

Explorativ dataanalys (EDA)

KERSTIN VÄNNMAN

Högskolan i Luleå

1. Inledning. I många situationer stöter man på siffror, ofta samlade i en stor hop. Det kan vara i tidningar eller böcker i form av tabeller eller diagram. Det kan vara mätvärden som man själv samlat in. Meningen är att man ska kunna läsa ut något från siffrorna. Eftersom vi människor tänker i bilder är det viktigt att man omformar information serverad i siffror till lättförståeliga bilder. Det gäller att skaffa sig enkla metoder som får informationen hos siffrorna att träda fram och som gör siffrornas budskap synligt.

John W. Tukey lanserade under 1970-talet EDA (explorativ dataanalys), som innehåller många nya och lättfattliga sätt att hantera siffermaterial och upptäcka mönster och samband. EDA-metoderna kompletterar de traditionella sätten. De lyfter ofta fram andra intressanta egenskaper hos material än vad de traditionella metoderna gör. Tukey beskriver själv EDA som ett numeriskt eller grafiskt detektivarbete, en jakt efter ledtrådar för att kunna upptäcka viktiga samband och strukturer. Tukey betonar vikten av grafiska metoder som gör att informationen tränger sig på.

Några av de enkla EDA-metoderna är stam-bladdiagram och låddiagram, användbara just när man vill få en samling siffror att berätta något. Dessa finns beskrivna i [1]. Metoderna går ut på att ordna siffror och rita okomplicerade figurer. De bygger inte på krångliga formler eller besvärliga uträkningar.

Utgående från [1] kan ett flertal specialarbeten göras. Här följer två förslag.

2. Att utforska en del av verkligheten med stam-bladdiagram, lådagram och jämförelser. Läs och arbeta igenom kapitel 1 - 3 i [1] för att få bakgrundskunskaper. Välj sedan den del av verkligheten som känns intressant att studera, t ex returpapper, skolmaten, bekämpningsmedel, vattenförbrukning, fel på bilar, kroppsmått förr och nu, idrottsresultat, etc. Med hjälp av de enkla EDA-metoderna gör du sedan ditt detektivarbete. Här får du träning i att arbeta med siffermaterial, se och upptäcka likheter och olikheter i form av läge och variation, göra jämförelser och presentera information överskådligt.

Om du vill ha tips på uppgifter att arbeta med så gå till kapitel 5 i [1]. Där finns förslag på många olika uppgifter att arbeta med. Dels finns uppgifter med givet siffermaterial, dels uppgifter till vilka man själv producerar eller söker material.

3. Att transformera till insikt. Läs och arbeta igenom kapitel 1 - 4 i [1]. I kapitel 4 lär du dig hur man med hjälp av logaritm- och potensfunktionen kan vrida och vända på sitt siffermaterial för att bättre kunna dra slutsatser. Att arbeta med transformationer på detta sätt ger fördjupade insikter om logaritm- och potensfunktionen. Lös övningarna i kapitel 4 och uppgifterna 521 och 522 om du inte har några intressanta material som behöver transformeras.

Litteratur

[1] Vännman, K. och Dunkels, A., *Boken om kreativ statistik med EDA*. Förlagshuset Gothia, Göteborg, 1984.