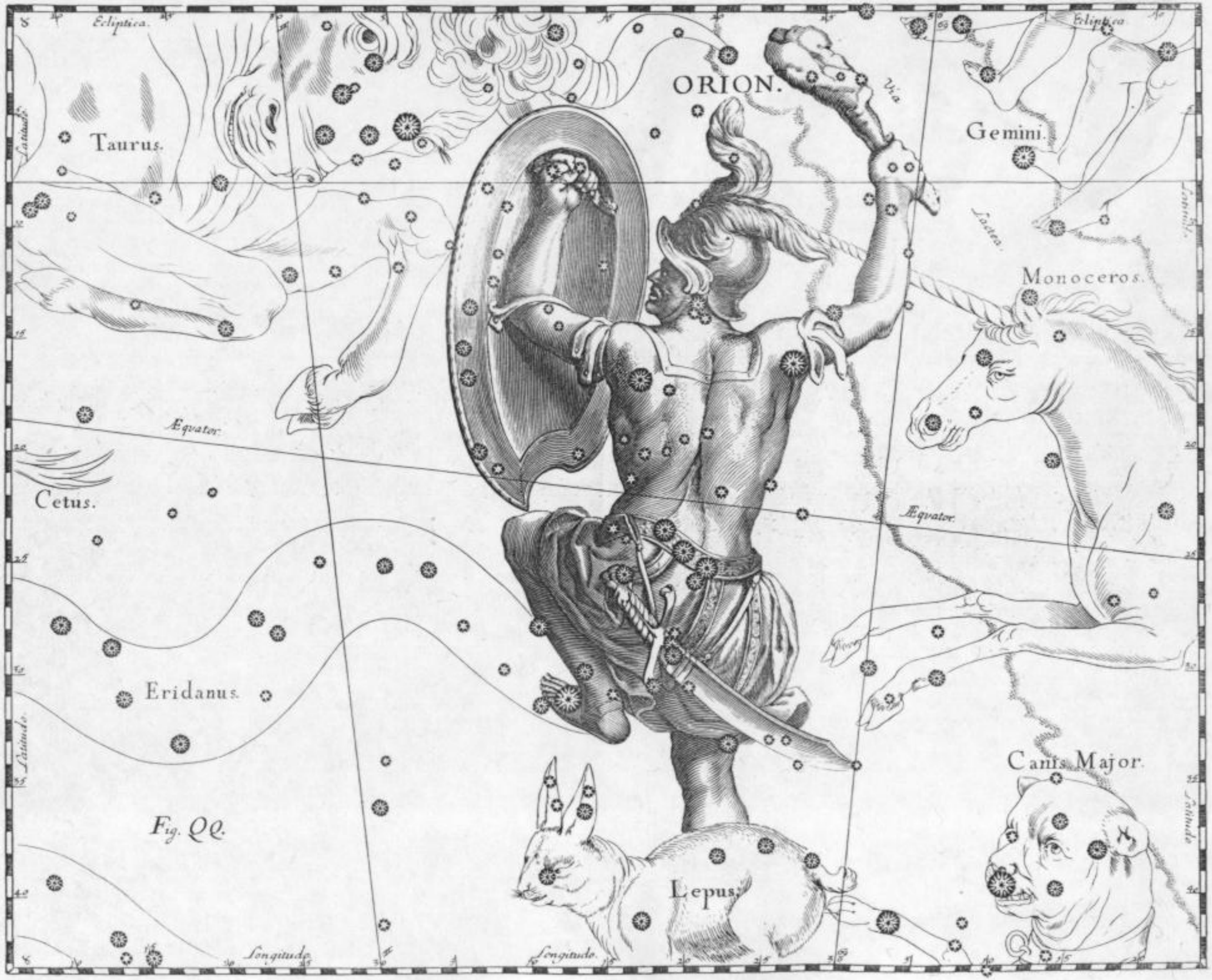
A machine learns to create constellations

Christian Gilbertson

STAT 557 Final Project

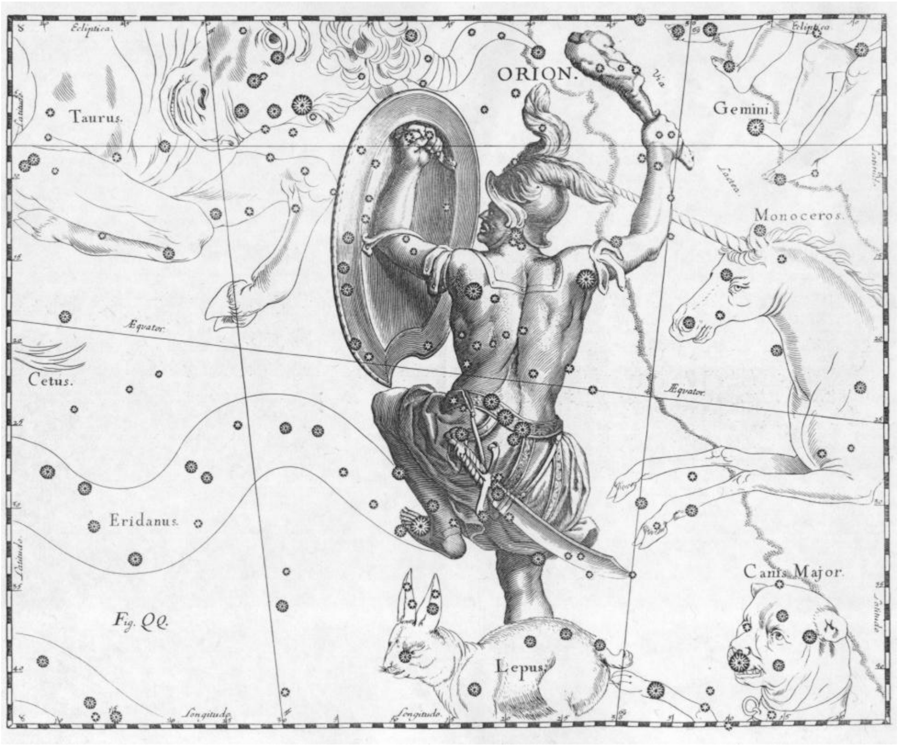
Introduction



ORION.
Taurus.
Gemini.
Monoceros.
Cetus.
Eridanus.
Canis Major.
Lepus.
Fig. QQ.
Ecliptica.
Æquator.
Via Lactea.
Longitudo.
Latitudo.

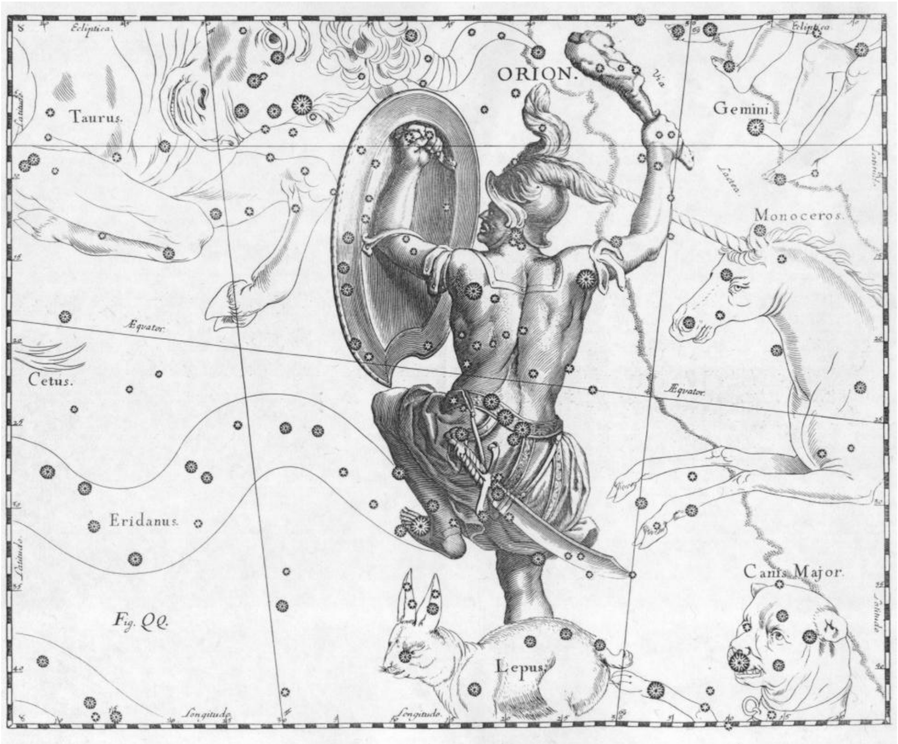
What makes a constellation?

- Group of nearby, usually bright stars



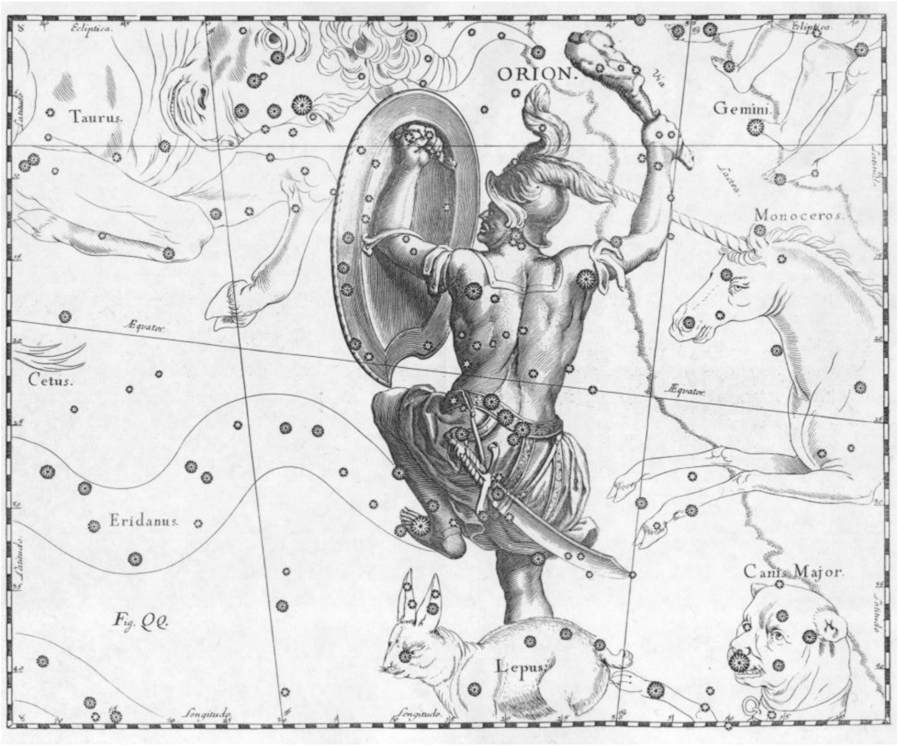
What makes a constellation?

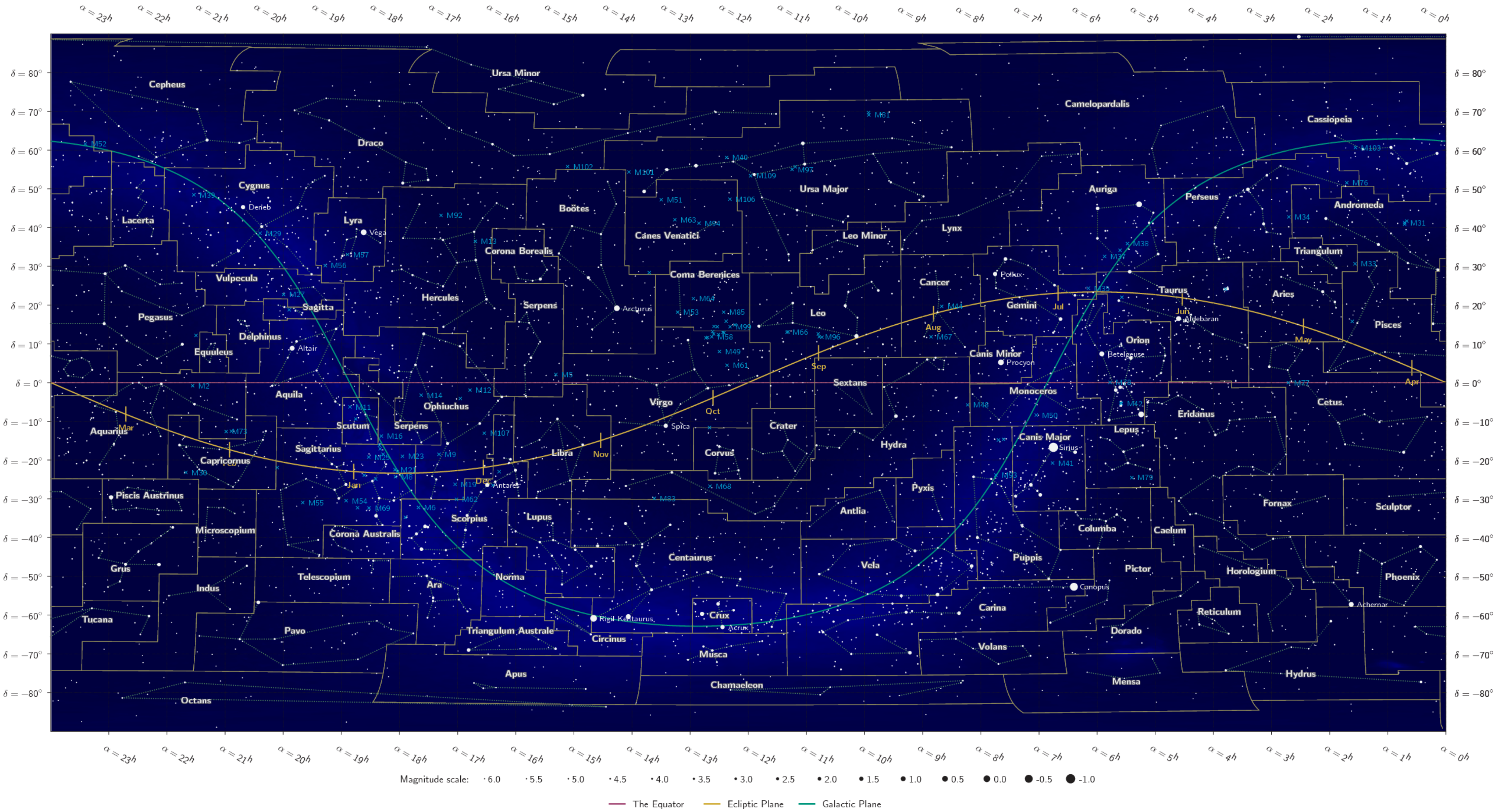
- Group of nearby, usually bright stars
- A human arbitrarily looking for “meaningful” patterns



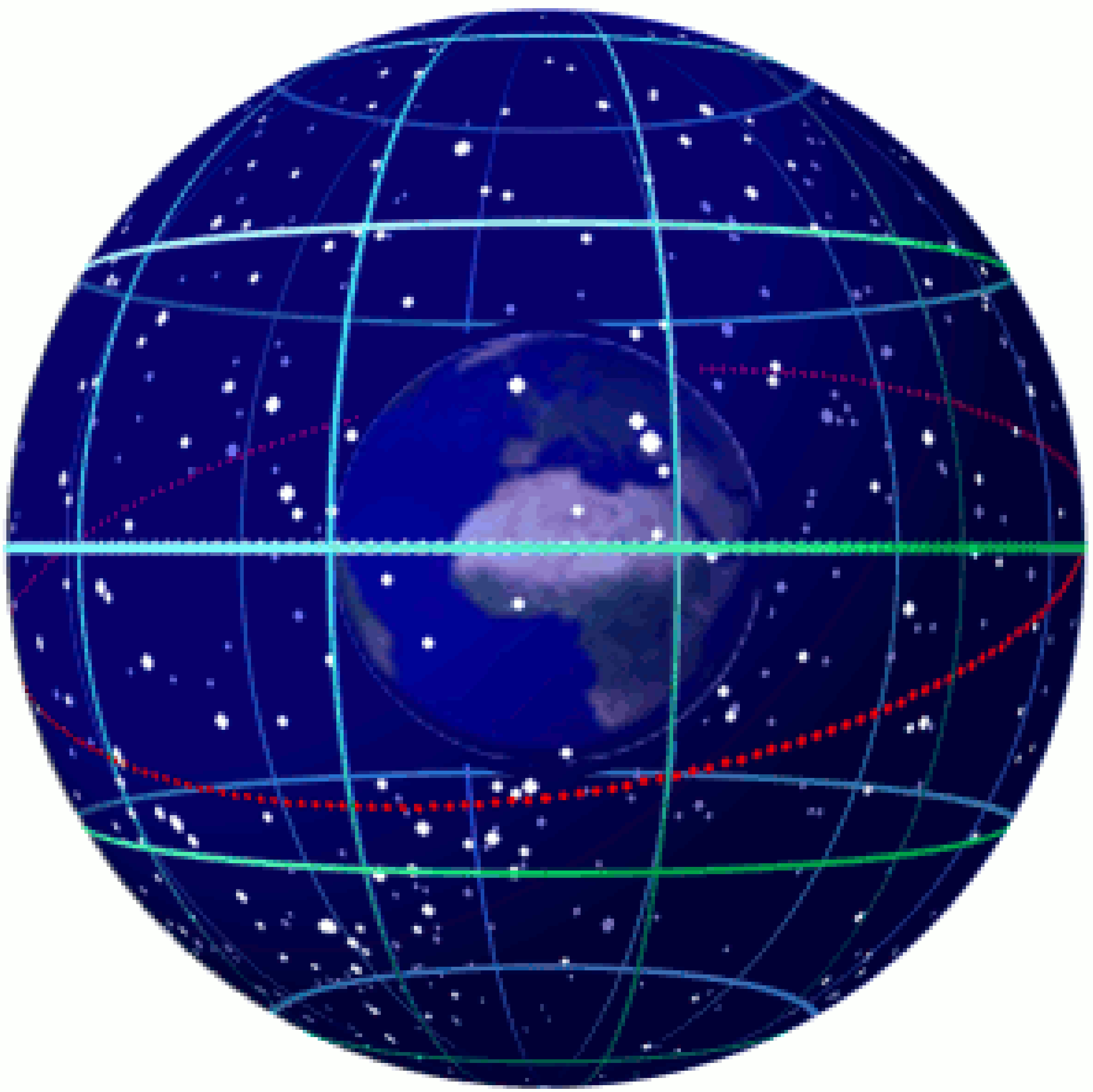
What makes a constellation?

- Group of nearby, usually bright stars
- A human arbitrarily looking for "meaningful" patterns
- That's basically it





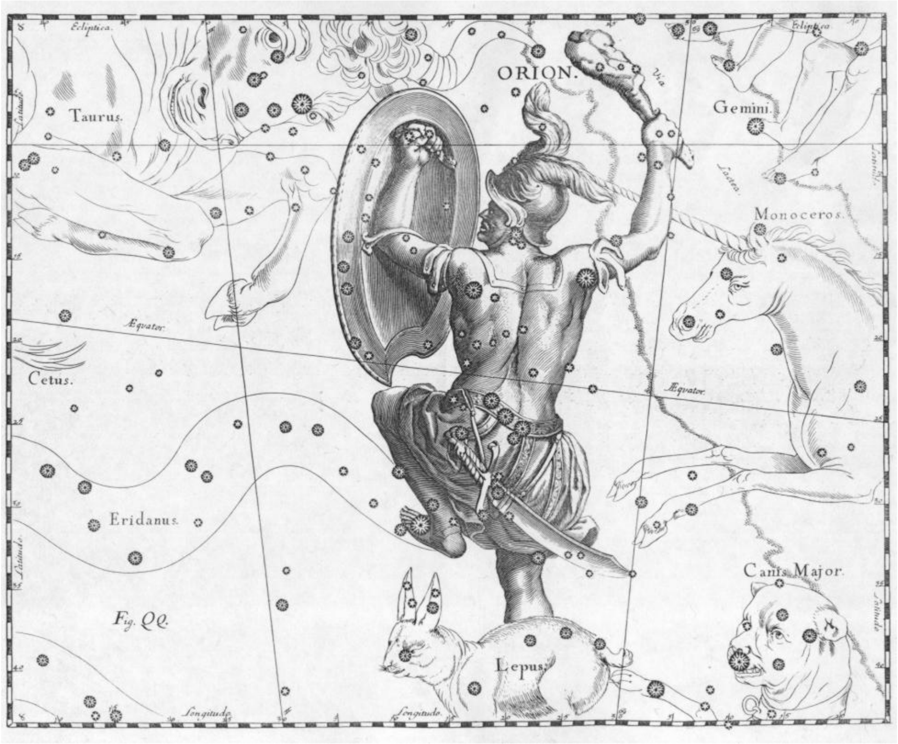
Cepheus
Ursa Minor
Draco
Camelopardalis
Cassiopeia
Cygnus
Deneb
Lacerta
Lyra
Vega
Boötes
Canes Venatici
Ursa Major
Leo Minor
Lynx
Auriga
Perseus
Andromeda
Triangulum
Corona Borealis
Coma Berenices
Vulpecula
Hercules
Serpens
Arcturus
Cancer
Pollux
Gemini
Taurus
Aldebaran
Aries
Pisces
Pegasus
Sagitta
Delphinus
Equuleus
Altair
Leo
Canis Minor
Procyon
Orion
Betelgeuse
Aquila
Ophiuchus
Virgo
Spica
Sextans
Monoceros
Cetus
Eridanus
Aquarius
Scutum
Capricornus
Sagittarius
Libra
Corvus
Crater
Hydra
Canis Major
Sirius
Lepus
Piscis Austrinus
Antares
Scorpius
Lupus
Antlia
Pyxis
Fornax
Sculptor
Microscopium
Corona Australis
Centaurus
Columba
Caelum
Grus
Telescopium
Norma
Vela
Puppis
Pictor
Horologium
Phoenix
Indus
Ara
Canopus
Tucana
Pavo
Triangulum Australe
Rigil Kentaurus
Crux
Acrux
Carina
Reticulum
Achernar
Circinus
Musca
Dorado
Volans
Apus
Chamaeleon
Mensa
Hydrus
Octans
Jan
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Magnitude scale: 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
The Equator
Ecliptic Plane
Galactic Plane



What makes a constellation?

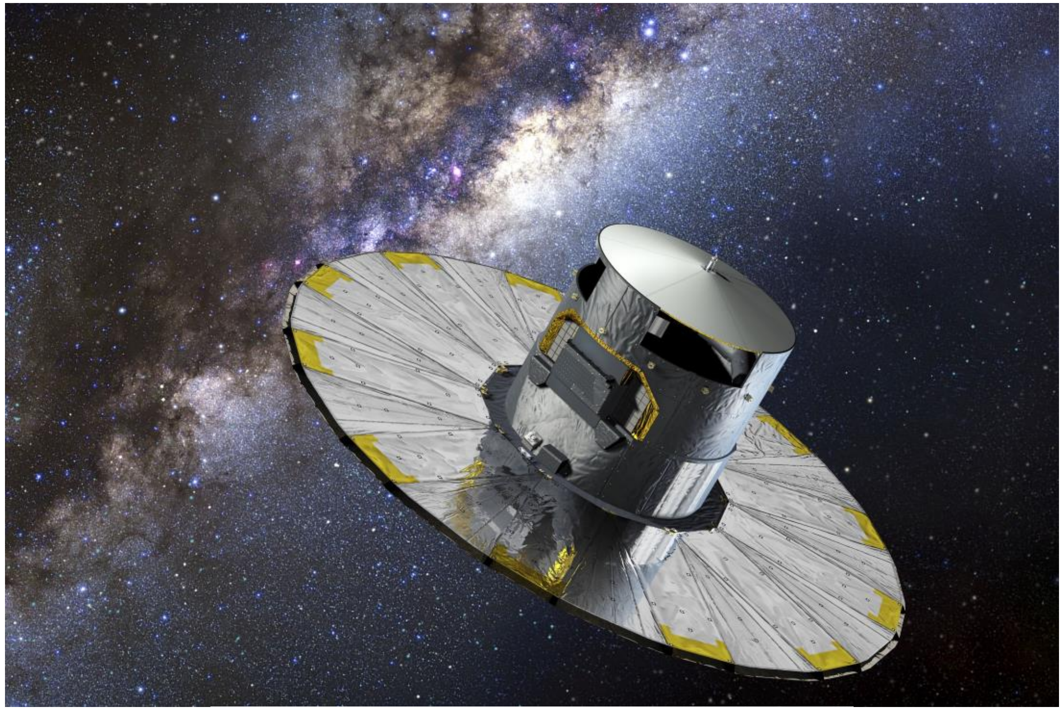
- Group of nearby, usually bright stars
- ~~A human arbitrarily looking for “meaningful” patterns~~
- A computer systematically clustering quantifiably similar stars?

- Lets see



Methods

Data



gaia

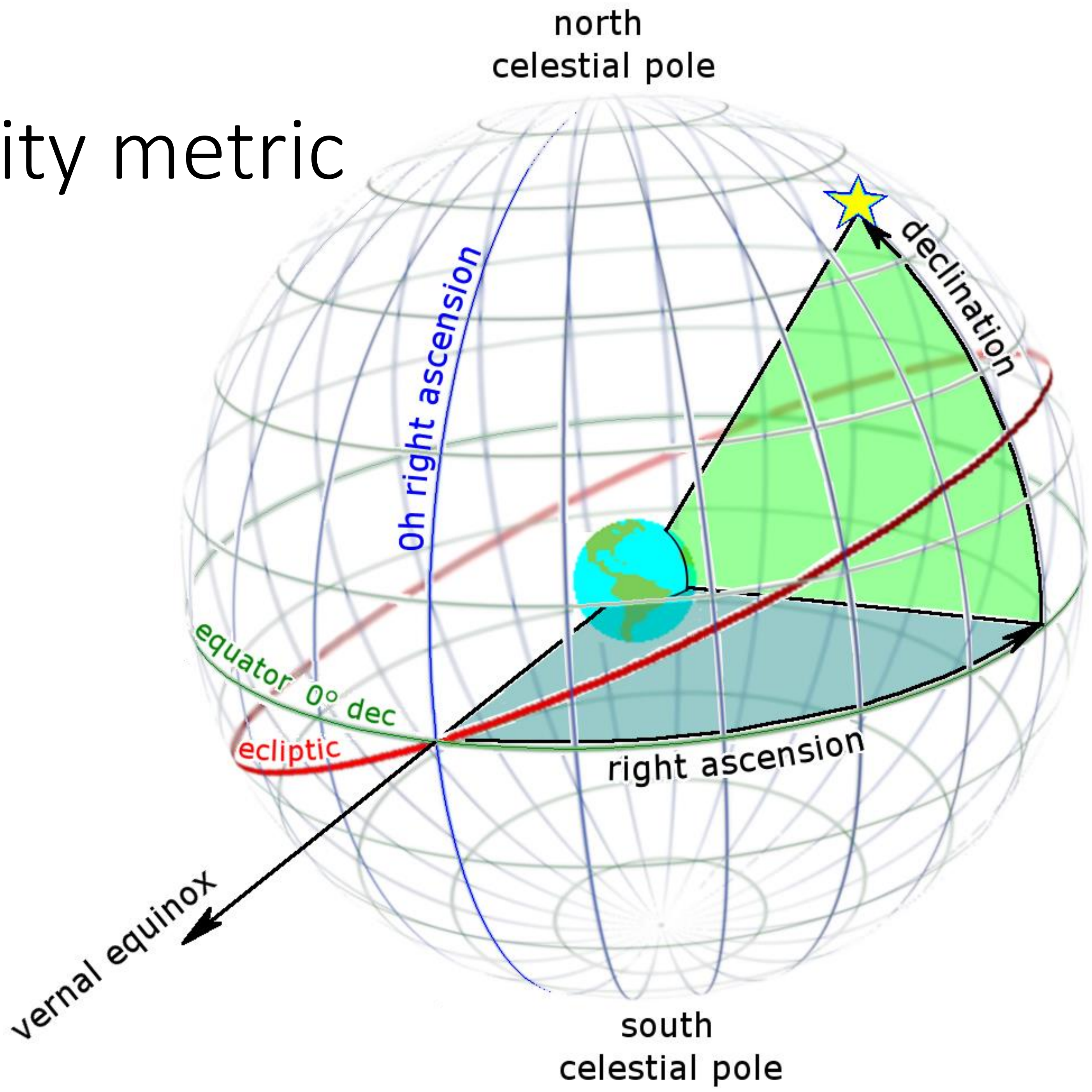
VIZIER

esa

European Space Agency

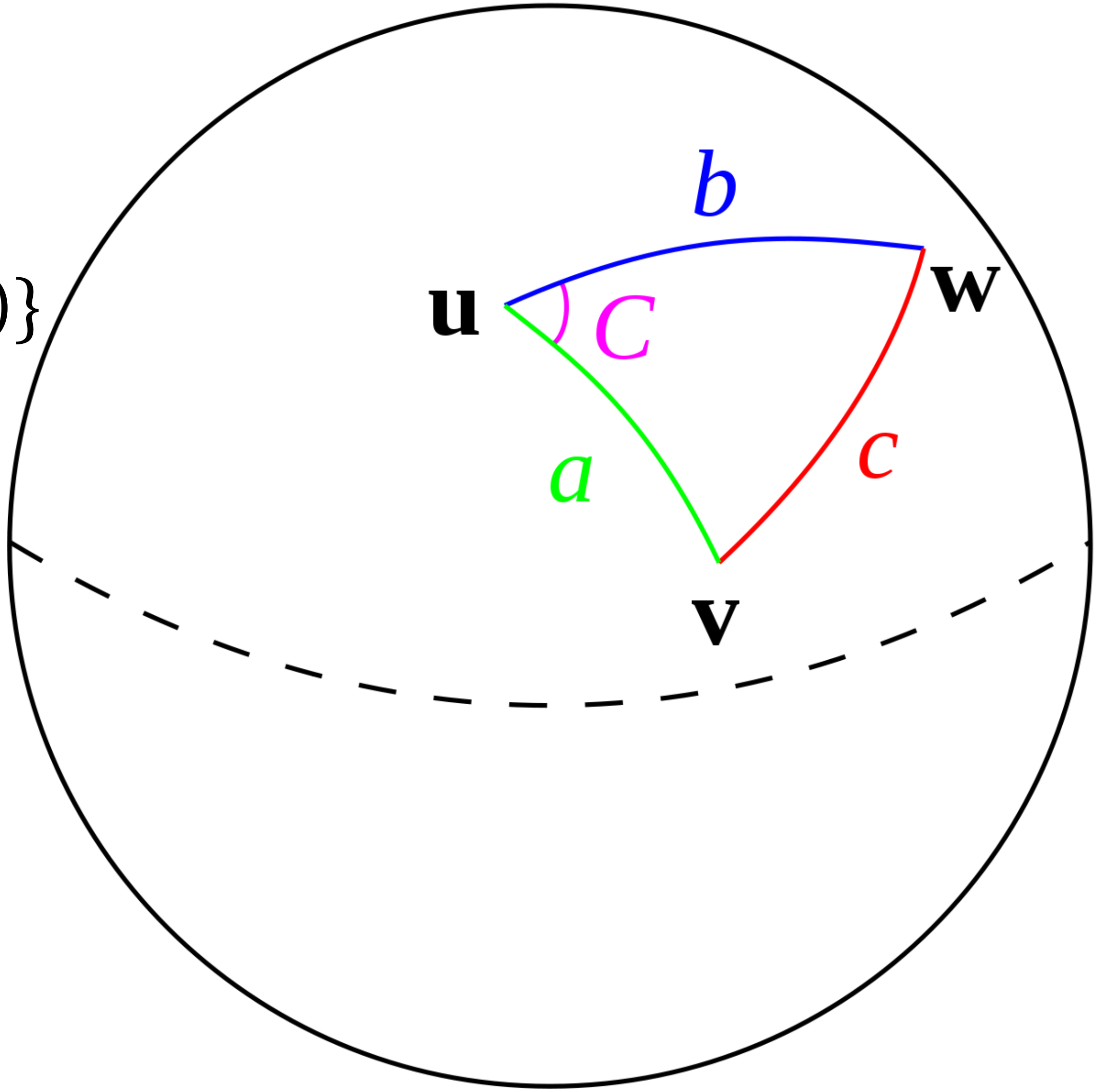
Generating a similarity metric

- Right ascension
- Declination
- Distance?
- Apparent magnitude



Distances between points

Positions: $\{\mathrm{x}, \mathrm{y}, \mathrm{z}\} =$
$\{\cos(\delta)\cos(\alpha),\ \cos(\delta)\sin(\alpha),\ \sin(\delta)\}$



Arc Length:

$$d_{ij} = \cos^{-1}(\langle x_i | p_j \rangle)$$

Cosine Dissimilarity:

$$d_{ij} = 1 - \langle x_i | p_j \rangle$$

Distances between points

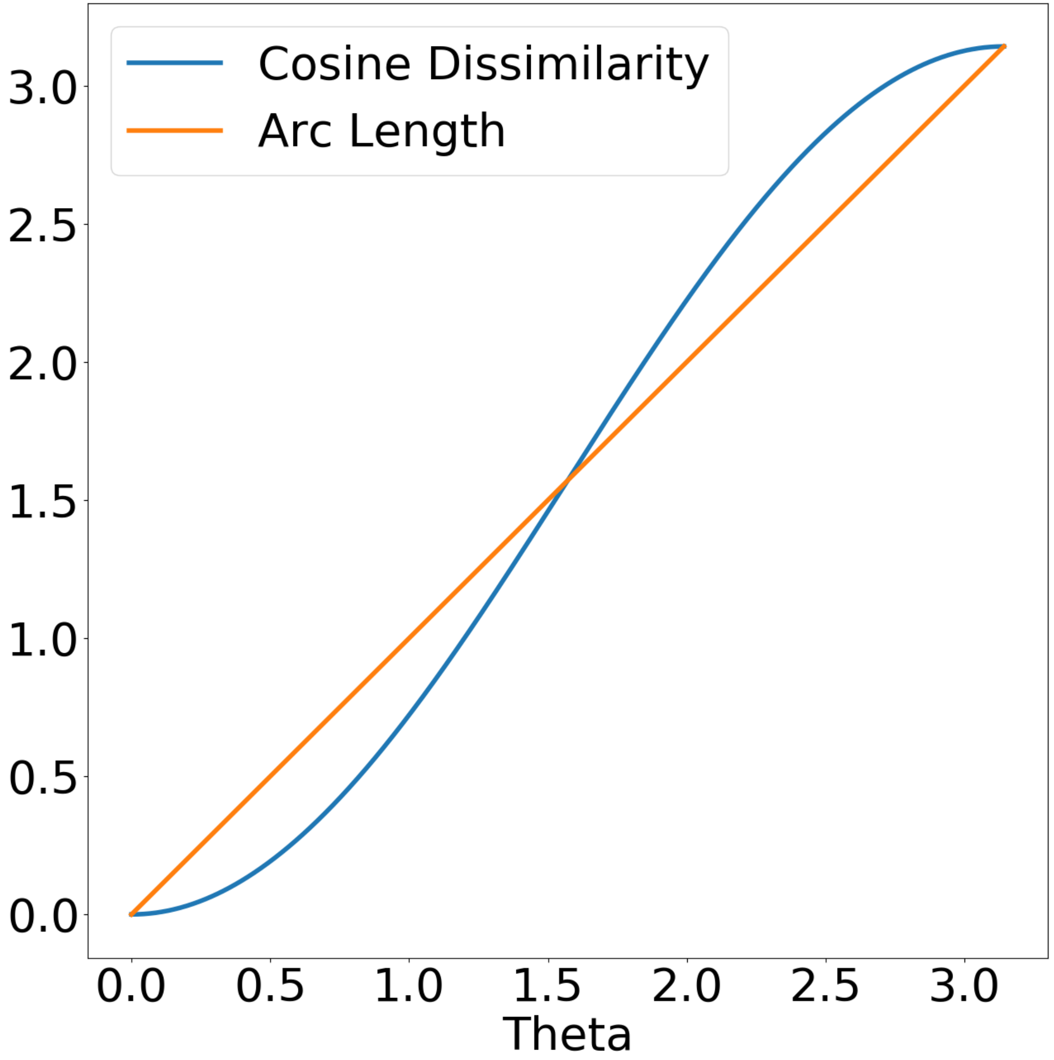
Positions: $\{\mathrm{x}, \mathrm{y}, \mathrm{z}\} =$
$\{\cos(\delta)\cos(\alpha),\ \cos(\delta)\sin(\alpha),\ \sin(\delta)\}$

Arc Length:

$$d_{ij} = \cos^{-1}(\langle x_i | p_j \rangle)$$

Cosine Dissimilarity:

$$d_{ij} = 1 - \langle x_i | p_j \rangle$$

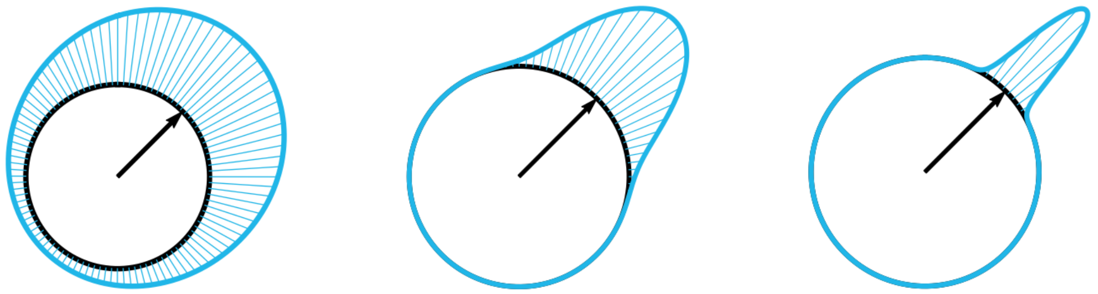


Heuristic cluster similarity

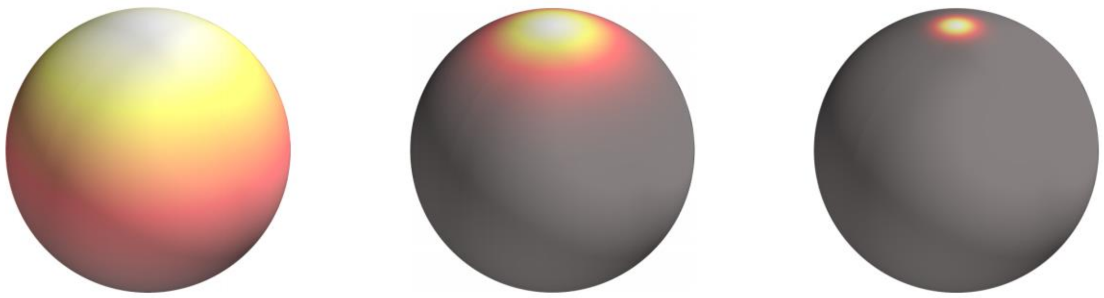
$$SG(x;\mu,\lambda) = \frac{\tau}{2\,\pi(1-e^{-2\tau})}\, e^{\tau\,(\langle\mu|x\rangle-1)}$$

aka 3D von Mises-Fisher distribution

- Create an approximate PDF by placing a SG at average location of every real constellation (spherical KDE)
- Evaluate at central locations of predicted constellations and see which method is has the highest combined “score”



(a) $\tau = 1$ (b) $\tau = 10$ (c) $\tau = 100$

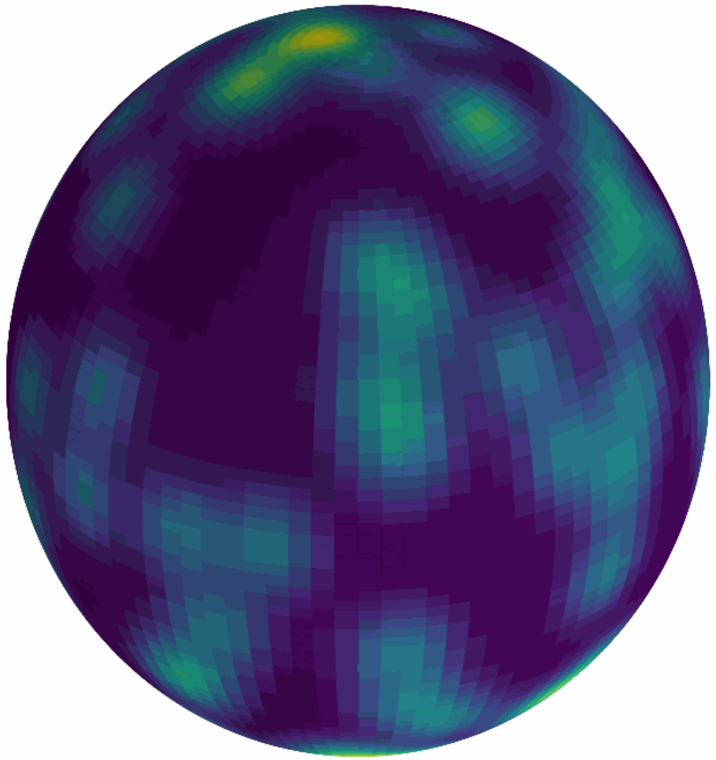


(a) $\tau = 1$ (b) $\tau = 10$ (c) $\tau = 100$

Straub 2017

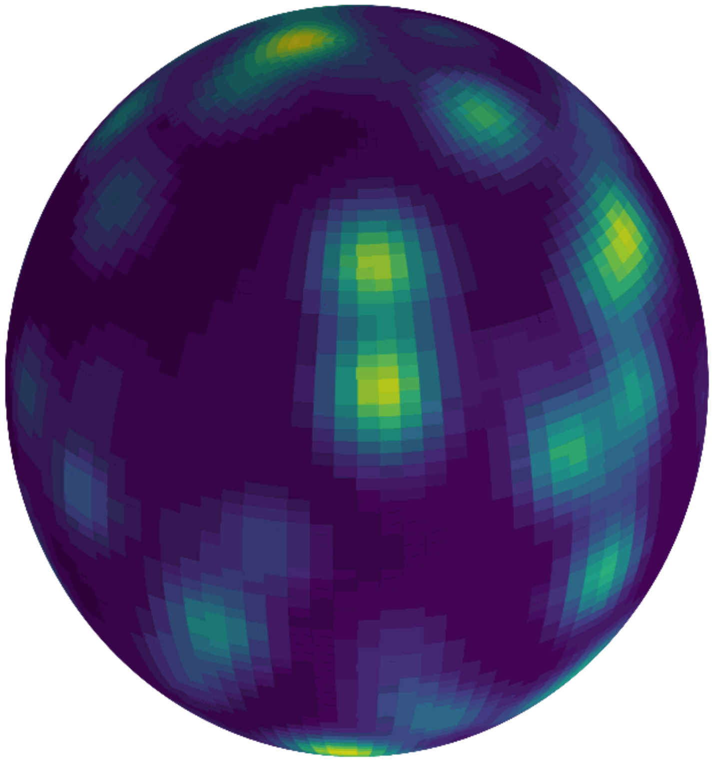
Cluster similarity heuristic

- Simple
  - Constant τ SG at all constellation locations
  - Constellation agnostic
  - Boring



Cluster similarity heuristic

- Simple
  - Constant τ SG at all constellation locations
  - Constellation agnostic
  - Boring
- Sky-area weighted
  - Add a multiplicative term based on sky area
  - Bigger constellations are more important

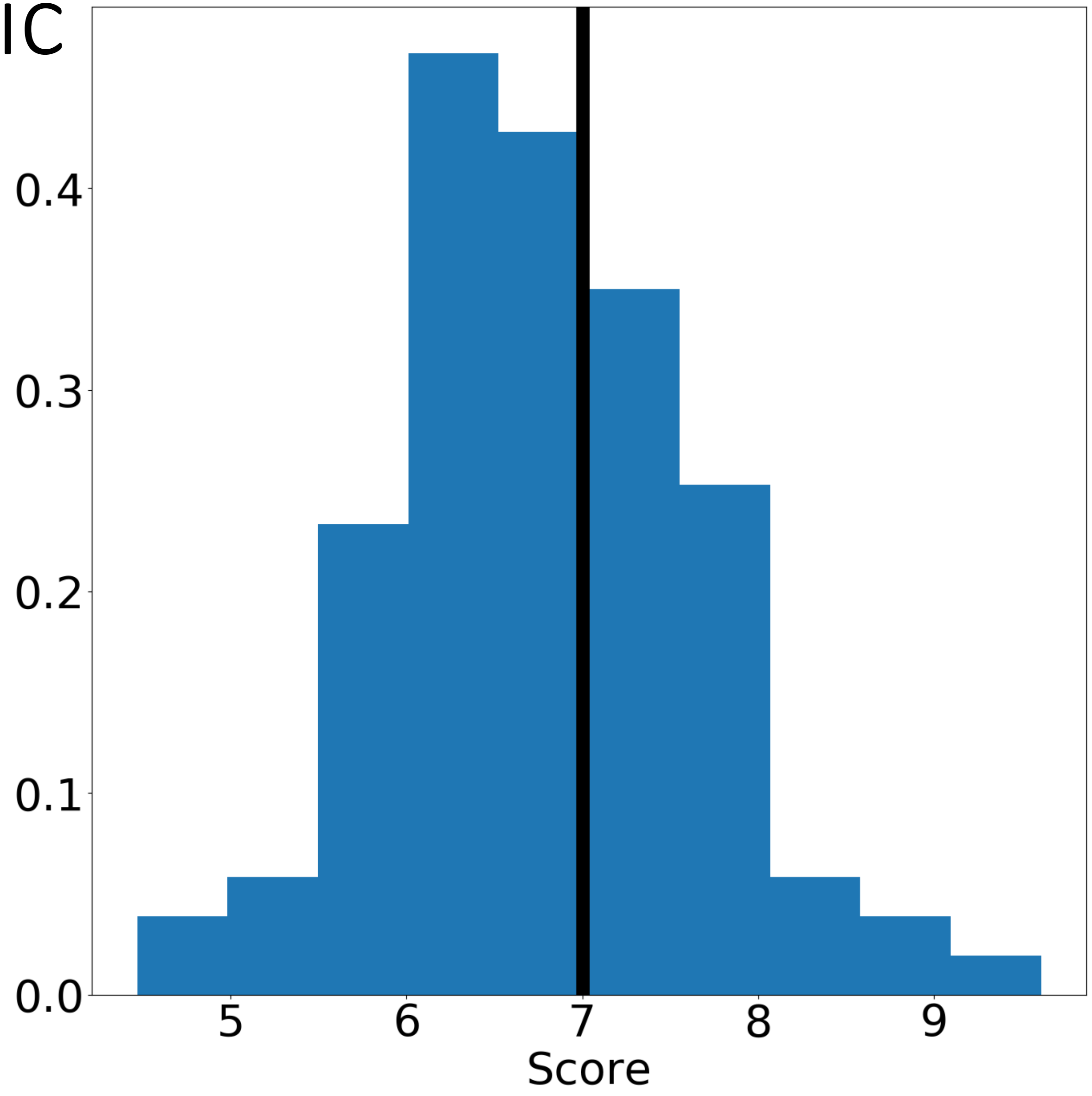


Cluster similarity heuristic

- Simple
  - Constant τ SG at all constellation locations
  - Constellation agnostic
  - Boring
- Sky-area weighted
  - Add a multiplicative term based on sky area
  - Bigger constellations are more important
- Sky-area spread
  - Smaller τ for larger constellations
  - Penalizes for missing smaller constellations
  - Feels worse

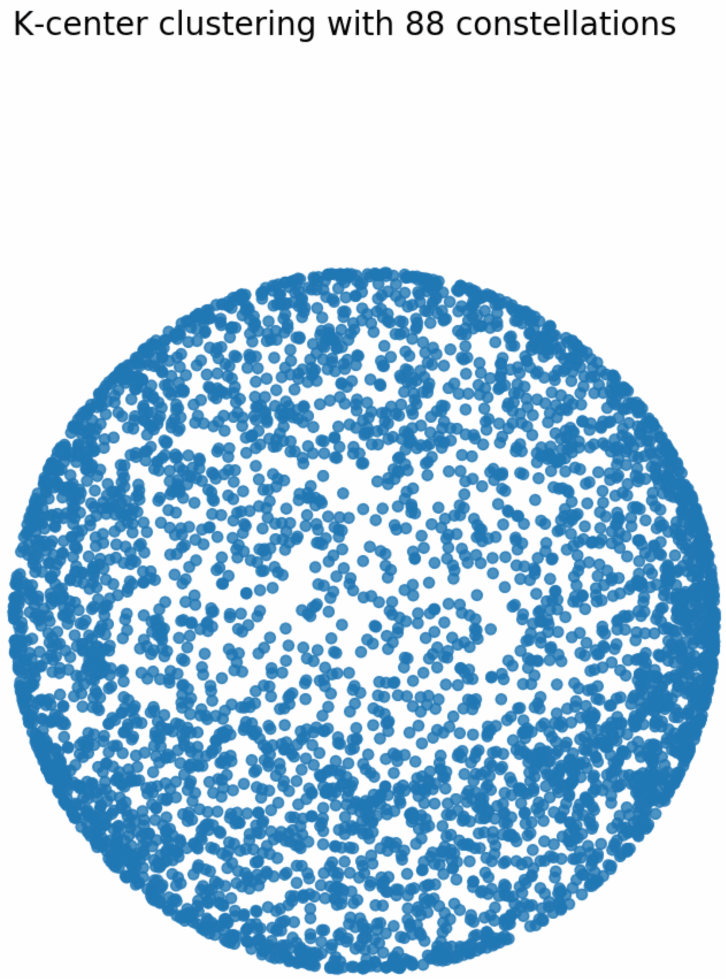
Cluster similarity heuristic

- Sky-area weighted heuristic
- 88 randomly drawn centers
- Distributed around $\frac{88}{4\,\pi}$ as expected

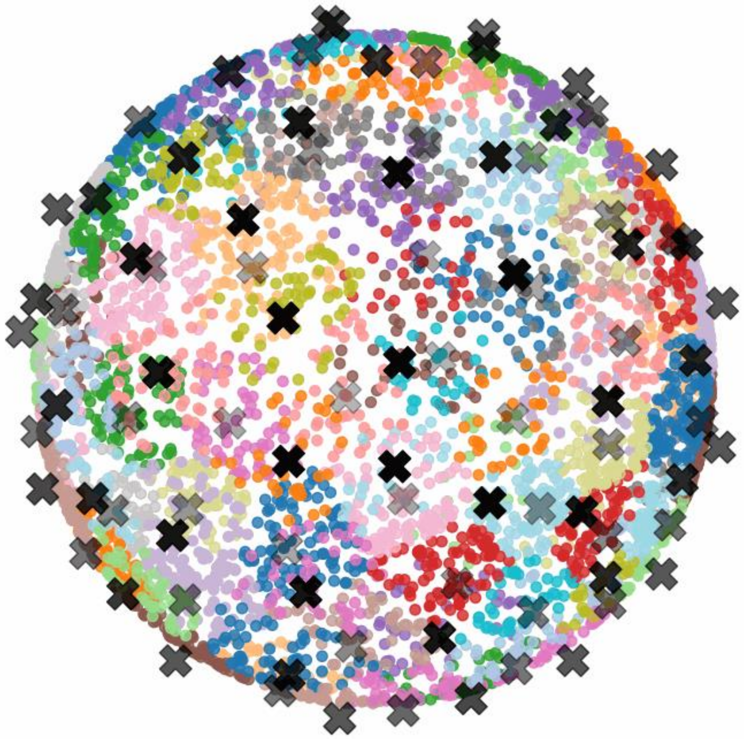


K-center clustering

- Randomly initialize first cluster at one star
- Find star “furthest” away from all existing cluster centers and make that a new center
- Repeat until the desired amount of clusters is found
- Highly dependent on initial star



K-center clustering with 88 constellations



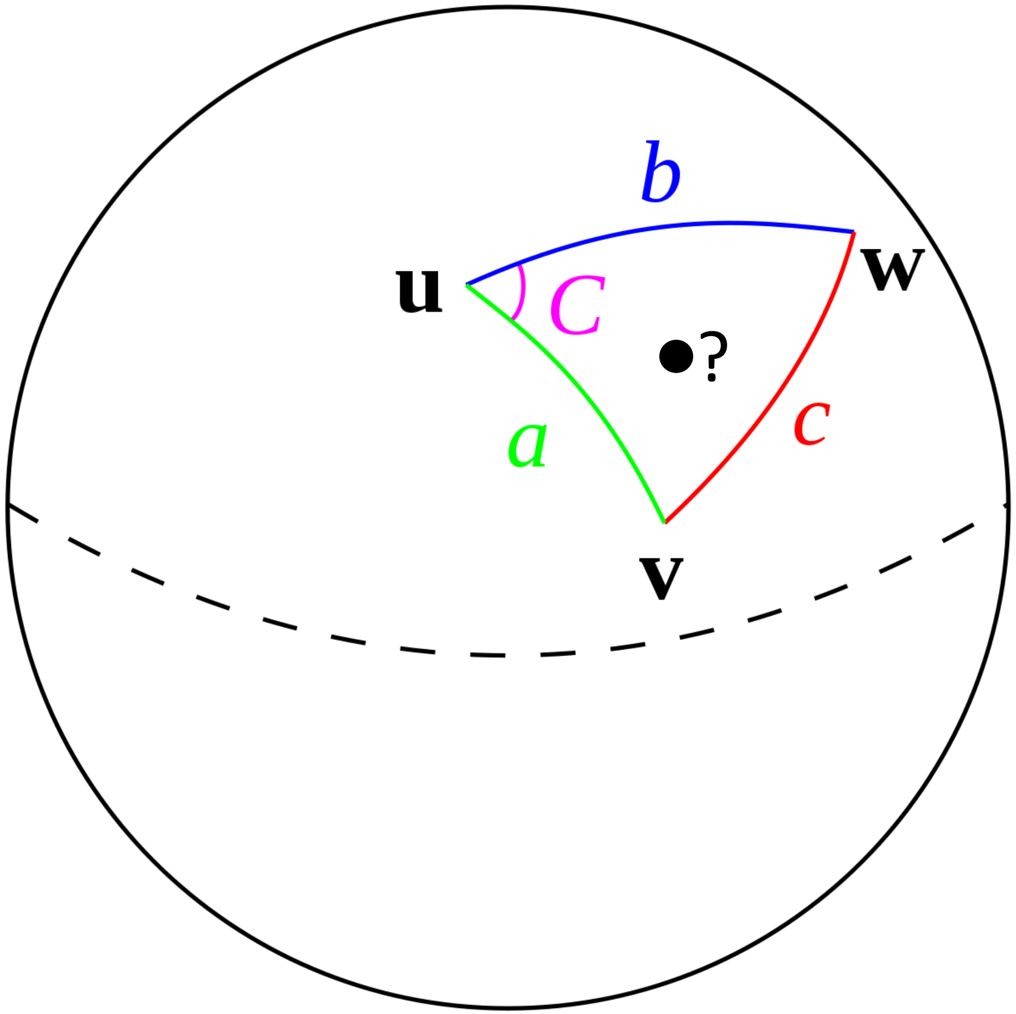
Spherical k-means

- Randomly initialize cluster centers s.t. $\alpha \in \{0°, 360°\}, \delta \in \{-90°, 90°\}$
- Assign points to nearest cluster centers
- Update cluster centers to be the "center" of their current members
- Repeat until the total measured clustering distances changes "insignificantly"

$$min \sum_{i,j} \mu_{ij}\, d_{ij}$$

$$\mu = \begin{array}{cc} \text{Classes} & \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} & \begin{array}{l}\text{Data} \\ \text{points}\end{array} \end{array}$$

Center of a cluster of spherical points?



Center of a cluster of spherical points?

$$p_j \text{ s.t. } min \sum_{i,j:\mu_{ij}=1} 1 - \langle x_i | p_j \rangle$$

$$\frac{d}{d\alpha} \sum_{i,j:\mu_{ij}=1} d_{ij} = \frac{d}{d\delta} \sum_{i,j:\mu_{ij}=1} d_{ij} = 0$$

$$\tan(\alpha) = \frac{\Sigma \mathrm{y}}{\Sigma \mathrm{x}} \text{ and } \tan(\delta) = \frac{\Sigma \mathrm{z}}{\sqrt{(\Sigma \mathrm{x})^2 + (\Sigma \mathrm{y})^2}}$$

Center of a cluster of spherical points?

$$p_j \text{ s.t. } min \sum_{i,j:\mu_{ij}=1} 1 - \langle x_i | p_j \rangle$$

$$p_j = \frac{\{\Sigma\mathrm{x},\ \Sigma\mathrm{y},\ \Sigma\mathrm{z}\}}{\sqrt{(\Sigma\mathrm{x})^2 + (\Sigma\mathrm{y})^2 + (\Sigma\mathrm{z})^2}}$$

$$p_j \propto \sum_{i:\mu_{ij}=1} x_i$$

Extended spherical k-means

- Randomly initialize cluster centers s.t. $\alpha \in \{0°, 360°\}, \delta \in \{-90°, 90°\}$
- Assign points to nearest cluster centers
- Update cluster centers (with weighted distances)
- Repeat

$$min \sum_{i,j} \mu_{ij}\, w_i\, d_{ij}$$

Classes

$$\mu = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Data points

Hornik et al. 2012

Weighted center of a cluster of spherical points?

$$p_j \text{ s.t. } min \sum_{i,j:\mu_{ij}=1} w_i \left(1 - \langle x_i | p_j \rangle\right)$$

$$p_j \propto \sum_{i:\mu_{ij}=1} w_i \, x_i$$

just normalize and we are good!

Hornik et al. 2012

Online spherical k-means

- Randomly initialize cluster centers s.t. $\alpha \in \{0°, 360°\}, \delta \in \{-90°, 90°\}$
- Assign points to nearest cluster centers
- For each data point, update its cluster center s.t. $p_{j,t+1} \propto p_{j,t} + \eta_t \, x_i$
- Repeat until tolerance is reached

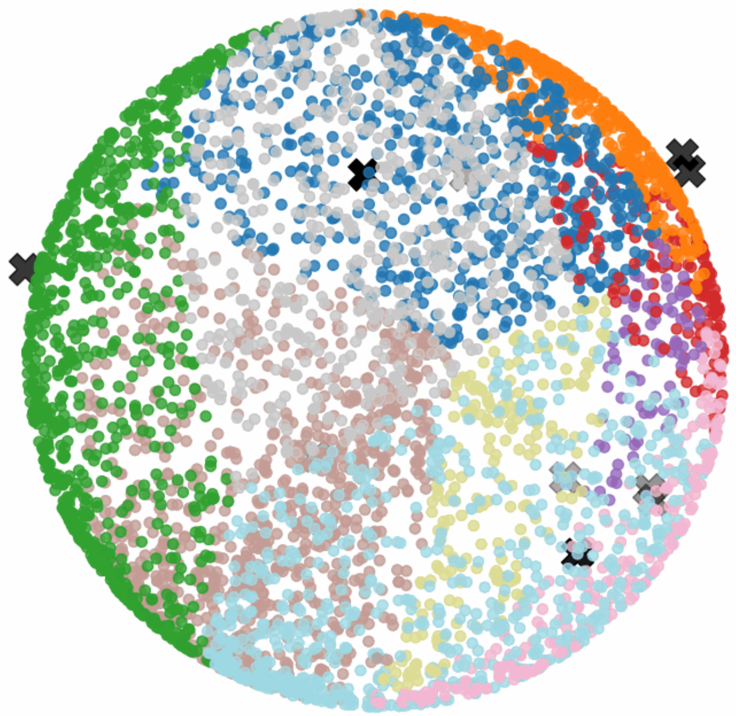
$$\eta_t = C$$

$$\eta_t = \frac{1}{|cluster\ size|}$$

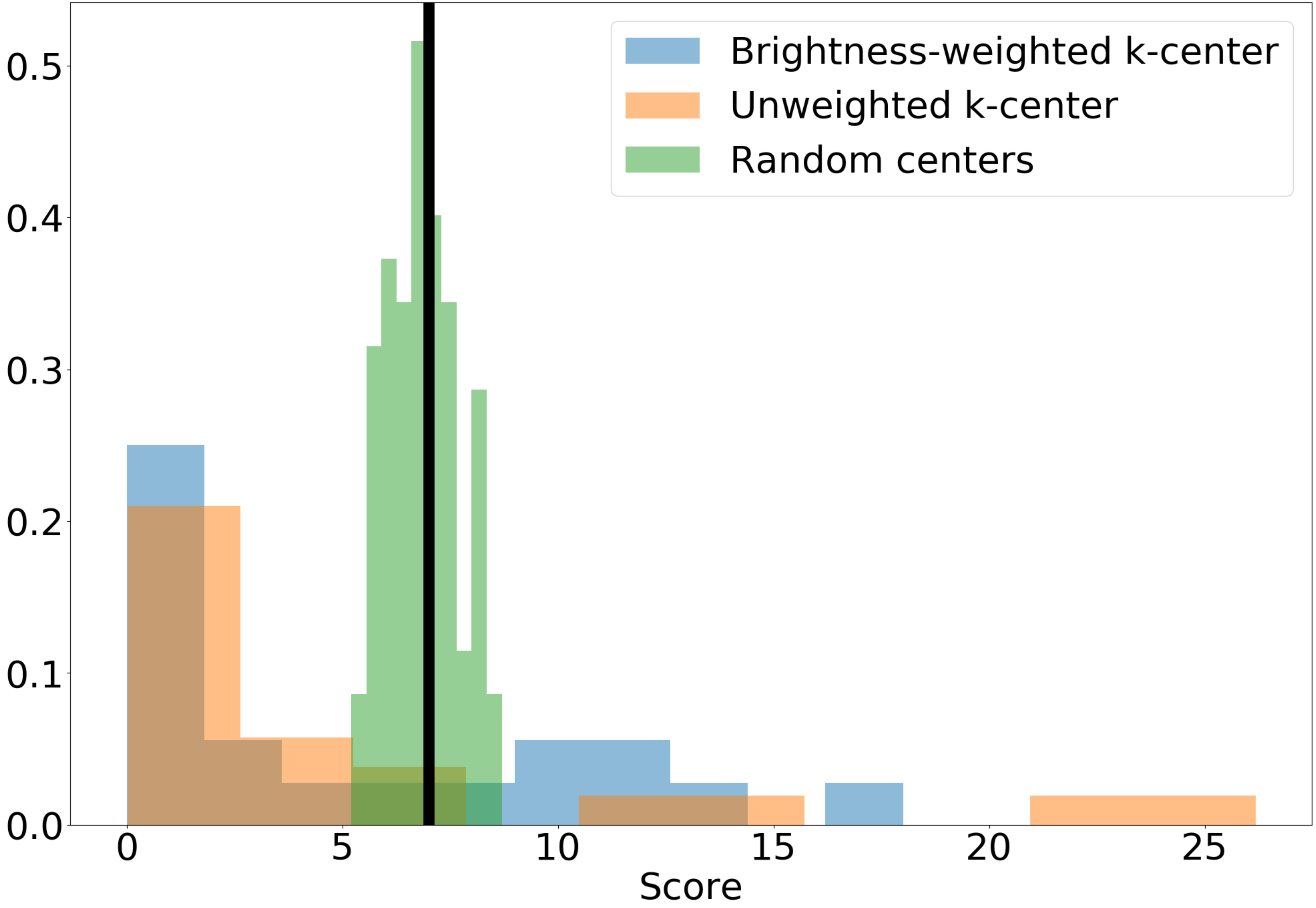
$$\eta_t = \eta_0 \left(\frac{\eta_f}{\eta_0}\right)^{\frac{t}{N\ M}}$$

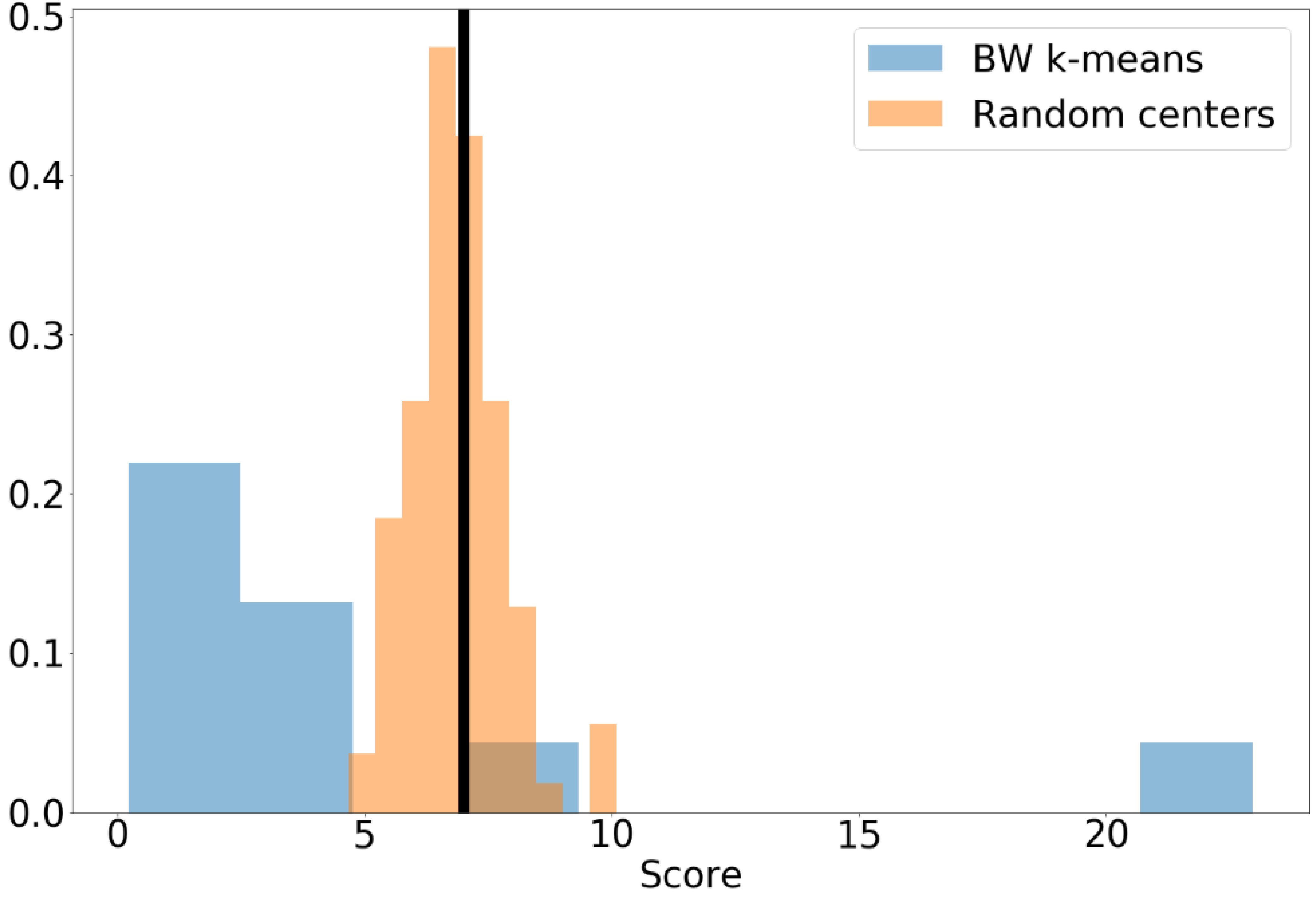
Zhong 2005

K-means clustering with 10 constellations



Results

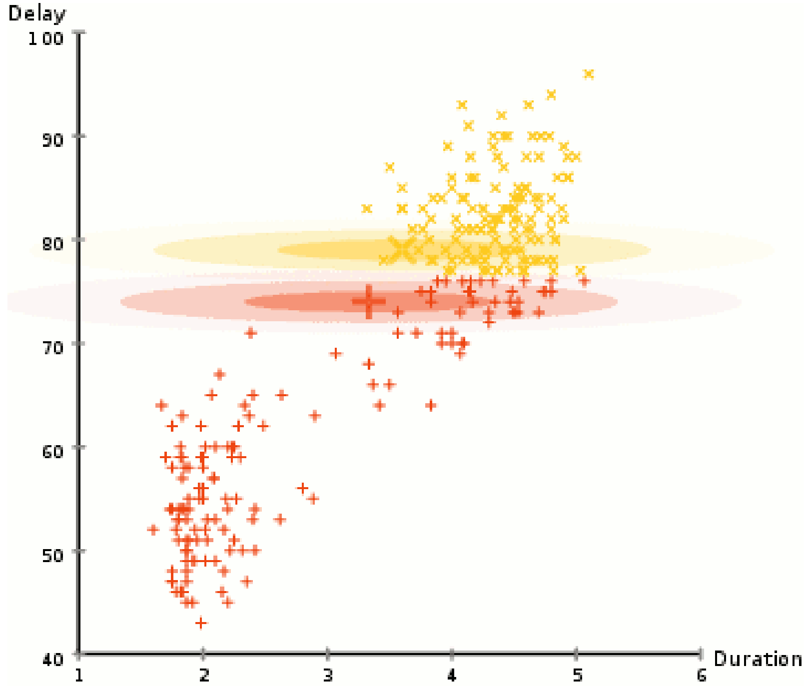

Brightness-weighted k-center
Unweighted k-center
Random centers
0.0
0.1
0.2
0.3
0.4
0.5
0
5
10
15
20
25
Score

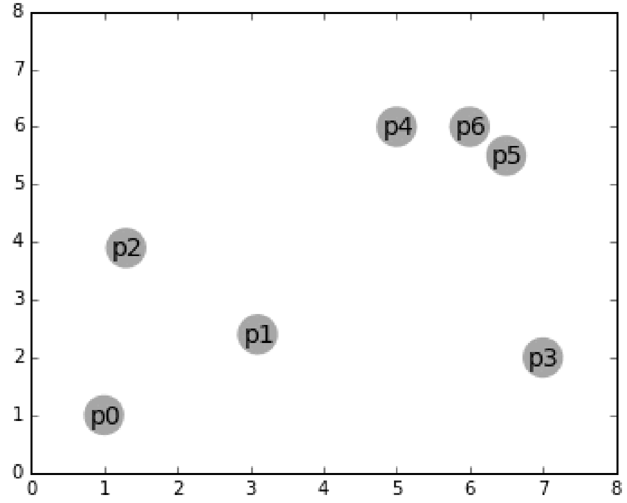


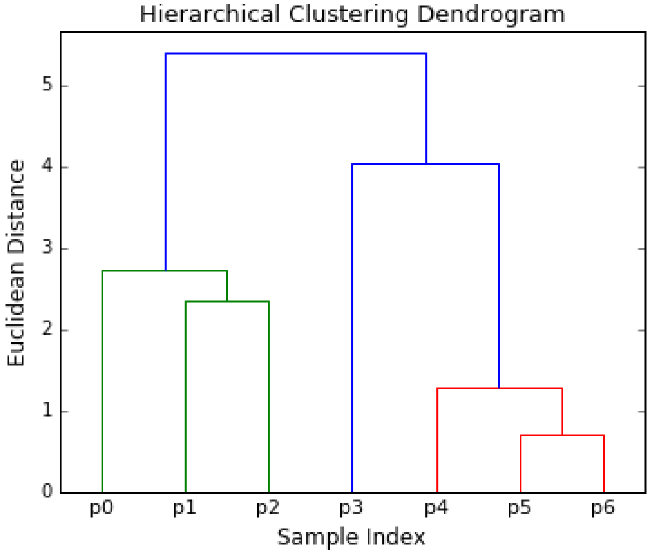
BW k-means
Random centers
0.0
0.1
0.2
0.3
0.4
0.5
0
5
10
15
20
Score

Future work

- More OSKM learning rates
- Additional clustering methods (e.g. hierarchical clustering (SLINK, CLINK), Gaussian mixture expectation maximization)
- Intuitive visualizations
- A machine learns to name constellations?
- Interactive web applet?







References

- Buchta, C., Kober, M., Feinerer, I., & Hornik, K. (2012). Spherical k-means clustering. *Journal of Statistical Software*, *50*(10), 1-22.
- Straub, J. Bayesian Inference with the von-Mises-Fisher Distribution in 3D.