

# Various presentation aids

For the Boston useR talk "Random mandalas deconstruction with R, Python, and Mathematica"

Anton Antonov

SimplifiedMachineLearningWorkflows-book at GitHub

February 2022

---

## Setup

```
Import["https://raw.githubusercontent.com/antononcube/NLP-Template-Engine/main/Packages/WL/NLPTemplateEngine.m"]
```

```
In[315]:= Clear[MyFromRForm];
MyFromRForm[RDataFrame[RNames[names__], RData[data__], RRowNames[rownames__]]] :=
  Dataset@AssociationThread[{rownames} → Map[AssociationThread[{names} → #] &, Transpose[MyFromRForm /@ {data}]]];
MyFromRForm[RObject[x_, RAttributes[attrs__]]] /; MemberQ[Lookup[{attrs}, "class"], "Date"] := DatePlus[DateObject[{1970, 1, 1}], #] & /@ x;
MyFromRForm[RObject[ls_List, RAttributes["names" → nms_List]]] := If[Length[nms] > 0 && Length[nms] == Length[ls], AssociationThread[nms, MyFromRForm /@ ls], ls];
MyFromRForm[args___] := args;
```

---

## NLP template engine demo


### WL

```
In[262]:= Concretize["Make LSA object over aAbstracts. Extract 20 topics using SVD."]
Out[262]:= Hold[lsaObj = LSAMonUnit[SVD] ⇒ LSAMonMakeDocumentTermMatrix[StemmingRules → Automatic, StopWords → Automatic] ⇒ LSAMonEchoDocumentTermMatrixStatistics[LogBase → 10] ⇒
  LSAMonApplyTermWeightFunctions[GlobalWeightFunction → IDF, LocalWeightFunction → None, NormalizerFunction → Cosine] ⇒ LSAMonExtractTopics[NumberOfTopics → 20,
  Method → SVD, MaxSteps → 16, MinNumberOfDocumentsPerTerm → 20] ⇒ LSAMonEchoTopicsTable[NumberOfTerms → 10] ⇒ LSAMonEchoStatisticalThesaurus[Words → {SVD}];]
```


### R

```
In[263]:= Concretize["Make LSA object over aAbstracts. Extract 20 topics using SVD.", "TargetLanguage" → "R"]
Out[263]:= parse( text = 'lsaObj <-
  LSAMonUnit(SVD) %>%
  LSAMonMakeDocumentTermMatrix(stemWordsQ = NULL, stopWords = NULL) %>%
  LSAMonEchoDocumentTermMatrixStatistics(logBase = 10) %>%
  LSAMonApplyTermWeightFunctions(globalWeightFunction = "IDF", localWeightFunction = "None", normalizerFunction = "Cosine") %>%
  LSAMonExtractTopics(numberOfTopics = 20, method = "SVD", maxSteps = 16, minNumberOfDocumentsPerTerm = 20) %>%
  LSAMonEchoTopicsTable(numberOfTerms = 10, wideFormQ = TRUE) %>%
  LSAMonEchoStatisticalThesaurus(words = c("SVD"))')
```

In[304]:= InstallR["RHomeLocation" → "/Library/Frameworks/R.framework/Versions/4.1-arm64/Resources", "RVersion" → "4.1.2"]


In[292]:=  R.version

Out[292]= RObject[{{aarch64-apple-darwin20}, {aarch64}, {darwin20}, {aarch64, darwin20}, {}, {4}, {1.2}, {2021}, {11}, {01}, {81115}, {R}, {R version 4.1.2 (2021-11-01)}, {Bird Hippie}},  
RAttributes[names → {platform, arch, os, system, status, major, minor, year, month, day, svn rev, language, version.string, nickname}, class → {simple.list}]]

In[328]:= 


```
{  
  library(LSAMon)  
  library(magrittr)  
}
```


Out[328]= {magrittr, LSAMon, stats, graphics, grDevices, utils, datasets, methods, base}

In[336]:= 

```
{  
  dsAbstracts <- read.csv('https://raw.githubusercontent.com/antononcube/SimplifiedMachineLearningWorkflows-book/master/Data/Wolfram-Technology-Conference-2016-to-2019-abstracts.csv')  
  aAbstracts <- setNames(dsAbstracts$Abstract, dsAbstracts$ID)  
}
```

In[337]:= MyFromRForm[%]


In[306]:=  textHamlet

In[344]:= 

```
lsaObj <-  
LSAMonUnit(aAbstracts) %>%  
LSAMonMakeDocumentTermMatrix(stemWordsQ = TRUE, stopWords = NULL) %>%  
LSAMonApplyTermWeightFunctions(globalWeightFunction = "IDF", localWeightFunction = "None", normalizerFunction = "Cosine") %>%  
LSAMonExtractTopics(numberOfTopics = 20, method = "NNMF", maxSteps = 16, minNumberOfDocumentsPerTerm = 4) %>%  
LSAMonEchoTopicsTable(numberOfTerms = 10, wideFormQ = TRUE)
```

Out[344]= RObject[ {RDataFrame[ RNames[TermRank, 01.none.br.clear, 02.point.spatial.pattern, 03.convex.optim.solver, 04.model.system.engin, 05.uniti.link.geo, 06.asymptot.equat.calculu,  
07.student.modul.think, 08.featur.audio.new, 09.financi.feed.market, 10.network.neural.workshop, 11.dive.latest.gain, 12.teacher.assess.grade, 13.deploi.cloud.excel,  
14.entiti.queri.properti, 15.pde.solv.ndsolv, 16.structur.cluster.unsupervis, 17.connect.human.<e2><80>, 18.imag.color.process, 19.label.placement.line, 20.t.face.chain],  
RData[{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}, ... 19 ..., {t, face, chain, notat, n, typeset, box, that, 1, conjectur}],  
RRowNames[1, 2, 3, 4, 5, 6, 7, 8, 9, 10], ... 12 ..., {NNMF}], RAttributes[... 1 ...] ]

large output | [show less](#) | [show more](#) | [show all](#) | [set size limit...](#)

In[345]:=  lsaObj %>% LSAMonTakeTopicsTable(wideFormQ=T)

```
Out[345]= RDataFrame[RNames[TermRank, 01.none.br.clear, 02.point.spatial.pattern, 03.convex.optim.solver, 04.model.system.engin, 05.uniti.link.geo, 06.asymptot.equat.calculu,
07.student.modul.think, 08.featur.audio.new, 09.financi.feed.market, 10.network.neural.workshop, 11.dive.latest.gain, 12.teacher.assess.grade, 13.deploi.cloud.excel,
14.entiti.queri.properti, 15.pde.solv.ndsolv, 16.structur.cluster.unsupervis, 17.connect.human.<e2><80>, 18.imag.color.process, 19.label.placement.line, 20.t.face.chain],
RData[{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}], {none, br, clear, solver, <e2><80>, not, step, disciplin, mathematica, you, viewer, educ},
{point, spatial, pattern, earthquak, deposit, crime, forest, gold, star, sight, anim, locat}, {convex, optim, solver, problem, focu, linear, new, overview, function, be, effici, applic},
{model, system, engin, more, develop, i, be, world, scienc, how, control, us}, {uniti, link, geo, game, languag, wolfram, function, intuit, compil, comput, we, new},
{asymptot, equat, calculu, algebra, differenti, relat, app, document, educ, solv, integr, sum}, {student, modul, think, topic, comput, our, my, cours, learn, i, we, experi},
{featur, audio, new, some, neural, classif, 3d, 11, graphic, network, version, extract}, {financi, feed, market, data, bloomberg, financ, live, enabl, retriev, kei, platform, access},
{network, neural, workshop, learn, dure, train, classifi, high, particip, project, framework, involv}, {dive, latest, gain, out, check, insight, into, upcom, geometri, your, analyz, you},
{teacher, assess, grade, materi, cours, student, assign, ani, sylva, test, school, question}, {deployi, cloud, excel, web, api, wolfram, local, your, code, librari, project, alpha},
{entiti, queri, properti, databas, data, chemic, defin, class, or, curat, type, store}, {pde, solv, ndsolv, equat, partial, variou, shown, differenti, numer, will, attend, region},
{structur, cluster, unsupervis, be, tool, intrins, soon, prove, anomali, classifi, power, learn}, {connect, human, <e2><80>, scene, we, ll, data, theorem, interfac, techniqu, a, relat},
{imag, color, process, filter, object, oper, visual, audio, mani, go, count, techniqu}, {label, placement, line, chart, manual, unit, fundament, plot, point, area, cover, quantiti},
{t, face, chain, notat, n, typeset, box, that, 1, conjectur, there, your}]], RRowNames[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]]
```

```
In[346]= MyFromRForm[%]
```

	TermRank	01.none.br.clear	02.point.spatial.patteri	03.convex.optim.solvei	04.model.system.engir	05.uniti.link.geo	06.asymptot.equat.calculu	07.student.modul.thinl	08.featur.audio.new	09.financi.feed.market
1	1	none	point	convex	model	uniti	asymptot	student	featur	financi
2	2	br	spatial	optim	system	link	equat	modul	audio	feed
3	3	clear	pattern	solver	engin	geo	calculu	think	new	market
4	4	solver	earthquak	problem	more	game	algebra	topic	some	data
5	5	<e2><80>	deposit	focu	develop	languag	differenti	comput	neural	bloomberg
6	6	not	crime	linear	i	wolfram	relat	our	classif	financ
7	7	step	forest	new	be	function	app	my	3d	live
8	8	disciplin	gold	overview	world	intuit	document	cours	11	enabl
9	9	mathematica	star	function	scienc	compil	educ	learn	graphic	retriev
10	10	you	sight	be	how	comput	solv	i	network	kei
11	11	viewer	anim	effici	control	we	integr	we	version	platform
12	12	educ	locat	applic	us	new	sum	experi	extract	access

K < columns 1–10 of 21 > >

Python

```
In[264]= Concretize["Make LSA object over aAbstracts. Extract 20 topics using SVD.", "TargetLanguage" → "Python"]
```

```
Out[264]= lsaObj = (LatentSemanticAnalyzer()
.make_document_term_matrix(docs=SVD, stop_words=None, stemming_rules=None,min_length=3)
.apply_term_weight_functions(global_weight_func='IDF', local_weight_func='None',normalizer_func='Cosine')
.extract_topics(number_of_topics=20, min_number_of_documents_per_term=20, method='SVD')
.echo_topics_interpretation(number_of_terms=10, wide_form=True)
.echo_statistical_thesaurus(terms=stemmerObj.stemWords(["SVD"]), wide_form=True, number_of_nearest_neighbors=12, method='cosine', echo_function=lambda x: print(x.to_string()))
```

```
In[265]:= Concretize["LatentSemanticAnalysis", "Make LSA object over aAbstracts. Extract 20 topics using SVD.", "TargetLanguage" → "Python"]

Out[265]= lsaObj = (LatentSemanticAnalyzer()
  .make_document_term_matrix(docs=SVD, stop_words=None, stemming_rules=None,min_length=3)
  .apply_term_weight_functions(global_weight_func='IDF', local_weight_func='None',normalizer_func='Cosine')
  .extract_topics(number_of_topics=20, min_number_of_documents_per_term=20, method='SVD')
  .echo_topics_interpretation(number_of_terms=10, wide_form=True)
  .echo_statistical_thesaurus(terms=stemmerObj.stemWords(["SVD"]), wide_form=True, number_of_nearest_neighbors=12, method='cosine', echo_function=lambda x: print(x.to_string())))
```

---

## DSL interpretation demo

```
In[571]:= dsAbstracts = ResourceFunction["ImportCSVToDataset"] [
  "https://raw.githubusercontent.com/antononcube/SimplifiedMachineLearningWorkflows-book/master/Data/Wolfram-Technology-Conference-2016-to-2019-abstracts.csv"];
aAbstracts = AssociationThread[Normal[dsAbstracts[All, "ID"]], Normal[dsAbstracts[All, "Abstract"]]];

In[580]:= LSAMonUnit[aAbstracts] ⇒
  LSAMonMakeDocumentTermMatrix["StemmingRules" → False, "StopWords" → Automatic] ⇒
  LSAMonApplyTermWeightFunctions["GlobalWeightFunction" → "IDF", "LocalWeightFunction" → "None", "NormalizerFunction" → "Cosine"] ⇒
  LSAMonExtractTopics["NumberOfTopics" → 12, Method → "SVD", "MaxSteps" → 12, "MinNumberOfDocumentsPerTerm" → 2] ⇒
  LSAMonEchoTopicsTable["NumberOfTerms" → 12] ⇒
  LSAMonEchoStatisticalThesaurus["Words" → {"notebook", "equation", "neural"}];
```

» topics table:

<b>1</b> 1.000 none 0.228 data 0.193 new 0.166 wolfram 0.157 language 0.141 functions 0.139 development 0.133 learning 0.131 talk 0.130 features 0.129 science 0.129 mathematica	<b>4</b> 1.000 science 0.550 math 0.327 data 0.183 analytics 0.119 information -0.089 new -0.071 neural -0.064 functions -0.058 audio 0.057 development 0.052 software -0.050 optimization	<b>7</b> 1.000 cloud 0.646 player 0.621 ios -0.534 neural -0.463 audio 0.382 apps 0.372 interact 0.365 websites 0.357 attachments -0.356 cloud/software 0.352 notebooks -0.333 image	<b>10</b> 1.000 students -0.797 education -0.687 cloud/software -0.638 new -0.586 cloud 0.539 course 0.531 computational 0.518 thinking -0.507 features -0.488 ios -0.476 audio -0.473 science
<b>2</b> 1.000 none -0.173 data -0.144 new -0.126 wolfram -0.121 language -0.110 functions -0.105 development -0.102 science -0.100 cloud -0.100 talk -0.100 features -0.099 software	<b>5</b> 1.000 spatial 0.937 point -0.571 education 0.325 data 0.288 patterns 0.286 points 0.278 modeling 0.270 positions 0.267 animal 0.259 crime 0.259 sightings 0.259 earthquakes	<b>8</b> 1.000 visualization/ip 2.587 × 10 <sup>-14</sup> education -2.466 × 10 <sup>-14</sup> optimization -2.204 × 10 <sup>-14</sup> convex 1.913 × 10 <sup>-14</sup> neural 1.432 × 10 <sup>-14</sup> audio -1.199 × 10 <sup>-14</sup> math 1.184 × 10 <sup>-14</sup> networks 1.167 × 10 <sup>-14</sup> data -1.116 × 10 <sup>-14</sup> solvers -1.037 × 10 <sup>-14</sup> problems -9.689 × 10 <sup>-15</sup> new	<b>11</b> 1.000 math -0.833 data -0.514 financial -0.487 abstract 0.471 neural -0.384 feeds -0.345 bloomberg -0.338 market 0.297 students -0.291 finance -0.286 providers 0.285 point
<b>3</b> 1.000 software 0.923 development 0.403 cloud/software 0.114 cloud -0.113 science -0.084 data -0.054 new -0.050 math -0.046 functions -0.046 neural -0.044 point -0.043 spatial	<b>6</b> 1.000 education 0.200 spatial 0.195 point -0.161 neural -0.119 audio -0.095 networks -0.089 image 0.089 students 0.083 calculus 0.072 asymptotic 0.068 equations 0.068 math	<b>9</b> 1.000 optimization 0.887 convex -0.803 education -0.585 neural 0.447 solvers -0.445 audio 0.393 problems -0.370 networks 0.365 new 0.340 asymptotic 0.327 math -0.326 data	<b>12</b> 1.000 cloud/software -0.528 software 0.252 unity 0.130 development 0.123 computational -0.114 education -0.110 neural 0.105 link 0.105 game -0.104 optimization 0.102 students 0.100 abstract

» statistical thesaurus:

term	statistical thesaurus entries
equation	{equation, integral, ndsolve, algebraic, special, form, nonlinear, pdes, numerical, included, precision, shown}
neural	{neural, audio, networks, image, network, classification, learning, workshop, framework, processing, net, machine}
notebook	{notebook, step, setting, educational, format, deployment, come, unique, document, enhance, product, cases}

## Dimension reduction 3D comparison demo

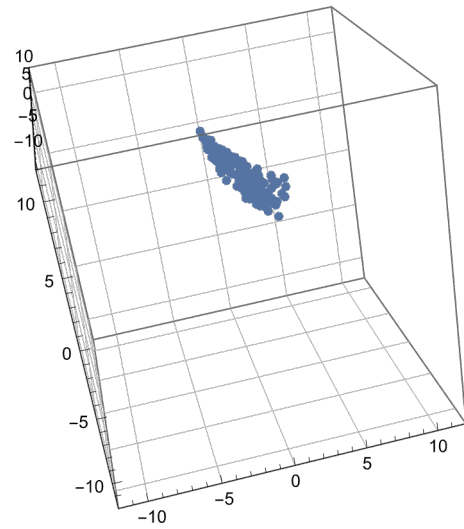
```
In[432]:= n = 200;  
c = 12;  
SeedRandom[232];  
points = Transpose[RandomVariate[NormalDistribution[0, #], n] & /@ {2, 4, 0.1}];  
points = points.RotationMatrix[{{0, 0, 1}, {-1, 0, 1}}];  
points = Map[{2, 8, 2} + # &, points];  
points = Clip[points, {0, Max[points]}];  
opts = {BoxRatios -> {1, 1, 1}, PlotRange -> Table[{c, -c}, 3], FaceGrids -> {{0, 0, -1}, {0, 1, 0}, {-1, 0, 0}}, ViewPoint -> {-1.1, -2.43, 2.09}, ImageSize -> 300};  
gr0 = ListPointPlot3D[points, opts];
```

```

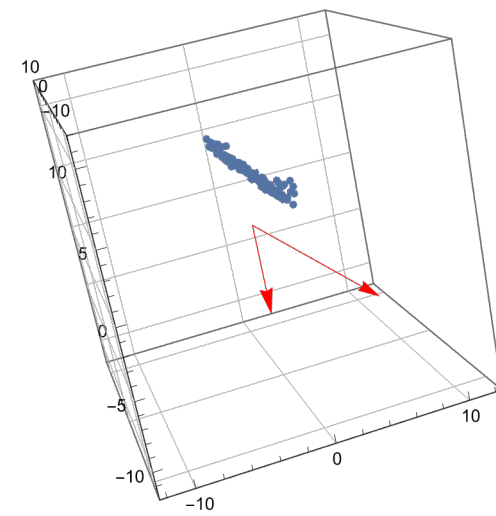
In[488]:= SeedRandom[232];
{W, H} = ResourceFunction["NonNegativeMatrixFactorization"][points, 2, "Normalization" → Right];
grNNMF = Show[{ListPointPlot3D[points], Graphics3D[{Red, Arrow[{0, 0, 0}, #]} & /@ (c * H / Norm[H])]}], opts, PlotLabel → "NNMF"];
{U, S, V} = SingularValueDecomposition[points, 2];
grSVD = Show[{ListPointPlot3D[points], Graphics3D[{Red, Arrow[{0, 0, 0}, #]} & /@ (c * Transpose[V])]}], opts, PlotLabel → "SVD"];
{A, S} = ResourceFunction["IndependentComponentAnalysis"][points, 2];
grICA = Show[{ListPointPlot3D[points], Graphics3D[{Red, Arrow[{0, 0, 0}, #]} & /@ (c * S / Norm[S])]}], opts, PlotLabel → "ICA"];
Grid[{{gr0, grSVD}, {grICA, grNNMF}}]

```

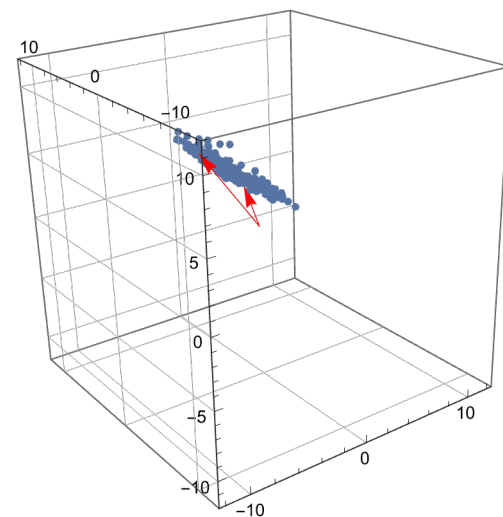
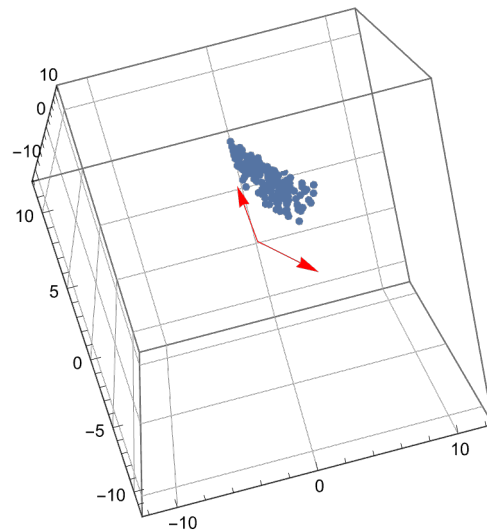
SVD



ICA



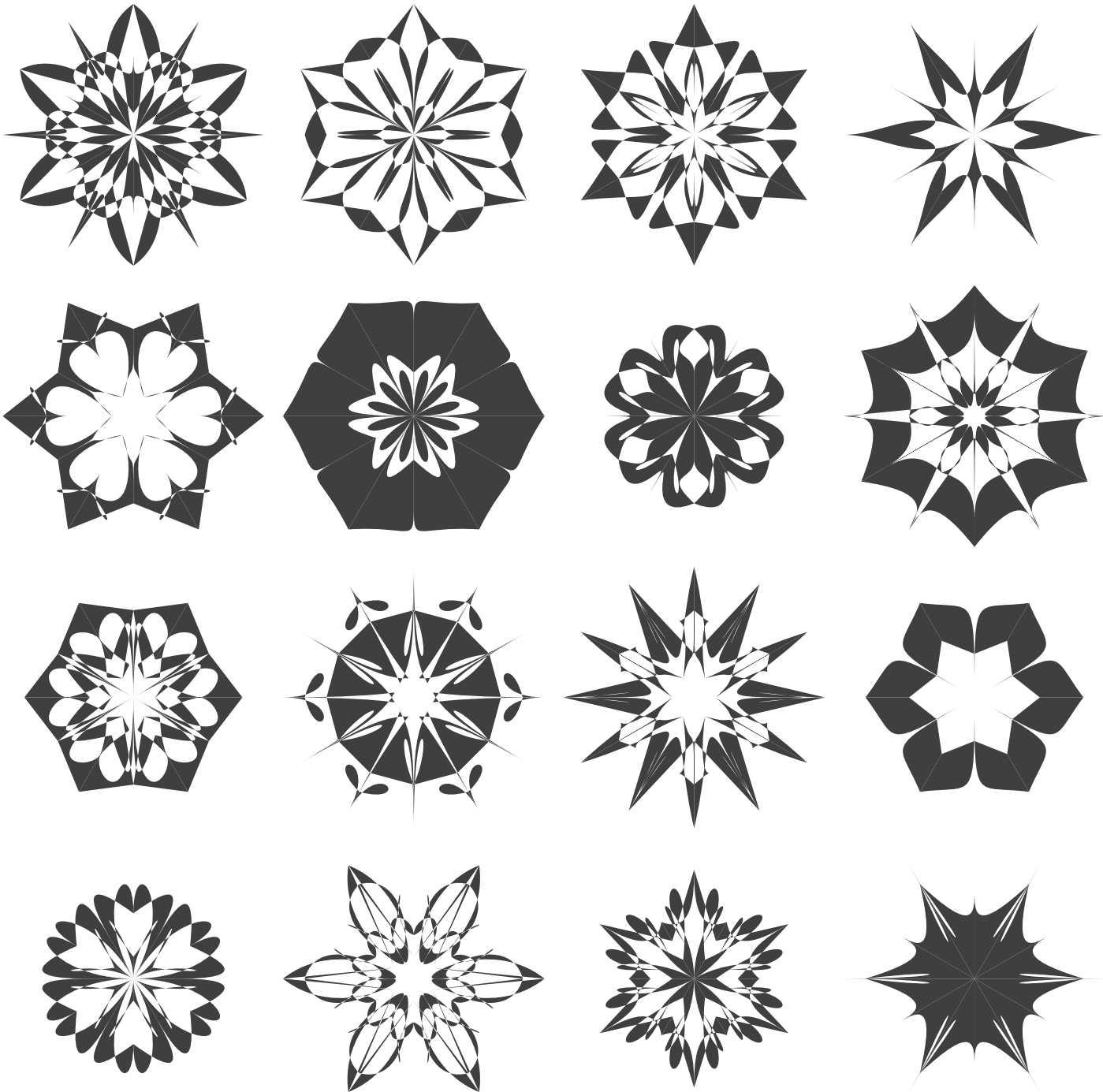
NNMF



# Random mandalas

```
In[582]:= Multicolumn@Table[ResourceFunction["RandomMandala"] ["ConnectingFunction" → FilledCurve@*BezierCurve], 16]
```

Out[582]=



In[590]:= Multicolumn@Table[ResourceFunction["RandomMandala"] ["Radius" → {10, 5, 2}, "ConnectingFunction" → FilledCurve@\*BezierCurve, ColorFunction → "DarkBands"], 16]

Out[590]=

