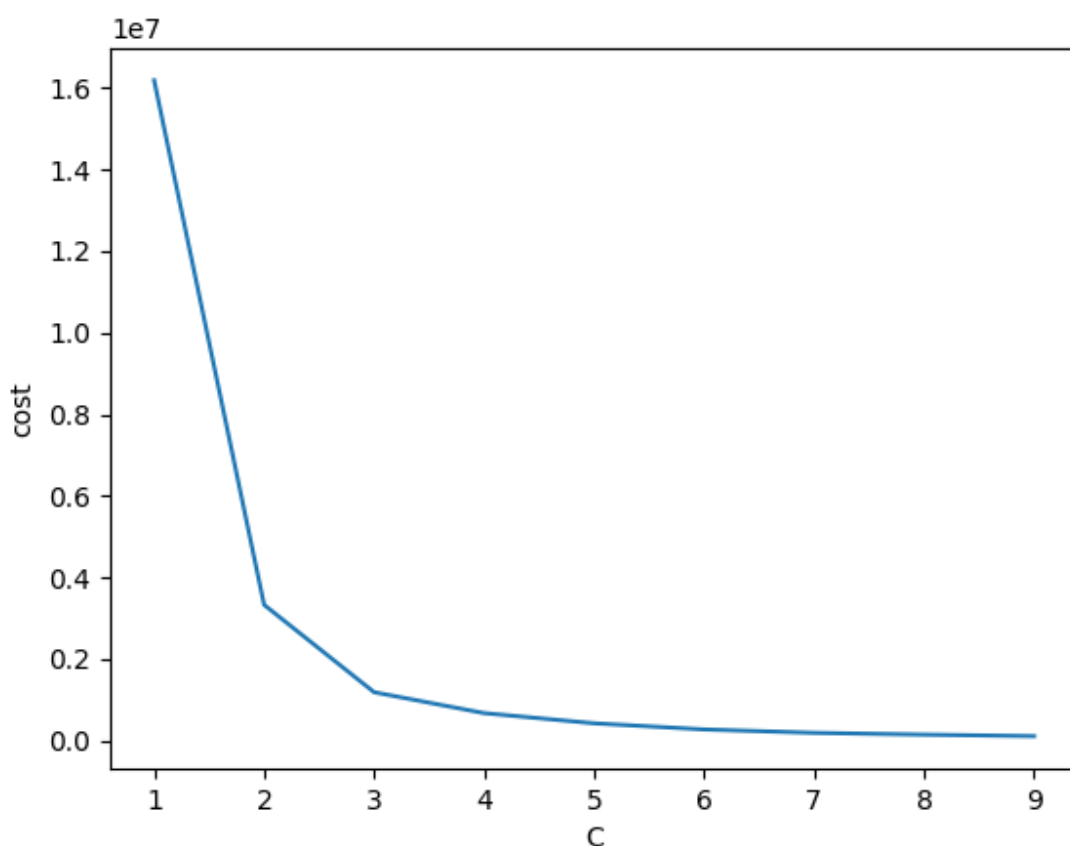


انتخاب C بهینه:

در تمام نمودارهای زیر $m=3$ در نظر گرفته شده است.

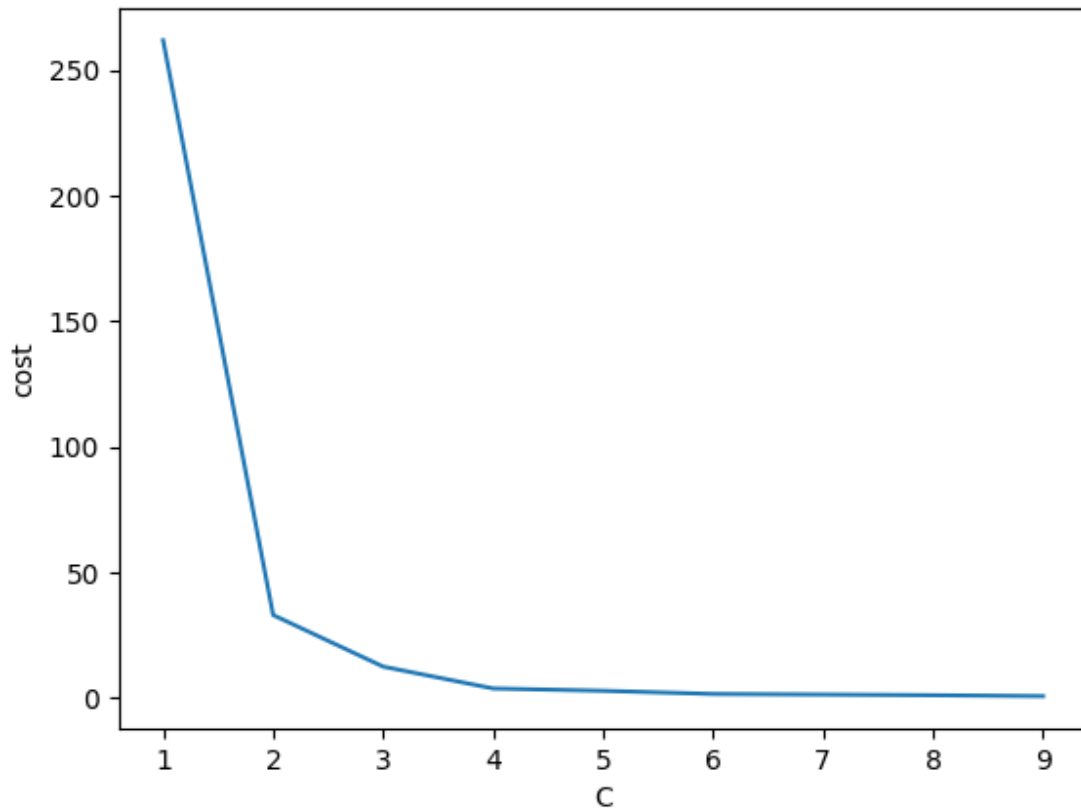
برای انتخاب C بهینه، کوچکترین C ای رو انتخاب می کنیم که خط واصل مقدار آن به مقدار $m=9$ منطبق بر باقی نمودار باشد.

● نمودار مجموعه داده اول به شرح زیر است:



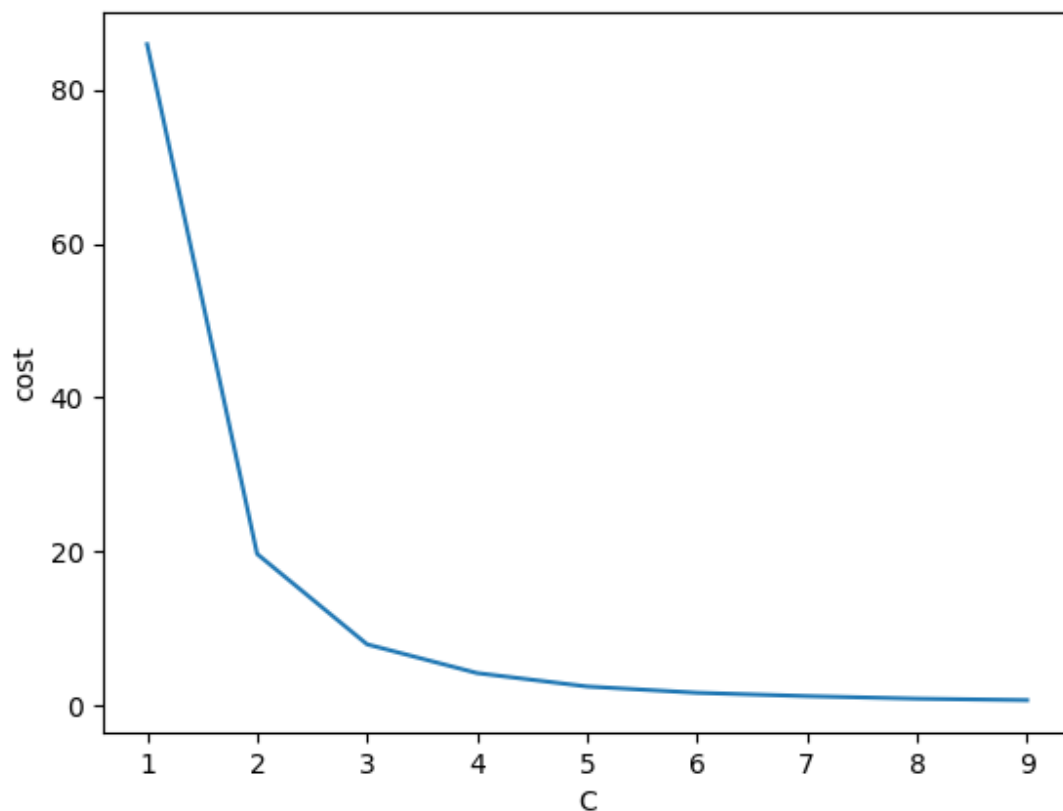
به این صورت بهترین تعداد مراکز طبق روش Elbow برابر با ۳ است.

● نمودار مجموعه داده دوم به شرح زیر است:



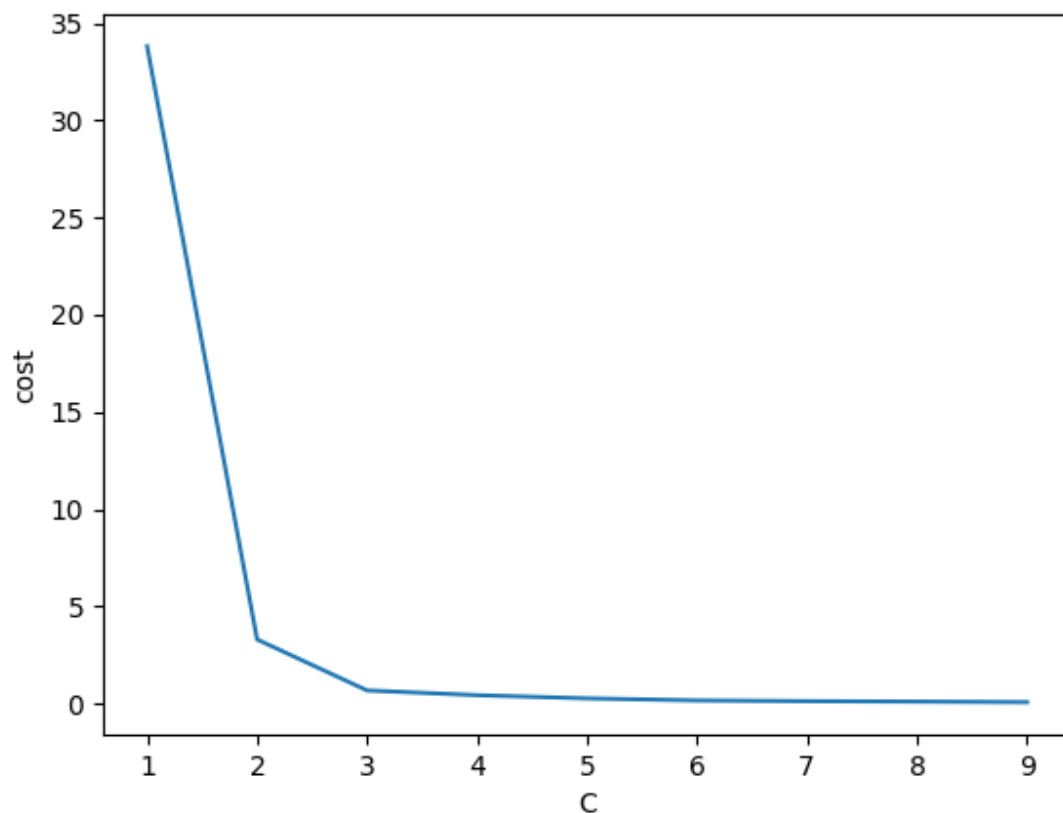
به این صورت بهترین تعداد مراکز طبق روش Elbow برابر با ۴ است.

● نمودار مجموعه داده سوم به شرح زیر است:



به این صورت بهترین تعداد مراکز طبق روش Elbow برابر با ۳ است.

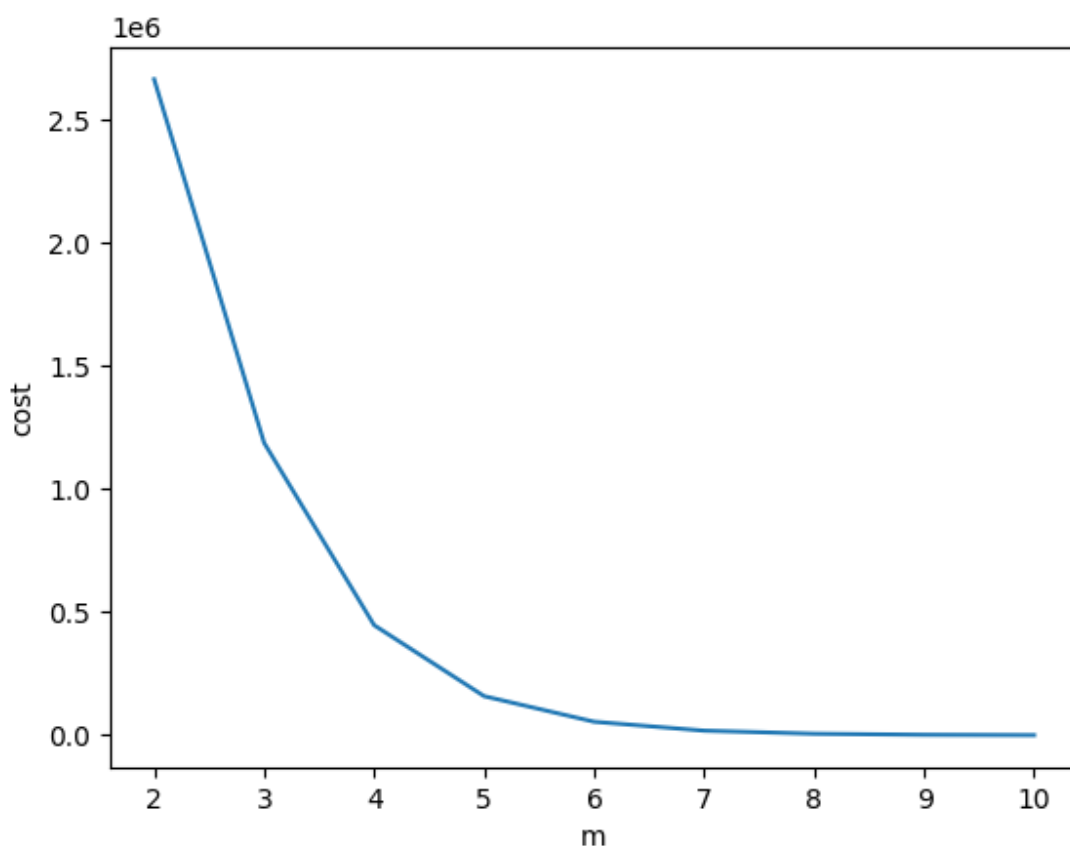
● نمودار مجموعه داده چهارم به شرح زیر است:



به این صورت بهترین تعداد مراکز طبق روش Elbow برابر با ۳ است.

بررسی تاثیر m :

ابتدا بر روی مجموعه داده‌های اول با $C=3$ ، مقدار m را از ۲ تا ۱۰ تغییر دادیم. نمودار تابع هزینه به شکل زیر شد:

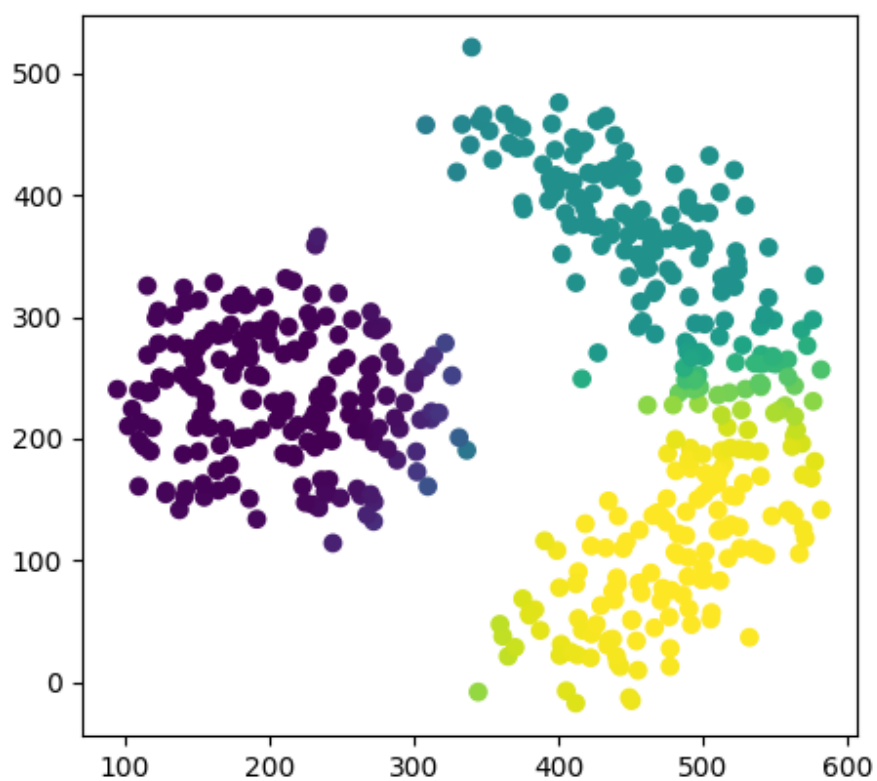


همانطور که مشاهده می‌کنید با افزایش m ، تابه هزینه نیز کاهش می‌باید.

همچنین بر روی مجموعه داده‌های اول با $C=3$ ، مقدار m را از ۲ تا ۵ تغییر داده و نقاط را بر اساس میزان تعلق آن‌ها به مراکز رنگ آمیزی کردیم.

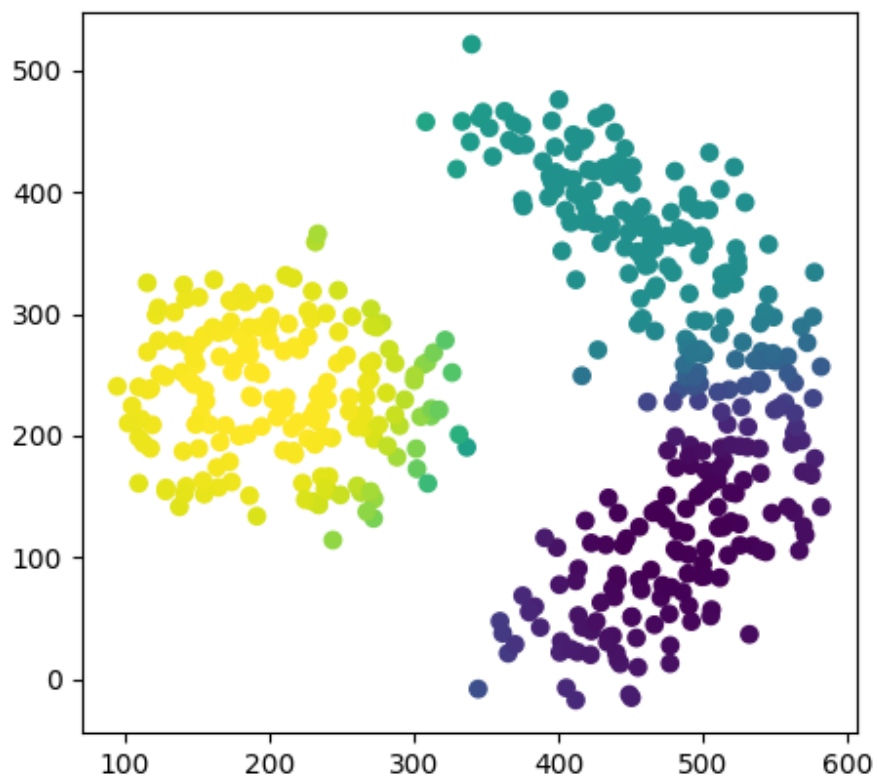
حاصل به شرح زیر است (تغییر رنگ‌های مراکز به دلیل تصادفی بودن آن‌هاست):

● مقدار m برابر با ۲:



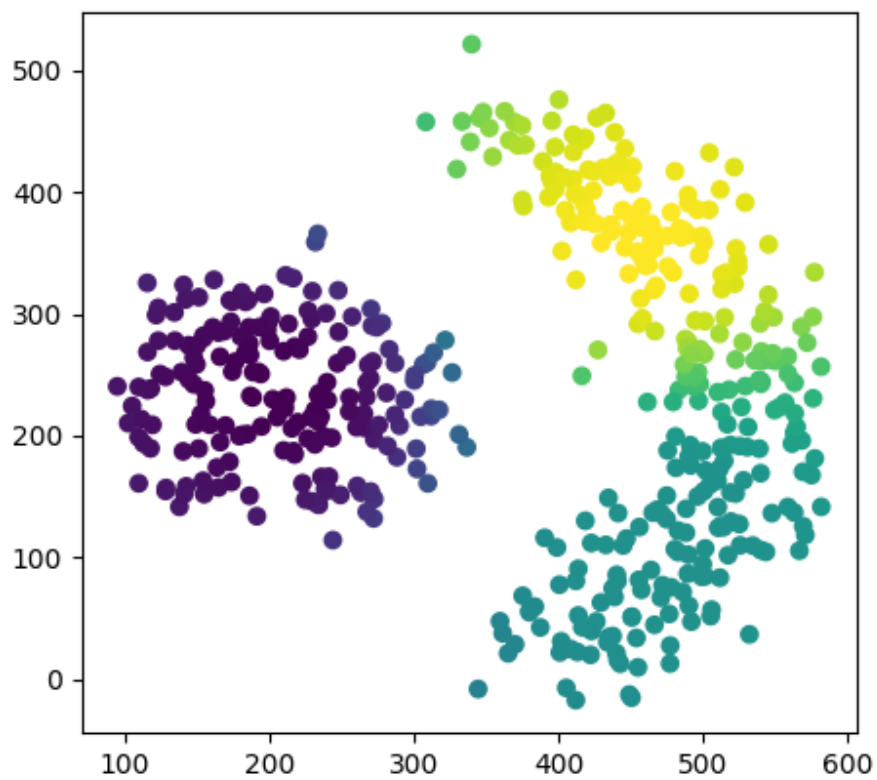
امیرحسین پاشایی هیر ۹۷۳۱۰۱۳

● مقدار m برابر با ۳:

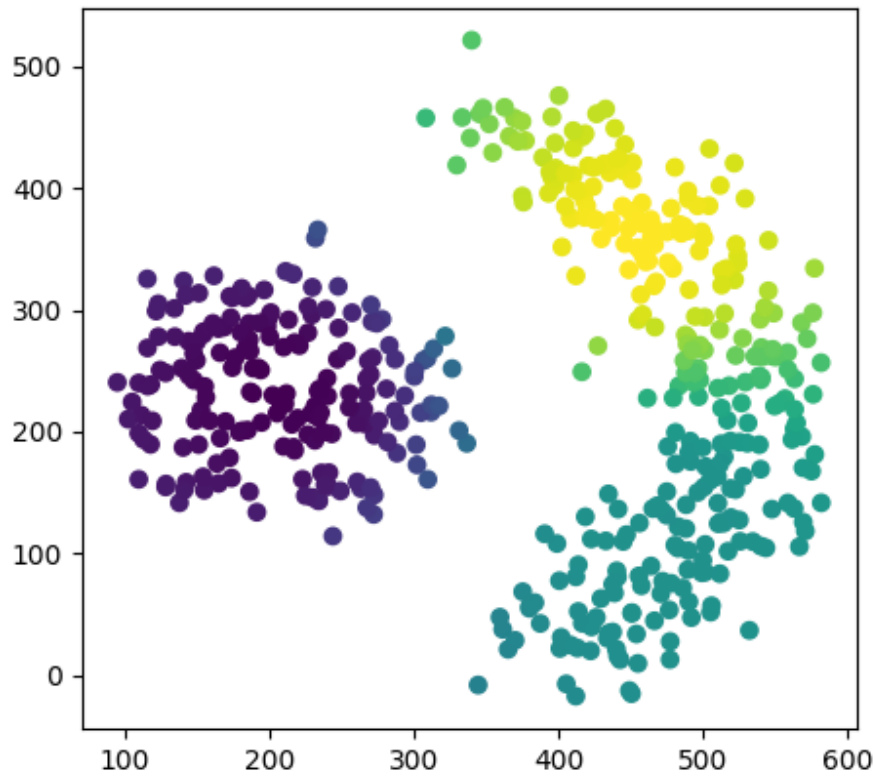


امیرحسین پاشایی هیر ۹۷۳۱۰۱۳

● مقدار m برابر با ۴:



● مقدار m برابر با ۵:



همانطور که مشاهده می کنید با افزایش مقدار m ، مرز بین خوشه ها از بین رفته و تعلق نقاط به اصطلاح فازی تر می شود.

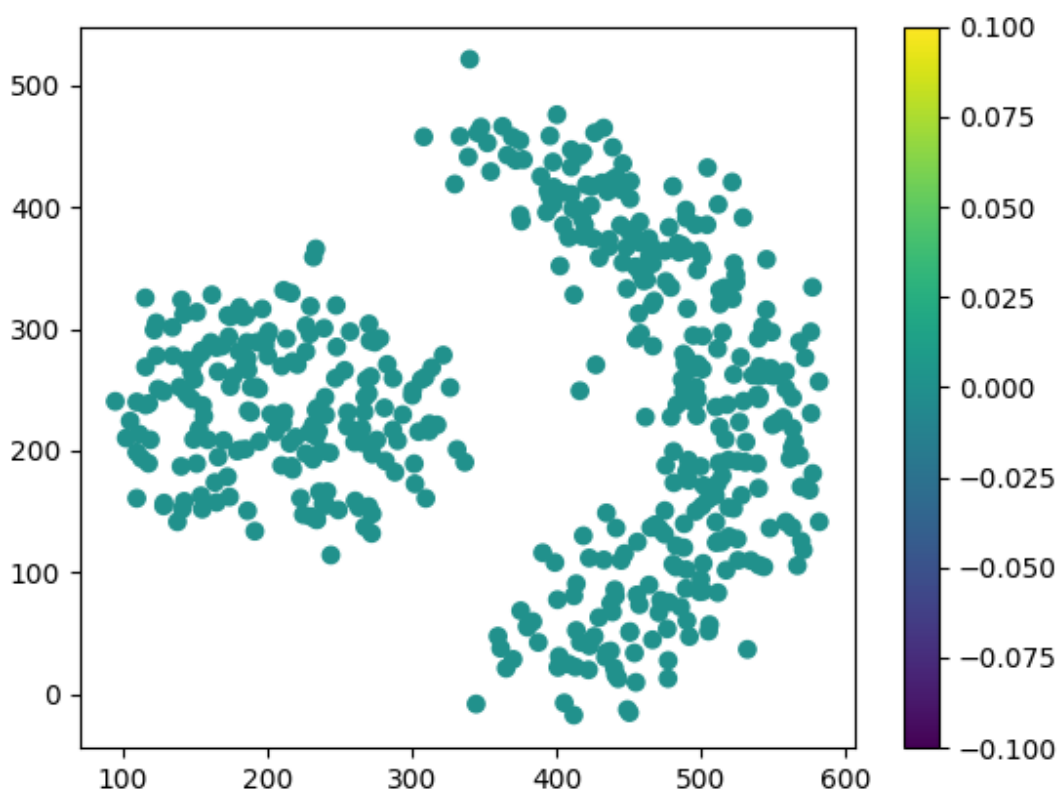
رسم خوشه بندی بر اساس C:

بر روی مجموعه داده‌های اول و با مقدار $m=3$ ، مقدار C را از ۱ تا ۵ تغییر داده و خوشی بندی را رسم

کرده‌ایم. (رنگ آمیزی به صورت crisp می‌باشد)

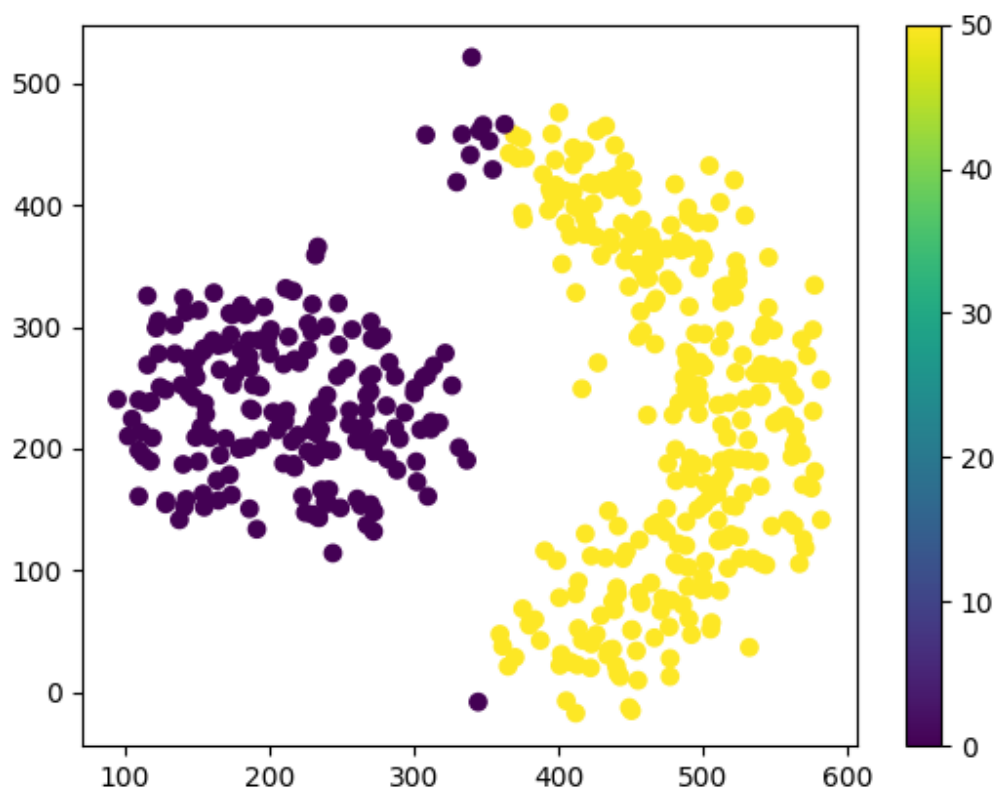
حاصل به شرح زیر است:

● مقدار C برابر با ۱:



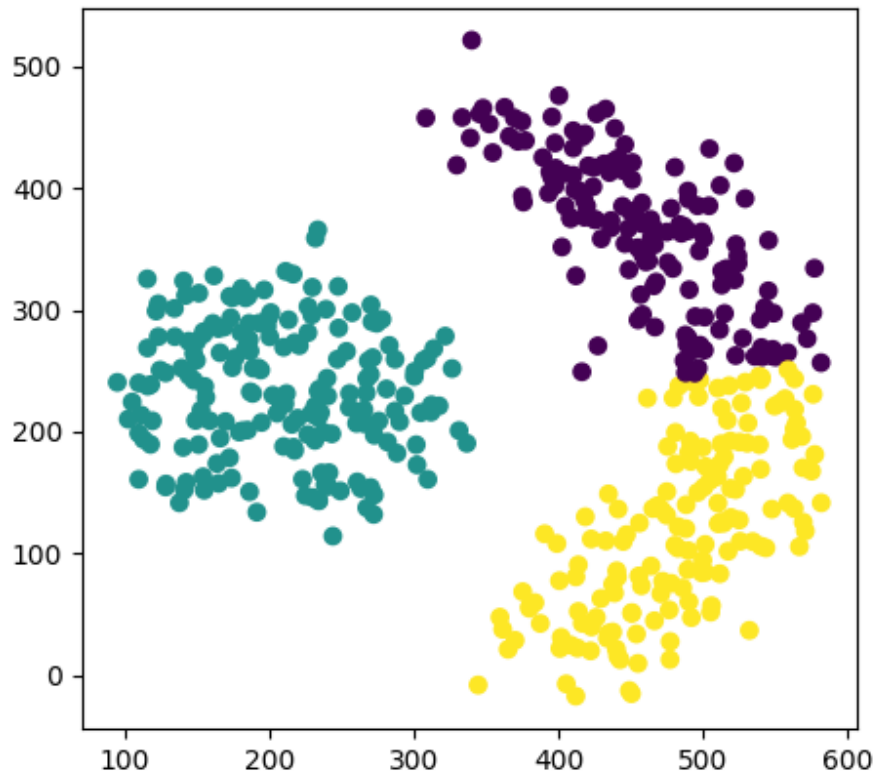
امیرحسین پاشایی هیر ۹۷۳۱۰۱۳

● مقدار C برابر با ۲:

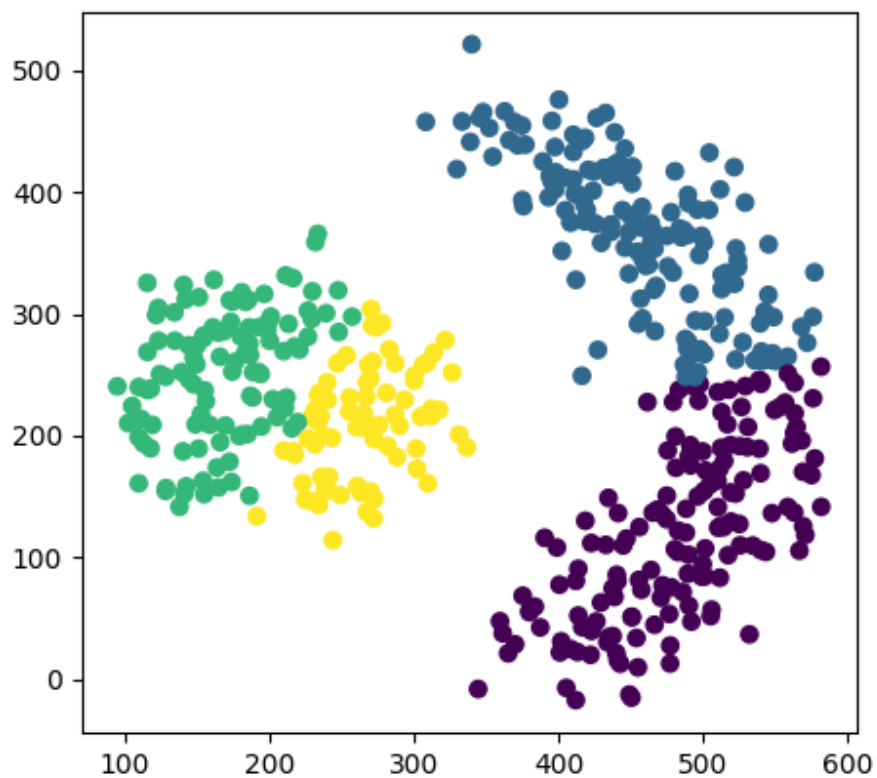


امیرحسین پاشایی هیر ۹۷۳۱۰۱۳

● مقدار C برابر با ۳:

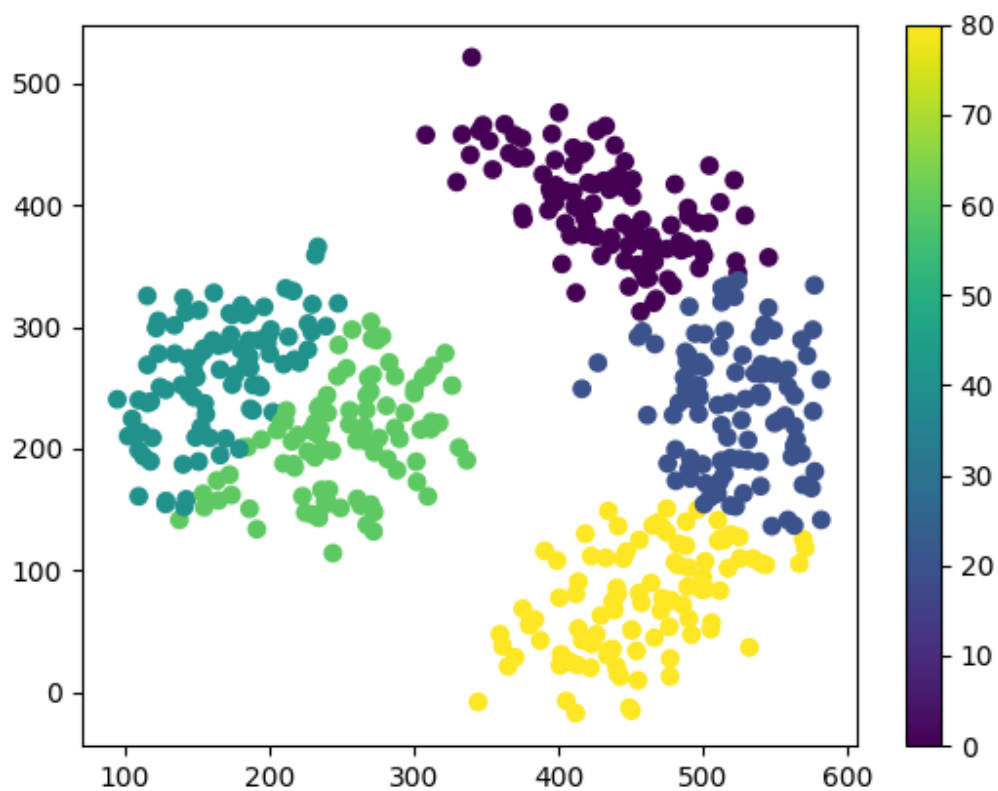


● مقدار C برابر با ۴:



امیرحسین پاشایی هیر ۹۷۳۱۰۱۳

● مقدار C برابر با ۵:



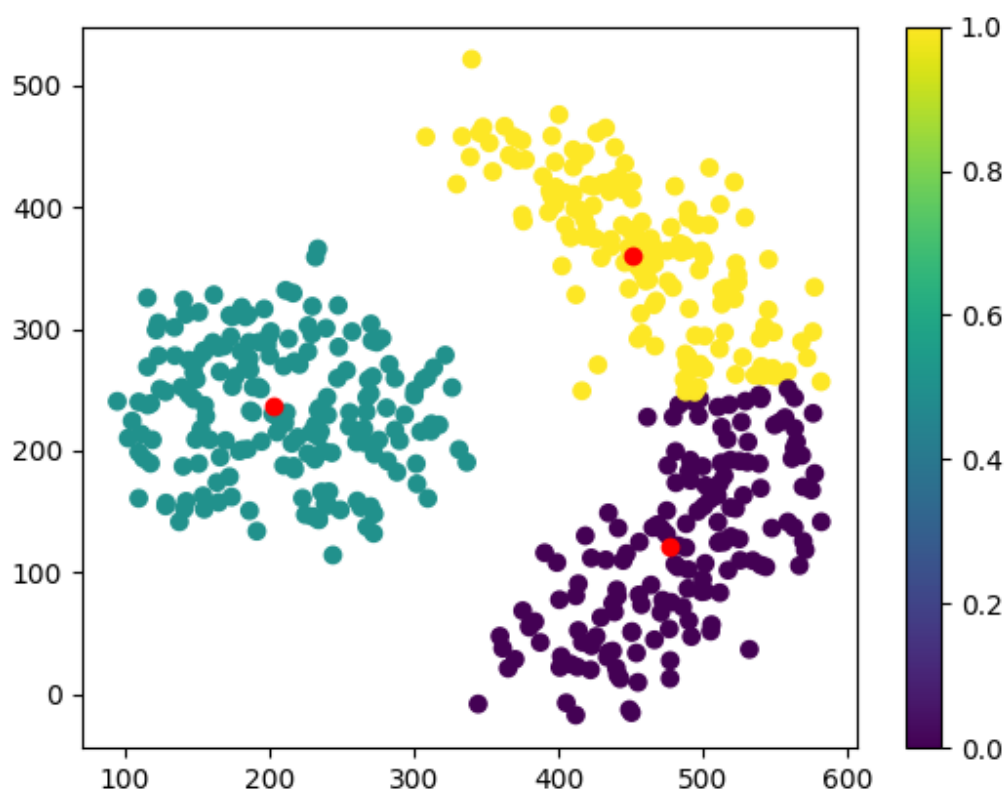
رسم خوشه بندی بر اساس C بهینه هر داده:

برای هر مجموعه داده (دو بعدی) خوشه بندی آن را براساس C بهینه آن رسم کرده ایم. (رنگ آمیزی به

صورت crisp می باشد)

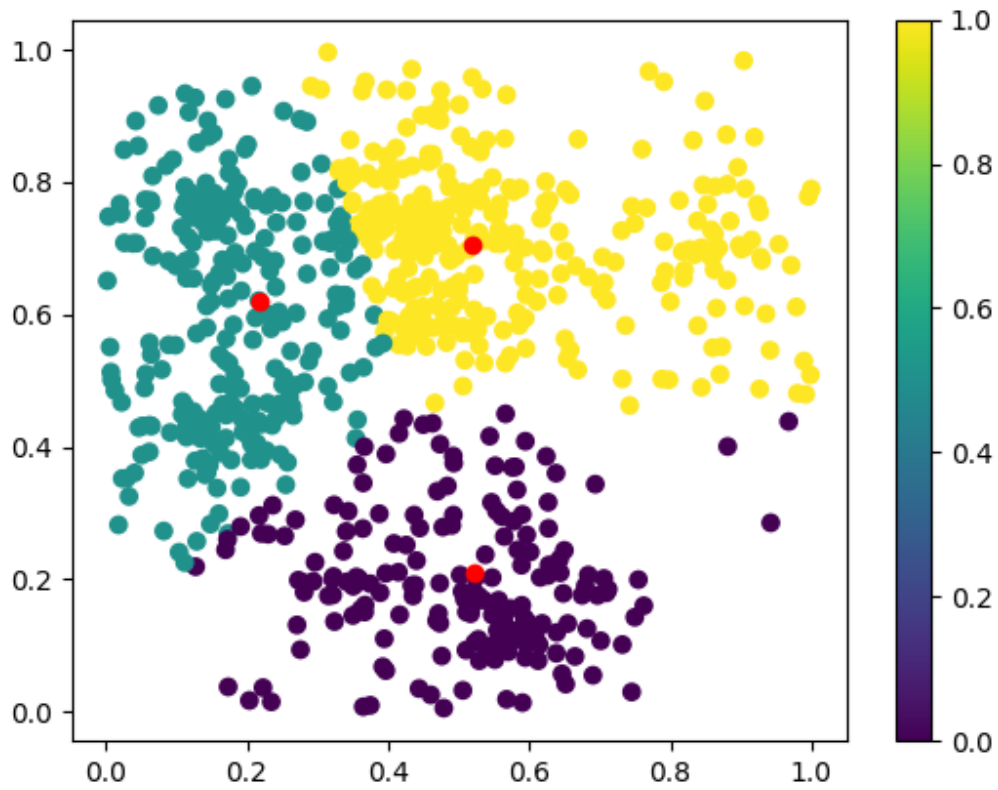
حاصل به شرح زیر است:

مجموعه داده های اول و مقدار C برابر با ۳:



امیرحسین پاشایی هیر ۹۷۳۱۰۱۳

مجموعه داده سوم و مقدار C برابر با ۳:



نحوه اجرای پروژه:

برای اجرای پروژه با پایتون از دستور زیر استفاده می کنیم:

```
python main.py [data set] [flags]
```

مقدار dataset باید path منتهی به مجموعه داده باشد (برای مثال data1.csv یا data2.csv)

مقدار flags به صورت زیر است:

- -plotc

مقادیر مختلف C را با $m=3$ بررسی می کند.

- -plotm

مقادیر مختلف m را با $C=3$ بررسی می کند.

- -crisp

رنگ آمیزی نقاط را به صورت crisp به جای فازی انجام می دهد.

- -center

مراکز خوشه ها را قرمز می کند.

یک مثال از اجرای پروژه:

```
python main.py data1.csv -plotc -crisp -center
```