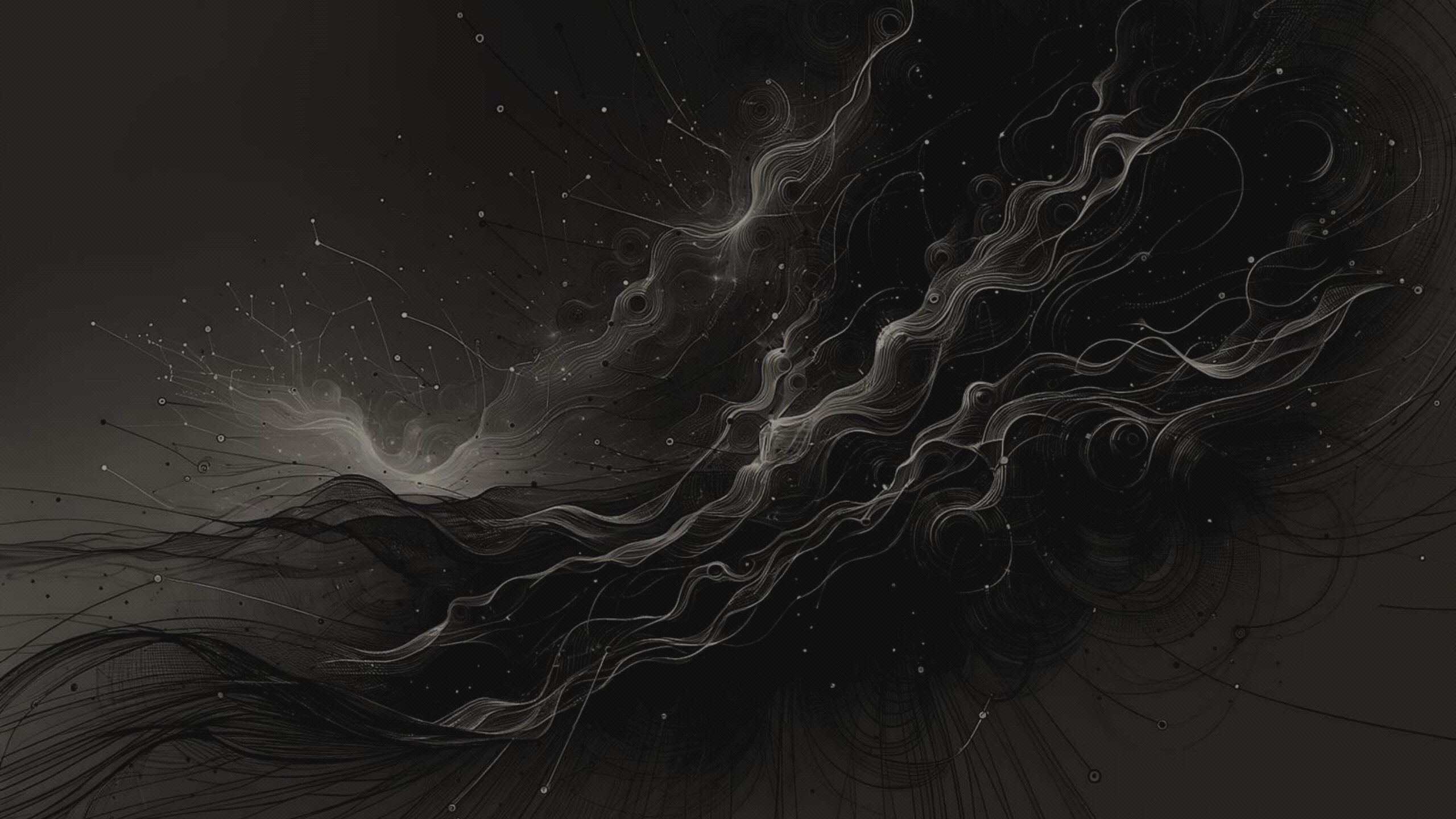

K-MEANS & FRIENDS

EXPLORING CLUSTERING WITH PHOTOGRAPHS

shaurya agarwal

Outline

- me motivate you
- me movie you
- me math you
- me method you
- you more me

me motivate you

what are we talking about again?

teach / learn

this dude

teach / learn

it’s hubris to think that anyone can come here and teach you

teach / learn

no one can teach you

teach / learn

but you can learn

teach / learn

the best anyone can hope for is to inspire you to learn

teach / learn

that’s the only greatness i aspire to today

wip

this is a larger workshop i am building and is a work-in-progress

wip

thank you for partnering in helping me build it

wip

we’ll switch around a lot of windows for now

wip

close your eyes if you see something you should not ;)

me movie you

a rather unusual use case for k-means

i got to do this

cinematography

on

a

budget

powered

by

python

first principles

the best cinematography has exceptional control on color

first principles

showcase what inspired the film’s cinematography

first principles - colors

- one, or two or a few

first principles - colors

- one, or two or a few but no more

first principles - colors

for each scene,

the director has ascribed meaning to these colors

and is using them to tell a story

first principles - colors

so, play as much as you like,

as long as these colors are (approximately) constant

pre-production

- there’s a lot of prep
- you select a color palette for each scene

pre-production

there are expensive cameras and studio setup that offer great control over the process

pre-production

a typical day may cost 10s of lakhs of rupees for camera and lighting

pre-production

for a 10 day shoot, interior and exterior

pre-production

we had 30K total

pre-production

so i had to leverage my technical knowhow to identify key colors,

tell a story around them

pre-production

or lose my limbs*

* we had a great director and a very dedicated team

enter Kandinsky

built this ages ago using javascript

enter Kandinsky

revived in Python

key insight

light is wavelengths

key insight

eyes respond to light

key insight

and to colors: long wavelength, medium wavelength, short wavelength

key insight

film and camera sensors do too, but differently

CIE XYZ 1931

they devised a standard based on how human eyes respond to various wavelengths

CIE Lab*

3 dimensional model (we have 3 cones remember) to represent how we see color

ACES, ARRI, SONY, sRGB, Rec.709, Rec.2020, etc. etc.

all models build up from the perceptual model – this is how all films manage color

color mastering

so the colors we decided on with the director, had to be identified at a perceptual level and kept consistent throughout the production as they information travelled from the camera sensor to the computer monitor to the cinema projector to the mobile phone and laptop screens...

(and TVs, screw TVs, but more on them later)

color mastering

long wavelength, medium wavelength, short wavelength

if i could capture the key colors perceptually, i can manage the rest of the process

how do you pick colors?

google ‘QUANTIZATION’

how do you pick colors?

i used k-means++ to do it

let’s see how

me math you

how clustering works

quantization, clustering and k-means

Image Compression, Customer Segmentation, Document Clustering, Anomaly Detection, Feature Learning and Dimensionality Reduction, Medical Imaging, Genomics and Bioinformatics, Speech Recognition, Astronomical Data Analysis, Pattern Recognition and Classification, even enabling efficient training for LLMs

naïve explainer

imagine you have a bunch of marbles of different colors scattered around, and you want to organize them into neat groups

the k-means method is like deciding to group these marbles based on how close they are to each other (in color)

the "k" part is you deciding into how many groups (or piles) you want to sort them.

technically

K-Means is a clustering algorithm used in machine learning and data mining to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. This results in a partitioning of the data space into Voronoi cells.

steps

Initialization

Start by selecting k initial centroids, where k is a predefined number of clusters

steps

Assignment

Assign each data point to the closest centroid, creating clusters

steps

Update

Recalculate the centroids as the mean of all points in each cluster

steps

Repeat

Repeat the assignment and update steps until the centroids no longer change significantly, indicating convergence

steps

Limit

Often the centroids do not converge, but dance around the ideal convergence points, this is when we stop the algorithm by specifying the maximum number of iterations it can cycle for

choosing K

Elbow Method

Plot the cost (e.g., within-cluster sum of squares) against different k values. The "elbow" point, where the rate of decrease sharply changes, can indicate a good k value.

choosing K

Silhouette Score

Measure how similar an object is to its own cluster compared to other clusters. A high silhouette score suggests the object is well matched to its own cluster and poorly matched to neighboring clusters. The k that maximizes the average silhouette score may be chosen.

choosing K

Gap Statistic

Compares the total within intra-cluster variation for different values of k with their expected values under null reference distribution of the data. The k with the highest gap statistic suggests the optimal clustering.

how?

let’s see this in action

me method you

let’s look at some code

tech stack

this is compute intensive

so leverage vectorized, just-in-time approach

tech stack

JAX parallelizes and optimizes the compute DAG over numpy

code

switch to code

YOU MORE ME

JOIN THE DISCUSSION, GITHUB, MAKE THIS BETTER

open-source

Eight Down Toofaan Mail is now on YouTube

https://www.youtube.com/watch?v=VnHPtozfhRU

this project on github: https://github.com/shauryashaurya/kandinsky

join me