

فهرست مطالب

۱۱	۱	روش نزدیک‌ترین فاصله
۱۲	۱.۱	مثال الف
۱۷	۲.۱	مثال دوم
۲۱	۲	روش دورترین فاصله
۲۲	۱.۲	مثال الف
۲۲	۲.۲	مثال دوم
۲۷	۳	روش پیوند میانگین
۲۸	۱.۳	مثال
۲۹	۲.۳	مثال دوم
۳۳	۴	روش پیوند میانه
۳۳	۱.۴	مثال الف
۳۶	۲.۴	مثال دوم
۳۹	۵	روش پیوند به مرکز
۳۹	۱.۵	مثال الف
۴۳	۲.۵	مثال دوم

لیست تصاویر

- ۱۰.۱ نمایش نمرات درس ریاضی گروه راسپینا بر روی صفحه جهت بررسی فاصله بین مشاهدات ۱۴
- ۲.۱ نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش نزدیک‌ترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۱۶
- ۳.۱ نمایش نمرات دروس ریاضی و آمار گروه راسپینا بر روی صفحه جهت بررسی فاصله بین مشاهدات ۱۸
- ۴.۱ نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش نزدیک‌ترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۱۹
- ۱.۲ نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش دورترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۲۴
- ۲.۲ نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش دورترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۲۵
- ۱.۳ نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش پیوند میانگین. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۳۰

- ۲.۳ نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش پیوند میانگین. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۳۱
- ۱.۴ نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش پیوند میانه. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۳۵
- ۲.۴ نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش پیوند میانه. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۳۷
- ۱.۵ نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش پیوند به مرکز. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۴۲
- ۲.۵ نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش پیوند به مرکز. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود. ۴۳

لیست جداول

- ۱۰.۱ نمره دانش آموزان گروه راسپینا از درس ریاضی ۱۳
- ۲.۱ ماتریس فاصله. محاسبه فاصله دو به دوی نمرات درس ریاضی گروه راسپینا با استفاده از متر اقلیدسی ۱۳
- ۳.۱ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش نزدیکترین فاصله، تشکیل یک خوشه جدید داده‌اند. ۱۴
- ۴.۱ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش نزدیکترین فاصله، تشکیل یک خوشه جدید داده‌اند. ۱۵
- ۵.۱ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳، ۱۴} به روش نزدیکترین فاصله، تشکیل یک خوشه جدید داده‌اند. ۱۵
- ۶.۱ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو خوشه {۱۳، ۱۴، ۱۶} و {۱۹، ۲۰} به روش نزدیکترین فاصله، تشکیل یک خوشه جدید داده‌اند. ۱۵
- ۷.۱ مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش نزدیکترین فاصله بر روی نمرات درس ریاضی گروه راسپینا ۱۵
- ۸.۱ نمره دانش آموزان گروه راسپینا از درس ریاضی و آمار ۱۷
- ۹.۱ ماتریس فاصله. محاسبه فاصله دو به دوی نمرات درس ریاضی و آمار گروه راسپینا با استفاده از متر اقلیدسی ۱۸
- ۱۰.۲ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند. ۲۲
- ۲.۲ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند ۲۳

- ۳.۲ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳، ۱۴} به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند ۲۳
- ۴.۲ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۰ و خوشه {۱۳، ۱۴، ۱۶} به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند ۲۳
- ۵.۲ مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش دورترین فاصله بر روی نمرات درس ریاضی گروه راسپینا ۲۳
- ۱.۳ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند. ۲۹
- ۲.۳ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند ۲۹
- ۳.۳ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳، ۱۴} به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند ۳۰
- ۴.۳ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۰ و خوشه {۱۳، ۱۴، ۱۶} به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند ۳۰
- ۵.۳ مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش پیوند میانگین بر روی نمرات درس ریاضی گروه راسپینا ۳۱
- ۱.۴ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند. ۳۴
- ۲.۴ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند ۳۴
- ۳.۴ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳، ۱۴} به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند ۳۵
- ۴.۴ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که خوشه {۱۹، ۲۰} و خوشه {۱۳، ۱۴، ۱۶} به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند ۳۵
- ۵.۴ مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش پیوند میانه بر روی نمرات درس ریاضی گروه راسپینا ۳۶
- ۱.۵ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند. ۴۰
- ۲.۵ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند ۴۱
- ۳.۵ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳، ۱۴} به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند ۴۱

- ۴۰.۵ محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۰ و خوشه {۱۳، ۱۴، ۱۶} به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند . . ۴۱
- ۵۰.۵ مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش پیوند به مرکز بر روی نمرات درس ریاضی گروه راسپینا . . ۴۱

مقدمه

در هر دانشگاه تعدادی دانشجو در رشته‌های مختلف مشغول به تحصیل هستند. در بانک اطلاعاتی هر دانشگاه سال ورود، نمرات هر درس، کم‌ترین نمره در هر ترم، بیشترین نمره در هر ترم، معدل هر ترم، نمرات دوره دبیرستان، و اطلاعات دیگری موجود می‌باشد. اگر هدف پژوهشگری تحقیق درباره دانشجویان رشته خاصی، از دانشگاهی مورد نظر، باشد، ابتدا دانشجویان را بر اساس رشته تحصیلی طبقه‌بندی کرده سپس دانشجویانی که در طبقه مورد نظر پژوهش قرار گرفته‌اند را مورد بررسی و تحقیق قرار می‌دهد.

حال فرض کنید با توجه به هدف پژوهش، نتوان ایده‌ای خاص برای طبقه‌بندی ارائه داد. به عنوان مثال اگر هدف از پژوهشی بررسی تأثیرات علاقه تحصیلی و اعتقادات عقلی و دینی بر پشتکار تحصیلی دانشجویان باشد، شاید نوع پژوهش، ایده‌ای را برای طبقه‌بندی و دسته‌بندی در اختیار پژوهشگر قرار ندهد. روش‌های مختلف خوشه‌بندی به پژوهشگران این امکان را می‌دهد که دسته‌بندی، طبقه‌بندی و سطح بندی مشاهدات را با کمک تکنیک‌ها و روش‌های خوشه‌بندی انجام دهد.

تکنیک‌های خوشه‌بندی قبل از زمان آدم و حوا به کار گرفته می‌شده است. واضح‌ترین مثال آن در طبیعت نیز خوشه‌بندی یا قوت‌های انگور است. اما به صورت فرمول بندی شده از سال ... بوده است. از جمله دانشمندان تأثیر گذار در این علم می‌توان به موجنا (۱۹۷۷) و ... اشاره کرد. در خلال این سال‌ها تکنیک‌ها و روش‌های خوشه‌بندی با توجه به نیاز پژوهشگران تنوع و گسترش یافتند. به کمک این تکنیک‌ها و روش‌ها دانش بشر در زمینه‌های مختلف بهداشت و درمان، فنی و مهندسی، علوم اجتماعی و انسانی رشد و توسعه یافته است. یکی از جالب‌ترین موارد استفاده از خوشه‌بندی و کمک آن به نسل بشر در سال ... بود، زمانی که دانشمندان بزرگ پزشکی به بررسی تأثیر بیماری جنون گاوی بر روی انسان بودند. خوشه‌بندی را می‌توان به دو قسمت روش‌های خوشه‌بندی، که بر اساس فلسفه‌ای خاص بنا نهاده شده‌اند، و وسایل اندازه گیری، که با رشد علوم ریاضی وسعت یافته‌اند، تقسیم کرد. روش‌ها و تکنیک‌های خوشه‌بندی

را می‌توان به سلسله مراتبی، بخش بندی و مدل پایه تقسیم کرد.

در روش سلسله مراتبی، قابلیت اندازه گیری نسبت به دو روش دیگر بالاتر است ولی برای داده‌های زیاد و با بعد بالا به دلیل زمان زیادی که صرف محاسبات می‌نماید، مقرون به صرفه و گاهی نیز قابل محاسبه نیست. در این کتاب روش‌های مختلف سلسله مراتبی را در فصل‌های ۱ تا ۵ مورد بررسی قرار می‌دهیم. در فصل ۱ خوشه‌بندی به روش نزدیک‌ترین فاصله، در فصل ۲ خوشه‌بندی به روش دورترین فاصله، در فصل ۳ خوشه‌بندی به روش پیوند به میانگین، در فصل ۴ خوشه‌بندی به روش پیوند به میانه و در فصل ۵ خوشه‌بندی به روش پیوند به مرکز با تفسیری ساده شرح داده شده است و نحوه محاسبه فاصله بین خوشه‌ها را با توضیحات کامل بیان کرده و در هر فصل با ۳ مثال سعی می‌شود خواننده به درک شهودی و قابل قبولی از روش‌های خوشه‌بندی سلسله مراتبی برای دسترسی به مناسب‌ترین نحوه دسته‌بندی مشاهدات در پژوهش‌هایش برسد.

فصل ۱

روش نزدیک‌ترین فاصله

در روش نزدیک‌ترین فاصله، خوشه جدیدی که در هر مرحله تشکیل می‌گردد، فاصله‌اش تا دیگر مشاهدات و یا خوشه‌های دیگر برابر است با کم‌ترین فاصله بین مشاهدات تشکیل دهنده خوشه جدید با مشاهدات و یا خوشه‌های دیگر. برای مشاهدات x^1, \dots, x^n ، ℓ متغیر $\ell = 1, \dots, m$ اندازه‌گیری شده، ثبت گردیده است. برای خوشه‌بندی این مشاهدات با توجه به متغیرهای اندازه‌گیری شده برای هر مشاهده، فاصله بین مشاهدات را با متری که معمولاً متر اقلیدسی است اندازه می‌گیرند. متر اقلیدسی $d(x^i, x^j)$ فاصله بین دو مشاهده x^i و x^j را محاسبه می‌کند. روش محاسبه فاصله دو مشاهده x^i و x^j با استفاده از متر اقلیدسی به صورت زیر می‌باشد.

$$d(x^i, x^j) = \sqrt{(x_1^i - x_1^j)^2 + \dots + (x_m^i - x_m^j)^2} \quad (1.1)$$

با استفاده از این روش فاصله بین مشاهدات را به صورت دو به دو محاسبه کرده و فاصله‌های بدست آمده را در ماتریس فاصله قرار می‌دهیم. بعد از تشکیل ماتریس فاصله، کوچک‌ترین عدد در ماتریس فاصله را پیدا می‌کنیم (به غیر از عناصر روی قطر اصلی) عدد فوق در درایه‌ای است که محل تقاطع دو مشاهده می‌باشد. این عدد نشان دهنده این است که دو مشاهده فوق کم‌ترین فاصله بین مشاهدات را از یکدیگر دارند، لذا می‌توان دو مشاهده را در قالب یک خوشه قرار داد.

سوال مهمی که مطرح می‌شود این است که زمانی که قرار است دو مشاهده با هم تشکیل یک خوشه را بدهند، خوشه جدید تشکیل شده فاصله‌اش تا دیگر مشاهدات به چه نحوی محاسبه می‌شود؟

فرض کنید در ماتریس فاصله کم‌ترین عدد مخصوص به درایه‌ای باشد که محل تقاطع دو مشاهده a و b است. حال سوال این است که هنگامی که دو مشاهده a و b تشکیل خوشه ab را می‌دهند، فاصله خوشه جدید ab تا مشاهدات دیگر، به عنوان مثال مشاهده d ، چگونه محاسبه می‌گردد؟ در روش نزدیک‌ترین فاصله، تعیین فاصله خوشه جدید با مشاهدات و یا خوشه‌های دیگر با استفاده از متر زیر صورت می‌گیرد:

$$D(A, B) = \min\{d(x^i, x^j), \forall x^i \text{ in } A \text{ and } x^j \text{ in } B\} \quad (2.1)$$

این بدین معنی است که هر گاه دو مشاهده تشکیل یک خوشه جدید را می‌دهند، فاصله خوشه جدید تا مشاهدات دیگر برابر است با کم‌ترین فاصله از بین فاصله هر یک از دو مشاهده‌ای که تشکیل خوشه را داده‌اند تا مشاهدات دیگر. به عنوان مثال اگر مشاهدات a و b تشکیل خوشه ab را بدهند و فاصله مشاهده a تا مشاهده d برابر با c_1 و فاصله مشاهده b تا مشاهده d برابر با c_2 باشد به طوری که $c_1 < c_2$ آنگاه فاصله ما بین خوشه ab تا مشاهده d برابر با مقدار c_1 می‌باشد.

۱.۱ مثال الف

با توجه به جدول ۱.۱ می‌خواهیم افراد را بر اساس نمره‌ای که از درس ریاضی گرفته‌اند دسته‌بندی کنیم. فرض کنید ایده خاصی برای دسته‌بندی نداشته باشیم، در این هنگام از تکنیک‌های خوشه‌بندی برای دسته‌بندی این افراد بهره می‌گیریم. در این مثال از روش خوشه‌بندی نزدیک‌ترین فاصله، برای خوشه‌بندی افراد استفاده می‌کنیم. در شکل ۱.۱ محل این نقاط را، برای درک بهتر، به نمایش گذاشته‌ایم. ابتدا با استفاده از رابطه ۱.۱ فاصله بین نمرات را در ماتریس فاصله (جدول ۲.۱) قرار می‌دهیم. در این ماتریس فاصله دو به دوی بین نمرات را قرار گرفته است. کوچک‌ترین عدد در این جدول عدد ۱ می‌باشد. این عدد نشان دهنده این است که کم‌ترین فاصله بین نمرات، مربوط به نمرات ۱۹، ۲۰ و ۱۳، ۱۴ می‌باشد.

سوال مهمی که اینجا مطرح می‌شود این است که ما ابتدا نمرات ۱۹ و ۲۰ را یک خوشه در نظر بگیریم و یا نمرات ۱۳ و ۱۴ را؟

در واقع تفاوتی وجود ندارد که ابتدا کدام دو نمره را یک خوشه در نظر بگیریم. ما در اینجا ابتدا دو نمره ۱۹ و ۲۰ را یک خوشه در نظر می‌گیریم.

در جدول ۳.۱ بعد از قرار دادن نمرات ۱۹ و ۲۰ در یک خوشه فاصله بین خوشه جدید ۱۹، ۲۰ تا مشاهدات دیگر را با استفاده از رابطه ۲.۱ محاسبه گردیده شده است.

جدول ۱.۱: نمره دانش آموزان گروه راسپینا از درس ریاضی

نام افراد	حسین	امیر	کیمیا	دل آرام	محمد	گلناز
نمره درسی	۲۰	۱۳	۱۹	۱۶	۱۰	۱۴

جدول ۲.۱: ماتریس فاصله. محاسبه فاصله دو به دوی نمرات درس ریاضی گروه راسپینا با استفاده از متر اقلیدسی

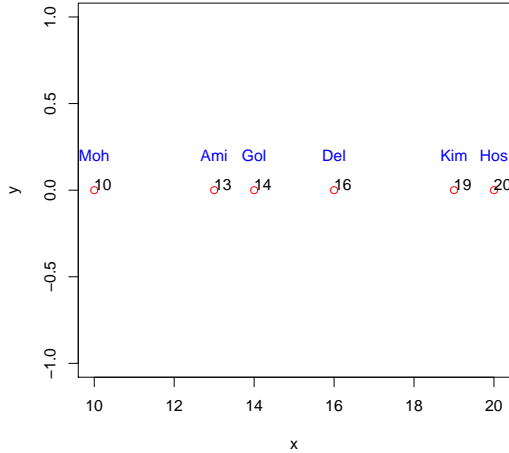
	۲۰	۱۳	۱۹	۱۶	۱۰	۱۴
۲۰	۰	۷	۱	۴	۱۰	۶
۱۳	۷	۰	۶	۳	۳	۱
۱۹	۱	۶	۰	۳	۹	۵
۱۶	۴	۳	۳	۰	۶	۲
۱۰	۱۰	۳	۹	۶	۰	۴
۱۴	۶	۱	۵	۲	۴	۰

نحوه محاسبه به صورت زیر می باشد:

$$\begin{aligned} d(\{19, 20\}, 13) &= \min [d(19, 13), d(20, 13)] \\ &= \min(6, 7) \\ &= 6 \end{aligned}$$

در جدول ۳.۱ کوچک ترین عدد، عدد ۱ می باشد که فاصله بین نمرات ۱۳ و ۱۴ می باشد. این بدین معنی است که از آنجایی که دو نمره ۱۳ و ۱۴ کم ترین فاصله دو به دوی بین مشاهدات را دارند می توانند خوشه جدیدی را تشکیل بدهند.

در جداول ۳.۱، ۴.۱، ۵.۱ و ۶.۱ ماتریس فاصله برای هر مرحله خوشه بندی محاسبه شده است. نمودار ۲.۱ مرحله های مختلف خوشه بندی به همراه فاصله بین خوشه ها را نمایش داده است. توضیحات عددی این نمودار در جدول ۷.۱ بیان گردیده است.



شکل ۱.۱: نمایش نمرات درس ریاضی گروه راسپینا بر روی صفحه جهت بررسی فاصله بین مشاهدات

جدول ۳.۱: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش نزدیک‌ترین فاصله، تشکیل یک خوشه جدید داده‌اند.

	۱۹، ۲۰	۱۳	۱۶	۱۰	۱۴
۱۹، ۲۰	۰	۶	۳	۹	۵
۱۳	۶	۰	۳	۳	۱
۱۶	۳	۳	۰	۶	۲
۱۰	۹	۳	۶	۰	۴
۱۴	۵	۱	۲	۴	۰

جدول ۴.۱: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش نزدیکترین فاصله، تشکیل یک خوشه جدید داده‌اند.

	۱۹،۲۰	۱۳،۱۴	۱۶	۱۰
۱۹،۲۰	۰	۵	۳	۹
۱۳،۱۴	۵	۰	۲	۳
۱۶	۳	۲	۰	۶
۱۰	۹	۳	۶	۰

جدول ۵.۱: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳،۱۴} به روش نزدیکترین فاصله، تشکیل یک خوشه جدید داده‌اند.

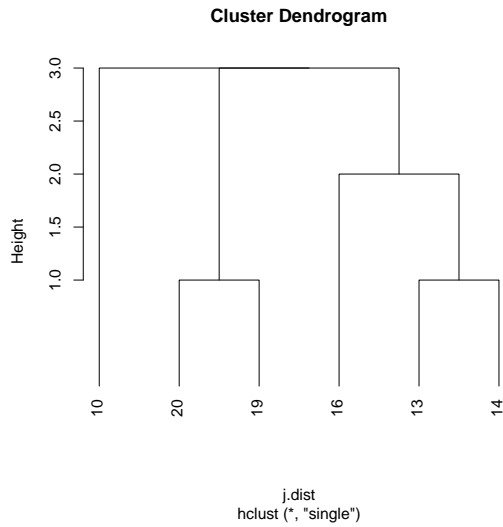
	۱۹،۲۰	۱۳،۱۴،۱۶	۱۰
۱۹،۲۰	۰	۳	۹
۱۳،۱۴،۱۶	۳	۰	۳
۱۰	۹	۳	۰

جدول ۶.۱: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو خوشه {۱۳،۱۴،۱۶} و {۱۹،۲۰} به روش نزدیکترین فاصله، تشکیل یک خوشه جدید داده‌اند.

	۱۳،۱۴،۱۶،۱۹،۲۰	۱۰
۱۳،۱۴،۱۶،۱۹،۲۰	۰	۳
۱۰	۳	۰

جدول ۷.۱: مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش نزدیکترین فاصله بر روی نمرات درس ریاضی گروه راسپینا

مراحل خوشه‌بندی	ادغام‌ها	فاصله بین خوشه‌ها
۱	۲۰، ۱۹	۱
۲	۱۴، ۱۳	۱
۳	۱۶، {۱۳، ۱۴}	۲
۴	{۱۳، ۱۴، ۱۶}، {۱۹، ۲۰}	۳
۵	{۱۳، ۱۴، ۱۶، ۱۹، ۲۰}، ۱۰	۳
۶	{۱۰، ۱۳، ۱۴، ۱۶، ۱۹، ۲۰}	۳



شکل ۲۰.۱: نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش نزدیک‌ترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.

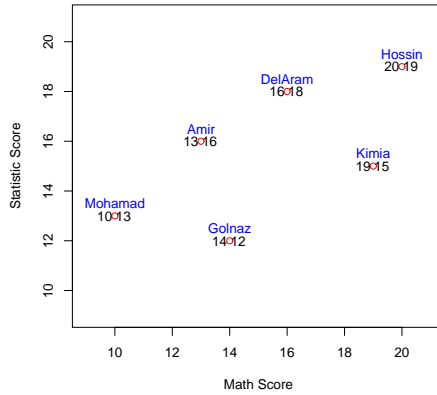
جدول ۸.۱: نمره دانش آموزان گروه راسپینا از درس ریاضی و آمار

نام افراد	حسین	امیر	کیمیا	دل آرام	محمد	گلناز
نمره درس ریاضی	۲۰	۱۳	۱۹	۱۶	۱۰	۱۴
نمره درس آمار	۱۹	۱۶	۱۵	۱۸	۱۳	۱۲

۲.۱ مثال دوم

با توجه به جدول ۸.۱ می‌خواهیم افراد را بر اساس نمره‌ای که از درس ریاضی و آمار گرفته‌اند دسته‌بندی کنیم. فرض کنید ایده خاصی برای دسته‌بندی نداشته باشیم، در این هنگام از تکنیک‌های خوشه‌بندی برای دسته‌بندی این افراد بهره می‌گیریم. در این مثال از روش خوشه‌بندی نزدیک‌ترین فاصله، برای خوشه‌بندی افراد استفاده می‌کنیم. در شکل ۳.۱ محل این نقاط را، برای درک بهتر، به نمایش گذاشته‌ایم. ابتدا با استفاده از رابطه ۱.۱ فاصله بین نمرات را در ماتریس فاصله (جدول ۹.۱) قرار می‌دهیم. در این ماتریس فاصله دو به دوی بین نمرات را قرار گرفته است. کوچک‌ترین عدد در این جدول عدد $\frac{3}{61}$ می‌باشد. این عدد نشان دهنده این است که کم‌ترین فاصله بین نمرات، مربوط به فاصله نمرات امیر و دل آرام می‌باشد.

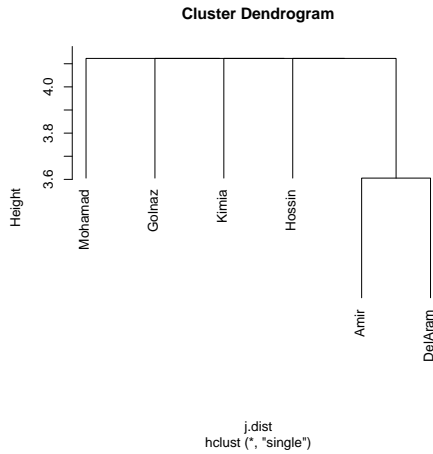
در شکل ۴.۱ مراحل خوشه‌بندی مشاهدات و فاصله بین خوشه‌ها نمایش داده شده است. بر اساس شکل ۴.۱ می‌توان گفت در مرحله اول خوشه‌بندی نمرات درس ریاضی و آمار گروه راسپینا به روش نزدیک‌ترین فاصله، ابتدا امیر و دل آرام در یک خوشه قرار گرفته و سپس در مرحله بعد در یک زمان و بدون فاصله بین خوشه‌ای دیگر اعضای گروه راسپینا حسین، محمد، کیمیا و گلناز در یک خوشه قرار گرفته‌اند. این شکل به ما این ایده را می‌دهد که بر اساس خوشه‌بندی به روش نزدیک‌ترین فاصله بر روی افراد گروه راسپینا، افراد گروه را می‌توان به دو دسته تقسیم کرد. دسته اول: امیر-دل آرام و دسته دوم: حسین-محمد-کیمیا-گلناز. دلیل این ایده این است که فاصله بین مشاهدات دسته اول و دوم در خوشه‌بندی به روش نزدیک‌ترین فاصله ثابت بوده و جهش فاصله‌ای بین دسته اول و دوم $\frac{5}{10}$ واحد، که بیشترین جهش بین خوشه‌ای است، می‌باشد.



شکل ۳.۱: نمایش نمرات دروس ریاضی و آمار گروه راسپینا بر روی صفحه جهت بررسی فاصله بین مشاهدات

جدول ۹.۱: ماتریس فاصله. محاسبه فاصله دو به دوی نمرات درس ریاضی و آمار گروه راسپینا با استفاده از متر اقلیدسی

	حسین	امیر	کیمیا	دل آرام	محمد	گلناز
حسین	۰/۰۰	۷/۶۲	۴/۱۲	۴/۱۲	۱۱/۶۶	۹/۲۲
امیر	۷/۶۲	۰/۰۰	۶/۰۸	۳/۶۱	۴/۲۴	۴/۱۲
کیمیا	۴/۱۲	۶/۰۸	۰/۰۰	۴/۲۴	۹/۲۲	۵/۸۳
دل آرام	۴/۱۲	۳/۶۱	۴/۲۴	۰/۰۰	۷/۸۱	۶/۳۲
محمد	۱۱/۶۶	۴/۲۴	۹/۲۲	۷/۸۱	۰/۰۰	۴/۱۲
گلناز	۹/۲۲	۴/۱۲	۵/۸۳	۶/۳۲	۴/۱۲	۰/۰۰



شکل ۴.۱: نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش نزدیک‌ترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.

فصل ۲

روش دورترین فاصله

در روش دورترین فاصله، خوشه جدیدی که در هر مرحله تشکیل می‌گردد، فاصله‌اش تا دیگر مشاهدات و یا خوشه‌های دیگر برابر است با دورترین فاصله بین مشاهدات تشکیل دهنده خوشه جدید با مشاهدات و یا خوشه‌های دیگر.

در این روش، پس از تعیین متری جهت محاسبه فاصله بین دو مشاهده، فاصله دو به دو بین مشاهدات را در ماتریس فاصله قرار می‌دهیم. سپس درایه‌ای که کوچک‌ترین عدد (به غیر از قطر اصلی) در آن قرار دارد را جهت یافتن دو مشاهده‌ای که کم‌ترین فاصله را از هم دارند مشخص می‌سازیم.

در روش دورترین فاصله، تعیین فاصله خوشه جدید با مشاهدات و یا خوشه‌های دیگر با استفاده از متر زیر صورت می‌گیرد:

$$D(A, B) = \max\{d(x^i, x^j), \forall x^i \text{ in } A \text{ and } x^j \text{ in } B\} \quad (۱.۲)$$

این بدین معنی است که هر گاه دو مشاهده تشکیل یک خوشه جدید را می‌دهند، فاصله خوشه جدید تا مشاهدات دیگر برابر است با بیشترین فاصله از بین فاصله هر یک از دو مشاهده‌ای که تشکیل خوشه را داده‌اند تا مشاهدات دیگر.

به عنوان مثال اگر مشاهدات a و b تشکیل خوشه ab را بدهند و فاصله مشاهده a تا مشاهده d برابر با c_1 و فاصله مشاهده b تا مشاهده d برابر با c_2 باشد به طوری که $c_1 < c_2$ آنگاه فاصله ما بین خوشه ab تا مشاهده d برابر با مقدار c_2 می‌باشد.

جدول ۱.۲: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند.

	۱۹،۲۰	۱۳	۱۶	۱۰	۱۴
۱۹،۲۰	۰	۷	۴	۱۰	۶
۱۳	۷	۰	۳	۳	۱
۱۶	۴	۳	۰	۶	۲
۱۰	۱۰	۳	۶	۰	۴
۱۴	۶	۱	۲	۴	۰

۱.۲ مثال الف

می‌خواهیم مثال ۱.۱ را به روش دورترین فاصله انجام دهیم. در اینجا نیز ابتدا دو مشاهده ۱۹ و ۲۰ تشکیل اولین خوشه را می‌دهند. بعد از قرار گرفتن نمرات ۱۹ و ۲۰ در یک خوشه فاصله بین خوشه جدید ۱۹،۲۰ تا مشاهدات دیگر را با استفاده از رابطه ۱.۲ محاسبه گردیده شده است و در جدول ۱.۲ قرار داده‌ایم. نحوه محاسبه به صورت زیر می‌باشد:

$$\begin{aligned} d(\{19, 20\}, 13) &= \max [d(19, 13), d(20, 13)] \\ &= \max(6, 7) \\ &= 7 \end{aligned}$$

در جداول ۱.۲، ۲.۲، ۳.۲ و ۴.۲ ماتریس فاصله برای هر مرحله خوشه‌بندی محاسبه شده است. نمودار ۱.۲ مرحله‌های مختلف خوشه‌بندی به همراه فاصله بین خوشه‌ها را نمایش داده است. توضیحات عددی این نمودار در جدول ۵.۲ بیان گردیده است.

۲.۲ مثال دوم

در این بخش می‌خواهیم مثال ۲.۱ را به روش دورترین فاصله انجام دهیم. نمودار درختی ۲.۲ خوشه‌بندی به روش دورترین فاصله را برای داده‌های جدول نمرات ۸.۱، با ماتریس فاصله ۱.۹ نمایش می‌دهد.

در نمودار درختی (شکل ۲.۲) خوشه‌بندی نمرات دروس ریاضی و آمار گروه راسپینا به روش دورترین فاصله، می‌توان ملاحظه نمود که ابتدا مشاهدات دو به دو به صورت حسین-کیمیا،

جدول ۲.۲: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند

	۱۹،۲۰	۱۳،۱۴	۱۶	۱۰
۱۹،۲۰	۰	۷	۴	۱۰
۱۳،۱۴	۷	۰	۳	۴
۱۶	۴	۳	۰	۶
۱۰	۱۰	۴	۶	۰

جدول ۳.۲: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳،۱۴} به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند

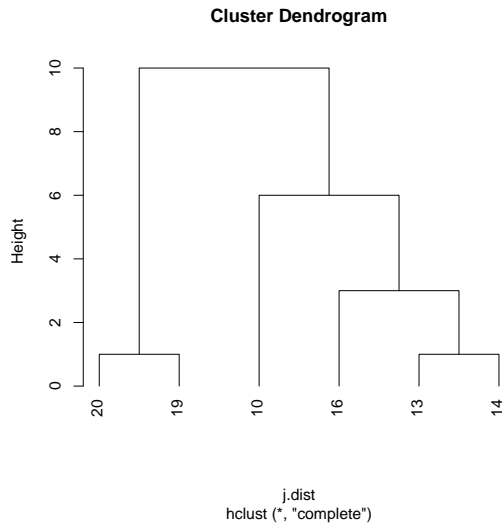
	۱۹،۲۰	۱۳،۱۴،۱۶	۱۰
۱۹،۲۰	۰	۷	۱۰
۱۳،۱۴،۱۶	۷	۰	۶
۱۰	۱۰	۶	۰

جدول ۴.۲: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۰ و خوشه {۱۳،۱۴،۱۶} به روش دورترین فاصله، تشکیل یک خوشه جدید داده‌اند

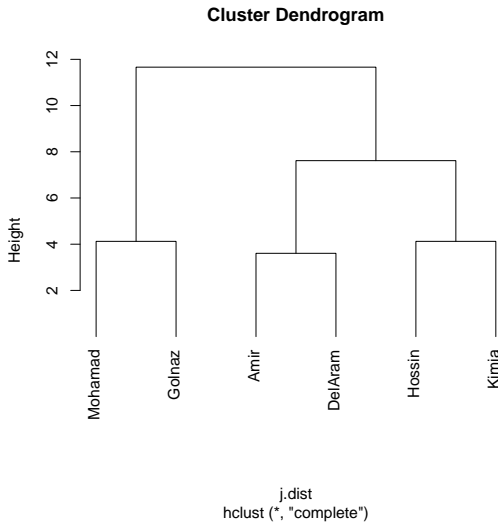
	۱۹،۲۰	۱۰،۱۳،۱۴،۱۶
۱۹،۲۰	۰	۱۰
۱۰،۱۳،۱۴،۱۶	۱۰	۰

جدول ۵.۲: مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش دورترین فاصله بر روی نمرات درس ریاضی گروه راسپینا

مراحل خوشه‌بندی	ادغام‌ها	فاصله بین خوشه‌ها
۱	۲۰، ۱۹	۱
۲	۱۴، ۱۳	۱
۳	۱۶، {۱۳، ۱۴}	۳
۴	۱۰، {۱۳، ۱۴، ۱۶}	۶
۵	{۱۹، ۲۰}، {۱۰، ۱۳، ۱۴، ۱۶}	۱۰
۶	{۱۰، ۱۳، ۱۴، ۱۶، ۲۰}	{۱۹، ۲۰}، {۱۰، ۱۳، ۱۴، ۱۶}



شکل ۱۰.۲: نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش دورترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را مشاهده نمود.



شکل ۲.۲: نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش دورترین فاصله. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.

امیر-دل آرام و محمد-گلناز با یکدیگر تشکیل خوشه داده‌اند و سپس حسین-کیمیا و امیر-دل آرام با هم تشکیل خوشه داده‌اند و در آخر تمام مشاهدات در فاصله‌ای به اندازه ۱۲ واحد، تشکیل خوشه واحد را داده‌اند. با توجه به نمودار می‌توان گفت فاصله بین خوشه امیر-دل آرام از حسین-کیمیا نزدیک به ۴ واحد می‌باشد و فاصله بین خوشه محمد-گلناز از مشاهدات دیگر ۷/۵ واحد می‌باشد.

تعاریف مهم

مشاهدات: تعداد واحدهای آماری که ℓ متغیر $(\ell = 1, \dots, m)$ از آن‌ها اندازه‌گیری شده است.

n_a, n_b : منظور از n_a و n_b تعداد مشاهداتی است که تا قبل از تبدیل خوشه A و B به خوشه AB ، خوشه‌های A و B را تشکیل داده‌اند.

c : کوچک‌ترین عدد در ماتریس فاصله که درایه‌ای که در آن قرار دارد محل تقاطع دو مشاهده است که دارای کوتاه‌ترین فاصله دو به دو از هم، در بین تمام فاصله‌های دو به دو، می‌باشد.

- هنگامی که می‌خواهیم فاصله خوشه AB را تا خوشه C محاسبه کنیم، a را فاصله خوشه A تا C و b را فاصله خوشه B تا C می‌نامیم.

فصل ۳

روش پیوند میانگین

در روش پیوند میانگین خوشه جدید AB فاصله‌اش تا دیگر مشاهدات و یا خوشه‌های دیگر به صورت وزنی از سوی خوشه‌های A و B محاسبه می‌گردد. این روش به ۲ بخش پیوند میانگین ساده و میانگین وزنی تقسیم می‌گردد.

روش پیوند میانگین ساده در این روش وزن اختصاص یافته از سوی خوشه‌های A و B برای محاسبه فاصله خوشه AB تا دیگر مشاهدات و یا خوشه‌های دیگر به صورت وزنی برابر می‌باشد که به هر یک از خوشه‌های A و B ارزش یکسانی را اختصاص می‌دهد. روش پیوند میانگین وزنی در این روش وزن اختصاص یافته، به صورت تلخیص متناسب، می‌باشد. یعنی وزن اختصاص یافته به هر یک از دو خوشه A و B بر اساس تعداد مشاهداتی است که تشکیل هر یک از دو خوشه A و B را داده‌اند.

متر به کار گرفته شده در روش پیوند میانگین در حالت کلی به صورت:

$$D(A, B) = \frac{1}{n_A n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} d(x^i, x^j) \quad (1.3)$$

ساده شده این متر برای روش پیوند میانگین ساده به صورت:
ساده شده این متر برای روش پیوند میانگین ساده، برای محاسبه فاصله خوشه AB تا خوشه C به صورت:

$$d(AB, C) = \frac{1}{2}a + \frac{1}{2}b \quad (2.3)$$

و برای روش پیوند میانگین وزنی به صورت:

$$d(AB, C) = \frac{n_A}{n_A + n_B} a + \frac{n_B}{n_A + n_B} b \quad (۳.۳)$$

می باشد. در رابطه ی (۳.۳) متر فوق، به این معنی است که در محاسبه فاصله خوشه AB تا مشاهدات و یا خوشه های دیگر، میانگین وزنی از فاصله های A و B در نظر گرفته می شود. به طوری که وزن اختصاص یافته شده به فاصله a تا دیگر مشاهدات و یا خوشه های دیگر به صورت $\frac{n_A}{n_A + n_B}$ و وزن اختصاص یافته به فاصله b نیز $\frac{n_B}{n_A + n_B}$ می باشد. لازم به ذکر است که در این کتاب منظور از روش پیوند میانگین، روش پیوند میانگین وزنی است.

۱.۳ مثال

می خواهیم مثال ۱.۱ را به روش دورترین فاصله انجام دهیم. در اینجا نیز ابتدا دو مشاهده ۱۹ و ۲۰ تشکیل اولین خوشه را می دهند. بعد از قرار گرفتن نمرات ۱۹ و ۲۰ در یک خوشه فاصله بین خوشه جدید ۱۹،۲۰ تا مشاهدات دیگر را با استفاده از رابطه ۳.۳ محاسبه گردیده شده است و در جدول ۱.۳ قرار داده ایم. نحوه محاسبه به صورت زیر می باشد:

$$\begin{aligned} d(\{19, 20\}, 13) &= \frac{n_A}{n_A + n_B} a + \frac{n_B}{n_A + n_B} b; \\ \text{If } n_A = n_B = 1 \text{ and } a = 7 \text{ } b = 6 \\ &= \frac{1}{2}(7) + \frac{1}{2}(6) \\ &= 6.5 \end{aligned}$$

دقت شود در این مرحله (جدول ۱.۳) $n_A = n_B = 1$ می باشند. ولی با توجه به جدول ۳.۳ داریم:

$$\begin{aligned} d(\{13, 14, 16\}, \{19, 20\}) &= \frac{n_A}{n_A + n_B} (a) + \frac{n_B}{n_A + n_B} (b); \\ \text{If } n_A = 2; n_B = 1 \text{ and } a = 6 \\ b = 3.5 &= \frac{2}{3}(6) + \frac{1}{3}(3.5) \\ &= 5.17 \end{aligned}$$

جدول ۱.۳: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند.

	۱۹،۲۰	۱۳	۱۶	۱۰	۱۴
۱۹،۲۰	۰	۶/۵	۳/۵	۹/۵	۵/۵
۱۳	۶/۵	۰	۶	۳	۱
۱۶	۳/۵	۳	۰	۶	۲
۱۰	۹/۵	۳	۶	۰	۴
۱۴	۵/۵	۱	۲	۴	۰

جدول ۲.۳: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند.

	۱۹،۲۰	۱۳،۱۴	۱۶	۱۰
۱۹،۲۰	۰	۶	۳/۵	۹/۵
۱۳،۱۴	۶	۰	۲/۵	۳/۵
۱۶	۳/۵	۲/۵	۰	۶
۱۰	۹/۵	۳/۵	۶	۰

در جداول ۱.۳، ۲.۳، ۳.۳ و ۴.۳ ماتریس فاصله برای هر مرحله خوشه‌بندی محاسبه شده است. نمودار ۱.۳ مرحله‌های مختلف خوشه‌بندی به همراه فاصله بین خوشه‌ها را نمایش داده است. توضیحات عددی این نمودار در جدول ۵.۳ بیان گردیده است.

۲.۳ مثال دوم

در این بخش می‌خواهیم مثال ۲.۱ را به روش پیوند میانگین انجام دهیم. نمودار درختی ۲.۳ خوشه‌بندی به روش پیوند میانگین را برای داده‌های جدول نمرات ۸.۱، با ماتریس فاصله ۹.۱ نمایش می‌دهد.

با توجه به داده‌های جدول ۸.۱ که مربوط به نمرات ریاضی و آمار گروه راسپینا است، تفسیر شکل ۲.۳ نیز مانند شکل ۲.۲ می‌باشد با این تفاوت که فاصله بین خوشه امیر-دل آرام از حسین-کیمیا ۵/۵ واحد می‌باشد و فاصله بین خوشه محمد-گلناز تا خوشه‌بندی تمام مشاهدات ۳/۱ واحد است.

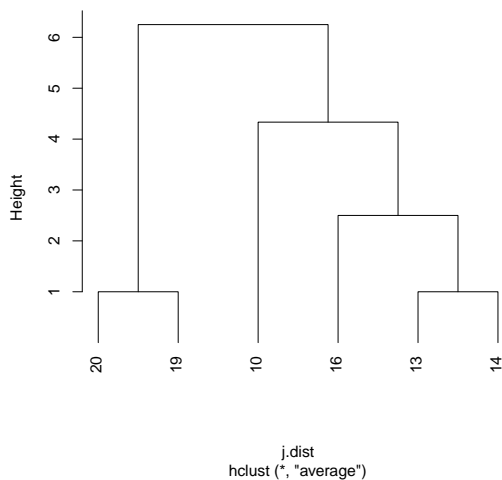
جدول ۳.۳: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳، ۱۴} به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند

	۱۹،۲۰	۱۳،۱۴،۱۶	۱۰
۱۹،۲۰	۰	۵/۱۷	۹/۵
۱۳،۱۴،۱۶	۵/۱۷	۰	۴/۳۳
۱۰	۹/۵	۴/۳۳	۰

جدول ۴.۳: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۰ و خوشه {۱۳، ۱۴، ۱۶} به روش پیوند میانگین، تشکیل یک خوشه جدید داده‌اند

	۱۹،۲۰	۱۰،۱۳،۱۴،۱۶
۱۹،۲۰	۰	۶/۲۵
۱۰،۱۳،۱۴،۱۶	۶/۲۵	۰

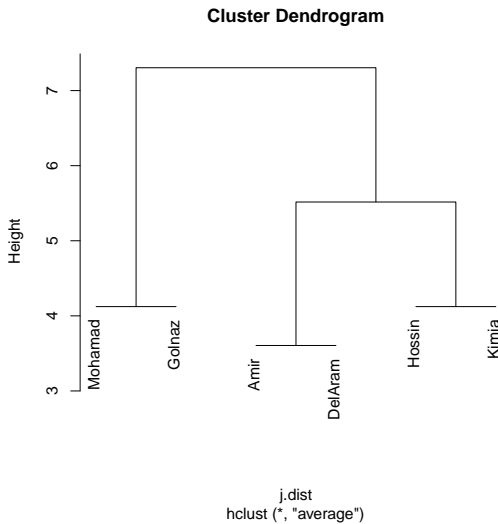
Cluster Dendrogram



شکل ۱.۳: نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش پیوند میانگین. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.

جدول ۵.۳: مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش پیوند میانگین بر روی نمرات درس ریاضی گروه راسپینا

فاصله بین خوشه‌ها	ادغام‌ها	مراحل خوشه‌بندی
۱	۲۰، ۱۹	۱
۱	۱۴، ۱۳	۲
۲/۵	۱۶، {۱۳، ۱۴}	۳
۴/۳۳	۱۰، {۱۳، ۱۴، ۱۶}	۴
۶/۲۵	{۱۹، ۲۰}، {۱۰، ۱۳، ۱۴، ۱۶}	۵
	{۱۰، ۱۳، ۱۴، ۱۶، ۲۰}	۶



شکل ۲.۳: نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش پیوند میانگین. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.

فصل ۴

روش پیوند میانه

در روش پیوند میانه، خوشه جدید AB فاصله‌اش تا دیگر مشاهدات و یا خوشه‌های دیگر با استفاده از متر زیر محاسبه می‌گردد:

$$m_{AB} = \frac{1}{2}d(\bar{x}_A + \bar{x}_B) \quad (1.4)$$

ساده شده این متر برای روش پیوند میانه به صورت:

$$d(AB, C) = \frac{1}{2}a + \frac{1}{2}b - \frac{1}{4}c \quad (2.4)$$

می‌باشد.

۱.۴ مثال الف

می‌خواهیم مثال ۱.۱ را به روش دورترین فاصله انجام دهیم. در اینجا نیز ابتدا دو مشاهده ۱۹ و ۲۰ تشکیل اولین خوشه را می‌دهند. بعد از قرار گرفتن نمرات ۱۹ و ۲۰ در یک خوشه فاصله بین خوشه جدید ۱۹،۲۰ تا مشاهدات دیگر را با استفاده از رابطه ۲.۴ محاسبه گردیده

جدول ۱.۴: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند.

	۱۹،۲۰	۱۳	۱۶	۱۰	۱۴
۱۹،۲۰	۰	۶/۲۵	۳/۲۵	۹/۲۵	۵/۲۵
۱۳	۶/۲۵	۰	۶	۳	۱
۱۶	۳/۲۵	۳	۰	۶	۲
۱۰	۹/۲۵	۳	۶	۰	۴
۱۴	۵/۲۵	۱	۲	۴	۰

جدول ۲.۴: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند

	۱۹،۲۰	۱۳،۱۴	۱۶	۱۰
۱۹،۲۰	۰/۰۰	۵/۵	۳/۲۵	۹/۲۵
۱۳،۱۴	۵/۵	۰/۰۰	۲/۲۵	۳/۲۵
۱۶	۳/۲۵	۲/۲۵	۰/۰۰	۶
۱۰	۹/۲۵	۳/۲۵	۶	//۰

شده است و در جدول ۱.۴ قرار داده‌ایم. نحوه محاسبه به صورت زیر می‌باشد:

$$\begin{aligned}
 d(\{19, 20\}, 13) &= \frac{1}{2}a + \frac{1}{2}b - \frac{1}{4}c \quad a = 7; b = 6; c = 1 \\
 &= \frac{1}{2}(7) + \frac{1}{2}(6) - \frac{1}{4}(1) \\
 &= 6.25
 \end{aligned}$$

در جداول ۱.۴، ۲.۴، ۳.۴ و ۴.۴ ماتریس فاصله برای هر مرحله خوشه‌بندی محاسبه شده است. نمودار ۱.۴ مرحله‌های مختلف خوشه‌بندی به همراه فاصله بین خوشه‌ها را نمایش داده است. توضیحات عددی این نمودار در جدول ۵.۴ بیان گردیده است.

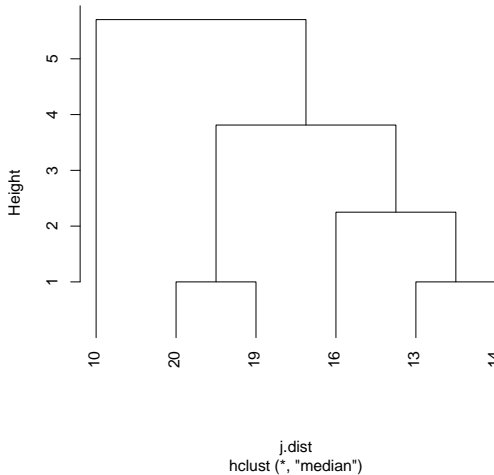
جدول ۳.۴: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳، ۱۴} به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند

	۱۹،۲۰	۱۳،۱۴،۱۶	۱۰
۱۹،۲۰	۰/۰۰	۳/۸۱	۹/۲۵
۱۳،۱۴،۱۶	۳/۸۱	۰/۰۰	۴/۰۶
۱۰	۹/۲۵	۴/۰۶	۰/۰۰

جدول ۴.۴: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که خوشه {۱۹، ۲۰} و خوشه {۱۳، ۱۴، ۱۶} به روش پیوند میانه، تشکیل یک خوشه جدید داده‌اند

	۱۳، ۱۴، ۱۶-۱۹، ۲۰	۱۰
۱۳، ۱۴، ۱۶-۱۹، ۲۰	۰	۵/۷۰
۱۰	۵/۷۰	۰

Cluster Dendrogram



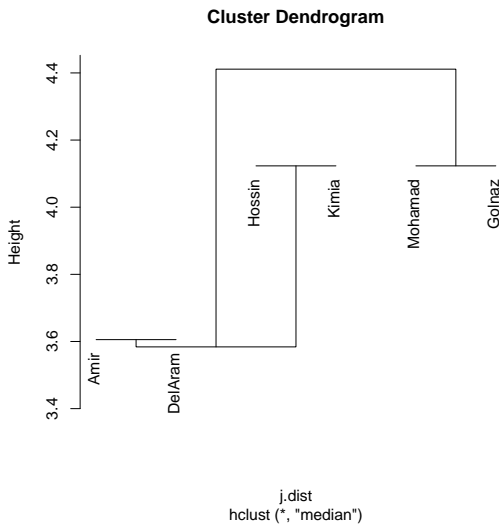
شکل ۱.۴: نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش پیوند میانه. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.

جدول ۵.۴: مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش پیوند میانه بر روی نمرات درس ریاضی گروه راسپینا

مراحل خوشه‌بندی	ادغام‌ها	فاصله بین خوشه‌ها
۱	۲۰، ۱۹	۱/۰۰
۲	۱۴، ۱۳	۱/۰۰
۳	۱۶، {۱۳، ۱۴}	۲/۲۵
۴	{۱۹، ۲۰}، {۱۳، ۱۴، ۱۶}	۳/۸۱
۵	۱۰، {۱۳، ۱۴، ۱۶، ۲۰}	۵/۷۰
۶	{۱۰، ۱۳، ۱۴، ۱۶، ۲۰}	

۲.۴ مثال دوم

در این بخش می‌خواهیم مثال ۲.۱ را به روش پیوند میانه انجام دهیم. نمودار درختی ۲.۴ خوشه‌بندی به روش پیوند میانه را برای داده‌های جدول نمرات ۸.۱، با ماتریس فاصله ۹.۱ نمایش می‌دهد. با توجه به نمودار ۲.۴ می‌توان گفت:



شکل ۲.۴: نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش پیوند میانه. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغامها و فاصله بین خوشه‌ها را ملاحظه نمود.

فصل ۵

روش پیوند به مرکز

در روش پیوند به مرکز، خوشه جدید AB فاصله‌اش تا دیگر مشاهدات و یا خوشه‌های دیگر با استفاده از متر زیر محاسبه می‌گردد:

$$D(A, B) = d(\bar{x}_A, \bar{x}_B) \quad (1.5)$$

ساده شده این متر برای روش پیوند به مرکز به صورت:

$$d(AB, D) = \frac{n_A}{n_A + n_B}(a) + \frac{n_B}{n_A + n_B}(b) - \frac{n_A n_B}{(n_A + n_B)^2}(c) \quad (2.5)$$

می‌باشد.

۱.۵ مثال الف

می‌خواهیم مثال ۱.۱ را به روش دورترین فاصله انجام دهیم. در اینجا نیز ابتدا دو مشاهده ۱۹ و ۲۰ تشکیل اولین خوشه را می‌دهند. بعد از قرار گرفتن نمرات ۱۹ و ۲۰ در یک خوشه فاصله بین خوشه جدید ۱۹،۲۰ تا مشاهدات دیگر را با استفاده از رابطه ۱.۵ محاسبه گردیده

جدول ۱.۵: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۹ و ۲۰، به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند.

	۱۹، ۲۰	۱۳	۱۶	۱۰	۱۴
۱۹، ۲۰	۰	۶/۲۵	۳/۲۵	۹/۲۵	۵/۲۵
۱۳	۶/۲۵	۰	۶	۳	۱
۱۶	۳/۲۵	۳	۰	۶	۲
۱۰	۹/۲۵	۳	۶	۰	۴
۱۴	۵/۲۵	۱	۲	۴	۰

شده است و در جدول ۱.۵ قرار داده‌ایم. نحوه محاسبه به صورت زیر می‌باشد:

$$d(\{19, 20\}, 13) = \frac{n_A}{n_A + n_B}(a) + \frac{n_B}{n_A + n_B}(b) - \frac{n_A n_B}{(n_A + n_B)^2}(c);$$

If $n_A = n_B = 1$ and $a = 7; b = 6; c = 1$

$$= \frac{1}{2}(7) + \frac{1}{2}(6) - \frac{1}{4}(1)$$

$$= 6.25$$

دقت شود در این مرحله (جدول ۱.۵) $n_A = n_B = 1$ می‌باشند. اما در جدول ۳.۵ داریم:

$$d(\{13, 14, 16\}, \{19, 20\}) = \frac{n_A}{n_A + n_B}(a) + \frac{n_B}{n_A + n_B}(b) - \frac{n_A n_B}{(n_A + n_B)^2}(c);$$

If $n_A = 2; n_B = 1$ and $a = 5.5; b = 3.25; c = 2.25$

$$= \frac{2}{3}(5.5) + \frac{1}{3}(3.25) - \frac{2}{(3)^2}(2.25)$$

$$= 4.25$$

در جداول ۱.۵، ۱.۵، ۱.۵ و ۱.۵ ماتریس فاصله برای هر مرحله خوشه‌بندی محاسبه شده است. نمودار ۱.۵ مرحله‌های مختلف خوشه‌بندی به همراه فاصله بین خوشه‌ها را نمایش داده است. توضیحات عددی این نمودار در جدول ۵.۵ بیان گردیده است.

جدول ۲.۵: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که دو مشاهده ۱۳ و ۱۴، به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند

	۱۹،۲۰	۱۳،۱۴	۱۶	۱۰
۱۹،۲۰	۰	۵/۵	۳/۲۵	۹/۲۵
۱۳،۱۴	۵/۵	۰	۲/۲۵	۳/۲۵
۱۶	۳/۲۵	۲/۲۵	۰	۶
۱۰	۹/۲۵	۳/۲۵	۶	۰

جدول ۳.۵: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۶ و خوشه {۱۳،۱۴} به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند

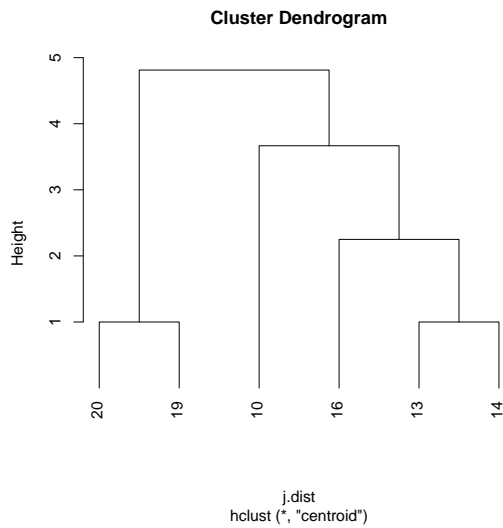
	۱۹،۲۰	۱۳،۱۴،۱۶	۱۰
۱۹،۲۰	۰	۴/۲۵	۹/۲۵
۱۳،۱۴،۱۶	۴/۲۵	۰	۳/۶۷
۱۰	۹/۲۵	۳/۶۷	۰

جدول ۴.۵: محاسبه ماتریس فاصله نمرات درس ریاضی گروه راسپینا زمانی که مشاهده ۱۰ و خوشه {۱۳،۱۴،۱۶} به روش پیوند به مرکز، تشکیل یک خوشه جدید داده‌اند

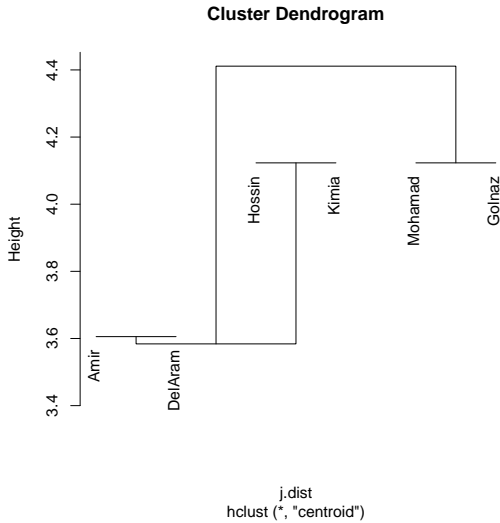
	۱۹،۲۰	۱۰،۱۳،۱۴،۱۶
۱۹،۲۰	۰	۴/۸۱
۱۰،۱۳،۱۴،۱۶	۴/۸۱	۰

جدول ۵.۵: مراحل خوشه‌بندی، ادغام مشاهدات و خوشه‌ها و فاصله بین خوشه‌ها بر اساس خوشه‌بندی به روش پیوند به مرکز بر روی نمرات درس ریاضی گروه راسپینا

فاصله بین خوشه‌ها	ادغام‌ها	مراحل خوشه‌بندی
۱/۰۰	۲۰، ۱۹	۱
۱/۰۰	۱۴، ۱۳	۲
۲/۲۵	۱۶، {۱۳، ۱۴}	۳
۳/۶۷	۱۰، {۱۳، ۱۴، ۱۶}	۴
۴/۸۱	{۱۹، ۲۰}، {۱۰، ۱۳، ۱۴، ۱۶}	۵
	{۱۰، ۱۳، ۱۴، ۱۶، ۲۰}	۶



شکل ۱.۵: نمایش درختی خوشه‌بندی نمرات درس ریاضی گروه راسپینا به روش پیوند به مرکز. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.



شکل ۲.۵: نمایش درختی خوشه‌بندی افراد گروه راسپینا بر اساس نمرات درس ریاضی و آمار به روش پیوند به مرکز. در این نمودار می‌توان مراحل خوشه‌بندی، نحوه ادغام‌ها و فاصله بین خوشه‌ها را ملاحظه نمود.

۲.۵ مثال دوم

در این بخش می‌خواهیم مثال ۲.۱ را به روش پیوند به مرکز انجام دهیم. نمودار درختی ۲.۵ خوشه‌بندی به روش پیوند به مرکز را برای داده‌های جدول نمرات ۸.۱، با ماتریس فاصله ۹.۱ نمایش می‌دهد.

با توجه به نمودار ۲.۵ می‌توان گفت: