



Städer och stora datamängder
SNS Analys - Edward Glaeser

Edward Glaeser

- Professor i nationalekonomi vid Harvard University
- Världsledande forskare inom urbanekonomi
- Författare av Stadens Triumf (The Triumph of the City)



Big Data och urbanekonomi

- Historisk skiljelinje mellan arkitektur och samhällsvetenskap; stora datamängder erbjuder nya möjligheter
- Rapporten belyser möjligheter som stora datamängder erbjuder, både för fysisk stadsplanering men även för policyunderlag
- Här är det viktigt att skilja på syfte!
 - Kausalitet eller korrelation?

Taxonomi av datakällor

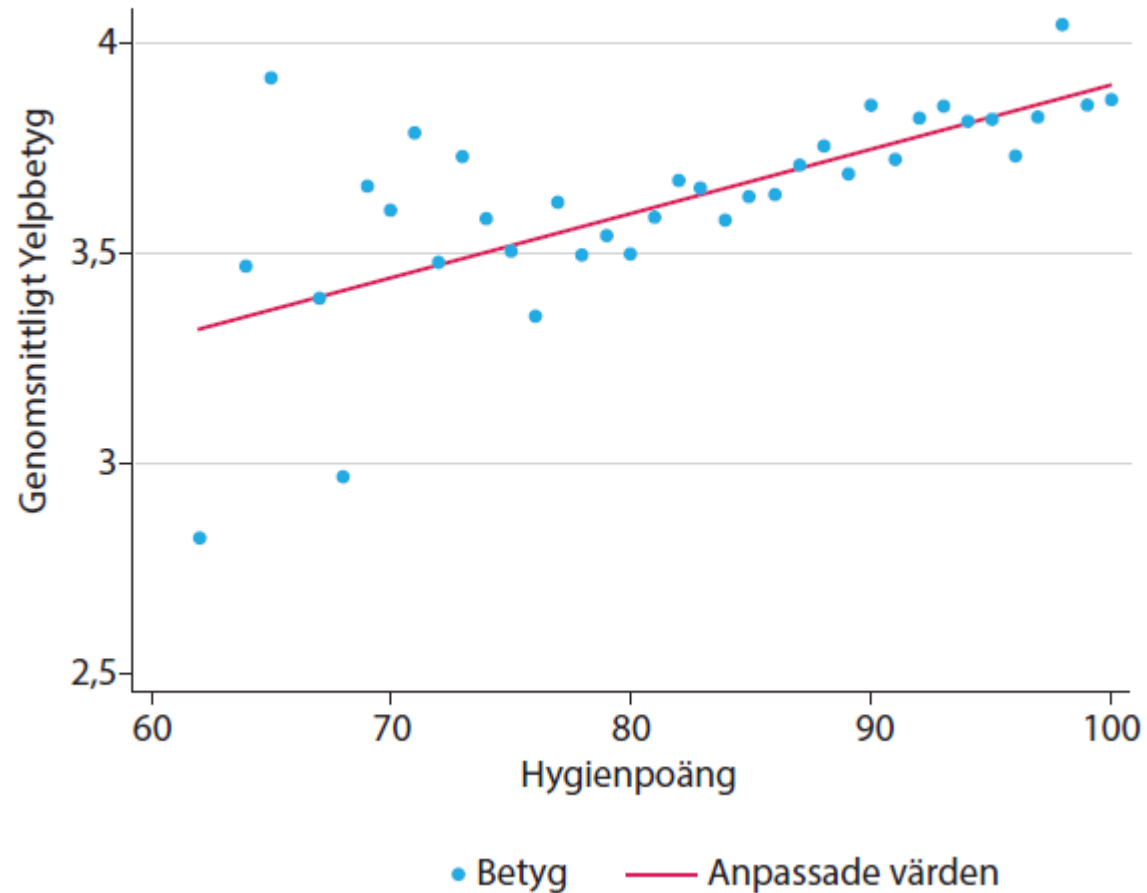
- Digitala utsläpp
 - Recensioner från Yelp kan hjälpa att rikta kontroller av hygienstandard hos restauranger.
 - Google sökningar kan prediktera influensautbrott
- Offentliga handlingar
 - Öppna datakällor
 - Ökad transparens
 - Uppmuntrar entreprenörer
- Företagsdata
 - Gymmedlemskap för att förstå hälsobeteenden
 - Data från teleoperatörer kan visa hur folk rör sig i en stad

Hur kan ny data användas i staden

- Tre aktiviteter där städer kan dra stor nytta av mer dataanalys
- Utvärdering av och fastställande av policy och reglering
 - Kan bredda frågeställningar, ex. skatteförändring hotell
 - Ökade mätfrekvenser, ex. arbetslöshet LinkedIn
- Prognostisering
 - Framtida ekonomisk aktivitet. Boende, turism, och arbetslöshet
- Drift av offentlig service
 - Vilka fall ska följas upp? Prediktionsproblem!
 - Ex. hälsoinspektioner

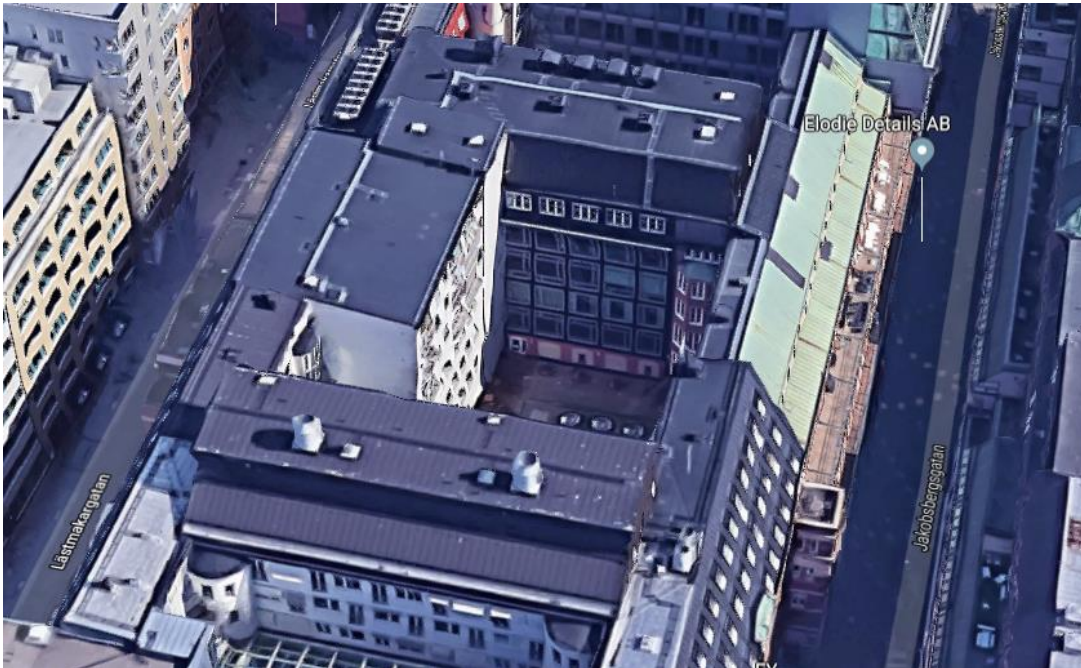
Hälsoinspektioner

Korrelation mellan Yelpbetyg och hygieninspektörers poäng.



Att mäta gatubilden

- Google street view i kombination med computer vision-algoritmer
- En del bilder går att länka ihop med gps-kodade attribut (till exempel bostadspriser eller inkomst)
- Träningsdatamängd → Utfallsdatamängd



- Kan gatubilder användas för att predicera inkomst?
- Kan gatubilden visa hur egenskaper i byggnader och omgivningar påverkar bostadspriserna? Hedonisk prissättning

Prediktion med pixlar

- Bilder och inkomst: **Kan gatubilder användas för att prediktera inkomster?**
 - Den tränade modellen fungerar bra i andra kontexter än tränings- och utfallsprovsdatamängderna. Boston och New York
 - Inkomstmättet som predikterats enbart av bilderna fångar 77 procent av variationen i det sanna inkomstmättet
- Bilder och bostadspriser: **Kan gatubilder predicera hur vissa egenskaper i byggnader påverkar bostadspriser?**
 - Sakerna som kan ses från gatan (alltså i bilderna) har ungefär lika stark förmåga att prediktera bostadspriserna som sakerna som inte kan ses från gatan

Nutidsprognosticering av gentrifiering

- Olika sätt att mäta gentrifiering, försäljningspriser av bostäder, demografiska faktorer och StreetScore (trygghetsmått)
- Bostadspriser
 - Priserna ökar i områden med till exempel fler caféer
- Demografiska faktorer
 - Korrelationer mellan andelen med eftergymnasial utbildning och antalet caféer, restauranger, frisörer, vinbarer, närbutiker, florister med mera
- StreetScore
 - Ökat antal vegetariska restauranger och Starbucks-caféer var associerat med högre trygghet

Värdet av stora datamängder

- Direkt förbättra beslutsunderlag
- Störst potential för stadsforskning när kopplas ihop med exogen variation
- Komplement till officiellt producerad statistik