	1	1
##	recoveri	smf
##	1	1
##	summaryobject	updateoldmodel
##	1	1
##	variancecovari	vcovobject
##	1	1
##	\024	\025
##	1	1
##	clue	encompass
##	1	1
##	fmfull	fy
##	1	1
##	invert	loos
##	1	1
##	m η	m-
##	1	1

##	nonnorm	tight
##	1	1
##	unifi	variat
##	1	1
##	$y\lambda\mu$	βi
##	1	1
##	βp	γ
##	1	1
##	$\lambda\mu$	τ
##	1	1
##	aegean	carri
##	1	1
##	congenit	datasal
##	1	1
##	diseas	eye
##	1	1
##	familyfamilygener	familygener
##	1	1
##	glmformula	instrument
##	1	1
##	inversegaussian	name”
##	1	1
##	sale	sight
##	1	1
##	suffer	xx
##	1	1
##	“famili	aris
##	1	1
##	binomiallinkprobit	bn
##	1	1
##	cbindkalythosi	chanc
##	1	1
##	cldp	dataframex
##	1	1
##	fml	fmod
##	1	1
##	fmp	functionb
##	1	1
##	genuin	grace
##	1	1
##	kalythosi	kalythosn
##	1	1
##	kalythosymat	ldcoeffml
##	1	1
##	ldcoeffmp	ldp
##	1	1
##	loglinear	nongaussian
##	1	1
##	overall	poissonlinksqrt
##	1	1
##	squareroot	summaryfml
##	1	1

##	summaryfmp	surrog
##	1	1
##	trial	wormcount
##	1	1
##	ϕz	'success'
##	1	1
##	$-\beta$	approach
##	1	1
##	arisen	biochem
##	1	1
##	characterist	guarante
##	1	1
##	incident	lackoffit
##	1	1
##	nlfit	plausibl
##	1	1
##	quasilinkinvers	satisfactori
##	1	1
##	sumi	upon
##	1	1
##	varianceconst	$-z$
##	1	1
##	adequ	correl
##	1	1
##	doserespons	freedom
##	1	1
##	michaelismenten	nlsi
##	1	1
##	pp	prt
##	1	1
##	solveouthessian	sqrtdiagoutminimumlengthi
##	1	1
##	std	subtract
##	1	1
##	summaryfit	sumofsquar
##	1	1
##	acepack	bifurc
##	1	1
##	conclud	ded
##	1	1
##	ension	ext
##	1	1
##	gam	heterogen
##	1	1
##	highlyresist	homogen
##	1	1
##	httpscran	httpscranrproject
##	1	1
##	httpscranrprojectorgpackageacepack	httpscranrprojectorgpackagegam
##	1	1
##	httpscranrprojectorgpackagemass	httpscranrprojectorgpackagemda
##	1	1

##	httpscranrprojectorgpackagenlm	logchoosen
##	1	1
##	mda	messi
##	1	1
##	mgcv	mixedeffect
##	1	1
##	nlogexpppx	orgpackagemgcv
##	1	1
##	pursuit	recommen
##	1	1
##	rprojectorgpackagemass	ses
##	1	1
##	sqrtdiagsolveouthessian	stateofart
##	1	1
##	trend	yield
##	1	1
##	yppx	cranrprojectorgpackagerpart
##	1	1
##	https	packagetre
##	1	1
##	rpart	akin
##	1	1
##	ation	coexist
##	1	1
##	devnew	eras
##	1	1
##	harder	multipanel
##	1	1
##	plotxi	timeseri
##	1	1
##	trelli	'base'
##	1	1
##	breakpoint	breaksb
##	1	1
##	cointerv	coplot-
##	1	1
##	distributioncomparison	enlighten
##	1	1
##	givenvalu	joint
##	1	1
##	nclass	nclassn
##	1	1
##	pairsx	pairwis
##	1	1
##	panelsmooth	plotdf
##	1	1
##	ploti	qqplotx
##	1	1
##	quartil	axestru
##	1	1
##	axes-us	dotchartx
##	1	1

##	highdens	ideal
##	1	1
##	logarithm	logx
##	1	1
##	logxi	mainstr
##	1	1
##	overlaid	perspx
##	1	1
##	probabilitytru	rectangl
##	1	1
##	stepfunct	typeb
##	1	1
##	typeh	typeo
##	1	1
##	typep	visual
##	1	1
##	xlabstr	yaxi
##	1	1
##	ylabstr	ablinea
##	1	1
##	ablinehi	ablinelmobj
##	1	1
##	ablinevx	colv
##	1	1
##	fillv	hatch
##	1	1
##	hy	labels
##	1	1
##	legendx	lmobj
##	1	1
##	ltyv	lwdv
##	1	1
##	modelfit	pchv
##	1	1
##	plot'	pointsx
##	1	1
##	polygonx	shade
##	1	1
##	vx	xcoordin
##	1	1
##	andor	astronom
##	1	1
##	atopn	axissid
##	1	1
##	cartograph	cyril
##	1	1
##	demohershey	demojapanes
##	1	1
##	demoplotmath	exampleplotmath
##	1	1
##	expressionpastebgroup	helpershey
##	1	1

##	helpjapanes	helpplotmath
##	1	1
##	japanes	kana
##	1	1
##	kanji	px
##	1	1
##	qnx	render
##	1	1
##	screen	titlemain
##	1	1
##	zodiac	absent
##	1	1
##	locatorn	out
##	1	1
##	public	textloc
##	1	1
##	'col'	behavior-usu
##	1	1
##	commonlyus	durat
##	1	1
##	namevalu	parameter'
##	1	1
##	parccol	parnoreadonlytru
##	1	1
##	reset	settabl
##	1	1
##	undesir	xaxt
##	1	1
##	yaxt	"default"
##	1	1
##	ascharacter	cex
##	1	1
##	colaxi	collab
##	1	1
##	colmain	colsub
##	1	1
##	dash	dot
##	1	1
##	duplic	face
##	1	1
##	fontaxi	fontlab
##	1	1
##	fontmain	fontsub
##	1	1
##	gap	ive
##	1	1
##	legendloc	lwd
##	1	1
##	onward	proport
##	1	1
##	solid	'.'
##	1	1

##	bound	cexaxi
##	1	1
##	cexlab	cexmain
##	1	1
##	cexsub	compris
##	1	1
##	divis	fall
##	1	1
##	labc	las
##	1	1
##	perpendicular	toosmal
##	1	1
##	xaxsr	yaxsi
##	1	1
##	maic	marc
##	1	1
##	corner	deal
##	1	1
##	figc	helpdevic
##	1	1
##	instruct	jump
##	1	1
##	line"	mfcoll
##	1	1
##	mfcollc	mfrow
##	1	1
##	newtru	omac
##	1	1
##	omic	outertru
##	1	1
##	packagelattic	pagewis
##	1	1
##	parcex	splitscreen
##	1	1
##	unequallys	"draw
##	1	1
##	encapsul	epsf
##	1	1
##	incorpor	landscap
##	1	1
##	postscriptfilep	postscriptplotep
##	1	1
##	scompat	sure
##	1	1
##	unusu	winmetafil
##	1	1
##	winprint	buja
##	1	1
##	cloud	cook
##	1	1
##	devcopydevic	devoffk
##	1	1

##	devprint	devprintdevic
##	1	1
##	devsetwhichk	ggobi
##	1	1
##	httpggobiorg	httpggobiorgrggobihtml
##	1	1
##	httpscranrprojectorgpackagergl	packagerggobi
##	1	1
##	rggobi	rgl
##	1	1
##	swayn	tiff
##	1	1
##	“brushing”	administr
##	1	1
##	aid	author
##	1	1
##	bioconductor	boot
##	1	1
##	download	dware
##	1	1
##	har	hide
##	1	1
##	httpswwwbioconductororg	installpackag
##	1	1
##	libraryboot	loadednamespac
##	1	1
##	meant	mirror
##	1	1
##	navig	packageboot
##	1	1
##	repositori	thousand
##	1	1
##	updatepackag	“creat
##	1	1
##	“instal	baset
##	1	1
##	interdepend	triplecolon
##	1	1
##	abil	ap
##	1	1
##	behaviour	caseinsensit
##	1	1
##	choosedir	choosefil
##	1	1
##	claim	complianc
##	1	1
##	consider	dir
##	1	1
##	dircreat	effort
##	1	1
##	fat	feasibl
##	1	1

##	fileaccess	filechoos
##	1	1
##	filecopi	filecr
##	1	1
##	fileedit	fileexist
##	1	1
##	fileinfo	filelink
##	1	1
##	fileremov	fileshow
##	1	1
##	gone	listdir
##	1	1
##	listfil	osdepend
##	1	1
##	platformindepend	propriat
##	1	1
##	reli	sysglob
##	1	1
##	sysreadlink	tcltk
##	1	1
##	tempfil	tkchoosedir
##	1	1
##	tkchoosefil	touch
##	1	1
##	trail	under
##	1	1
##	unlink	wildcard
##	1	1
##	'exists'	abcdef
##	1	1
##	acl	api
##	1	1
##	assembl	attempt
##	1	1
##	basenam	behav
##	1	1
##	bzip	compromis
##	1	1
##	dabcfoobar	decompress
##	1	1
##	dfoobar	dirnam
##	1	1
##	employ	expansion'
##	1	1
##	expens	extent
##	1	1
##	gzip	ownergroupal
##	1	1
##	partial	readon
##	1	1
##	readwriteexecut	serverdirdirfil
##	1	1

##	shortpathnam	shquot
##	1	1
##	substitut	syschmod
##	1	1
##	syswhich	transpar
##	1	1
##	unc	uncserverdirdirfil
##	1	1
##	'magic'	'tild
##	1	1
##	httpsenwikipediaorg	httpsenwikipediaorgwikilzma
##	1	1
##	lzma	unpack
##	1	1
##	untar	unzip
##	1	1
##	wikixz	'tarballs'
##	1	1
##	attachdummi	banner
##	1	1
##	disappear	iconifi
##	1	1
##	lowessx	lrfi
##	1	1
##	plane	pseudorandom
##	1	1
##	puzzlement	rnormx
##	1	1
##	rnormxw	soon
##	1	1
##	sqrtx	tild
##	1	1
##	unfamiliar	unweight
##	1	1
##	weightw	'weight'
##	1	1
##	anovafm	aovspe
##	1	1
##	attachmm	datamm
##	1	1
##	factormmexpt	factormmrun
##	1	1
##	fileshowfilepath	heteroscedast
##	1	1
##	kurtosi	mainspe
##	1	1
##	mmexpt	mmrun
##	1	1
##	morley	morleytab
##	1	1
##	plotexpt	plotfittedfm
##	1	1

##	qqnormresidfm	rankit
##	1	1
##	readtablefilepath	red
##	1	1
##	residfm	skew
##	1	1
##	sl	systemfiledata
##	1	1
##	vs	xlabexperi
##	1	1
##	xlabfit	ylabresidu
##	1	1
##	π	'experiments'
##	1	1
##	$-\pi$	disc
##	1	1
##	expith	exploratori
##	1	1
##	ftf	ifelsemodw
##	1	1
##	parnoreadon	part"
##	1	1
##	pchxlabx	plotz
##	1	1
##	rmth	rnormi
##	1	1
##	sqrtrunifexppirunifi	th
##	1	1
##	xlabx	"asymmetr
##	1	1
##	"square"	helpstartup
##	1	1
##	infil	outfil
##	1	1
##	papers	renvirobus
##	1	1
##	rhomeetcrenvironsit	rhomeetcrprofilesit
##	1	1
##	rlib	rprintcmd
##	1	1
##	windowsspecif	wrapper
##	1	1
##	'memory'	'startup'
##	1	1
##	apart	directory"
##	1	1
##	enc	encodingenc
##	1	1
##	filefil	french
##	1	1
##	iconv	latter'
##	1	1

##	norestorehistori	rhistfil
##	1	1
##	utf	"home
##	1	1
##	archnam	backcompat
##	1	1
##	buffer	con
##	1	1
##	debuggernam	deduc
##	1	1
##	disregard	fed
##	1	1
##	fifo	garbag
##	1	1
##	gdb	guityp
##	1	1
##	heap	maxppsizen
##	1	1
##	megabyt	million
##	1	1
##	minnsizen	minvsizen
##	1	1
##	pipe	progress
##	1	1
##	pti	rinferiormod
##	1	1
##	rnsizen	rterm
##	1	1
##	rvsizen	stack
##	1	1
##	statistics"	subarchitecture
##	1	1
##	suffix	tildeexpans
##	1	1
##	valgrind	'm'
##	1	1
##	'tcltk'	"emacs
##	1	1
##	arch	channel
##	1	1
##	config	conjunct
##	1	1
##	diff	invoc
##	1	1
##	java	javareconf
##	1	1
##	postprocess	rdiff
##	1	1
##	rdpdf	rprof
##	1	1
##	rtag	shorthand
##	1	1

##	sr	stangl
##	1	1
##	tag	txt
##	1	1
##	bat	bstinput
##	1	1
##	child	clarifi
##	1	1
##	cmdex	debug
##	1	1
##	debugger"	doubleclick
##	1	1
##	draganddrop	exe
##	1	1
##	homedr	homepath
##	1	1
##	ldd	macro
##	1	1
##	miktex	mydoctex
##	1	1
##	nomdi	overrid
##	1	1
##	pager	pdflatex
##	1	1
##	rex	rostyp
##	1	1
##	ruser	sharetexmf
##	1	1
##	terminalapp	tmp
##	1	1
##	usercontrol	"break
##	1	1
##	bash	binsh
##	1	1
##	bourm	chem
##	1	1
##	chmod	commonplac
##	1	1
##	françoi	hardcod
##	1	1
##	helprscript	pathtorscript
##	1	1
##	pinard	process'
##	1	1
##	qstatusexit	rframework
##	1	1
##	scann	scanstdin
##	1	1
##	segment	tradi
##	1	1
##	usrlocalbin	filenam
##	1	1

##	goe	usabl
##	1	1
##	alt	cc
##	1	1
##	cg	cm
##	1	1
##	controlm	cr
##	1	1
##	ctrl	displac
##	1	1
##	editlin	erron
##	1	1
##	escb	extsoftvers
##	1	1
##	httpsessrprojectorg	info
##	1	1
##	libedit	ma
##	1	1
##	mac	metab
##	1	1
##	mi	netbsd'
##	1	1
##	pc	phase
##	1	1
##	readmerterm	resiz
##	1	1
##	statistics'	twocharact
##	1	1
##	'console'	'emacs
##	1	1
##	'help'	'windows'
##	1	1
##	ce	ck
##	1	1
##	ct	cx
##	1	1
##	cy	endif
##	1	1
##	inputrc	mc
##	1	1
##	md	mf
##	1	1
##	ml	q''n
##	1	1
##	yank	"saved"
##	1	1
##	aperm	devset
##	1	1
##	getsmethod	isna
##	1	1
##	isnan	kstest
##	1	1

```
##          nrow          prod
##          1          1
##          qqline      qqnorm
##          1          1
##          shapirotest    vartest
##          1          1
##          wilcoxtest      tabul
##          1          1
##          annett         belmont
##          1          1
##          bootstrap    cambridg
##          1          1
##          chap          duxburi
##          1          1
##          ed            penguin
##          1          1
##          peter         son
##          1          1
##          springer      wiley
##          1          1
##          "blue         "green
##          1          1
##          "white
##          1
```

```
new_freq<-freq[freq>50]
new_freq
```

```
## function      use      r      vector      x      can      data      plot
##      420      397      363      254      239      200      194      175
##      model      y      valu      file      exampl      list      graphic      name
##      165      164      155      146      144      135      134      134
##      argument    will    object    command    array      may      variabl    matrix
##      126      120      116      114      112      109      103      101
##      packag      b      c      charact    number      see      line      one
##      93      90      86      86      82      82      81      80
##      form      way    factor      also    paramet    chapter      fit      set
##      80      80      79      78      77      77      76      75
##      frame      two    devic    distribut    first      oper      case    express
##      74      73      68      67      66      65      64      64
##      point    result    given    compon    window    default    avail      •
##      63      62      62      61      60      60      59      58
##      assign      mean      m      defin    specifi      index      column    statist
##      56      56      55      55      55      54      54      53
##      differ      give      page    current      make      length    directori    call
##      53      52      52      52      51      51      51      51
```

```
# Extract words (terms) and their frequencies
words <- names(new_freq)
words
```

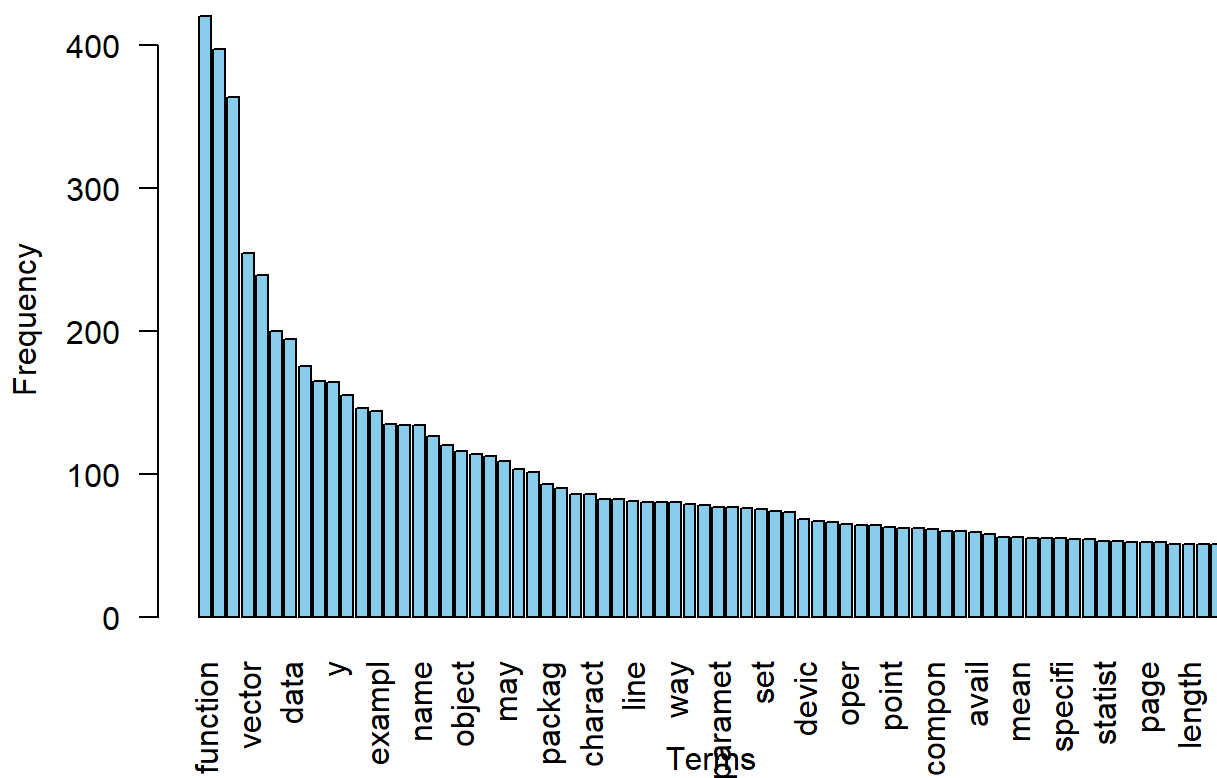
```
## [1] "function" "use" "r" "vector" "x" "can"
## [7] "data" "plot" "model" "y" "valu" "file"
## [13] "exampl" "list" "graphic" "name" "argument" "will"
## [19] "object" "command" "array" "may" "variabl" "matrix"
## [25] "packag" "b" "c" "charact" "number" "see"
## [31] "line" "one" "form" "way" "factor" "also"
## [37] "paramet" "chapter" "fit" "set" "frame" "two"
## [43] "devic" "distribut" "first" "oper" "case" "express"
## [49] "point" "result" "given" "compon" "window" "default"
## [55] "avail" "•" "assign" "mean" "m" "defin"
## [61] "specifi" "index" "column" "statist" "differ" "give"
## [67] "page" "current" "make" "length" "director" "call"
```

```
max(freq)
```

```
## [1] 420
```

```
# Create barplot #Better
barplot(new_freq,
  main = "Frequency of Most Frequent Terms",
  xlab = "Terms",
  ylab = "Frequency",
  names.arg = words, # Display words on x-axis
  las = 2, # Rotate x-axis labels if needed
  col = "skyblue") # Specify color if desired
```

Frequency of Most Frequent Terms



8

```
library(car)
```

```
## Warning: package 'car' was built under R version 4.3.3
```

```
## Loading required package: carData
```

```
## Warning: package 'carData' was built under R version 4.3.3
```

```
data(airquality)
```

```
# a) shapiro-Wilk test of Normality Test  
aq <- airquality  
str(aq)
```

```
## 'data.frame': 153 obs. of 6 variables:
## $ Ozone : int 41 36 12 18 NA 28 23 19 8 NA ...
## $ Solar.R: int 190 118 149 313 NA NA 299 99 19 194 ...
## $ Wind : num 7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6 ...
## $ Temp : int 67 72 74 62 56 66 65 59 61 69 ...
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...
## $ Day : int 1 2 3 4 5 6 7 8 9 10 ...
```

```
#Here the number of observations is greater than 100, so we do ks test for
#normality
shapiro_test_result<-shapiro.test(aq$Wind)
shapiro_test_result
```

```
##
## Shapiro-Wilk normality test
##
## data: aq$Wind
## W = 0.98575, p-value = 0.1178
```

```
# Here, p value is greater than 0.05 so, it follows normal distribution.

# b) Variance Test
# for equality of variance
# Perform a Bartlett test to compare variances of wind data by month
bartlett.test(Wind ~ Month, data = aq)
```

```
##
## Bartlett test of homogeneity of variances
##
## data: Wind by Month
## Bartlett's K-squared = 1.6178, df = 4, p-value = 0.8056
```

```
#Here, p value is greater than 0.05, thus the variance of wind with respect to
#months are equal

#c
#Fitting one way anova
anova_test<- oneway.test(aq$Wind~factor(aq$Month), var.equal = TRUE)
anova_test
```

```
##
## One-way analysis of means
##
## data: aq$Wind and factor(aq$Month)
## F = 3.529, num df = 4, denom df = 148, p-value = 0.00879
```

```
(anova_model <- aov(Wind ~ as.factor(Month), data = aq))
```

```
## Call:
##   aov(formula = Wind ~ as.factor(Month), data = aq)
##
## Terms:
##           as.factor(Month) Residuals
## Sum of Squares      164.2708 1722.2831
## Deg. of Freedom           4      148
##
## Residual standard error: 3.411312
## Estimated effects may be unbalanced
```

#Interpretation: Here, the p value is less than 0.05, thus the mean wind speed in various months are different is not equal.

```
#d
# Apply TukeyHSD to the ANOVA model
tukey_result <- TukeyHSD(anova_model)
print(tukey_result)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Wind ~ as.factor(Month), data = aq)
##
## $`as.factor(Month)`
##           diff          lwr          upr      p adj
## 6-5 -1.35591398 -3.768713  1.0568846 0.5305524
## 7-5 -2.68064516 -5.073585 -0.2877054 0.0197174
## 8-5 -2.82903226 -5.221972 -0.4360925 0.0117066
## 9-5 -1.44258065 -3.855379  0.9702179 0.4674045
## 7-6 -1.32473118 -3.737530  1.0880674 0.5535894
## 8-6 -1.47311828 -3.885917  0.9396803 0.4456532
## 9-6 -0.08666667 -2.519162  2.3458285 0.9999786
## 8-7 -0.14838710 -2.541327  2.2445527 0.9998052
## 9-7  1.23806452 -1.174734  3.6508631 0.6176733
## 9-8  1.38645161 -1.026347  3.7992502 0.5081147
```

Interpretation: The pairs in which p value is less than 0.05 ie. 7-5 month and 8-5 month has difference in mean wind. ALL other pairs have similar mean wind values.

9

```
data<-USArrests
head(data)
```

```
##           Murder Assault UrbanPop Rape
## Alabama    13.2    236      58 21.2
## Alaska     10.0    263      48 44.5
## Arizona     8.1    294      80 31.0
## Arkansas    8.8    190      50 19.5
## California  9.0    276      91 40.6
## Colorado   7.9    204      78 38.7
```

```
str(data)
```

```
## 'data.frame':  50 obs. of  4 variables:
## $ Murder   : num  13.2 10 8.1 8.8 9 7.9 3.3 5.9 15.4 17.4 ...
## $ Assault  : int  236 263 294 190 276 204 110 238 335 211 ...
## $ UrbanPop: int  58 48 80 50 91 78 77 72 80 60 ...
## $ Rape     : num  21.2 44.5 31 19.5 40.6 38.7 11.1 15.8 31.9 25.8 ...
```

```
# a Split 70 30
set.seed(13)
ind <- sample(2, nrow(data),

              replace=T, prob = c(0.7, 0.3))

train <- data[ind==1,]

test <- data[ind==2,]

# b) fit linear and KNN
# Load necessary libraries
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```
##
## Attaching package: 'ggplot2'
```

```
## The following object is masked from 'package:NLP':
##
## annotate
```

```
# For plotting
library(caret)
```

```
## Warning: package 'caret' was built under R version 4.3.3
```

```
## Loading required package: lattice
```



```
## Warning: package 'lattice' was built under R version 4.3.3
```

```
linear_model <- lm(UrbanPop ~ Murder + Assault + Rape, data=data)
# Fit the linear regression using lm function.

summary(linear_model)
```

```
##
## Call:
## lm(formula = UrbanPop ~ Murder + Assault + Rape, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -35.456  -6.950   0.077   7.770  25.221
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  52.84187    4.82483   10.952 2.09e-14 ***
## Murder       -1.41154    0.71954   -1.962  0.0559 .
## Assault       0.05190    0.04161    1.247  0.2186
## Rape         0.69841    0.26776    2.608  0.0122 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.08 on 46 degrees of freedom
## Multiple R-squared:  0.2337, Adjusted R-squared:  0.1837
## F-statistic: 4.676 on 3 and 46 DF, p-value: 0.006208
```

```
# Check the model summary
# Load necessary libraries
library(caret)

# Standardize the data
preProc <- preProcess(data[, -3], method=c("center", "scale"))
data_std <- predict(preProc, data[, -3])

# Add the dependent variable back
data_std$UrbanPop <- data$UrbanPop

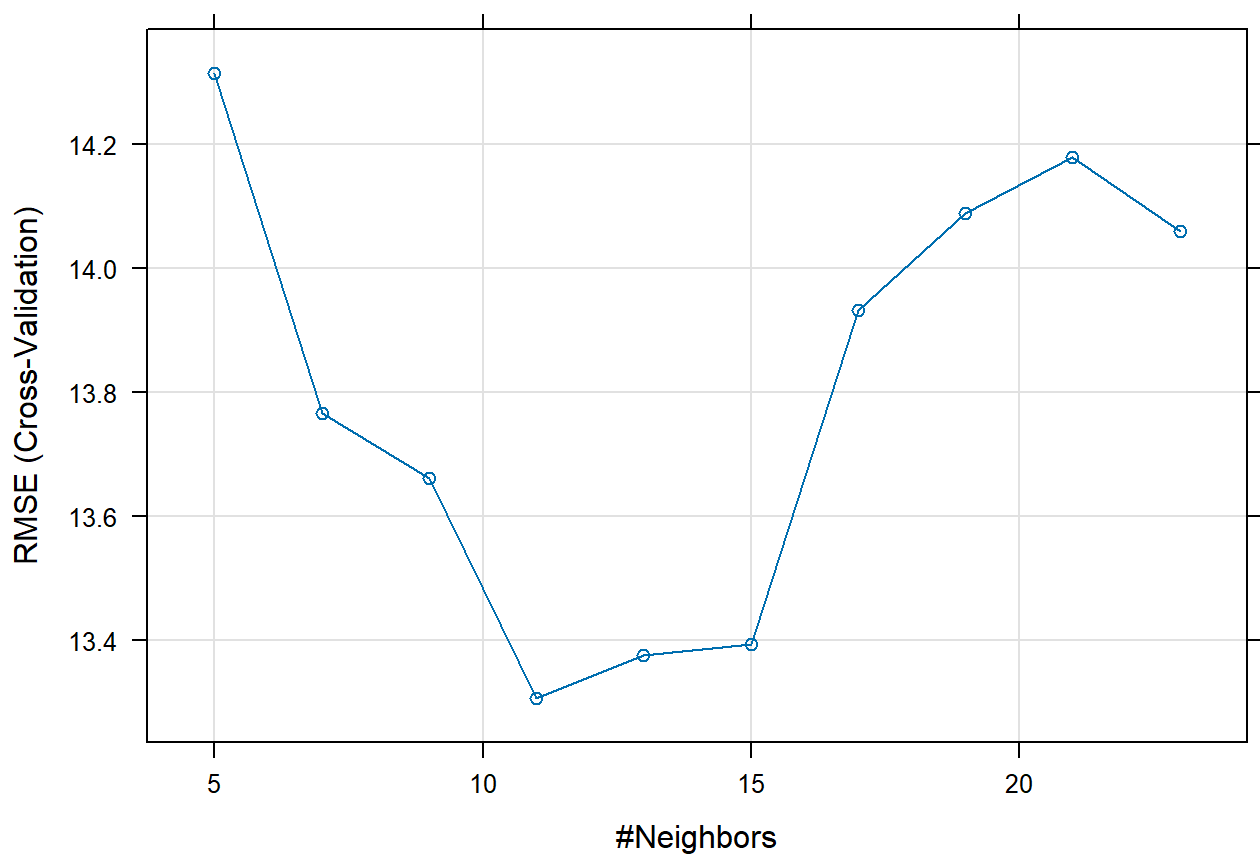
# Define the control method for training
train_control <- trainControl(method="cv", number=10)

# Train the KNN model
knn_model <- train(UrbanPop ~ Murder + Assault + Rape, data=data_std, method="knn", trControl=train_control, tuneLength=10)

# Check the results
print(knn_model)
```

```
## k-Nearest Neighbors
##
## 50 samples
## 3 predictor
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 44, 45, 45, 46, 45, 46, ...
## Resampling results across tuning parameters:
##
##  k    RMSE      Rsquared    MAE
##   5  14.31598  0.2751990  12.40538
##   7  13.76603  0.3280303  11.35714
##   9  13.66134  0.3037106  11.31421
##  11  13.30576  0.3831001  10.91229
##  13  13.37505  0.3749772  10.93414
##  15  13.39414  0.4383720  11.02686
##  17  13.93350  0.3323776  11.60188
##  19  14.08870  0.1857393  11.74160
##  21  14.18037  0.2819866  11.98591
##  23  14.05966  0.2752558  11.83362
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was k = 11.
```

```
plot(knn_model)
```



```
# c) predict in test data
linear_predictions <- predict(linear_model, test)
linear_predictions
```

```
##      Alabama      California      Delaware      Florida      Idaho
##      61.26440      82.81805      67.90100      70.77016      65.31736
##      Illinois      Maine North Carolina      North Dakota      Rhode Island
##      67.84696      59.63298      63.22675      59.14656      62.87013
##      Texas      Virginia      Wisconsin
##      63.15682      63.39738      59.46543
```

```
knn_predictions <- predict(knn_model, test)
knn_predictions
```

```
## [1] 66.18182 70.00000 66.18182 67.83333 70.00000 67.83333 70.00000 66.18182
## [9] 70.00000 66.18182 67.83333 70.00000 70.00000
```

```
#d) comparing the model
linear_mse <- mean((test$UrbanPop - linear_predictions)^2)
linear_rmse <- sqrt(linear_mse)
linear_r2 <- 1 - (sum((test$UrbanPop - linear_predictions)^2) / sum((test$UrbanPop - mean(test$UrbanPop))^2))
linear_mse
```

```
## [1] 160.1733
```

```
linear_rmse
```

```
## [1] 12.65596
```

```
linear_r2
```

```
## [1] 0.3348089
```

```
knn_mse <- mean((test$UrbanPop - knn_predictions)^2)
knn_rmse <- sqrt(knn_mse)
knn_r2 <- 1 - (sum((test$UrbanPop - knn_predictions)^2) / sum((test$UrbanPop - mean(test$UrbanPop))^2))

knn_mse
```

```
## [1] 254.4571
```

```
knn_rmse
```

```
## [1] 15.95171
```

```
knn_r2
```

```
## [1] -0.05674686
```

```
#from the above value we can conclude that the we can choose linear_mse
#linear mse is less so we can choose linear model
#at last linear model is better than knn from this result
```

10

```
ir_label <- iris$Species
ir_data <- iris[,-5]
head(ir_data)
```

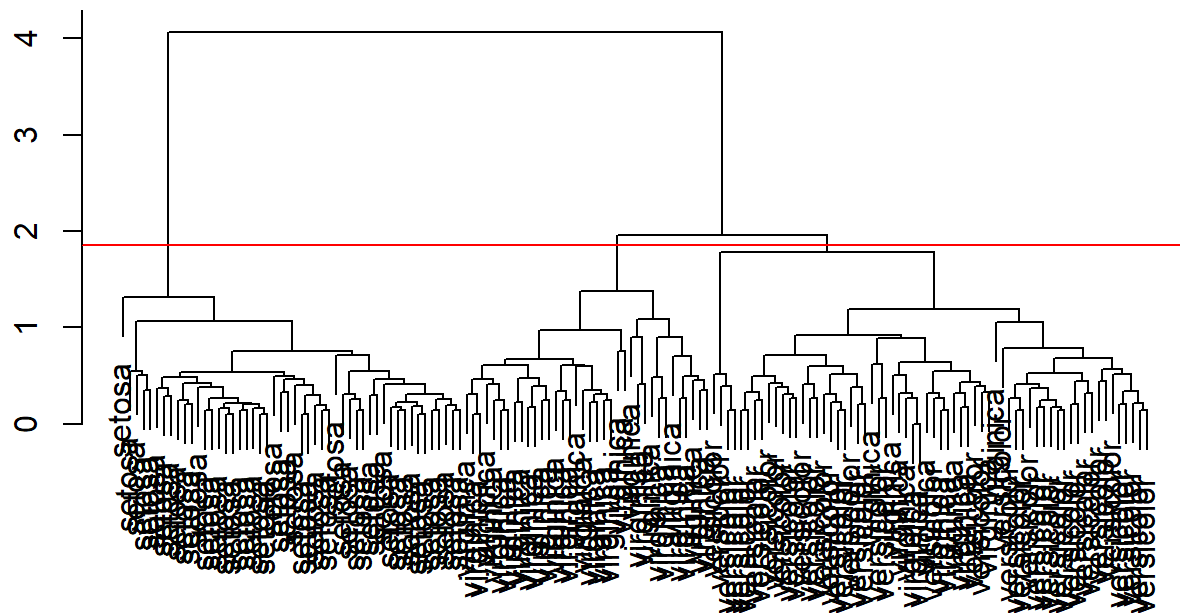
```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1          5.1          3.5          1.4          0.2
## 2          4.9          3.0          1.4          0.2
## 3          4.7          3.2          1.3          0.2
## 4          4.6          3.1          1.5          0.2
## 5          5.0          3.6          1.4          0.2
## 6          5.4          3.9          1.7          0.4
```

```
## A) Hierarchical clustering using average
data.dist <- dist(ir_data)
plot(hclust(data.dist, method = "average"), xlab = "", sub = "", ylab = "",
     labels = ir_label, main = "Average Linkage")

## B) Best value is provided by average Linkage
plot(hclust(data.dist, method = "average"), xlab = "", sub = "", ylab = "",
     labels = ir_label, main = "Average Linkage")

abline(h = 1.85, col = "red")
```

Average Linkage



```
## C) fit k-mean clustering
kmeans.c3<-kmeans(ir_data,centers = 3,nstart = 20)
kmeans.c3
```

```
## K-means clustering with 3 clusters of sizes 38, 62, 50
##
## Cluster means:
##   Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1    6.850000    3.073684    5.742105    2.071053
## 2    5.901613    2.748387    4.393548    1.433871
## 3    5.006000    3.428000    1.462000    0.246000
##
## Clustering vector:
##   [1] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
##  [38] 3 3 3 3 3 3 3 3 3 3 3 3 3 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
##  [75] 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 1 1 1 2 1 1 1
## [112] 1 1 2 2 1 1 1 1 2 1 2 1 2 1 1 2 2 1 1 1 1 2 1 1 1 1 2 1 1 1 2 1
## [149] 1 2
##
## Within cluster sum of squares by cluster:
## [1] 23.87947 39.82097 15.15100
## (between_SS / total_SS =  88.4 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

```
#d) compare using confusion matrix
cm<-table(iris$Species,kmeans.c3$cluster)
cm
```

```
##
##           1  2  3
## setosa    0  0 50
## versicolor 2 48  0
## virginica 36 14  0
```

```
(accuracy<-
  sum(diag(cm))/sum(cm))
```

```
## [1] 0.32
```

```
# Accuracy of k-mean clustering is 0.893 while size of cluster is 3. so we can use it.
```