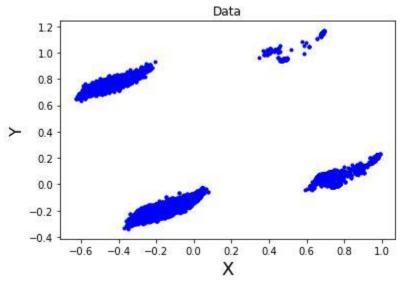
# **HR Project**

## **Clustering and Outliers Analysis:**

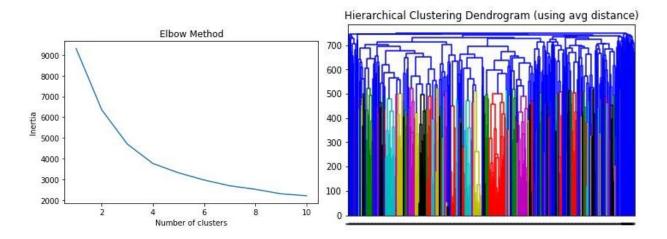
## **Clustering Analysis:**

In clustering analysis, the most important thing to decide is, how many optimal number of clusters are in data. It can either be decided by visualizing data or elbow method.

For **visualizing data**, it was converted into **2-dimensional data using PCA technique**, and from the figure below, it can be clearly seen that number of optimal clusters for data is 4.



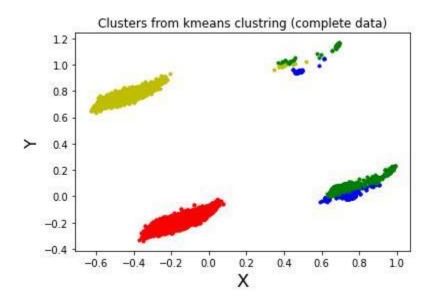
Other techniques like elbow technique and visualization of hierarchical tree of the data, also showed the same results. The figures of both techniques mentioned earlier are given below.



Graphs, data statistics and results of the clusters, made using different clustering techniques are shown below:

## K-means Clustering:

As, number of optimal number of clusters are known, we ran K-means clustering on data and got following results:



It is visible that clusters are well formed but there are some outliers in them.

Now let us analyze the data by grouping it on base of K-means labels.

	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	salary	left
count	1632.000000	1632.000000	1632.000000	1632.000000	1632.000000	1632.000000	1632.000000	1632.000000	1632.0
mean	0.415025	0.520705	2.077819	147.297181	3.051471	0.046569	0.009191	0.419118	1.0
std	0.070913	0.063042	0.392384	20.823436	0.342813	0.210778	0.095458	0.555516	0.0
min	0.100000	0.450000	2.000000	126.000000	2.000000	0.000000	0.000000	0.000000	1.0
25%	0.380000	0.480000	2.000000	135.000000	3.000000	0.000000	0.000000	0.000000	1.0
50%	0.410000	0.510000	2.000000	145.000000	3.000000	0.000000	0.000000	0.000000	1.0
75%	0.440000	0.550000	2.000000	154.000000	3.000000	0.000000	0.000000	1.000000	1.0
max	0.890000	1.000000	6.000000	301.000000	6.000000	1.000000	1.000000	2.000000	1.0

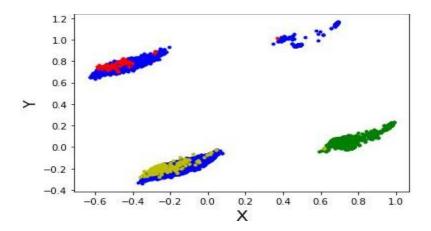
	${\tt satisfaction\_level}$	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5year	rs salaı	y left
count	1927.000000	1927.000000	1927.000000	1927.000000	1927.000000	1927.000000	1927.00000	00 1927.00000	0 1927.0
mean	0.459175	0.884759	5.359107	258.280228	4.570835	0.042034	0.0020	76 0.41100	1.0
std	0.350847	0.090415	1.027528	30.319334	0.776097	0.200719	0.0455	25 0.52184	12 0.0
min	0.090000	0.450000	2.000000	132.000000	2.000000	0.000000	0.0000	0.0000	00 1.0
25%	0.100000	0.840000	5.000000	243.000000	4.000000	0.000000	0.0000	0.0000	00 1.0
50%	0.460000	0.890000	5.000000	259.000000	5.000000	0.000000	0.0000	0.0000	00 1.0
75%	0.810000	0.950000	6.000000	278.000000	5.000000	0.000000	0.0000	00 1.00000	00 1.0
max	0.920000	1.000000	7.000000	310.000000	6.000000	1.000000	1.00000	2.00000	00 1.0
s	atisfaction_level	last_evaluation r	number_project a	verage_montly_hours ti	ime_spend_company Wo	ork_accident p	romotion_last_5years	salary	left
count	2012.000000	2012.000000	2012.000000	2012.000000	2012.000000	2012.0	2012.000000 2	2012.000000 2	012.000000
mean	0.666059	0.713415	3.791252	199.488569	3.479125	1.0	0.035785	0.622266	0.005964
std	0.218819	0.163720	0.997841	45.903097	1.695362	0.0	0.185801	0.652777	0.077017
min	0.120000	0.360000	2.000000	96.000000	2.000000	1.0	0.000000	0.000000	0.000000
25%	0.540000	0.580000	3.000000	161.000000	2.000000	1.0	0.000000	0.000000	0.000000
50%	0.680000	0.720000	4.000000	199.000000	3.000000	1.0	0.000000	1.000000	0.000000
75%	0.830000	0.850000	4.000000	238.000000	4.000000	1.0	0.000000	1.000000	0.000000
max	1.000000	1.000000	6.000000	287.000000	10.000000	1.0	1.000000	2.000000	1.000000
	${\tt satisfaction\_level}$	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5year	s salar	y left
count	9428.000000	9428.000000	9428.000000	9428.000000	9428.000000	9428.0	9428.00000	00 9428.00000	0 9428.0
mean	0.667122	0.716027	3.786169	198.991090	3.360416	0.0	0.02418	0.65676	7 0.0
std	0.216710	0.161592	0.975619	45.624039	1.531858	0.0	0.15362	0.65555	7 0.0
min	0.120000	0.360000	2.000000	96.000000	2.000000	0.0	0.00000	0.00000	0.0
25%	0.540000	0.580000	3.000000	162.000000	2.000000	0.0	0.00000	0.00000	0.0
50%	0.690000	0.720000	4.000000	198.000000	3.000000	0.0	0.00000	00 1.00000	0.0
75%	0.840000	0.850000	4.000000	238.000000	4.000000	0.0	0.00000	1.00000	0.0
max	1.000000	1.000000	6.000000	287.000000	10.000000	0.0	1.00000	2.00000	0.0

From the above statistics, we can define K-means clusters of data as:

- Employee with **low** satisfaction level, evaluation score, number of projects, average monthly hours and salary, who **left**,
- Employee with **low** satisfaction level and salary but **high** number of projects, average monthly hours, who **left**,
- Employee with **high** satisfaction level, evaluation score and salary but with **lower** number of projects and average monthly hours, but **no work accident** and **stayed**,
- Employee with **high** satisfaction level, evaluation score and salary but with **lower** number of projects and average monthly hours, but **had work accident** and **stayed**.

Similar results were achieved from Birch clustering technique, but **Birch** gave high importance to **promotion** and **work accident**.

Figure of clustering is given below:



It can be seen that, clusters are well formed but there are some outliers in them.

Now let us analyze the data by grouping it on base of Birch labels.

	satisfaction_level	last_evaluation	number_project	${\tt average\_montly\_hours}$	time_spend_company	Work_accident	promotion_last_5years	salar	y left	
count	3399.000000	3399.000000	3399.000000	3399.000000	3399.00000	3399.0	3399.000000	3399.00000	0 3399.0	
mean	0.439444	0.718308	3.858782	207.603119	3.87614	0.0	0.003530	0.41394	5 1.0	
std	0.263392	0.197760	1.824415	61.317451	0.97871	0.0	0.059321	0.53937	9 0.0	
min	0.090000	0.450000	2.000000	126.000000	2.00000	0.0	0.000000	0.00000	0 1.0	
25%	0.120000	0.520000	2.000000	146.000000	3.00000	0.0	0.000000	0.00000	0 1.0	
50%	0.410000	0.790000	4.000000	225.000000	4.00000	0.0	0.000000	0.00000	0 1.0	
75%	0.725000	0.900000	6.000000	262.000000	5.00000	0.0	0.000000	1.00000	0 1.0	
max	0.920000	1.000000	7.000000	310.000000	6.00000	0.0	1.000000	2.00000	0 1.0	
	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	salary	left	
count	satisfaction_level 73.000000	last_evaluation 73.000000	number_project 73.000000	average_montly_hours 73.000000	time_spend_company 73.000000	Work_accident 73.0	promotion_last_5years 73.0	73.000000	73.000000	
count	N <del>T</del> 8		770 7			(478)	t	(f.)		)
	73.000000	73.000000	73.000000	73.000000	73.000000	73.0	73.0	73.000000	73.000000	)
mean	73.000000 0.665890	73.000000 0.724521	73.000000 3.684932	73.000000 216.095890	73.000000 4.178082	73.0	73.0 1.0	73.000000 1.000000	73.000000	)
mean std	73.000000 0.665890 0.194785	73.000000 0.724521 0.206337	73.000000 3.684932 1.165198	73.000000 216.095890 47.622988	73.000000 4.178082 1.924567	73.0 1.0 0.0	73.0 1.0 0.0	73.000000 1.000000 0.645497	73.000000 0.013699 0.117041	)
mean std min	73.000000 0.665890 0.194785 0.160000	73.000000 0.724521 0.206337 0.380000	73.000000 3.684932 1.165198 2.000000	73.000000 216.095890 47.622988 100.000000	73.000000 4.178082 1.924567 2.000000	73.0 1.0 0.0 1.0	73.0 1.0 0.0 1.0	73.000000 1.000000 0.645497 0.000000	73.000000 0.013699 0.117041 0.000000	) ) )
mean std min 25%	73.000000 0.665890 0.194785 0.160000 0.570000	73.000000 0.724521 0.206337 0.380000 0.550000	73.000000 3.684932 1.165198 2.000000 3.000000	73.00000 216.095890 47.622988 100.00000 183.000000	73.000000 4.178082 1.924567 2.000000 3.000000	73.0 1.0 0.0 1.0	73.0 1.0 0.0 1.0 1.0	73.000000 1.000000 0.645497 0.000000 1.000000	73.000000 0.013699 0.117041 0.000000 0.000000	) ) ) )

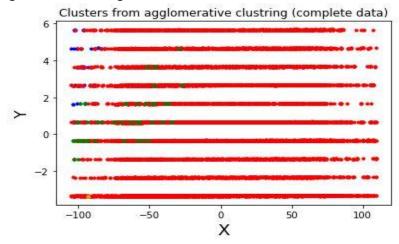
	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	salary	lef
count	11296.000000	11296.000000	11296.000000	11296.000000	11296.000000	11296.000000	11296.000000	11296.000000	11296.00000
mean	0.663461	0.715516	3.786385	199.084720	3.365262	0.185552	0.000266	0.636686	0.01487
std	0.219957	0.162406	0.994075	45.864575	1.529145	0.388762	0.016295	0.650464	0.12104
min	0.090000	0.360000	2.000000	96.000000	2.000000	0.000000	0.000000	0.000000	0.00000
25%	0.530000	0.580000	3.000000	161.000000	2.000000	0.000000	0.000000	0.000000	0.00000
50%	0.690000	0.720000	4.000000	198.000000	3.000000	0.000000	0.000000	1.000000	0.00000
75%	0.840000	0.850000	4.000000	238.000000	4.000000	0.000000	0.000000	1.000000	0.00000
max	1.000000	1.000000	7.000000	309.000000	10.000000	1.000000	1.000000	2.000000	1.00000
	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	salary	left
count		=	= 1	N = 1=		1000			
count	231.000000	231.000000	231.000000	231.000000	231.000000	231.0	) 231.0	231.000000	
	231.000000	231.000000	231.000000 3.835498	231.000000 195.995671	231.000000 4.225108	231.0	) 231.0	231.000000 1.073593	231.000000
mean	231.000000 0.671688	231.000000 0.709610 0.154915	231.000000 3.835498 0.936573	231.000000 195.995671 46.568930	231.000000 4.225108 2.286344	231.0 0.0 0.0	231.0	231.000000 1.073593 0.645213	231.000000 0.012987
mean std	231.000000 0.671688 0.209079	231.000000 0.709610 0.154915 0.370000	231.000000 3.835498 0.936573 2.000000	231.000000 195.995671 46.568930 102.000000	231.000000 4.225108 2.286344 2.000000	231.0 0.0 0.0	231.0 ) 1.0 ) 0.0 ) 1.0	231.000000 1.073593 0.645213 0.000000	231.000000 0.012987 0.113464
mean std min	231.000000 0.671688 0.209079 0.150000	231.000000 0.709610 0.154915 0.370000 0.590000	231.000000 3.835498 0.936573 2.000000 3.000000	231.000000 195.995671 46.568930 102.000000	231.000000 4.225108 2.286344 2.000000 3.000000	231.0 0.0 0.0 0.0	231.0 1.0 0.0 0.0 1.0 1.0	231.000000 1.073593 0.645213 0.000000 1.000000	231.000000 0.012987 0.113464 0.000000
mean std min 25%	231.000000 0.671688 0.209079 0.150000	231.000000 0.709610 0.154915 0.370000 0.590000 0.710000	231.000000 3.835498 0.936573 2.000000 3.000000	231.000000 195.995671 46.568930 102.000000 152.000000	231.000000 4.225108 2.286344 2.000000 3.000000	231.0 0.0 0.0 0.0 0.0	231.0 1.0 0.0 0.0 1.0 1.0 1.0	231.000000 1.073593 0.645213 0.000000 1.000000	231.000000 0.012987 0.113464 0.000000 0.000000

From the above statistics, we can define clusters from Birch of data as:

- Employee with **low** satisfaction level, evaluation score and salary. And had **no work accident** and **left**,
- Employee with **high** satisfaction level, evaluation score and salary but with **lower** average monthly hours. But had **no work accident** and **no promotion** and **stayed**,
- Employee with **high** satisfaction level, evaluation score and salary but with **lower** number of projects, average monthly hours. And **got promotion** but **had work accident** and **stayed**.
- Employee with **high** satisfaction level, evaluation score and **very high salary** and **got promotion** and **stayed**.

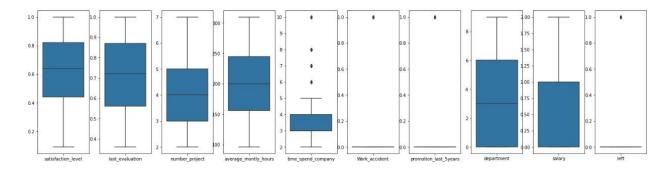
Agglomerative clustering was also tried, but no meaningful clusters were extracted from Agglomerative clustering.

Figure of clusters is given below:



## **Outliers Analysis:**

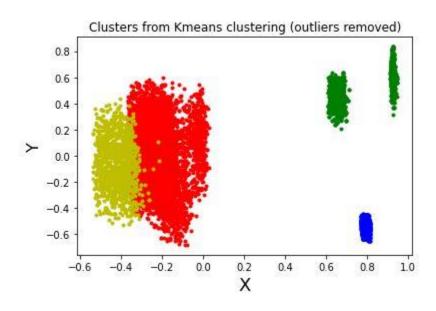
Outliers in the data depends upon the objective we are trying to achieve, like from the below figure we can see that in attribute "left" of the data, value '1' is outlier. But since our objective is to figure the reasons behind, why employee left and how can they be retained in future. So, it makes it clear that we can not declare outliers by looking on a single attribute, so we cannot use boxplot, z score or IQR for outliers analysis, since they will declare employee who left, had promotion or work accident as outliers.



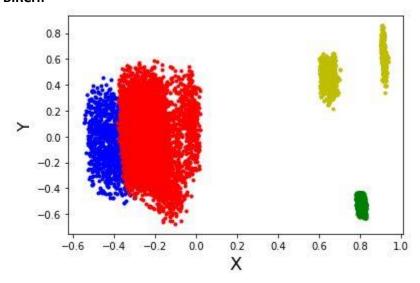
Therefore, outlier detection technique which will be useful for our objective is "Isolation Forest".

The clusters extracted after applying "Isolation Forest" are shown in the figures below:

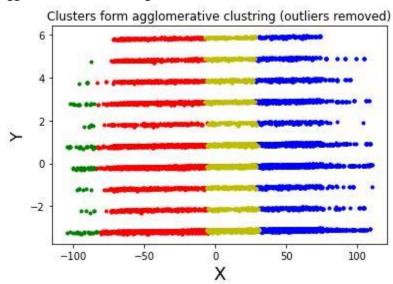
#### K-means:



#### **BIRCH:**



## **Agglomerative Clustering:**



After removal of outliers, it is visible that K-means and Birch gave very clear clusters with almost no outliers or overlapping. But agglomerative clustering were not clearly separated.

Now, let us check statistics of clusters after outlier removal.

## K-means:

After removal outliers, it can be seen the employees despite having all the problems like **low** salary, satisfaction level and high average monthly hours **did not leave** are considered as **outlier** or we can also say **exception** and **removed**.

As a result, we got perfect clusters with zero standard deviation in "left" column, which are shown below:

	${\tt satisfaction\_level}$	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	left
count	731.000000	731.000000	731.000000	731.000000	731.000000	731.0	731.0	731.0
mean	0.797177	0.879980	0.526402	0.703766	0.391245	0.0	0.0	1.0
std	0.062423	0.089779	0.096518	0.073695	0.042061	0.0	0.0	0.0
min	0.626374	0.656250	0.400000	0.565421	0.375000	0.0	0.0	1.0
25%	0.747253	0.796875	0.400000	0.644860	0.375000	0.0	0.0	1.0
50%	0.802198	0.875000	0.600000	0.700935	0.375000	0.0	0.0	1.0
75%	0.846154	0.968750	0.600000	0.761682	0.375000	0.0	0.0	1.0
max	0.912088	1.000000	0.600000	0.841121	0.500000	0.0	0.0	1.0
	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	left
count	1432.000000	1432.000000	1432.0	1432.000000	1432.000	1432.0	1432.0	1432.0
mean	0.349408	0.235815	0.0	0.222390	0.125	0.0	0.0	1.0
std	0.033114	0.056305	0.0	0.047300	0.000	0.0	0.0	0.0
min	0.296703	0.140625	0.0	0.140187	0.125	0.0	0.0	1.0
25%	0.318681	0.187500	0.0	0.182243	0.125	0.0	0.0	1.0
50%	0.351648	0.234375	0.0	0.224299	0.125	0.0	0.0	1.0
75%	0.373626	0.281250	0.0	0.261682	0.125	0.0	0.0	1.0
max	0.406593	0.328125	0.0	0.303738	0.125	0.0	0.0	1.0
	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	left
count	satisfaction_level 1566.000000	1566.000000	1566.000000	average_montly_hours 1566.000000	time_spend_company	Work_accident 1566.0	promotion_last_5years 1566.0	left 1566.0
count mean								
	1566.000000	1566.000000	1566.000000	1566.000000	1566.000000	1566.0	1566.0	1566.0
mean	1566.000000 0.717493	1566.000000 0.813388	1566.000000 0.339336	1566.000000 0.333915	1566.000000 0.122925	1566.0 0.0	1566.0 0.0	1566.0
mean std	1566.000000 0.717493 0.158684	1566.000000 0.813388 0.111796	1566.000000 0.339336 0.158501	1566.000000 0.333915 0.102939	1566.000000 0.122925 0.123577	1566.0 0.0 0.0	1566.0 0.0 0.0	1566.0 0.0 0.0
mean std min	1566.000000 0.717493 0.158684 0.340659	1566.000000 0.813388 0.111796 0.578125	1566.000000 0.339336 0.158501 0.000000	1566.00000 0.333915 0.102939 0.009346	1566.000000 0.122925 0.123577 0.000000	1566.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0
mean std min 25%	1566.000000 0.717493 0.158684 0.340659 0.593407	1566.000000 0.813388 0.111796 0.578125 0.718750	1566.000000 0.339336 0.158501 0.000000 0.200000	1566.000000 0.333915 0.102939 0.009346 0.252336	1566.000000 0.122925 0.123577 0.000000 0.000000	1566.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0
mean std min 25% 50%	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500	1566.00000 0.339336 0.158501 0.000000 0.200000 0.400000	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000	1566.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75%	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.400000	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.750000	1566.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75%	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.400000	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.750000	1566.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75% max	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000 satisfaction_level	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000 last_evaluation	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.800000 number_project	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037 average_montly_hours	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.125000 0.750000 time_spend_company	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 promotion_last_5years	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75% max	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000 satisfaction_level	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000 last_evaluation	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.400000 0.800000 number_project	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037 average_montly_hours 1077.000000	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.125000 0.750000 time_spend_company	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75% max count	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000 satisfaction_level 1077.000000 0.685645	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000 last_evaluation 1077.000000 0.549661	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.800000 number_project 1077.000000 0.340019	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037 average_montly_hours 1077.000000 0.488342	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.125000 0.750000 time_spend_company 1077.000000 0.102484	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 promotion_last_5years 1077.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75% max count mean std	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000 satisfaction_level 1077.000000 0.685645 0.160499	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000 last_evaluation 1077.000000 0.549661 0.218006	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.400000 0.800000 number_project 1077.000000 0.340019 0.140118	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037 average_montly_hours 1077.000000 0.488342 0.180474	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.125000 0.750000 time_spend_company 1077.000000 0.102484 0.081111	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 promotion_last_5years 1077.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75% max  count mean std min	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000 satisfaction_level 1077.000000 0.685645 0.160499 0.153846	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000 last_evaluation 1077.000000 0.549661 0.218006 0.140625	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.800000 number_project 1077.000000 0.340019 0.140118 0.200000	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037 average_montly_hours 1077.000000 0.488342 0.180474 0.098131	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.750000 time_spend_company 1077.000000 0.102484 0.081111 0.0000000	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 promotion_last_Syears 1077.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
mean std min 25% 50% 75% max count mean std min 25%	1566.000000 0.717493 0.158684 0.340659 0.593407 0.714286 0.846154 1.000000 satisfaction_level 1077.000000 0.685645 0.160499 0.153846 0.560440	1566.000000 0.813388 0.111796 0.578125 0.718750 0.812500 0.906250 1.000000 last_evaluation 1077.000000 0.549661 0.218006 0.140625 0.359375	1566.000000 0.339336 0.158501 0.000000 0.200000 0.400000 0.800000 number_project 1077.000000 0.340019 0.140118 0.200000 0.200000	1566.000000 0.333915 0.102939 0.009346 0.252336 0.336449 0.420561 0.528037 average_montly_hours 1077.000000 0.488342 0.180474 0.098131 0.331776	1566.000000 0.122925 0.123577 0.000000 0.000000 0.125000 0.750000 time_spend_company 1077.000000 0.102484 0.081111 0.0000000 0.0000000	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 promotion_last_5years 1077.0 0.0 0.0 0.0 0.0	1566.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

## Birch:

Similar effect of **outlier removal** was observed in Birch's cluster output, Employees who despite having similar properties decided opposite on whether to leave or not, as compare to majority, were considered **outlier or exception** and was **removed**.

Statistics of clusters are shown in figures below:

	${\tt satisfaction\_level}$	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5year	s salar	ry left
count	3399.000000	3399.000000	3399.000000	3399.000000	3399.00000	3399.0	3399.00000	0 3399.00000	00 3399.0
mean	0.439444	0.718308	3.858782	207.603119	3.87614	0.0	0.00353	0.41394	45 1.0
std	0.263392	0.197760	1.824415	61.317451	0.97871	0.0	0.05932	1 0.53937	79 0.0
min	0.090000	0.450000	2.000000	126.000000	2.00000	0.0	0.00000	0.00000	00 1.0
25%	0.120000	0.520000	2.000000	146.000000	3.00000	0.0	0.00000	0.00000	00 1.0
50%	0.410000	0.790000	4.000000	225.000000	4.00000	0.0	0.00000	0.00000	00 1.0
75%	0.725000	0.900000	6.000000	262.000000	5.00000	0.0	0.00000	0 1.00000	00 1.0
max	0.920000	1.000000	7.000000	310.000000	6.00000	0.0	1.00000	0 2.00000	00 1.0
	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	salary	left
count	73.000000	73.000000	73.000000	73.000000	73.000000	73.0	73.0	73.000000	73.000000
mean	0.665890	0.724521	3.684932	216.095890	4.178082	1.0	1.0	1.000000	0.013699
std	0.194785	0.206337	1.165198	47.622988	1.924567	0.0	0.0	0.645497	0.117041
min	0.160000	0.380000	2.000000	100.000000	2.000000	1.0	1.0	0.000000	0.000000
25%	0.570000	0.550000	3.000000	183.000000	3.000000	1.0	1.0	1.000000	0.000000
50%	0.650000	0.800000	4.000000	222.000000	3.000000	1.0	1.0	1.000000	0.000000
75%	0.790000	0.880000	4.000000	259.000000	6.000000	1.0	1.0	1.000000	0.000000
max	1.000000	1.000000	6.000000	275.000000	10.000000	1.0	1.0	2.000000	1.000000
satisfa	action_level last_e	valuation number	_project averag	e_montly_hours time_s	spend_company Work_	accident promo	tion_last_5years	salary	left
			r_project averag 96.000000	e_montly_hours time_s 11296.000000		accident promo 06.000000		salary 6.000000 112	
					11296.000000 1129		11296.000000 1129		
	11296.000000 112	96.000000 1129	96.000000	11296.000000	11296.000000 1129 3.365262	96.000000	11296.000000 1129 0.000266	6.000000 112	296.000000
	11296.000000 112 0.663461	296.000000 1129 0.715516	96.000000 3.786385	11296.000000 199.084720	11296.000000 1129 3.365262	06.000000 0.185552	11296.000000 1129 0.000266 0.016295	6.000000 112 0.636686	0.014873
	11296.000000 112 0.663461 0.219957	96.000000 1129 0.715516 0.162406	96.000000 3.786385 0.994075	11296.000000 199.084720 45.864575	11296.000000 1129 3.365262 1.529145 2.000000	06.000000 0.185552 0.388762	11296.000000 1129 0.000266 0.016295 0.000000	6.000000 112 0.636686 0.650464	0.014873 0.121048
	11296.000000 112 0.663461 0.219957 0.090000	0.715516 0.162406 0.360000	96.000000 3.786385 0.994075 2.000000	11296.000000 199.084720 45.864575 96.000000	11296.000000 1129 3.365262 1.529145 2.000000 2.000000	0.185552 0.388762 0.000000	11296.000000 1129 0.000266 0.016295 0.000000 0.000000	6.000000 112 0.636686 0.650464 0.000000	0.014873 0.121048 0.000000
	11296.000000 112 0.663461 0.219957 0.090000 0.530000	96.000000 1129 0.715516 0.162406 0.360000 0.580000	96.000000 3.786385 0.994075 2.000000 3.000000	11296.000000 199.084720 45.864575 96.000000 161.000000	11296.000000 1129 3.365262 1.529145 2.000000 2.000000	06.000000 0.185552 0.388762 0.000000 0.000000	11296.000000 1129 0.000266 0.016295 0.000000 0.000000	6.000000 112 0.636686 0.650464 0.000000	0.014873 0.121048 0.000000 0.000000
	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000	96.00000 3.786385 0.994075 2.000000 3.000000	11296.000000 199.084720 45.864575 96.000000 161.000000	11296.000000 1129 3,365262 1,529145 2,000000 2,000000 3,000000	0.000000 0.185552 0.388762 0.000000 0.000000	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 0.000000	6.000000 112 0.636686 0.650464 0.000000 0.000000	0.014873 0.121048 0.000000 0.000000 0.000000 0.000000 1.000000
	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 0.840000 1.000000	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 7.000000	11296.000000 199.084720 45.864575 96.000000 161.000000 198.0000000 238.000000	11296.000000 1129 3.365262 1.529145 2.000000 2.000000 4.000000 10.000000	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 0.000000	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 0.000000 1.000000	6,000000 112 0,636686 0,650464 0,000000 0,000000 1,000000	0.014873 0.121048 0.000000 0.000000 0.000000 0.000000
	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 0.840000 1.000000	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 7.000000	11296.000000 199.084720 45.864575 96.000000 161.000000 198.000000 238.000000 309.000000	11296.000000 1129 3.365262 1.529145 2.000000 2.000000 4.000000 10.000000	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 0.000000	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 1.000000 promotion_last_5years	6.000000 112 0.636686 0.650464 0.000000 1.000000 1.000000 2.000000	0.014873 0.121048 0.000000 0.000000 0.000000 0.000000 1.000000 left
	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 0.840000 1.0000000 satisfaction_level	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000 last_evaluation	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 4.000000 7.000000 number_project	11296.000000 199.084720 45.864575 96.000000 161.000000 198.000000 238.000000 309.000000 average_montly_hours	11296.000000 1129 3.365262 1.529145 2.000000 2.000000 4.000000 4.000000 10.0000000 time_spend_company	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 0.000000 1.000000 Work_accident	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 1.000000 promotion_last_5years	6.000000 112 0.636686 0.650464 0.000000 0.000000 1.000000 1.000000 2.000000 salary	0.000000 0.014873 0.121048 0.000000 0.000000 0.000000 0.000000 1.000000 left
count mean std	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 0.840000 1.000000 satisfaction_level 231.000000 0.671688 0.209079	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000 last_evaluation	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 7.000000 number_project 231.000000 3.835498 0.936573	11296.000000 199.084720 45.864575 96.000000 161.000000 238.000000 309.000000 average_montly_hours 231.000000 195.995671 46.568930	11296.000000 1129 3,365262 1,529145 2,000000 2,000000 3,000000 4,000000 10,000000 time_spend_company 231,000000 4,225108 2,286344	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 1.000000 Work_accident 231.0 0.0	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 1.000000 promotion_last_5years 231.0 1.0 0.0	6.000000 112 0.636686 0.650464 0.000000 0.000000 1.000000 1.000000 2.000000 salary 231.000000 1.073593 0.645213	296.000000 0.014873 0.121048 0.000000 0.000000 0.000000 1.000000 1.000000 1.000000 1.000000 0.012987 0.113464
count mean std min	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 1.000000 satisfaction_level 231.000000 0.671688 0.209079 0.150000	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000 last_evaluation 231.000000 0.709610 0.154915 0.370000	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 4.000000 7.000000 number_project 231.000000 3.835498 0.936573 2.000000	11296.000000 199.084720 45.864575 96.000000 161.000000 198.0000000 309.0000000 average_montly_hours 231.000000 195.995671 46.568930 102.000000	11296.000000 1129 3,365262 1,529145 2,000000 2,000000 3,000000 4,000000 10,000000 time_spend_company 231,000000 4,225108 2,286344 2,000000	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 1.000000 Work_accident 231.0 0.0 0.0	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 1.000000 promotion_last_5years 231.0 1.0 0.0	6.000000 112 0.636686 0.650464 0.000000 0.000000 1.000000 1.000000 2.000000 salary 231.000000 1.073593 0.645213 0.000000	296.000000 0.014873 0.121048 0.000000 0.000000 0.000000 1.000000 1.000000 1.000000 1.000000 1.000000 0.012987 0.113464 0.000000
count mean std min 25%	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 1.000000 satisfaction_level 231.000000 0.671688 0.209079 0.150000 0.530000	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000 last_evaluation 231.000000 0.709610 0.154915 0.370000 0.590000	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 4.000000 7.000000 number_project 231.000000 3.835498 0.936573 2.000000 3.000000	11296.000000 199.084720 45.864575 96.000000 161.000000 198.0000000 309.000000 average_montly_hours 231.000000 195.995671 46.568930 102.000000 152.000000	11296.000000 1129 3.365262 1.529145 2.000000 2.000000 3.000000 4.000000 10.000000 time_spend_company 231.000000 4.225108 2.286344 2.0000000 3.0000000	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 1.000000 Work_accident 231.0 0.0 0.0 0.0	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 1.000000 promotion_last_5years 231.0 1.0 0.0 1.0	6.000000 112 0.636686 0.650464 0.000000 1.000000 1.000000 2.000000 3alary 231.000000 1.073593 0.645213 0.000000 1.000000	296.000000 0.014873 0.121048 0.000000 0.0000000 0.0000000 1.0000000 1eft 231.000000 0.012987 0.113464 0.000000 0.0000000
count mean std min 25% 50%	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 1.000000 satisfaction_level 231.000000 0.671688 0.209079 0.150000 0.530000 0.720000	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000 1ast_evaluation 231.000000 0.709610 0.154915 0.370000 0.590000 0.710000	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 7.000000 number_project 231.000000 3.835498 0.936573 2.000000 3.000000 4.000000 4.000000	11296.000000 199.084720 45.864575 96.000000 161.000000 198.000000 309.000000 average_montly_hours 231.000000 195.995671 46.568930 102.000000 152.0000000 196.0000000	11296.000000 1129 3.365262 1.529145 2.000000 2.000000 3.000000 4.000000 10.0000000 time_spend_company 231.000000 4.225108 2.286344 2.000000 3.000000 3.000000	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 1.000000 Work_accident 231.0 0.0 0.0 0.0 0.0	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 1.000000 promotion_last_5years 231.0 1.0 0.0 1.0 1.0 1.0	6.000000 112 0.636686 0.650464 0.000000 0.000000 1.000000 1.000000 salary 231.000000 1.073593 0.645213 0.000000 1.000000 1.000000	296.000000 0.014873 0.121048 0.000000 0.000000 0.000000 1.000000 1eft 231.000000 0.012987 0.113464 0.000000 0.000000 0.000000
count mean std min 25%	11296.000000 112 0.663461 0.219957 0.090000 0.530000 0.690000 1.000000 satisfaction_level 231.000000 0.671688 0.209079 0.150000 0.530000	96.000000 1129 0.715516 0.162406 0.360000 0.580000 0.720000 0.850000 1.000000 last_evaluation 231.000000 0.709610 0.154915 0.370000 0.590000	96.000000 3.786385 0.994075 2.000000 3.000000 4.000000 4.000000 7.000000 number_project 231.000000 3.835498 0.936573 2.000000 3.000000	11296.000000 199.084720 45.864575 96.000000 161.000000 198.0000000 309.000000 average_montly_hours 231.000000 195.995671 46.568930 102.000000 152.000000	11296.000000 1129 3.365262 1.529145 2.000000 2.000000 3.000000 4.000000 10.000000 time_spend_company 231.000000 4.225108 2.286344 2.0000000 3.0000000	0.000000 0.185552 0.388762 0.000000 0.000000 0.000000 1.000000 Work_accident 231.0 0.0 0.0 0.0	11296.000000 1129 0.000266 0.016295 0.000000 0.000000 0.000000 1.000000 promotion_last_5years 231.0 1.0 0.0 1.0	6.000000 112 0.636686 0.650464 0.000000 1.000000 1.000000 2.000000 3alary 231.000000 1.073593 0.645213 0.000000 1.000000	296.000000 0.014873 0.121048 0.000000 0.000000 0.0000000 1.0000000 1eft 231.000000 0.012987 0.113464 0.000000 0.0000000

## **Agglomerative Clustering:**

Even after the removal of outliers, no meaningful clusters (for the objective) were obtained using agglomerative clusters.