	Hierarchical Clustering
au Sob	Hierarchical alestering is an unsupervised machine
V	leaving algorithm used to group data points into
Syrid 250	clusters based on similarity. It seeks to build
3. 3 313	a hierarchy of clusters. Hierarchical clustering does
Was	not require us to pre-specify the number of clusters.
	There are two main types of hierarchical chutering-
(1)	Agglomerative Hierarchical Clustering Bottom-upapproach
1.1.4	Agglomerative clustering is a bottom-up approach
g last ti	where each data pointé starte as its own cluster,
5 1	and pairs of clusters are merged step by step based
204	on distance metric until all points are merged into
7	One Clustered at the Stephen Travalued Constitution
TŲ.	The state of the s
(1)- 2	Assign each point to its own cluster.
(6)	Calculate the distance between all pairs of clueters.
(d)	Recombute the distance believes.
(e)	Recompute the distances between the remaining clusters. Repeat steps (c) & (d) until all points are merged into a
(2)	Divisive Hierarchical Clustering (Tob-Down abbund
	To a single a lister To
	assure mency untileach data hourt
	and Consider
(0)	Steps Involved:
(d)	Assign all data points to a single cluster. Split the cluster into la
(C)	pared on distance in the
	Contain only one data point or
*	Linkage Criteria: Todesile l
	Linkage Criteria: To decide how to compute distances between two clusters, different linkage criteria can be used. These define how to linkage criteria can
	clusters is calculated, influencing the final shape of
	Jan State of



	Page 02
	hierarchy. The main types are
IAI	Single linkage. The minimum distance between points
The state of the s	in each cluster.
(8)	hierarchy: The main types are Single Linkage: The minimum distance between points in each cluster. Complete Linkage: The maximum distance between points in two clusters. Average Linkage: The average distance between all points in two clusters.
	Accrose Linkage: The average die tours had
_0	in two clusters.
(0)	Centroid Linkage: The distance between the centraid
	(entroid Linkage: The distance between the centroids (mean points) of two clusters.
(E)	Ward's Method: I method that minimizes the variance
	of the clusters being merged, leading to a compact and
	ipherical cluster. This is the default method in scikit-law
**	Dendogram: The result of hierarchical chestering can
	be defit visualized using a dendogram, a tree-like.
	diagram that records the sequence of cluster merges or
	splits. The ventical axis represents the distance or
	dissimilarity between clusters, while the horizontal axis
	shows the data points
	In the dendogram:
>	The height of the merge represents the distance between
Wa	the clusters.
<u>→</u>	A cut-off threshold can be applied to divide the dendogram
	A cut-off threshold can be applied to divide the dendogram into clusters by cutting the tree at a certain level.
جهز ٢	BRE A
	Distance lines
_	192 tinez
	P ₂
	P1 P2 P3 P4 P5 Points
	In this example, we are taking
	5 random points from P1 toPs and using agglomenative
	Tf ourse threshold is such that it cuts line 1, then we will
	have 2 cluster (P. D. P. D. I. I. P. Il and I at 17 and we
	have 2 clusters (P1,P2,P3,P4) and P5, if we cut at line 2, we will have 3 clusters [(P1,P2), (P2,P4) & P5]

12-	Date Page 03
-	
12.	Advantages of Hierarchical Chustering: Unlike K-Means, we don't have to pre-define the number
(a)	Unlike K-Means, we don't have to pre-define the number
	of clusters to be assigned.
(b)	Dendogram provides a visual representation of the
	data's hierarchical structure
(c)	It works well with clusters of different shapes and
	It works well with clusters of different shapes and sizes and densities without assuring spherical clusters. Dis advantages:
3/2/03	Dis advantages:
	It is computationally expensive (usually O(n3)) and
22.12.120	doesn't ecale well for large datasets.
161	Que a marca accellities had a sit could be a
	Once a merge or split is made, it can't be undone.
	Noise and outliers can affect the structure of
	hierarchy significantly.
- 5c 5c	Look of magabash of private har it many fifel and
	com retails for example the square of the stander many side
	Lieber The resting resident river half the following to the first following
y : = 1. 1	o - i wad all the state of the
	_ atal at all and a sugar
	: more yet with attention
To the standard	The Reight of the manger represents to the his or
-	Assatis With
	with a history of british and you blanks well the houter so
41.07	I sideson to said all pritted and estado of in
	Table 1
· · · · · · · · · · · · · · · · · · ·	
i will to	Sand 501 29
5 mil	192 192
me to a	
	3 it remakes the bar Tot of 1
	of all the country in the country of the control of the country of
0)	and by a first and
and the second	and the desired of the state of
	Falso of the state

classmate