

به نام خداوند بخشنده‌ی مهربان

تحلیل داده‌ها

پروپوزال پروژه‌ی درس

اساتید درس: دکتر احیائی - دکتر میرصادقی

محمد مهدی فاریابی میلاد آقا جوهری
۹۳۱۰۱۹۵۱ ۹۴۱۰۵۴۷۴

اسفند ماه ۱۳۹۶



عنوان پروژه

تعیین موقعیت مکانی مناسب ایجاد کسب و کار جدید با استفاده از داده‌های مکانی شهر، توزیع جمعیت و کسب و کارهای فعلی

هدف پروژه

هدف از انجام این پروژه تولید راهکاری نرم‌افزاری داده‌محور برای ارائه‌ی مشاوره‌ی تجاری به متولیان کسب و کارهای جدید است. محصول نهایی باید قابلیت پیشنهاد موقعیت‌های مناسب و ناب تجاری را به مشتریان داشته باشد. در پروژه‌ی این درس تمرکز ما بر روی تحلیل وضع مکانی فعلی کسب و کارها و مراکز درمانی، تفریحی و غیره است تا بتوانیم شهود مناسبی از چیدمان فعلی آن‌ها ارایه کنیم.

مقدمه

امروزه با توجه به ظهور انواع و اقسام کسب و کارهای نو در کنار کسب و کارهای قدیمی‌تر، ایجاد شهرک‌های مسکونی جدید و متمرکز شدن تعداد زیادی از کسب و کارهای فعلی در نقاطی مشخص از شهرها، میل به یافتن موقعیت مکانی مناسب برای ایجاد کسب و کاری جدید یا شعبه‌ی جدیدی از یک کسب و کار فعلی موجود به شدت احساس می‌شود. مکانی که در آن تعداد کسب و کارهای مشابه و در نتیجه رقابت کمینه بوده و تقاضای مردم برای ایجاد چنین کسب و کاری بیشینه باشد. از این رو نیاز به سیستمی هوشمند و داده‌محور برای پیشبینی و معرفی چنین موقعیت‌ها و ظرفیت‌هایی احساس می‌شود.

چکیده‌ی فعالیت‌های پیش رو

به منظور دستیابی به این هدف و تولید چنین سیستمی، اطلاعات کسب و کارهای فعلی و همچنین توزیع مکانی آنها و همین‌طور توزیع جمعیتی شهر استخراج خواهد شد. در ابتدا سعی ما بر تحلیل هرچه بهتر و دقیق‌تر و مشخص‌تر چیدمان فعلی کسب و کارها در فضای شهر تهران است تا بتوانیم اطلاعات مفید و خلاصه‌ای در مورد آن‌ها ارایه کنیم. در صورتی که فاز اول با موفقیت اجرا شد سعی ما بر این است که و سپس‌های مناسب کسب و کار جدید با استفاده از الگوریتم‌هایی استخراج شده و به متولیان کسب و کار توصیه خواهد شد. سعی می‌شود این موقعیت‌ها، با توجه به نیاز مردم با بازدید بالایی همراه باشند و اطلاعات مالی و حقوقی تملک آن مکان نیز در نظر گرفته شود. در ادامه و به عنوان مسیری برای یک راهکار تجاری می‌توان سیستم را توسعه داد به گونه‌ای که داده‌هایی مانند مسیرهای حمل و نقل و وضع ترافیک و کوتاه‌ترین مسیرهای دسترسی به موقعیت را نیز در نظر بگیرد. البته در این درس تمرکز اصلی ما بر روی گام اول و تحلیل هرچه دقیق‌تر داده‌های فعلی است و بخش دوم را انشالله به صورت مفصل‌تر بعدها در شرکتی که به کمک هم تاسیس خواهیم کرد پی می‌گیریم.

چالش‌ها

۱. استخراج اطلاعات مشاغل:

برای استخراج این اطلاعات از سرویس رایگان Google Places API استفاده خواهیم کرد. با استفاده از این سرویس می‌توان اطلاعات کامل ۱۵۰ هزار موقعیت را در روز استخراج کرد.

۲. استخراج اطلاعات جمعیت:

برای استخراج این اطلاعات از گزارش‌های مرکز آمار ایران و همچنین داده‌های موجود در وبگاه "آماربستا" استفاده خواهیم کرد. همچنین می‌توان از کتابخانه‌ی Metadata در R برای به دست آوردن این اطلاعات استفاده کرد.

۳. کار با اطلاعات مکانی در R:

برای کار با اطلاعات مکانی از package‌هایی مانند rgeos، maptools، raster و sp استفاده خواهیم کرد.

۴. تحلیل داده‌ها:

برای تحلیل داده‌ها و به دست آوردن نتایج مورد نظر از ابزارهایی همچون spatstat spatial و spatgraphs بهره خواهیم برد.

۵. نمایش اطلاعات به دست آمده:

برای نمایش اطلاعات به دست آمده به صورت بصری از بسته‌های R ای مانند RgoogleMaps dismo و googleVis استفاده خواهیم کرد.

- [1] <http://amarista.ir/content/statistics/386/>
- [2] <https://developers.google.com/places/>
- [3] https://www.rdocumentation.org/packages/googleway/versions/2.2.0/topics/google_places
- [4] <https://github.com/SymbolixAU/googleway>
- [5] <https://stochasticcoder.com/2016/04/12/using-r-and-google-places-api-geocode-locations/>
<http://www.nickeubank.com/gis-in-r/>
- [6] <http://leafletjs.com/>
- [7] <https://cran.r-project.org/web/packages/dismo/dismo.pdf>
- [8] <https://cran.r-project.org/web/packages/RgoogleMaps/RgoogleMaps.pdf>
- [9] <http://pakillo.github.io/R-GIS-tutorial/>
- [10] <https://cran.r-project.org/doc/contrib/intro-spatial-rl.pdf>