

Anvisningar för makrogranskning

- med praktiska exempel i SAS VA

Innehåll

Anvisningar för makrogranskning	1
- med praktiska exempel i SAS VA	1
Introduktion till outputgranskning	1
Förutsättningar	1
Allmänt	1
Tekniska förutsättningar.....	2
Övrigt.....	2
Principer och antaganden.....	2
Vilka metoder finns tillgängliga?.....	3
Utvärdera och återkoppla.....	4
Praktisk vägledning för SAS VA.....	5
Beställa uppsättning av en ny undersökning i SAS VA.....	5
Beställa behörighet till befintlig undersökning i SAS VA.....	5
Kort beskrivning av datamaterial och SAS program	5
Skapa en rapport i SAS VA	6
Aggregatmetoden.....	8
Övrigt stöd	12

Introduktion till makrogranskning

Makrogranskning innebär kontroll av resultatet av statistikproduktionen, det vill säga granskning av den statistik som ska levereras eller publiceras. Granskningen kan avse kontroller av storheter i form av punkt- och osäkerhetsskattningar alternativt fördelningar av mikrodata.

De fel som upptäcks i makrogranskningen kan härröra från insamlingen (kvarvarande fel efter uppgiftslämnar- och mikrogranskningen) eller från fel som uppstått vid bearbetningar (t.ex. att fel datafiler, vägningstal, metoder etc. har använts). Makrogranskning bidrar även till kunskap om datamaterialet som är viktig för de statistikansvariga inför publicering och som förklaring till förändringar. Resultatet av granskningen kan då utgöra ett underlag till analysen av produktionsomgången och underlätta tolkningen av statistiken.

Det är emellertid viktigt att skilja på makrogranskningen och analysen av statistikvärden som fyller olika syften. Detta beskrivs vidare i Verksamhetsstödet 6.5 *Tolka och förklara* och 6.6 *Fastställ innehåll för redovisning och kommunikation*. När punkt- och/eller osäkerhetsskattningar i makrogranskningen identifierats som misstänkta gäller det att hitta det eller de objekt eller bearbetningssteg som kan ha gjort att estimatet blivit misstänkt.

Makrogranskning är nödvändig för att hitta mätfel med stor effekt som inte upptäckts tidigare i granskningsprocessen och för att upptäcka fel i bearbetningen. Att fokus ligger på prioriterade kvalitetskrav är viktigt ur effektivitetssynpunkt. Det minskar risken för att alltför mycket arbete läggs ned på enskilda objekt som inte har så stor inverkan på den slutgiltiga statistiken.

En utgångspunkt är att granskning bör ske så nära källan som möjligt. Detta innebär att man i första hand ska satsa på en effektiv uppgiftslämnargranskning, mikrogranskning och/eller selektiv granskning för att åtgärda fel i datamaterialet. Makrogranskningen ska alltså endast hitta fel som kvarstår efter tidigare granskningsprocesser och som påverkar statistikens kvalitet.

Förutsättningar

Allmänt

Makrogranskning sker i många produkter under en mycket begränsad tidsperiod mellan insamling och publicering, vilket gör det än viktigare att den är effektiv. Makrogranskningen ska till så liten del som möjligt vara personberoende och den ska utföras på jämförbart sätt mellan produktionsomgångarna. Förutsättningarna för hur makrogranskningen kan utföras skiljer sig åt beroende på typ av undersökning och vilken data som finns tillgänglig. Exempelvis råder skilda förutsättningar för undersökningar som har historik som går att använda i granskningsprocessen jämfört med engångsundersökningar. Att granskningen ska vara effektiv och ske på ett standardiserat sätt gäller dock alltid. Eftersom makrogranskning bör ta en relativt liten del av resurserna ska inte implementeringskostnader och förvaltning av vare sig metoder eller verktyg vara för resurskrävande.

Om insamlingen sker löpande under en längre period bör en större del av granskningsresurserna läggas på mikrogranskning, som sker löpande under insamlingens gång. Om däremot större delen av mikrodata kommer in mer eller mindre samtidigt, och om tiden mellan avslutad insamling och publicering är kort, får makrogranskningen större betydelse.

För makrogranskning gäller precis som för selektiv granskning av mikrodata att man måste kunna ta fram förväntade värden för de skattningar som ska granskas. Vid makrogranskning behöver man alltså ta fram lämpliga värden att jämföra med. I många fall är det naturligt att jämföra med skattningar i tidsserier, men ibland kan det också vara värt att samgranska med produkter med jämförbar output.

Granskningens omfattning måste anpassas till prioriterade kvalitetskrav för respektive undersökning. Om outputen är så stor att en prioritering måste göras mellan skattningarna behöver denna förstås samordnas med den prioritering som görs vid den selektiva mikrogranskningen om sådan utförs.

Tekniska förutsättningar

För en effektiv makrogranskning bör uppgiftslämnar- och mikrogranskningen vara dokumenterad, så att det går att följa de tidigare stegen i granskningsprocessen och att utnyttja information från tidigare insamlingsomgångar. Kommentarer och noteringar på objektsnivå som sparats i t.ex. Triton kan tas med till makrogranskningen. När man utreder ett aggregat ska man kunna se alla observationer som bidrar, inklusive eventuella kommentarer från tidigare granskningsaktiviteter.

Vi rekommenderar att flaggningen av aggregat är regelstyrd och att den görs i förväg innan materialet tas in i till makrogranskningen. Att dokumentera sin makrogranskning löpande är nödvändigt för att minimera risken att granska redan genomgången material.

Övrigt

Det är viktigt att ha dokumenterade regler för flaggning samt att det finns stöd för att dokumentera makrogranskningen när den utförs. På så sätt kan man följa vad som flaggats och vad som konstaterats vara korrekt/acceptabelt respektive korrigerats mellan olika tillfällen. Om flera personer arbetar med makrogranskningen måste denna information vara tillgänglig för samtliga.

I den löpande produktionen ska sedan checklisten för outputgranskning användas för att säkerställa att alla relevanta aktiviteter genomförs. Den hittas i verksamhetsstödet 6.2. I avsnitt 3 i checklisten finns det aktiviteter som försöker hitta olika typer av bearbetningsfel som kan vara svåra att upptäcka i exempelvis SAS VA. Detta är en viktig del i outputgranskning som kan vara svår att hantera enbart i en strikt visuell miljö.

Principer och antaganden

Att känna till användarnas behov och deras bedömning av vilka aggregat som är särskilt viktiga är en förutsättning för att kunna göra prioriteringar i granskningen. Detta är avgörande för att kunna konstruera en effektiv granskningsprocess. Det kan även finnas specificerade kvalitetskrav för vilken

kvalitet statistikvärdena måste uppnå. Dessa ska då finnas med i utformandet av processen.

Om kvalitetskraven inte är specificerade kan man dock anta att dessa är högre för mer övergripande aggregat, för att sedan minska i takt med att allt mindre delar av populationen redovisas. Det kan därför vara klokt att ordna sina aggregat på ett sådant sätt att man kan prioritera de mest övergripande aggregaten först, sedan de näst mest övergripande aggregaten o.s.v. Hur detta kan automatiseras beskrivs i dokumentet *Ordna redovisningsgrupper metodbeskrivning* som hittas på analyseras [SharePoint ytan granska makrodata under filen arbetsdokument](#) i verksamhetsstödet.

Vilka metoder finns tillgängliga?

Det finns framförallt två övergripande metoder för makrogranskning, fördelnings- och aggregeringsmetoden. Bägge metoderna kan hitta fel i mikrodata men endast aggregatmetoden kan hitta fel i bearbetningsprocessen.

Fördelningsmetoden

Med fördelningsmetoden jämför man observationerna i inkommen mikrodata för en period med varandra för att se om något värde avviker från den fördelning datamaterialet i övrigt har. De värden som avviker utreds sedan vidare. Detta förfarande har uppenbar likhet med mikrogranskning. Skillnaden är att det är lättare att hitta avvikelser när majoriteten av all data har kommit in. Lådagram kan vara ett sätt att illustrera detta, samt spridningsplottar mellan två variabler eller samma variabel för två tidpunkter. Det ska helst vara enkelt att gå från de olika diagrammen till underliggande mikrodata genom att direkt klicka i diagrammen.

Aggregeringsmetoden

Aggregeringsmetoden innebär att man ser över de siffror som ska publiceras, dvs. skattningar i tabellerna, och jämför med förväntade värden. Det kan exempelvis vara motsvarande skattning från tidigare publiceringsomgångar i kombination med ett förväntat utvecklingstal. Granskningen kan utföras en tid före publicering med hjälp av preliminära uppräkningsvikter och ett inte helt fullständigt datamaterial. Om ett misstänkt aggregat upptäcks är nästa steg att undersöka underliggande mikrodata för att hitta orsaken till det misstänka värdet. "Top down" betecknar förfarandet att man borrar sig ned från aggregatet till avvikande aggregat på finare nivå och slutligen ned till eventuellt felaktiga mikrovärden.

Båda dessa metoder, i synnerhet aggregeringsmetoden, förutsätter att en stor del av datamaterialet har kommit in innan makrogranskningen kan börja. De två metoderna kompletterar i viss mån varandra men de kan också vara två olika vägar att gå för att hitta samma fel i ett enskilt datavärde.

Granskning bör i så stor utsträckning som möjligt vara styrd av förbestämda regler och det gäller även makrogranskning. Detta för att säkerställa en pålitlig och jämn process med så få subjektiva bedömningar som möjligt. De förutbestämda reglerna skapar fellistor på ett likande sätt som i mikrogranskningen där misstänkta aggregat finns med någon typ av prioritetsordning. Ett lämpligt sätt att rangordna misstänkta aggregat är med hjälp av en poängfunktion där funktionen ska ta i beaktning misstankegraden

av aggregatet och aggregatets prioritet. Ett exempel på en enkel poängfunktion är att jämföra ett aggregat för referenstiden med ett förväntat värde i kombination med ett osäkerhetsintervall. Det förväntade värdet kan vara baserat på en framskrivning av föregående referenstids skattade värde för motsvarande aggregat. En enkel framskrivning kan vara ett medelvärde av historisk förändring för aggregatet ett antal perioder bakåt.

Om $\hat{X}_t > \tilde{X}_{t,\text{övre gräns}}$ eller $\hat{X}_t < \tilde{X}_{t,\text{undre gräns}}$ beräknas poängfunktionen som:

$$\text{Poängfunktion} = \frac{|\hat{X}_t - \tilde{X}_t|}{|\tilde{X}_{t,\text{övre gräns}} - \tilde{X}_{t,\text{undre gräns}}|} \times PF$$

\hat{X}_t = Aggregat för referenstid t

\tilde{X}_t = Förväntat värde för aggregatet vid referenstid t , där genomsnittlig förändring för aggregatet under r perioder används i kombination med värdet för period $t-1$ (föregående period). Exempelvis,

$$\tilde{X}_t = \hat{X}_{t-1} * \frac{\sum_{n=1}^{n=r} \hat{X}_{t-n}}{r}$$

$\tilde{X}_{t,\text{övre gräns}}$ & $\tilde{X}_{t,\text{undre gräns}}$ = förväntade värden för aggregatets referenstids övre och undre intervallgränser för vad som anses vara misstänkt. Här kan exempelvis medelfelens storlek för aggregatet \hat{X}_{t-1} användas för att beräkna en övre och undre gräns för det förväntade värdet \tilde{X}_t .

PF = prioritetsfunktion för hur viktigt aggregatet är för publicering.

Med en poängfunktion är det möjligt att prioritera vilka aggregat som ska utredas först eller mer noga. Poängfunktionen skulle även kunna användas till att flagga aggregat som överstiger ett framtaget tröskelvärde.

Utvärdera och återkoppla

Granskning är en process med delaktiviteter som sträcker sig över flera avdelningar där de olika aktiviteterna ska komplettera varandra. För att stödja arbetet med att ständigt förbättra processen behövs underlag som ska hjälpa till med analysen. Det gäller att på ett systematiskt sätt samla in information och där så är möjligt återkoppla och åtgärda problemen tidigare i processen.

Utvärdera

Det som ska kunna utvärderas är om granskningsaktiviteten är effektiv och om det är möjligt att fånga felanvändningen tidigare i processen. Beroende på hur outputgranskningen är konstruerad får man logga aktiviteten på olika sätt. Används regler för att flagga olika aggregat går det att jämföra antalet flaggningar före och efter outputgranskningen. Okulär besiktning av aggregat bör undvikas i den mån som går både på grund av subjektiviteten i bedömning av misstänkta fel men även för att kunna sammanställa granskningsaktiviteten systematiskt med hjälp av processdata.

Återkoppla

För att ständigt förbättra både undersökningen och granskningen gäller det att skapa systematiska arbetssätt för att kunna tillvarata lärdomar för att förbättra nästkommande produktionsomgång. Hittas bearbetningsfel ska dessa åtgärdas i den pågående produktionen, och om möjligt ska man åtgärda källan till felet. Ett exempel på det skulle kunna vara att ersätta manuella steg där det lätt kan bli fel med automatiserade procedurer.

Praktisk vägledning för SAS VA

Syftet med den här vägledningen är att demonstrera hur makrogranskning kan göras med aggregatmetoden och hur man går tillväga för att skapa en enkel granskningsrapport i SAS VA. Vägledningen förutsätter egentligen ingen bakgrundskunskap i SAS VA men det är heller ingen detaljerad beskrivning av samtliga moment som krävs för att skapa rapporten.

För att kunna återskapa rapporten måste först en beställning av en ny undersökning i SAS VA göras. Alternativt, för att bara titta på granskningsrapporten i exemplet, går det att söka läsrättigheter i beställarportalen.

Beställa uppsättning av en ny undersökning i SAS VA

I beställarportalen finns tjänsten "Upplägg av ny undersökning i SASVA", en ny undersökning i SAS VA sätts upp genom att fylla i och skicka in formuläret.

Beställa behörighet till befintlig undersökning i SAS VA

Görs i beställarportalen av den person som vill ha behörighet.

Beställarportalen => Ansök om behörigheter

Klicka in i fältet *Grupp* och sök efter *sasva*. Välj sedan aktuell undersökning.

OBS: För varje undersökning finns det två grupper. En för läsrätt (=väljs av rapportanvändare) och en för skrivrätt (=väljs av rapportskapare och dataladdare, dvs. PowerUsers).

För att se exempelrapporten sök följande behörighet.

PMU/PPA Läs på \\sasvaviya\fa\data\OutpGr

Kort beskrivning av datamaterial och SAS program

Det datamaterial som används är kvartalsdata från ett urval av fiktiva företag och innehåller både granskade och ogranskade data. Det kvartal som är tänkt att makrogranskas är fjärde kvartalet 2016. I det här exemplet har vi valt att låta variabeln omsättning (oms) för några företag i Bransch E gå ogranskad genom tidigare delprocesser i granskningen.

Totalskattningar av omsättning tas fram för varje kvartal och bransch. Totalt finns sju redovisningsgrupper, Domän A, B, C, D, E, F-G och A-G. I A-G ingår samtliga branscher. Skattningarna tas fram med hjälp av ETOS.

I stort sett samtliga bearbetningar som behövs för att färdigställa tabellerna som kontroller, flaggning av observationer, utredning och ändringar i data och tabellberäkningar görs i SAS. En del av detta kan göras direkt i SAS VA istället.

För att kunna borra oss ner i datamaterialet så måste vi ordna en koppling mellan aggregat och mikrodata med lämpliga kategoriska variabler. De observationer som ingår i flera redovisningsgrupper behöver förekomma med en rad för varje redovisningsgrupp i mikrodatatabellen för att kopplingen ska bli korrekt. Makro- och mikrodata koppas ihop med variablerna *kvartal* och *Bransch*.

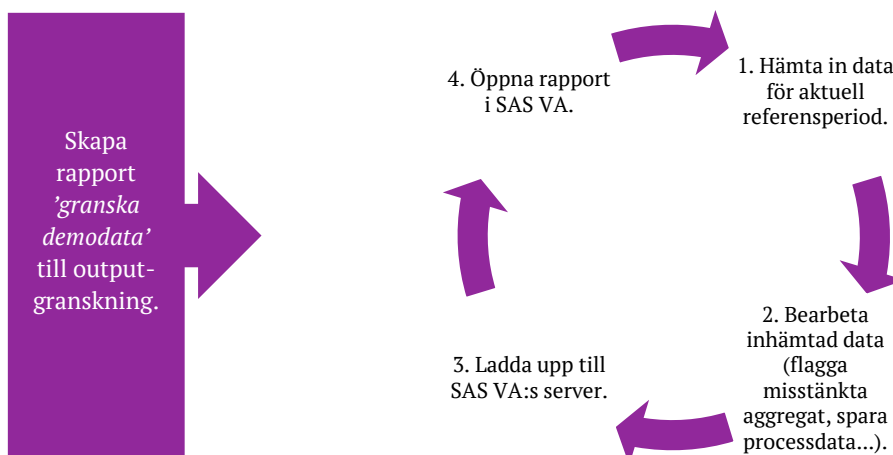
Oavsett vilken metod som används, aggregatmetoden eller fördelningsmetoden, så är det viktigt att vi loggar processen så att vi kan se vad som flaggas när makrogranskningen börjar och vad som flaggas vid publicering. I exemplet har vi valt att skapa en regelstyrd indikator, flagga, så att vi kan jämföra om makrogranskningen har gett någon effekt. När outputgranskningen börjar är det därför lämpligt att spara en första datummärkt version som möjliggör utvärdering av outputgranskningsprocessen.

SAS program och testdata finns tillgängliga under analyseras [SharePoint yta granska makrodata under filen arbetsdokument](#) i SPS.

Skapa en rapport i SAS VA

För att använda SAS VA som verktyg i makrogranskningen är det första steget att bygga upp en rapport via *Utforska och visualisera* i SAS VA. Denna rapport kan sedan återanvändas löpande i produktionen. Ett exempel på hur rapporten hör ihop med produktionsprocessen kan ses i den schematiska bilden nedan.

Schematisk bild över outputgranskning med SAS VA i produktionen.



För att ladda upp tabeller till SAS VA skapar vi ett libname i SAS på följande sätt.

```
libname Demo '\\A387\f\data\&undersokning.\autoload';
```

```
libname Demo '\\TSTA232\F\data\&undersokning.\Autoload';
```

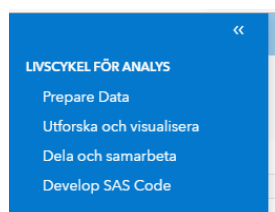
SAS VA kommer vi åt med följande länkar.

Produktionsmiljön <https://sasvaviya.scb.intra>

Testmiljön <https://sasvaviyatest.scb.intra>

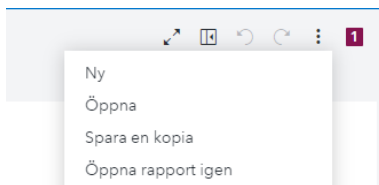
Inloggning sker via Single-Sign-On, genom att besöka webblänken sker inloggning automatiskt.

För att skapa eller redigera en redan befintlig rapport går vi in via Utforska och visualisera.



Figur 1

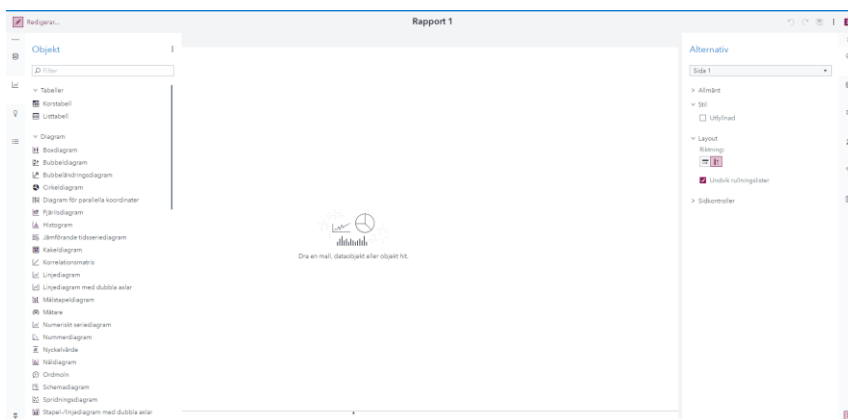
För att titta på en befintligt rapport går vi in via *Öppna* i menyns högra sida.



Figur 2

Aggregatmetoden

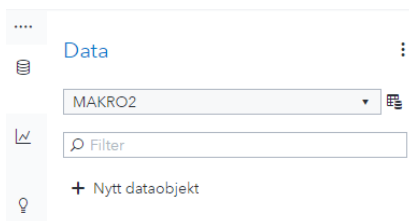
Vi börjar med att skapa en sida med listtabell och tidsseriediagram. När man går in via *Utforska* och *visualisera* kommer man till följande vy.



Figur 3

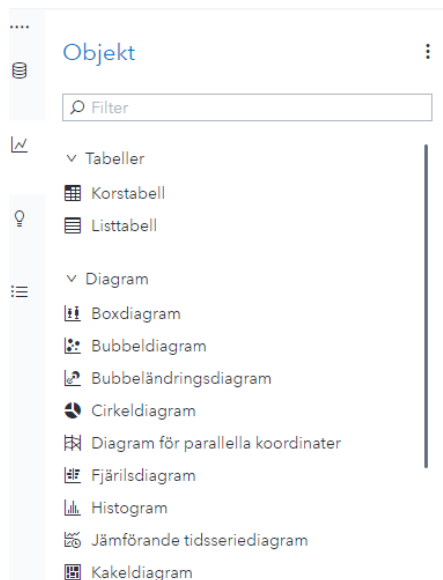
Rapporter byggs upp av sidor där varje sida kan innehålla ett eller flera objekt. Ett objekt kan exempelvis vara en tabell eller ett diagram.

I menyn till vänster hanterar vi datakällor, och väljer objekt. I menyn till höger finns olika alternativ för att redigera ett enskilt objekt, som t.ex. utseende, filtreringar, skapa länkar mellan objekt i en sida, länkar mellan sidor och visningsregler osv.



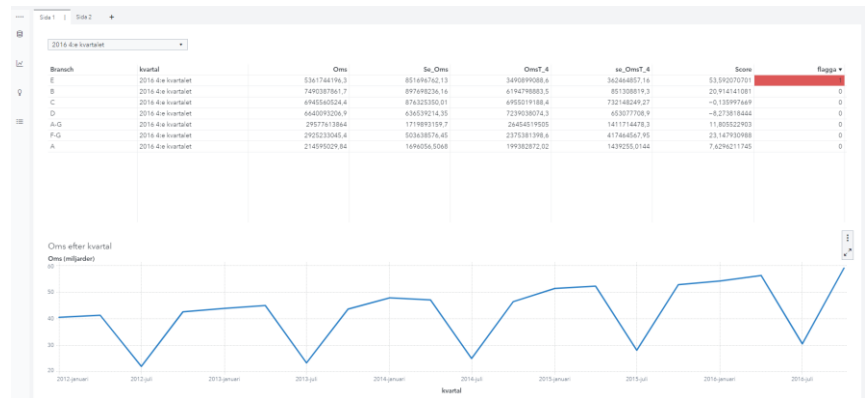
Figur 4

Innan vi börjar bygga upp en ny granskningsrapport måste vi hämta de datakällor vi behöver, via *Data* symbolen, i den vänstra menyn, och väljer sedan att lägga till en datakälla. De datakällor som används i detta exempel är *MAKRO1*, *MAKRO2* och *MIKRO*.



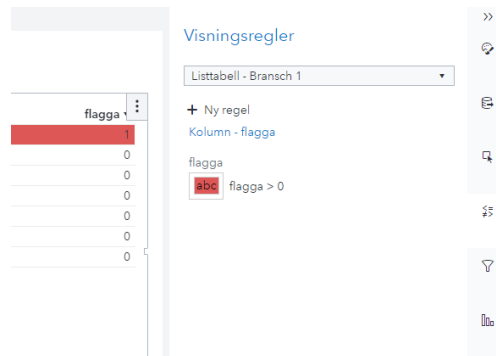
Figur 1

I den vänstra menyn väljer vi *objekt* och sedan en *listtabell*, ett *tidsseriediagram* och en *lustruta* (den lilla rutan uppe till höger i bilden). För att skapa Figur 2 nedan behöver vi två olika datakällor, *MAKRO1* och *MAKRO2*. Förutom att de har olika namn så är tabellerna identiska. I vissa fall krävs flera identiska datakällor för att kunna koppla data inom ett avsnitt och filtrera objekten oberoende av varandra. I fliken *Data* kan vi sen välja vilka variabler vi vill använda och dra in dem i valda objekt. *Listrutan* används för att filtrera *listtabellen* efter *kvartal*.



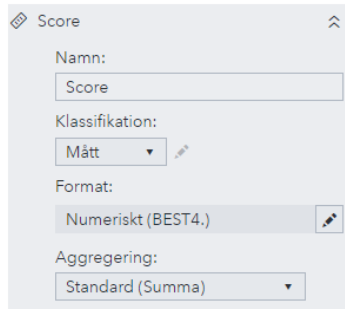
Figur 2

För att upptäcka misstänkta aggregat har vi i SAS skapat en poängfunktion, *Score*. *Score* är procentuell förändring av oms aktuellt kvartal jämfört med samma kvartal ett år tidigare. I tabellen sorteras *Score* efter storleksordning. Flagga = 1 indikerar att aggregatet är misstänkt och att *Score* överstiger vår godtyckligt valda tröskel på 30. *Bransch E* sticker ut från de övriga och är det enda aggregatet som flaggas. Från Figur 2 kan vi borra oss vidare ner till mikrodata som ingår i *Bransch E*.



Figur 3

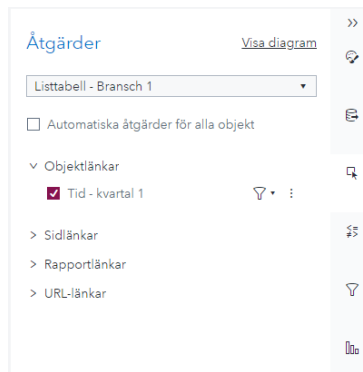
Förutom att vi har variabeln *flagga* som indikerar ett misstänkt aggregat eller observation går det i SAS VA att färgmarkera rader eller celler. För att göra det går vi till fliken *visningsregler* i menyn till höger.



Figur 4

För att formatera variabler går vi till *Datafilen* och markerar en variabel och väljer sedan ett lämpligt format.

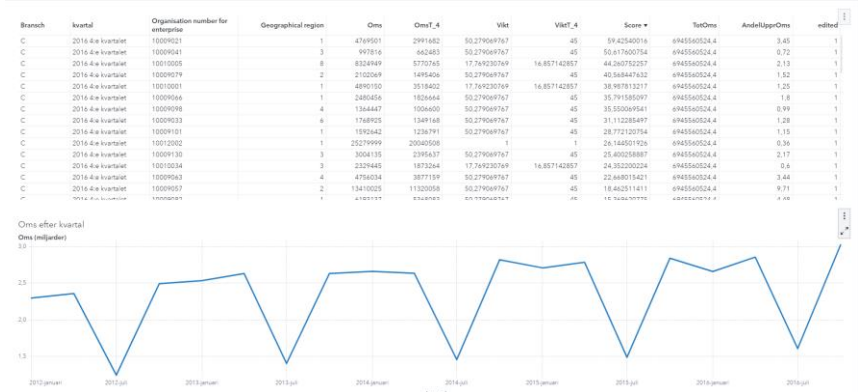
Listtabellen och *tidsseriediagrammet* i Figur 2 kopplas ihop med hjälp av menyn *Åtgärder* nedan.



Figur 9

I det här fallet baseras länkningen på *Bransch*, eftersom vi vill se samtliga kvartal för vald bransch i *tidseriediagrammet*

En koppling mellan två sidor görs på liknande sätt via *Åtgärder*. Kopplingen görs via *Sidlänkar* och genom att mappa på *kvartal* och *bransch*. I rapporten länkas *listtabellen* i sida 1 med *listtabellen* i sida 2. Tabellerna i avsnitten måste färdigställas innan kopplingen kan göras. Variablerna som ska användas i länkningen måste finnas i den tabell man kopplar från.



Figur 10. Mikrodata kopplat till selekterat aggregat (Bransch)

I Figur 10 använder vi *Mikro1* som datakälla. För att på ett överskådligt vis utreda vilka företag det är som påverkar det misstänkta aggregatet tas olika mått fram. I exemplet ovan används *Score*, kolumnerna i tabellen går att sortera på, vidare syns några valda stödvariabler uppräkningsstat (*Vikt*), andelar av skattning (*AndelUpprOms*). Listtabellen länkas till mikrodata, på så sätt finns möjlighet att analysera ett objekts rapportering över tid.

Övrigt stöd

Dessa exempel som har tagits fram i SAS VA är ett försök till att visa hur verktyget kan användas vid regelstyrd outputgranskning. Viktigt att komma ihåg är att spara undan nödvändig data för att löpande kunna utvärdera sin outputgranskning. Via SAS finns ytterligare [dokumentation](#).

Viktigt att komma ihåg är att SAS VA är ett verktyg som möjliggör både outputgranskning och analys. Att använda samma verktyg för dessa aktiviteter kan medföra vissa synergieffekter som exempelvis lägre förvaltningskostnader för produkten. Med det sagt så kan det finnas andra lämpliga verktyg som kan användas till outputgranskning.