

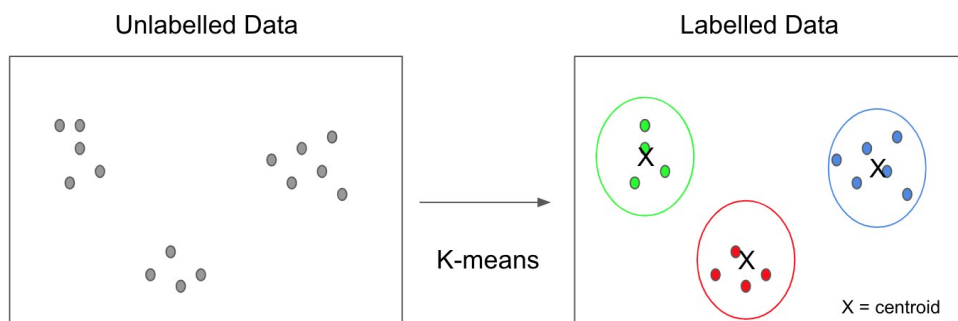
## Práctica No. 9

### Algoritmo K-Means

K-Means es un algoritmo de agrupamiento (clustering), no supervisado, que tiene como objetivo la partición de un conjunto  $n$  instancias u observaciones en  $K$  grupos en el que cada instancia pertenece al grupo cuyo valor medio es más cercano.

#### Cómo funciona

- El primer paso es definir en cuántos  $K$  grupos se dividirán las observaciones.
- Una vez definido  $K$ , se eligen  $K$  puntos al azar que serán los centroides iniciales.
- Cada instancia se asigna al centroide que esté más cerca, para ello se calcula la distancia entre cada instancia y cada centroide, asignándose al clúster cuyo centroide tenga la menor distancia.
- Se recalculan los centroides con base en la media de las instancias del clúster.
- Se repite (c) y (d).



El módulo **scikit-learn** es un módulo para aprendizaje de máquina para lenguaje Python, incluye una implementación de K-Means.

```
class sklearn.cluster.KMeans(n_clusters=8, *, init='k-means++', n_init='warn', max_iter=300, tol=0.0001, verbose=0, random_state=None, copy_x=True, algorithm='lloyd')
```

Puede consultar los parámetros del constructor en:

<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>

Asimismo, puede ver los métodos más importantes como *fit* y *predict*,

## Ejercicio:

Compilar los datos de 500 equipos de futbol con base en la página:

<https://projects.fivethirtyeight.com/soccer-predictions/global-club-rankings/>

RANK	1-WEEK CHANGE	TEAM	LEAGUE	LEAGUE COUNTRY	OFF.	DEF.
1		 <b>Man. City</b>	 Premier League	England	2.8	0.3
2		 <b>Bayern Munich</b>	 Bundesliga	Germany	3.0	0.7
3		 <b>Barcelona</b>	 La Liga	Spain	2.4	0.4
4		 <b>Real Madrid</b>	 La Liga	Spain	2.6	0.6
5		 <b>Liverpool</b>	 Premier League	England	2.6	0.7
6		 <b>Arsenal</b>	 Premier League	England	2.5	0.6

Construya su conjunto con los atributos:

- Equipo
- Calificación Ofensiva (goles que se espera que anote en una cancha neutral)
- Calificación defensiva (goles que se espera que reciba en una cancha neutral)

Equipo	Cal_Ofensiva	Cal_Defensiva
Manchester City	2.8	0.3
Bayern Munich	3.0	0.7
Barcelona	2.4	0.4

Elabore su informe que incluya:

Introducción

Desarrollo (incluya gráficas de codo y dispersión)

Resultados

Discusión de resultados

Conclusiones

Suba su informe a la plataforma en formato **PDF**.

## Ejemplo:

Conjunto de datos de entrada

Nombre	Salario	Experiencia
Yadira	2.3	1.5
Yael	3	1.5
Yafa	2.6	1.75
Yaira	2.45	2
Yaiza	2.65	2.25
Yakira	2.7	2

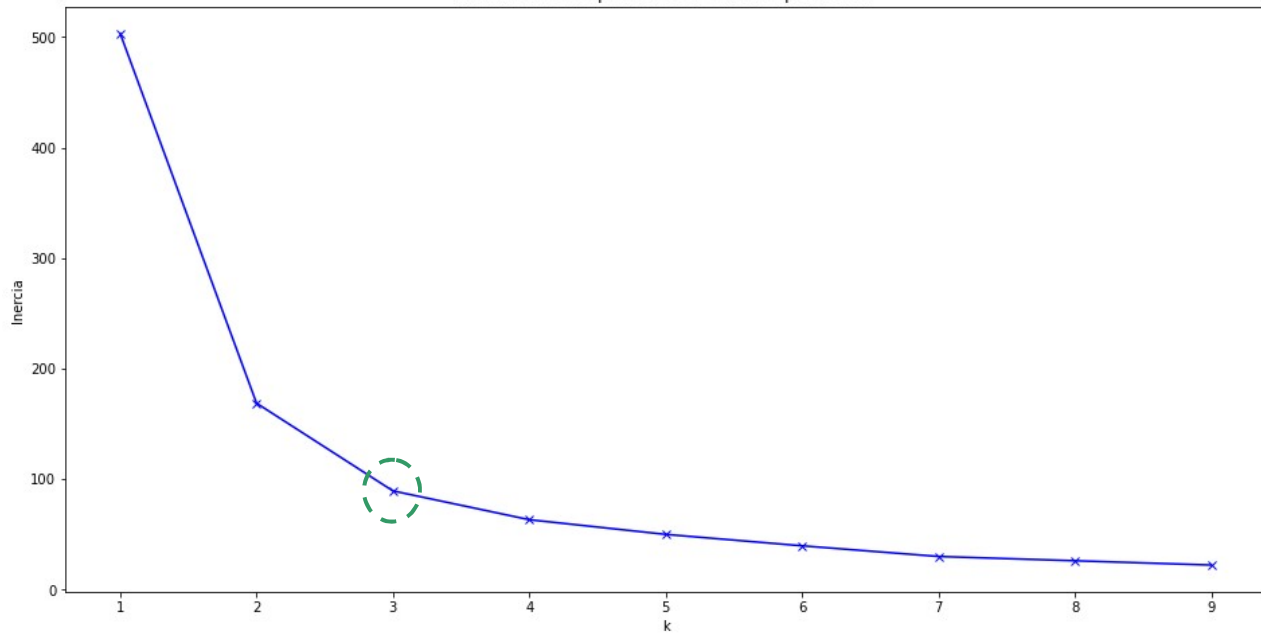
Yalens	2.7	2.25
Yalid	3.5	2
Yalika	2.8	2.5
Yalitza	3.1	2.5
Yamal	3.5	2.5
Yamanik	4	2.5
Yamelia	1.8	2.75
Yamila	2.5	2.75
Yamileth	3.7	2.75
Yamina	3.9	2.75

Yanai	4	3
Yanara	4.2	3.25
Yanel	4.9	3
Yaneli	6	3
Yanela	1.5	3.5
Yaneth	2.6	3.5
Yanil	3.2	3.5
Yanina	3.85	3.5
Yankiray	4.8	3.75
Yannel	5.2	3.75
Yannely	4	4.5
Yannick	5.3	4.5
Yannis	4.3	4.75
Yaquelina	4.7	4.75
Yarah	4.6	5.75
Yarazeth	5.7	5.75
Yarden	5.4	6
Yarely	6.3	6.25
Yaretzi	5.7	6.75
Yari	6.8	7
Variela	6.35	7.25
Yaritza	6.6	7.25
Yarmila	5.8	7.5
Yashila	6.2	7.5
Yashira	2.6	1.15
Yasmin	3.3	1.5
Yasmina	2.3	1.75
Yasna	2.7	2
Yasu	2.65	2.25
Yatzary	2.61	2
Yatzil	2.8	2.25
Yatziri	2.9	2
Yaxha	3.1	2.5
Yedda	2.8	2.5
Yei	3.9	2.5
Yeira	4.2	2.5
Yelitza	2	2.75
Yeni	2.6	2.75
Yerma	3.3	2.75
Yeruti	3.5	2.75
Yesenia	3.8	3
Yessica	4.4	3.25
Yetlanezi	4.7	3

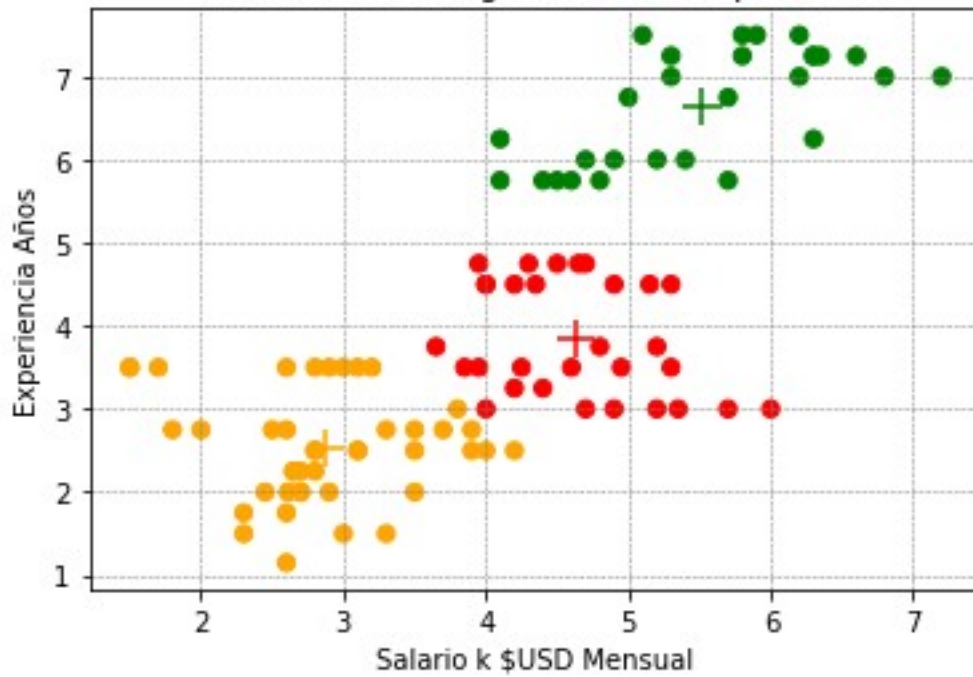
## Resultados:

Yetzali	5.2	3
Yetzye	1.7	3.5
Yexalen	2.8	3.5
Yeva	3	3.5
Yeyetzi	3.95	3.5
Yilda	4.6	3.5
Yildiz	5.3	3.5
Yin	4.2	4.5
Ylenia	4.9	4.5
Yoana	4	4.5
Yohualli	4.5	4.75
Yoi	4.8	5.75
Yoki	4.5	5.75
Yoko	4.9	6
Yola	4.1	6.25
Yolanda	5	6.75
Yolima	6.2	7
Yoliztli	5.8	7.25
Yolotzin	6.3	7.25
Yon	5.1	7.5
Yona	5.9	7.5
Yone	5.7	3
Yordan	5.35	3
Yordana	1.5	3.5
Yoriko	2.9	3.5
Yoshiko	3.1	3.5
Yoshimo	3.65	3.75
Yovela	4.25	3.5
Yua	4.95	3.5
Yuana	4.35	4.5
Yuga	5.15	4.5
Yuji	3.95	4.75
Yukie	4.65	4.75
Yukiko	4.1	5.75
Yoko	4.4	5.75
Yulene	4.7	6
Yulia	5.2	6
Yuliana	5.3	7
Yuna	7.2	7
Yunara	6.3	7.25
Yumei	5.3	7.25

Método del codo para mostrar el valor óptimo de K



K-Means Clustering Salarios vs Experiencia



### Datos agrupados

Nombre	Salario	Experiencia	clase_K-means
Yanai	4	3	0
Yanara	4.2	3.25	0
Yanel	4.9	3	0
Yaneli	6	3	0
Yanina	3.85	3.5	0
Yankiray	4.8	3.75	0
Yannel	5.2	3.75	0
Yannely	4	4.5	0
Yannick	5.3	4.5	0
Yannis	4.3	4.75	0
Yaquelina	4.7	4.75	0
Yessica	4.4	3.25	0
Yetlanezi	4.7	3	0
Yetzali	5.2	3	0
Yeyetzi	3.95	3.5	0
Yilda	4.6	3.5	0
Yildiz	5.3	3.5	0
Yin	4.2	4.5	0
Ylenia	4.9	4.5	0
Yoana	4	4.5	0
Yohualli	4.5	4.75	0
Yone	5.7	3	0
Yordan	5.35	3	0
Yoshimo	3.65	3.75	0
Yovela	4.25	3.5	0
Yua	4.95	3.5	0
Yuana	4.35	4.5	0
Yuga	5.15	4.5	0
Yuji	3.95	4.75	0
Yukie	4.65	4.75	0
Yadira	2.3	1.5	1
Yael	3	1.5	1
Yafa	2.6	1.75	1
Yaira	2.45	2	1
Yaiza	2.65	2.25	1
Yakira	2.7	2	1
Yalens	2.7	2.25	1
Yalid	3.5	2	1
Yalika	2.8	2.5	1
Yalitza	3.1	2.5	1
Yamal	3.5	2.5	1
Yamanik	4	2.5	1
Yamelia	1.8	2.75	1
Yamila	2.5	2.75	1
Yamileth	3.7	2.75	1
Yamina	3.9	2.75	1
Yanela	1.5	3.5	1
Yaneth	2.6	3.5	1
Yanil	3.2	3.5	1

Yashira	2.6	1.15	1
Yasmin	3.3	1.5	1
Yasmina	2.3	1.75	1
Yasna	2.7	2	1
Yasu	2.65	2.25	1
Yatzary	2.61	2	1
Yatzil	2.8	2.25	1
Yatziri	2.9	2	1
Yaxha	3.1	2.5	1
Yedda	2.8	2.5	1
Yei	3.9	2.5	1
Yeira	4.2	2.5	1
Yelitza	2	2.75	1
Yeni	2.6	2.75	1
Yerma	3.3	2.75	1
Yeruti	3.5	2.75	1
Yesenia	3.8	3	1
Yetzye	1.7	3.5	1
Yexalen	2.8	3.5	1
Yeva	3	3.5	1
Yordana	1.5	3.5	1
Yoriko	2.9	3.5	1
Yoshiko	3.1	3.5	1
Yarah	4.6	5.75	2
Yarazeth	5.7	5.75	2
Yarden	5.4	6	2
Yarely	6.3	6.25	2
Yaretzi	5.7	6.75	2
Yari	6.8	7	2
Yariela	6.35	7.25	2
Yaritza	6.6	7.25	2
Yarmila	5.8	7.5	2
Yashila	6.2	7.5	2
Yoi	4.8	5.75	2
Yoki	4.5	5.75	2
Yoko	4.9	6	2
Yola	4.1	6.25	2
Yolanda	5	6.75	2
Yolima	6.2	7	2
Yoliztli	5.8	7.25	2
Yolotzin	6.3	7.25	2
Yon	5.1	7.5	2
Yona	5.9	7.5	2
Yukiko	4.1	5.75	2
Yoko	4.4	5.75	2
Yulene	4.7	6	2
Yulia	5.2	6	2
Yuliana	5.3	7	2
Yuna	7.2	7	2
Yunara	6.3	7.25	2
Yumei	5.3	7.25	2