

# بررسی روشهای طبقه‌بندی داده‌ها در نقشه‌های کروپلت<sup>(۱)</sup>

## (مطالعه موردی: تهیه نقشه تراکم جمعیت ایران)

دکتر مهران مقصودی

استادیار دانشکده جغرافیا - دانشگاه تهران

سپیده زندیه

کارشناس سازمان نقشه برداری کشور

### چکیده

امروزه نقشه‌های موضوعی بویژه نقشه‌های کروپلت در اغلب گزارشهای دولتی، تجاری، تحقیقاتی و علمی مورد استفاده قرار می‌گیرد اما کمتر به نوع طبقه‌بندی داده توجه می‌شود. در واقع روشهای مختلفی برای طبقه‌بندی داده‌های آماری و استفاده آنها در نقشه‌های کروپلت وجود دارد که هر یک برای نوع خاصی از آمار و اطلاعات مناسب می‌باشند.

نمایش بهتر داده‌ها وقتی میسر است که طبقه‌بندی بنحوی انجام گیرد که کاربر براحتی بتواند اطلاعات مورد نیاز خود را از نقشه استخراج نماید. در ایران تهیه نقشه‌های موضوعی به سبک و سیاق امروزی آن سابقه چندانی طولانی ندارد. اما علیرغم این موضوع در سالهای اخیر با تهیه اطلسها و استفاده گسترده از نرم افزارهای کامپیوتری رشد چشمگیری در خصوص کارتوگرافی موضوعی و استفاده از نقشه‌های موضوعی بوجود آمده است.

### مواد و روشها

در این تحقیق سعی شده تا از همه روشهای تجربی و آماری موجود در زمینه طبقه‌بندی داده‌ها به منظور تهیه نقشه‌های کروپلت استفاده شود. با وجود اینکه نوع آمار و اطلاعات مورد استفاده در این تحقیق از اجزای اصلی تحقیق نبود اما با استفاده از اطلاعات مربوط به تراکم جمعیت بعنوان شاخصی که توسط بسیاری از متخصصان در زمینه جغرافیا، علوم اجتماعی، اقتصادی و حتی سیاسی شناخته شده است سعی در ملموس تر شدن نتایج تحقیق داشتیم. به همین دلیل از آمار سرشماری عمومی نفوس و مسکن سال ۱۳۷۵ مرکز آمار ایران استفاده گردید. برای نمایش اطلاعات بر روی نقشه‌ها نیز از نقشه مبناء ایران با تقسیمات استانی سال ۱۳۷۵ استفاده گردید. همچنین از نرم افزار Arcview نیز بمنظور تهیه نقشه‌های نهایی بهره گرفته شد.

### طبقه‌بندی داده‌ها در کارتوگرافی موضوعی

داده‌های آماری که از منابع مختلف جمع آوری می‌شود باید پردازش شده و سپس به روش مناسبی طبقه‌بندی شوند.

امروزه تقریباً در تمامی گزارشها، پژوهشهای علمی، طرح‌ها و پروژه‌های کاربردی که به نوعی با پراکنندگی پدیده‌های جغرافیایی سروکار دارند از نقشه‌های موضوعی استفاده می‌شود.

در کشورهای اروپایی تهیه نقشه‌های موضوعی به قرن ۱۷ بازمی‌گردد اما در کشور ما تهیه این نقشه‌ها با تدوین اطلس‌ها توسط برخی مؤسسات در دهه ۴۰ شروع شد. نوع خاصی از نقشه‌های موضوعی که کاربرد وسیعی نیز پیدا کرده، نقشه‌های کروپلت است. نقشه‌های کروپلت برای نمایش توزیع پدیده‌های کمی در سطح یک محدوده سیاسی یا اداری مورد استفاده قرار می‌گیرند. محدوده استان، شهرستان، بخش و دهستان از این قبیل محدوده‌ها می‌باشند.

نقشه‌های کروپلت معمولاً با گام رنگی (۲) یا هاشور نشان داده می‌شوند. نکته مهم در تهیه نقشه‌های کروپلت یافتن سیستم طبقه‌بندی مناسب برای داده‌های مورد استفاده است. در واقع سیستم‌های طبقه‌بندی گوناگونی در این زمینه وجود دارند که چنانچه برای نمایش اطلاعات سیستم طبقه‌بندی مناسبی انتخاب نشود چه بسا کاربرد چهار سردرگمی شود.

در این تحقیق سعی شده تا ضمن بررسی روشهای طبقه‌بندی در نقشه‌های کروپلت، ویژگی‌های هر یک با ذکر مثال مشخص گردد.

### واژه‌های کلیدی: کارتوگرافی، نقشه موضوعی، کروپلت، طبقه‌بندی

### مقدمه

بطور کلی تهیه نقشه‌های موضوعی و نمایش گرافیکی اطلاعات نسبت به تهیه نقشه‌های عمومی از قدمت کمتری برخوردار است و به سده‌های اخیر باز می‌گردد.

در واقع نمایش گرافیکی اطلاعات به تاریخ کارتوگرافی موضوعی<sup>(۳)</sup>، گرافهای آماری، نمایش تصویری داده‌ها و... مرتبط می‌شود. با افزایش استفاده از آمار در رشته‌های مختلف در قرن ۱۹ و همچنین در قرن بیستم، نمایش گرافیکی اطلاعات رشد سریعی یافت. در حقیقت نمایش اطلاعات به شکل مدرن دستاورد پیشرفت آمار و احتمالات و کارتوگرافی می‌باشد.

جدول شماره ۱: فاصله طبقات و فراوانی هر طبقه در روشهای مختلف طبقه بندی

فراوانی	فاصله مساوی	فراوانی	مساحت مساوی	فراوانی	انحراف معیار	فراوانی	کوانتال	فراوانی	شکست طبیعی	فراوانی	تصادف هندسی	فراوانی	تصادف حسابی
۲۶	۵-۱۱۲	۴	۵-۱۱	۲۱	۰-۶۸	۶	۲۰-۵	۷	۵-۲۴	۴	۵-۱۳	۱۱	۵-۴۱
۱	۱۱۲-۲۱۹	۱	۱۱-۱۵	-	۶۸	۶	۲۴-۴۲	۱۱	۲۴-۵۹	۵	۱۳-۳۳	۱۵	۴۱-۱۱۲
-	۲۱۹-۳۲۵	۳	۲۰-۳۱	۶	۶۸-۱۶۴	۶	۴۷-۵۹	۸	۵۹-۱۰۹	۱۵	۳۳-۸۳	۱	۱۱۲-۲۱۹
-	۳۲۵-۴۳۲	۱۰	۳۲-۵۹	-	۱۶۴-۲۶۱	۵	۶۵-۷۳	۱	۱۰۹-۱۶۱	۳	۸۳-۲۱۲	-	۲۱۹-۲۶۱
۱	۴۳۲-۵۳۹	۱۰	۶۵-۵۳۹	-	۲۶۱-۳۵۷	۵	۷۶-۵۳۹	۱	۱۶۱-۵۳۹	۱	۲۱۲-۵۳۹	۱	۲۶۱-۵۳۹
۲۸		۲۸		۱	۳۵۷-۵۳۹	۲۸		۲۸		۲۸		۲۸	جمع
				۲۸									

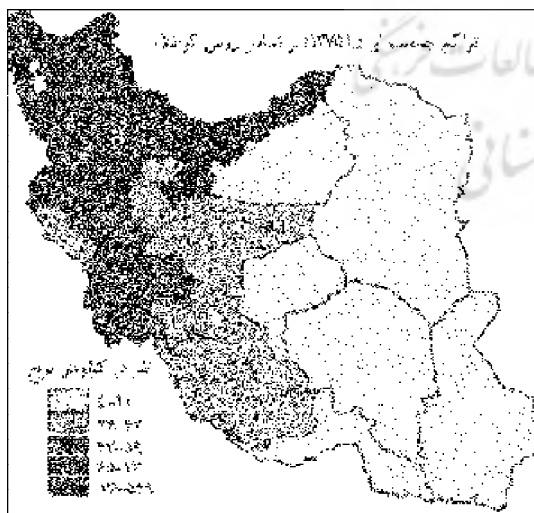
مناسبی در بین طبقات داشته باشند. یکی از اهداف در کارتوگرافی این است که اطلاعات بنحوی در اختیار کاربر قرار گیرد که کاربر بتواند تجزیه و تحلیل درستی از اطلاعات داشته باشد. روشهای متفاوتی به منظور تعیین فاصله طبقات وجود دارد که آنها را می توان به سه گروه عمده تقسیم نمود.

#### ۱- روش فواصل ثابت طبقات<sup>(۵)</sup>

در این روش داده ها به طبقات منظم و یکسان تقسیم می شوند و فاصله طبقات در هر طبقه دقیقاً برابر اندازه طبقه دیگر است. این روش خود به دو دسته تقسیم می شوند:

**الف: فواصل مساوی<sup>(۶)</sup>** (در این روش فاصله طبقات بر اساس دامنه تغییرات تعیین می گردد).

**ب: تعیین فواصل طبقات بر اساس انحراف معیار** (در این روش فاصله طبقات بر اساس انحراف معیار و میانگین طبقات تعیین می گردد)



نگاره ۱: تراکم جمعیت ایران (۱۳۷۵) بر اساس روش کوانتال

#### ۲- روش فواصل نامنظم

در واقع روشهایی را شامل می شود که با اعمال آنها، در فواصل طبقات بدست آمده هیچ نظم و رابطه مشخصی وجود ندارد.

در حقیقت نوع طبقه بندی می تواند ظاهری کاملاً متفاوت به یک نقشه موضوعی بدهد و اطلاعات درست را در اختیار کاربر قرار دهد و یا او را دچار سردرگمی کند. بطور کلی در کارتوگرافی از روشهای مختلفی به منظور طبقه بندی داده های آماری استفاده می شود. با استفاده از هر یک از روشها در نهایت نقشه جدیدی تولید می شود که از نظر ظاهر از هم متفاوت می باشند. تعدادی از روشها بسیار ساده بوده و فاصله طبقات در آنها به آسانی مشخص می شود. این در حالی است که بعضی دیگر نیاز به محاسبات پیچیده تری دارند. اغلب افراد مبتدی سعی می کنند از روشهای ساده جهت طبقه بندی استفاده کنند اما یک کارتوگراف ماهر سعی در انتخاب بهترین روش برای طبقه بندی داده ها دارد.

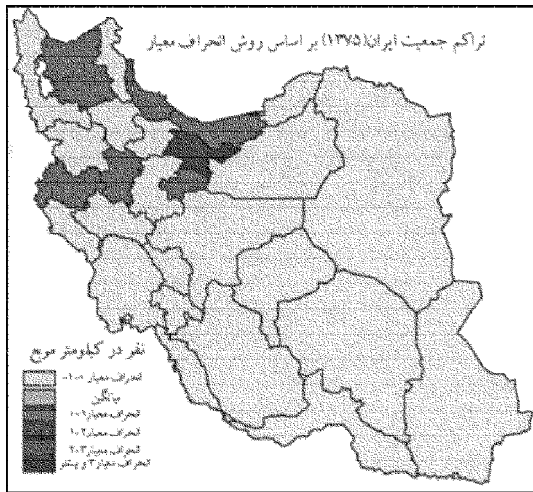
#### اصول کلی در طبقه بندی داده ها

بطور کلی به منظور طبقه بندی و تعیین فاصله طبقات، کارتوگراف باید با اصول و قوانین خاصی اقدام به طبقه بندی آمار و اطلاعات نماید که این اصول عبارتند از:

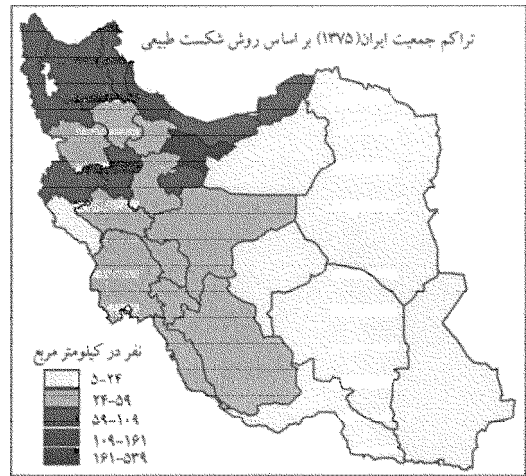
- ۱- طبقات باید تمام داده ها را شامل شود.
- ۲- بین دو طبقه متوالی نباید فصل مشترک وجود داشته باشد.
- ۳- داده هایی که در هر طبقه قرار داده می شوند تا حد امکان همگن باشند. به عبارت دیگر داده های هر طبقه با یکدیگر همگن و با داده های کلاسهای دیگر نا همگن باشند.
- ۴- در صورت لزوم باید طبقاتی را که فاقد داده ها هستند در راهنما گنجانند. با این عمل کاربر به وجود فواصل خالی در مجموعه داده ها پی می برد.
- ۵- دقت حدود محاسباتی کلاسها نباید بیش از دقت داده های اصلی باشد.
- ۶- تا حد امکان از اعداد رُند برای فاصله طبقات استفاده شود زیرا اعداد رُند به سهولت به خاطر سپرده می شوند و کاربر همواره ترجیح می دهد از نقشه هایی استفاده کند که حدود کلاسهای آن رُند باشد.
- ۷- در صورت امکان رابطه ریاضی بین طبقات وجود داشته باشد.
- ۸- پیچیدگی نقشه تا حد امکان به حداقل برسد (غضنفری، نقشه برداری، شماره ۲۴)

#### تعیین فاصله طبقات<sup>(۴)</sup>

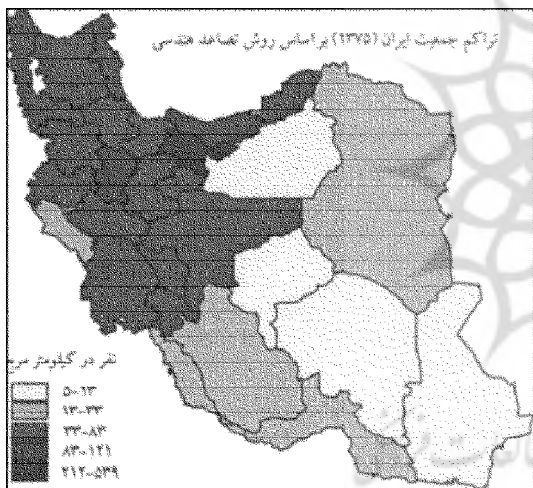
بطور کلی تعیین فواصل طبقات مناسب سبب می شود که داده ها توزیع



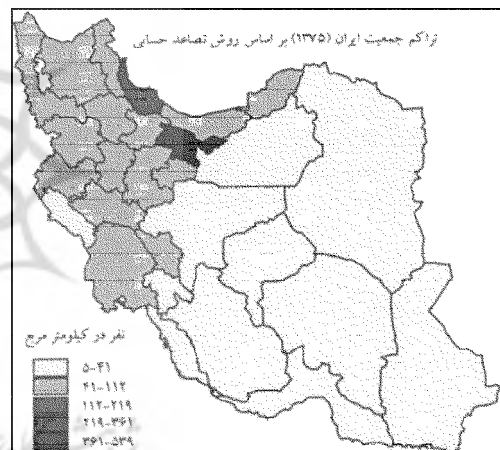
نگاره ۵: تراکم جمعیت ایران (۱۳۷۵) بر اساس روش انحراف معیار



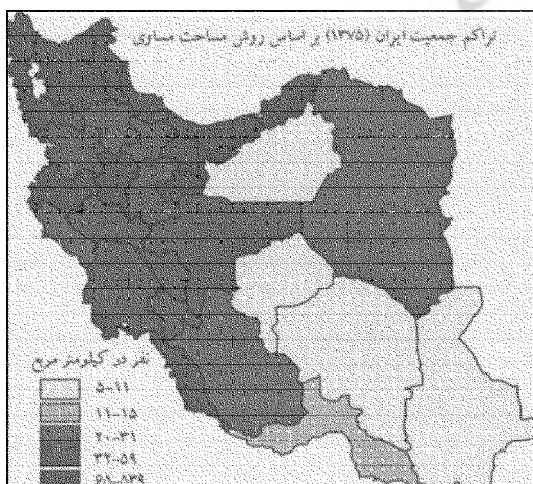
نگاره ۲: تراکم جمعیت ایران (۱۳۷۵) بر اساس روش شکست طبیعی



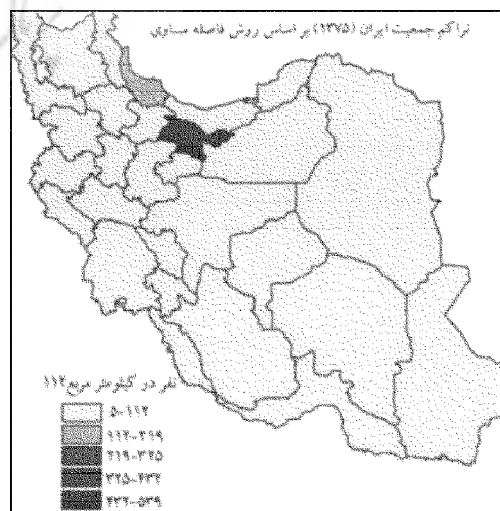
نگاره ۶: تراکم جمعیت ایران (۱۳۷۵) بر اساس روش تصاعد هندسی



نگاره ۳: تراکم جمعیت ایران (۱۳۷۵) بر اساس روش تصاعد حسابی



نگاره ۷: تراکم جمعیت ایران (۱۳۷۵) بر اساس روش مساحت مساوی



نگاره ۴: تراکم جمعیت ایران (۱۳۷۵) بر اساس روش فاصله مساوی

به عبارت دیگر فواصل طبقات در این روش بصورت نامنظم بدست می‌آیند. این روش شامل:

- تعیین فواصل طبقات با استفاده از روش کوانتال<sup>(۷)</sup> (در سیستم تعیین فواصل با استفاده از روش کوانتال هر طبقه دارای فراوانی مساوی می‌باشد)

- روش شکست طبیعی<sup>(۸)</sup> (اساس این روش فواصل طبیعی است که بین مقادیر وجود دارد)

- روش مساحت‌های برابر<sup>(۹)</sup> (در این روش تعیین فواصل بین طبقات بر اساس مساحت محدوده‌ها مشخص می‌شود).

### ۳- روش تعیین فواصل بطور سیستماتیک

این روش شامل روشهایی است که با استفاده از آنها فواصل طبقات بطور سیستماتیک و با استفاده از یک رابطه منطقی ریاضی تعیین می‌گردد. روشهای مورد استفاده بشرح زیر می‌باشد:

الف: روش تعیین فاصله طبقات بر اساس تصاعد حسابی

ب: روش تعیین فاصله طبقات بر اساس تصاعد هندسی

### بحث و نتیجه گیری

با توجه به نقشه‌های ترسیم شده و همچنین جدول شماره ۵ ملاحظه می‌گردد که با استفاده از هر روش طبقه بندی، نقشه‌ای نسبتاً متفاوت تهیه شده است. در حقیقت تفاوت موجود در فراوانی طبقات هر سیستم طبقه بندی، بخوبی ویژگی‌های هر روش را نشان می‌دهد.

از طرفی ویژگی آمارهای جمعیتی و سایر آمارهای اجتماعی و اقتصادی ایران نشان دهنده این واقعیت است که در یک مجموعه داده آماری، تعداد زیادی از داده‌ها دارای ارزش پائین می‌باشند و تعداد بسیار کمی از داده‌ها دارای ارزش بالا هستند. در خصوص داده‌های مورد استفاده در این تحقیق نیز این ویژگی دیده می‌شود، بطوریکه تراکم جمعیت در برخی مناطق مانند تهران فاصله زیادی با نزدیکترین استان یعنی گیلان دارد. بر این اساس تهیه نقشه‌ای که مناطق با تراکم بسیار کم و بسیار زیاد را بخوبی نشان دهد نسبتاً مشکل است. بعنوان مثال در روش فواصل مساوی، گل داده‌ها در سه طبقه قرار گرفته‌اند بطوریکه ۲۶ استان آن در یک طبقه قرار دارد و دو طبقه نیز بدون فراوانی می‌باشند. این روش علیرغم ساده بودن به علت پراکندگی و ناهمگن بودن داده‌ها (در مثال این تحقیق) روش مناسبی برای نمایش اطلاعات نیست. در عین حال در روش شکست طبیعی بدلیل اینکه طبقات از فواصل واقعی موجود بین داده‌ها تبعیت می‌کنند روش مناسب‌تری بنظر می‌رسد. در این روش بخوبی فواصل واقعی بین پرتراکم‌ترین استانها یعنی استانهای مازندران، گیلان و در نهایت تهران نمایش داده شده است. روش مساحت‌های مساوی نیز بدلیل آنکه اساس آن مساحت محدوده‌ها است و نه خود داده‌ها، نمی‌تواند داده‌ها را به طرز مناسبی نمایش دهد، علیرغم اینکه به ظاهر فراوانی داده‌ها در بین طبقات از توزیع خوبی برخوردار است. در روشهای تصاعد حسابی و هندسی بدلیل اینکه فاصله طبقات سریعاً افزایش می‌یابد شاهد طبقه بندی نسبتاً مناسبی

هستیم زیرا همانطور که گفته شد داده‌های دارای ارزش کم، بیشتر از داده‌های دارای ارزش زیاد می‌باشند. وجود فراوانی صفر نیز در تصاعد حسابی به دلیل افزایش سریعتر فاصله طبقات در این روش می‌باشد. بنابراین ما شاهد توزیع مناسبتری از داده‌ها در روش تصاعد هندسی هستیم. وجود فراوانی‌های مساوی و بی ارتباط در روش طبقه‌بندی کوانتال نشان از ناتوانی این روش در نمایش داده‌ها در این تحقیق است. تنها حسن این روش توزیع مناسب داده‌ها در بین طبقات است. اما قرار گرفتن داده‌های بین ۷۶ تا ۵۳۹ در یک طبقه از نقاط ضعف این روش می‌باشد. همانطور که در نگاره شماره ۵ ملاحظه می‌گردد در روش انحراف معیار سه طبقه دارای فراوانی صفر است. بدین ترتیب ملاحظه می‌گردد که روش انحراف معیار نیز روش مناسبی برای نمایش داده‌های این تحقیق نیست.

### منابع

- ۱- آلدنر، هنری ال؛ راسلر، ادوارد بی. مقدمه‌ای بر احتمالات و آمار، ترجمه عباسعلی زالی، جمشید جعفری شبستری، دانشگاه تهران: تهران، ۱۳۷۳.
- ۲- رضوانی، علی اصغر. نقشه‌ها و نمودارهای موضوعی در جغرافیای انسانی. تهران، پیام‌نور، ۱۳۷۴.
- ۳- غضنفری، سید بهداد. ترکیب معلومات کاربران معیارهای دقت و پیچیدگی نقشه در کار توگرافی تماتیک، نشریه نقشه برداری، سال ششم، شماره ۲۴.
- ۴- معافی، محمود؛ وحدانی تبار، حسن. مهارت‌های جغرافیایی، تهران: مدرسه، ۱۳۷۹.
- ۵- مهدوی، مسعود. آمار و روشهای تجزیه و تحلیل داده‌ها در جغرافیا. قومنس: تهران، ۱۳۷۳.
- ۶- مرکز آمار ایران، سالنامه آماری ایران، ۱۳۸۲، تهران: مرکز آمار ایران، ۱۳۸۳.
- 7- Anson. R. W, Basic Cartography, U.S.A, elsevier, 1988.
- 8- Friendly, Michael. Denis, Daniel. Milestone in the history of thematic Cartography, Statistical Graphics, and data Visualization, 2005.
- 9- Robinson, Arthur H and Others, Element of Cartography, Canada, john Wiley & Sons, 1995.
- 10- Stams, W (1980). National and Regional atlases, Netherland Ica.

### پی‌نوشت

- 1) Choropleth
- 2) Graduated Color
- 3) Thematic Cartography
- 4) Class Interval
- 5) Constant Class Interval
- 6) Equal Class intervals
- 7) Quantile
- 8) Natural Break
- 9) Equal Area