الف) پردازش داده‌ها

۴-۶- مقدمه:

پردازش داده‌ها مرحله‌ای است که طی آن به حجم زیاد اطلاعات گردآوری شده سامان داده می‌شود و با اعمال محاسبات آماری و زمین‌آماری گوناگون به شکل قابل تفسیر درمی‌آیند. از جمله عملیاتی که در این مرحله صورت می‌گیرد، می‌توان به طبقه‌بندی داده‌ها، ورود داده‌ها در بانکهای اطلاعاتی، رسم نمودارها و تنظیم جداول اشاره کرد و در طی این مراحل کنترلهای مختلفی صورت می‌گیرد تا از بروز خطاهای احتمالی جلوگیری شود. به علت اینکه نقشه زمین شناسی ۱/۲۰۰۰۰ منطقه تهیه نشد و همچنین بدلیل کوچکی مناطق آنومال معرفی شده از مرحله اکتشافات ژئوشیمیایی ۱:۱۰۰۰۰۰ و در نتیجه محدودیت جامعه نمونه‌برداری، پردازش جوامع سنگی و محاسبه شاخص غنی شدگی انجام نگرفت. داده ها بعد از محاسبات پارامترهای آماری داده‌های خام، نرمال شدند و بر اساس این داده ها، مناطق آنومالی نهائی محاسبه و معرفی گردیدند.

۴-۷-محاسبات پارامترهای آماری داده‌های خام:

در پردازش آماری داده‌های اولیه (داده‌های خام) که از آزمایشگاه دریافت می‌شود برای اینکه این داده‌ها با استفاده از روشهای آماری مورد آنالیز قرار گیرند باید ماهیت توزیع آنها مشخص گردد. بنابراین گام اول قبل از پردازش داده‌ها، محاسبه پارامترهای آماری داده خام و شناخت ماهیت تابع توزیع مربوط به عناصر Cs, Ba, La, Nd, Sm, Hf, Pb, U, P, Ti, S, As, Ce, Mn, Rb, Sc, Au, V, Cr, Co, Ni, Cu, Zn, Ga, Sr, Y, Zr, Nb می‌باشد. به این منظور پارامترهای آماری مهم نظیر میانگین، میانه، انحراف معیار، واریانس، چولگی، کشیدگی، مینیمم مقدار و ماکزیمم مقدار مربوط به هر عنصر به همراه هیستوگرام مربوط به آن در اشکال (۱-۴) در بخش ضمیمه آورده شده است. بر اساس این اشکال و داده‌های مربوط به آن نتایج مربوط به ماهیت جوامع آماری مختلف در ذیل آورده می‌شود.

عنصر آرسنیک (As) با توجه به چولگی ۰/۸۸۱ کشیدگی ۱/۳۱۲ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر طلا (Au) با توجه به چولگی ۱/۵۵۲ و کشیدگی۲/۱۶۷و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر باریم (Ba) با توجه به چولگی۰/۸۷۹- و کشیدگی ۳/۶۸ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر سریم (Ce) دارای چولگی ۰/۰۸۵ و کشیدگی ۰/۹۶۸- و شکل هیستوگرام آن دارای تابع توزیع نزدیک به متقارن می‌باشد.

عنصر کبالت (Co) دارای چولگی ۰/۶۱۲ و کشیدگی ۰/۲۷۹- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر کروم (Cr) با توجه به چولگی۰/۴۲۲ و کشیدگی ۰/۶۴۰- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر سزیم (Cs) با توجه به چولگی۵/۴۹۰ و کشیدگی ۳۱/۰۲۹ و شکل هیستوگرام آن دارای تابع توزیع تقریباً نرمال می‌باشد.

عنصر مس (Cu) با توجه به چولگی۴/۸۶۷ و کشیدگی ۲۵/۳۳۲ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر گالیوم (Ga) دارای چولگی ۱/۰۰۵ و کشیدگی۱/۳۲۳و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر هافنیوم (Hf) با توجه به چولگی ۰/۰۱۸ و کشیدگی ۰/۱۵۴- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر لانتانیم (La) با توجه به چولگی ۰/۱۷۸- و کشیدگی ۰/۳۳۷- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر منگنز (Mn) با توجه به چولگی ۱/۳۵۵ و کشیدگی ۲/۴۷۲و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر نیوبیوم (Nb) با توجه به چولگی ۱/۳۶۶ و کشیدگی ۳/۵۹۲ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر اندیوم (Nd) با توجه به چولگی ۰/۱۲۹- و کشیدگی ۰/۱۹۶- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر نیکل (Ni) با توجه به چولگی ۰/۲۰۸ و کشیدگی ۱/۱۷۶- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر فسفر (P) با توجه به چولگی ۰/۹۸۳ و کشیدگی ۰/۵۸۰ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر نیوبیوم (Pb) با توجه به چولگی ۱/۰۳۶ و کشیدگی ۱/۷۵۴ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر روبیدیوم (Rb) با توجه به چولگی ۱/۲۹۳ و کشیدگی ۱/۷۳۵ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر گوگرد (S) با توجه به چولگی ۰/۷۹۲ و کشیدگی ۰/۷۰۸- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصراسکاندیوم (Sc) با توجه به چولگی ۰/۱۶۴- و کشیدگی ۰/۵۰۲- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر اسمیوم (Sm) با توجه به چولگی ۰/۰۹۴- و کشیدگی ۰/۷۳۶ و شکل هیستوگرام آن دارای تابع توزیع تقریباً متقارن می‌باشد.

عنصر استرانسیوم (Sr) با توجه به چولگی ۰/۳۸۷ و کشیدگی ۰/۶۰۰- و شکل هیستوگرام آن دارای تابع توزیع تقریباً متقارن می‌باشد.

عنصر تیتانیوم (Ti) با توجه به چولگی ۱/۶۶۴ و کشیدگی ۳/۰۵۲ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر اورانیوم (U) با توجه به چولگی ۰/۴۳۷ و کشیدگی ۰/۵۱۴- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر وانادیوم (V) با توجه به چولگی ۰/۴۹۲ و کشیدگی ۰/۱۶۴- و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر ایتریوم (Y) با توجه به چولگی ۰/۹۴۶ و کشیدگی ۰/۵۲۶ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر روی (Zn) با توجه به چولگی ۱/۲۴۴ و کشیدگی ۳/۱۹۰ و شکل هیستوگرام آن دارای تابع توزیع نامتقارن می‌باشد.

عنصر زیرکونیوم (Zr) با توجه به چولگی ۱/۱۷۳ و کشیدگی ۳/۵۲۰ و شکل هیستوگرام آن دارای تابع توزیع تقریباً متقارن می‌باشد.

۴-۸- بررسی مقادیر خارج از رده (Outliers Samples):

هنگام بررسی مقادیرداده های خام به نمونه‌هایی برمی‌خوریم که در کرانه‌های بالا و پائین جامعه داده‌ها قرارگرفته‌اند و از جامعهٔ اصلی جدا افتاده‌اند. اگر نمودار جعبه‌ای (Box plot) آنها ترسیم شود این نمونه‌ها به نحو بارزی خودشان را از بقیه جدا می‌کنند(اشکال۲-۴).

مقادیر خارج از رده به سه حالت مختلف ممکن است بوجود آیند:

حالت اول: ممکن است از یک خطای سیستماتیک به هنگام نمونه‌برداری، آماده‌سازی یا تجزیهٔ شیمیایی نمونه‌ها ناشی شده باشند که باید از مرحلهٔ پردازش داده‌ها حذف یا اصلاح شوند.

حالت دوم: مشاهداتی که بصورت یک پدیدهٔ فوق العاده، نمود پیدا می‌کنند که باید پس از بررسی اعتبار آنها در مورد حفظ یا حذف آنها تصمیم گرفت.

حالت سوم : مشاهدات فوق العاده‌ای که هیچگونه توضیح مناسبی برای آنها وجود ندارد و کارشناس اگر احساس کند که آنها به عنوان گوشه‌ای از جامعهٔ مورد بررسی هستند می‌تواند آنها را حفظ کند. وجود مقادیر خارج از رده در جامعهٔ نمونه‌ها موجب افزایش واریانس جامعه و نیز همبستگی بین متغیرها و همچنین افزایش چولگی در نمودار توزیع عناصر می‌شود. برای کاهش این تأثیر سه راه وجود دارد:

۱- محاسبهٔ ضریب همبستگی با استفاده از روشهای ناپارامتری مانند روش اسپیرمن (Spearman)

۲- حذف نمودن این مقادیر از جامعهٔ شاخص غنی‌شدگی هر عنصر می‌باشد

تعدیل داده های خارج از رده است.

در روش تعدیل با توجه به نمودار های ترسیم شده در (Box plot) مرز عددی بین مقادیر خارج از رده و سایر داده ها تعیین گردیده و داده های خارج از رده به عدد فوق با یک روند کاهشی نزدیک می شود؛ در این جا برای تمام عناصر از روش تعدیل استفاده شده است. نمونه های دارای مقادیر خارج از رده برای عناصر مختلف بصورت جدول (۴-۱) آورده شده است.

جدول(۴-۱): نمونه های دارای مقادیر خارج از رده

شماره نمونه مقادیر خارج از رده	عنصر	شماره نمونه مقادیر خارج از رده	عنصر
-	Ni	109,108A	As
116	P	109	Au
109,108A	Pb	-	Ba
-	Rb	-	Ce
-	S	-	Co
-	Sc	-	Cr
113,104	Sm	113	Cs
-	Sr	105,114,111,120	Cu
-	Ti	-	Ga
-	U	-	Hf
-	V	-	La
113	Y	-	Mn
109,110	Zn	-	Nb
-	Zr	-	Nd

۴-۹- نرمال‌سازی داده های خام:

استفاده از برخی روشهای آماری منوط به نرمال بودن تابع توزیع متغیرهای مورد مطالعه است در حالیکه توابع توزیع از نوع لاگ‌نرمال است، به همین دلیل قبل از استفاده از این روش‌ها داده ها باید نرمال شوند. در این بخش از نوعی تبدیلات جهت نرمال کردن تابع توزیع مقادیر داده خام استفاده شده است. این کار شرط لازم کاربرد برخی روشهای آماری مانند تعیین نمونه‌های آنومالی با استفاده از اضافه کردن ضرایبی از

انحراف معیار به حد آستانه‌ای و یا محاسبه ضرایب همبستگی پیرسون می‌باشد. روش لاگ‌نرمال بصورت یک روش توزیعی برای نرمال کردن تابع توزیع جوامعی که دارای چولگی در نمودار خود هستند، به کار می‌رود. در اینجا از لگاریتم طبیعی مقادیر داده خام به اضافه یا منهای یک مقدار ثابت λ مطابق رابطهٔ تبدیلی زیر استفاده شده است.

$$Z=LnEI\pm \lambda$$

مقدارλ بگونه‌ای انتخاب می‌شود که پس از انتخاب داده‌ها به یک مقدار بهینه از چولگی و کشیدگی در منحنی توزیع نرمال دست پیدا کنیم. پارامترهای آماری و هیستوگرام‌های ترسیم شده برای داده‌های نرمال در شکل‌های (۴-۱) تا (۴-۱۱) در ضمیمه همین فصل آورده شده است. با توجه به این پارامترهای آماری می‌توان گفت که مقادیر چولگی و کشیدگی متغیرها در مقایسه با مقادیر متناظر مربوط به داده های خام نرمال نشده تا چه اندازه کاهش یافته و منحنی توزیع تجمعی آنها به صورت یک خط راست که مبین توزیع نرمال می‌باشد ظاهر شده است. هیستوگرام مقادیر نرمال شده نیز به شکل منحنی کاملاً متفاوتی در آمده است.

۴-۱۰ : بررسی‌های آماری چند متغیره

هر تجزیه وتحلیل چند متغیره که بر روی بیش از دو متغیر انجام می‌گیرد می‌تواند در قالب آنالیزهای چند متغیره بیان شود. غالب تکنیکهای چند متغیره در اصل بسط و توسعه آنالیزهای تک‌متغیره می‌باشند و البته بعضی از روش‌های چند متغیره تنها برای پاسخ‌گوئی به مقاصد چند متغیره طراحی شده‌اند که از جمله این روش‌ها می‌توان به تجزیه‌عاملی اشاره کرد. تجربه نشان داده است که چنانچه ترکیبی از متغیرها به جای یک متغیر بکار گرفته شوند و از نتایج ترکیبی آنها استفاده شود امکان تشخیص هاله‌های مرکب ژئوشیمیائی در اطراف توده‌های کانساری به مراتب افزایش می‌یابد و از طرفی اثرات خطاهای تصادفی در بکارگیری ترکیبی متغیرها نسبتاً کاهش می‌یابد. از دیگر مزایای استفاده از روشهای چند متغیره کاهش تعداد متغیرها

در مباحث داده‌پردازی و در نتیجه کاستن از تعداد نقشه‌ها است. با استفاده از این روش‌ها امکان مقایسه متغیرها و کسب نتایج راحت‌تر خواهد بود. البته استفاده بهینه از روش‌های چند متغیره در حالتی صادق خواهدبود که در پردازش داده‌ها با تعداد زیادی متغیر روبرو باشیم و تا حدودی امکان اخذ نتیجه از متغیرها به گونهٔ منفرد غیرممکن و یا توأم با خطای زیاد باشد. در این پروژه از روشهای چند متغیره فقط از روش آنالیز خوشه‌ای استفاده شده است.

آنالیز خوشه‌ای و تفسیر آن:

چون هر گروه معین از عناصر نسبت به یکسری از شرایط محیطی کم و بیش به طور مشابه حساسیت نشان می‌دهند، شناخت ارتباط و همبستگی ژنتیکی متقابل بین عناصر مختلف می‌تواند در شناخت دقیق‌تر تغییرات موجود در محیطهای ژئوشیمیایی بکارگرفته شود. ضمناً تجمع ژنتیکی بعضی از عناصر ممکن است بعنوان راهنمای مستقیم در تفسیر نوع نهشته‌ای که احتمالاً در ناحیه وجود دارد بکار رود. رویهم رفته شناخت همبستگی‌های ژنتیکی که در بین عناصر وجود دارد اطلاعات لازم را برای تفسیر هر چه صحیح‌تر داده‌های ژئوشیمیائی در اختیار می‌گذارد.

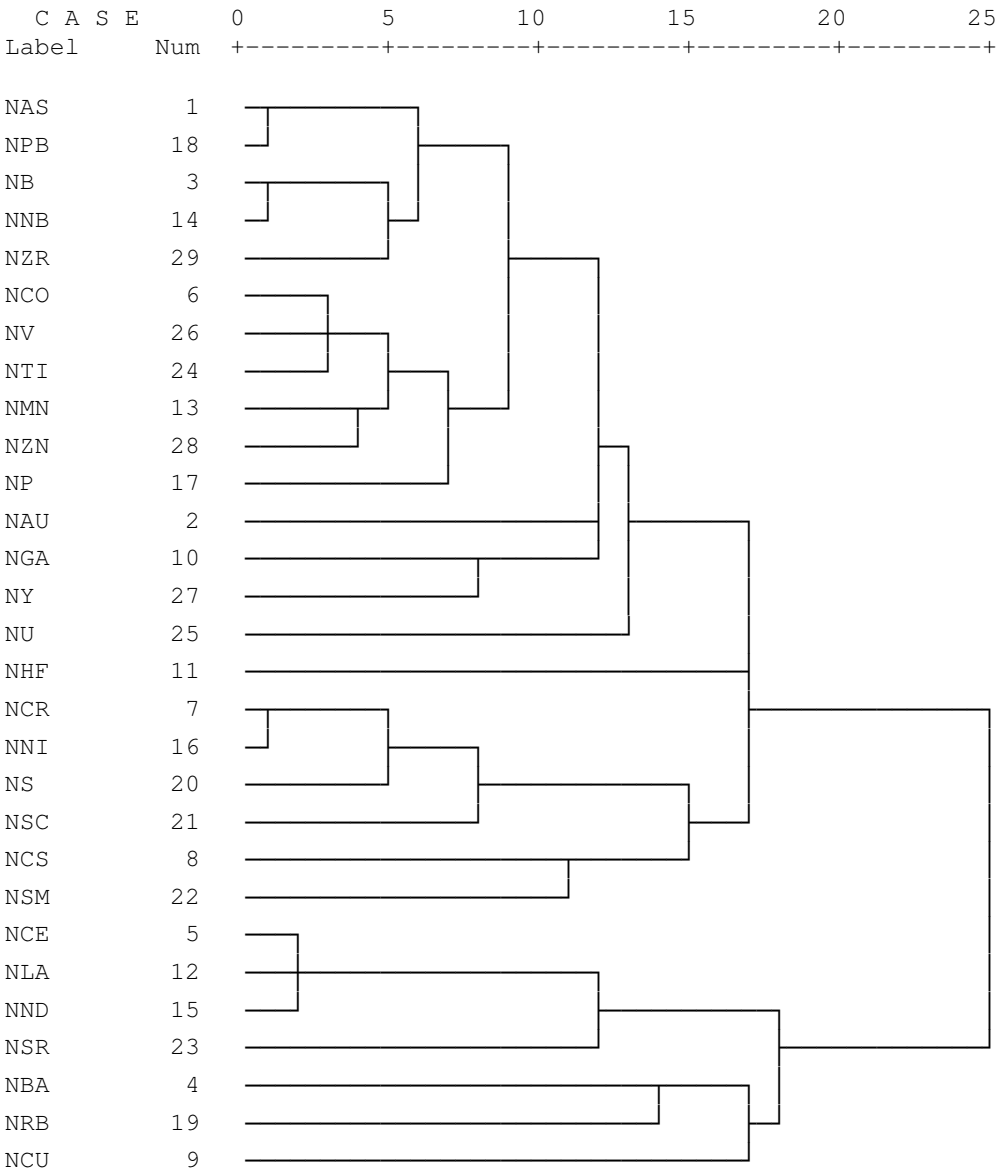
آنالیز خوشه‌ای یک روش آماری چند متغیره است که عناصر را بر اساس شباهت تغییرپذیری بین آنها در قالب دسته‌ها یا گروه‌هایی طبقه‌بندی می‌کند. دلایل زیادی برای ارزشمند بودن آنالیز خوشه‌ای وجود دارد از جمله اینکه آنالیز خوشه‌ای می‌تواند در پیدا کردن گروه‌های واقعی کمک کند و همچنین از تراکم داده‌ها بکاهد. البته باید توجه داشت که آنالیز خوشه‌ای می‌تواند گروه‌های غیرقابل انتظاری را ایجاد نمایدکه بیانگر روابط جدیدی خواهند بود و باید مورد بررسی قرارگیرند. نتایج حاصل از آنالیز خوشه‌ای عناصر مورد مطالعه در شکل(۳-۴) آورده شده است، با توجه به شکل می‌توان پنج گروه اصلی را جدا نمود که بیانگر ارتباط پاراژنزی بین متغیرها باشد .

گروه اول: شامل عناصر As,Pb,Nb,B,Zr می‌باشد.

گروه دوم: شامل عناصر Co,V,Ti,Mn,Zn,P می‌باشد

گروه سوم: شامل عناصر Cr,Ni,S,Sc,Cs,Sm می‌باشد..

گروه چهارم: شامل عناصر Ce,La,Nd,Sr,Ba,Pb,Cu می‌باشد.



شکل (۳-۴) :نتایج حاصل از آنالیز خوشه‌ای عناصر منطقه مورد مطالعه

۴-۱۱ - آنومالی عناصر مختلف:

در این بخش به شرح آنومالیهای بدست آمده از عناصر مختلف می‌پردازیم. قبل از توصیف آنومالی عناصر مختلف ذکر چند نکته الزامی است. در مورد جدایش آنومالیها از فرمول $x + 4S$ و $x + 3S$ و $x + 2S$ استفاده شده است.این کار هم در مورد داده‌های نرمال شده مورد استفاده قرار گرفته است. مقادیر بیشتر از $x + 4S$ به عنوان آنومالی درجه یک و مقادیر بین $x + 4S$ و $x + 3S$ به عنوان آنومالی درجه دو لحاظ شده‌اند (میانگین داده ها $=X$ ، انحراف معیار $= S$ و ضریب $= n$) دراین منطقه جدایش آنومالی ها بر اساس داده‌های نرمال صورت گرفته است. نقشه نمونه‌برداری منطقه نیز به صورت نقشه‌ای جداگانه آورده شده است (نقشه شماره یک) همچنین آنومالی‌های مربوط به هر عنصر نیز بصورت نقشه‌ای جداگانه آورده شده است که در آن حدود آنومالی‌های ممکن، احتمالی و قطعی و همچنین حدود زمینه هر عنصر آورده شده است (نقشه‌های دو تا بیست و نه).