



```

e[b]({}))var c=function(b){this.element=a(d),d=d&&d.replace(/.*(?=#["']\S*)$/g,""),  
dropdown-menu"},d=b.data("target");if(d||!(d=b.attr("href"),d=d&&d.replace(/.*(?=#["']\S*)$/g,"")),g=a.Event("show.bs.tab",{relatedTarget:e[0]  
at a"},f=a.Event("hide.bs.tab",{relatedTarget:b[0]})),g=a.Event("show.bs.tab",{relatedTarget:e[0]}),c.prototype.activate=function(b,d,e){fun  
faultPrevented()}{var h=a(d);this.activate(b.closest("li"),c),this.activate(h,h.parent(),function()  
-trigger({type:"shown.bs.tab",relatedTarget:e[0]}))}}},c.prototype.activate=function(b,d,e){fun  
> .active").removeClass("active").end().find("[data-toggle="tab"]').attr("aria-expanded",!1),  
ia-expanded",!0),h?(b[0].offsetWidth,b.addClass("in")):b.removeClass("fade"),b.parent(".dropdown  
.find("[data-toggle="tab"]').attr("aria-expanded",!0),e&&e())}var g=d.find("> .active"),h=e&&  
e")||||d.find("> .fade").length;g.length&&h?g.one("bsTransitionEnd",f).emulateTransitionEnd  
var d=a.fn.tab,a.fn.tab=b,a.fn.tab.Constructor=c,a.fn.tab.noConflict=function(){return a.fn.t  
show"));a(document).on("click.bs.tab.data-api","[data-toggle="tab"]',e).on("click.bs.tab.data  
strict";function b(b){return this.each(function(){var d=a(this),e=d.data("bs.affix"),f="ob  
typeof b&&b[0]()))var c=function(b,d){this.options=a.extend({},c.DEFAULTS,d),this.$target=a  
null,this.pinnedOffset=null}).on("click.bs.affix.data-api",a.proxy(this.checkPositionW  
State=function(a,b,c,d){var e=this.checkPosition();c.VERSION="3.3.7",c.RESET="affix affix-top  
bottom"--this.affixed)return null!=c?!(e+this.unpin<=f.top)&&"bottom":!(e+g<=a-d)&&"bottom"  
RESET).addClass("affix");var a=this.$target.scrollTop(),b=this.$element.offset().top;if(e+g<=a-d)  
withEventLoop=function(){setTimeout(a.proxy(this.checkPosition,this),0)}},c.prototype.  
nt.height(),d=this.options.offset,e=d.top,f=d.bottom,g=this.$element.offset().top,h=this.  
eof e&&(e=d.top(this.$element.offset(),d.top(f,d)))})
```

MATRÍCULA: 1797222

INTRODUCCIÓN

El data clustering, también conocido como agrupamiento de datos, es una técnica de análisis exploratorio que tiene como objetivo identificar grupos o clústeres naturales en conjuntos de datos. Consiste en dividir un conjunto de datos en subconjuntos más pequeños y homogéneos, donde los objetos dentro de un mismo clúster son similares entre sí y difieren de los objetos en otros clústeres.

El data clustering es una técnica ampliamente utilizada en diversas disciplinas, como la minería de datos, la inteligencia artificial, la bioinformática y la investigación de mercado. Permite descubrir patrones, estructuras y relaciones ocultas en los datos, lo que facilita la comprensión y el análisis de grandes volúmenes de información.

Para esta práctica, se utilizó la herramienta de Kaggle para obtener los datos, en este caso se usó los datos de "Popular Video Oranges vs. Grapefruit", con el compilador PyCharm ejecutamos el código.

El código anterior muestra un ejemplo de implementación de data clustering utilizando el algoritmo de k-means en Python. El objetivo es agrupar los datos de ventas en clústeres basados en las variables proporcionadas en el conjunto de datos.

Se seleccionan las variables relevantes para el clustering, que en este caso son "QUANTITYORDERED" y "SALES". Estas variables se extraen del DataFrame original y se asignan a una nueva variable X que se utilizará en el algoritmo de clustering.

Luego, se instancia el algoritmo de k-means con n_clusters=3, lo que significa que se buscarán 3 clústeres en los datos. Se entrena el modelo utilizando los datos de entrada X y se obtienen las etiquetas de clúster para cada muestra utilizando la función fit_predict().



