

Tema 3: Statistikproduktionen: Checklista för outputgranskning

Martin Ribe och Jörgen Svensson

martin.ribe@scb.se, jorgen.svensson@scb.se

Statistiska centralbyrån

Bearbeta & Analysera

Abstract. Statistiska centralbyrån (SCB) har tagit fram en checklista för outputgranskning. Den avser makrogranskning av statistikresultat som ska publiceras eller levereras till kund. Grundsytet med checklistan är att minska risken för fel i statistiken. Checklistan omfattar dels moment i förberedande skede före produktion, dels moment i genomförandet av produktionen. Fokus ligger på att kunna identifiera och följa upp misstänkt avvikande värden i framställd statistik. Effektivast är dock att kunna förebygga så att det inte uppstår utförandefel som påverkar värden, och checklistan har även detta syfte. Checklistan innefattar kontroller enligt krav för uppfyllande av standarden ISO 20252.

1. Outputgranskningens roll

Outputgranskning är den delprocess i statistikproduktionen som granskar statistikresultat som ska publiceras eller levereras till kund. Den ska säkra statistikens avsedda tillförlitlighet genom att upptäcka fel i statistikresultaten med orsaker som:

- Felaktiga värden i indata genom mätfel.
- Misstag eller felaktigt processutförande i statistikproduktionen.

Outputgranskningen ger underlag för flera slags åtgärder:

- Korrigering av upptäckta fel i den aktuella produktionsomgången.
- Kommentarer i redovisningen av statistikresultaten.
- Kvalitetsbeskrivning av statistiken.
- Förbättring av produktionsprocesser för att minska risker för fel.

Engelsk benämning är ”output validation” eller ”output editing”.

Plats på processkartan

På ”processkartan” över statistikproduktionen enligt modellen av METIS (2009, sida 9 och 18f.) ingår denna process som delprocess 6.2 Validate output. Outputgranskningen görs efter beräkningen av statistikresultaten (delprocess 6.1). Den kan sägas komplettera mikrogranskningen (delprocess 5.3), som görs före beräkningen av statistikresultaten (jfr Biemer & Lyberg, 2003, avsnitt 7.4.2). Statistiska centralbyrån (SCB) tillämpar för övrigt en modell snarlik den av METIS i utveckling, förvaltning och dokumentation av processer.

Metodlitteraturen

Makrogranskning är ibland benämning på outputgranskning men avser ofta *selektiv granskning* i samband med mikrogranskningen (jfr t.ex. Granquist, 1995; Hedlin, 2003). I den internationella metodlitteraturen synes outputgranskning som sådan allmänt sett mindre utförligt behandlad än mikro-

granskning, där just selektiv granskning har blivit uppmärksammas. Man kan även notera användningen av s.k. output-orienterade mått på effekt av granskning (Manzari & Della Rocca, 1999; Nordbotten, 2000). Genom parallellerna i syfte och metodik mellan mikro- och outputgranskning bör granskningslitteraturen vara indirekt relevant även för outputgranskning. Yu (2003) framhåller outputgranskningens roll:

Many survey areas in business statistics have established clearance process to assure the statistical output prior to publication but there is wide variation in the coverage and presentation of quality information presented to support the process. ... Similarly, for household surveys there is a need for strengthening the output validation process to ensure a timely focus on the quality of outputs.

2. Kvalitetskrav

Behovet av en checklista för säkrad systematik i outputgranskningen accentuerades för SCB genom två faktorer:

- SCB:s plan för certifiering enligt standarden ISO 20252.
- Uppmärksammas fel genom misstag i publicerad statistik.

ISO-standard SCB anpassar nu sina rutiner enligt den internationella standarden ISO 20252 för marknads-, opinions- och samhällsundersökningar (ISO, 2007). Där ingår krav på hur statistikresultat och tabeller ska kontrolleras.

Uppdrag av regeringen Efter ett programfel som påverkade konsumentprisindex (KPI) uppdrog regeringen åt SCB (2008) att redovisa åtgärder för att stärka säkerhetsrutinerna. Åtgärderna innefattar bland annat testrutiner, loggning av systemändringar, uppföljning av fel och uppföljning av avvikande värden, varav det sistnämnda berör outputgranskning.

3. Checklista för outputgranskning – utformning

3.a. Checklistans idé

Checklistan som har tagits fram av SCB (2009) avser att säkra systematiken i outputgranskningen av statistikresultaten. Outputgranskningen ska genom checklistan vara förberedd till att omfatta behövliga moment, så att den kan fånga in felaktiga resultat och korrigera dem innan de hinner redovisas och nå kunderna. Checklistan finns tillgänglig för processutförarna på SCB genom *Verksamhetsstödet*, ett intranätbaserat system för att tillhandahålla stöddokument för statistikproduktionen. Den ingår där som ett av verktygen för att säkra kvaliteten i statistiken och processerna.

Struktur De moment som enligt checklistan ska gås igenom omfattar åtgärder både före och efter produktion:

- Före produktion – Skapa stöd
- I produktion – Identifiera misstänkt avvikande värden
- I produktion – Följ upp misstänkt avvikande värden
- I produktion – Hantera slutsatser

Misstänkt avvikande värden Det centrala för checklistan är att ge stöd för att identifiera och följa upp *misstänkt avvikande värden* i framställd statistik. Med *misstänkt avvikande värden* menas här statistikvärden som så avviker från vad som kan anses förväntat eller normalt, att det kan misstänkas föreligga antingen svarsfel i indata eller fel/misstag i utförande av någon process.

Dessutom innehåller checklistan vissa övriga kontroller, såsom kontroll av fullständighet. Dessa kan delvis ses som en ytterligare säkring av leveranskontroller som hör till efterföljande processer, närmast delprocess 7.2.

3.b. Före produktion – Skapa stöd

Momenten före produktion inom varje statistikprodukt går ut på att utforma ett produktspecifikt stöd för outputgranskningen av produkten. Detta stöd kan ha formen av en produktspecifik checklista för outputgranskningens genomförande och mera specifikt peka ut kontroller som ska gås igenom.

Punkter ur checklistan Exempel på punkter inom detta avsnitt av checklistan är följande:

- ...
- 1.2 Specificera eventuella saklogiska samband som alltid ska råda mellan värden i statistiken. Exempel: En summapost som alltid ska stämma med summan av delposter.
- ...
- 1.5 För rimlighetskontroll av värden: Avgör om resultat från tidigare produktionsomgångar ska användas att jämföra med. Ange i så fall hur dessa resultat och produktionsomgångar ska väljas, och vilka värden som ska jämföras med dem.
- ...
- 1.8 Om möjligt, formulera kriterier för när värden ska bedömas som *misstänkt avvikande* (se definition i inledningen till denna checklista).
- ...

3.c. I produktion – Identifiera *misstänkt avvikande värden*

För varje produktionsomgång utförs outputgranskning genom att ett antal punkter gås igenom. Därvid nyttjas det produktspecifika stödet som har tagits fram för produkten, enligt det föregående.

Punkter ur checklistan Exempel på punkter inom detta avsnitt av checklistan är följande:

- ...
- 2.3 Genomför outputgranskningen med hjälp av det produktspecifika stödet och identifiera eventuella *misstänkt avvikande värden* (jfr punkt 1.8).
- 2.4 Bedöm översiktligt om resultaten i output verkar rimliga utifrån ämneskännedom, allmän erfarenhet och sunt förnuft (dvs. här bedömt mera fritt, ej direkt bundet av specifika kriterier).
- ...

3.d. I produktion – Följ upp *misstänkt avvikande värden*

För varje *misstänkt avvikande värde* som identifierats i outputgranskningen gås ett antal punkter igenom, med sikte på endera av slutsatserna:

- att processerna bedöms ha utförts korrekt
- att fel eller misstag i processutförande har identifierats.

Punkter ur checklistan Exempel på punkter inom detta avsnitt av checklistan är följande:

3.1 Om det går att förmoda någon särskilt trolig orsak till att värdet avviker: Utred i så fall om denna orsak kan verifieras, utifrån information från tidigare processer.

3.2 Ta reda på om det har förekommit någon driftstörning. Om så är fallet, utred då driftstörningens möjliga konsekvenser för processerna i fråga, och gör eventuellt om databehandlingssteg.

...

3.8 Ta reda på om någon ändring nyligen har gjorts i produktionsrutiner eller system. Om så är fallet, utred om ändringen har utförts korrekt. Anlita vid behov metod- och IT-kompetens.

...

3.e. I produktion – Hantera slutsatser

Efter genomförd outputgranskning i en produktionsomgång gås några punkter igenom, med tanke på eventuella kvarstående frågor.

Punkter ur checklisten Exempel på punkter inom detta avsnitt av checklisten är följande:

4.1 Om det har identifierats något fel eller misstag i hur processer har utförts, eller något annat oväntat fel: Utred felets orsak och hur det ska undvikas i fortsättningen.

...

4.4 Bedöm om tillräcklig klarhet har uppnåtts om orsakerna till eventuella *misstänkt avvikande* värden. Om så inte är fallet, dokumentera kommentarer om osäkerhet i statistiken på grund av detta. I eventuell kommentar för publicering, tänk på att skydda enskilda uppgiftslämnarens integritet.

4. Uppföljning och vidare utveckling

Checklisten för outputgranskning infördes i juni 2009 och ska tillämpas på alla statistikprodukter i SCB:s produktion. Erfarenheter från tillämpningen i praktiken avses tas till vara för checklistans vidare utveckling och lansering.

5. Referenser

- P.P. Biemer, L.E. Lyberg (2003), Introduction to Survey Quality, Hoboken: Wiley.
- L. Granquist (1995), Improving the traditional data editing process, in Business Survey Methods, Ed. by B.G. Cox et al., New York: Wiley, sida 385-401.
- D. Hedlin (2003), Score functions to reduce business survey editing at the U.K. Office for National Statistics, Journal of Official Statistics, 19, 177-199.
- ISO (2007), Market, Opinion and Social Research – Vocabulary and Service Requirements (ISO 20252:2006, IDT), International Standards Organisation.
- A. Manzari, G. Della Rocca (1999), A generalized system based on simulation approach to test the quality of editing and imputation procedures, Working Paper No. 13, UNECE Work Session on Statistical Data Editing, Rome.
- METIS (2009), Generic Statistical Business Process Model, Version 4.0, Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata.
- S. Nordbotten (2000), Evaluating efficiency of statistical data editing: General framework, återgiven i Statistical Data Editing, Volume No. 3 Impact on Data Quality, Geneva: UN Statistical Commission and UNECE, 2006, sida 3-16.
- SCB (2008), Redogörelse för hur säkerhetsrutinerna kan förbättras i statistikproduktionen, Bilaga 2: Säkerhetsrutiner i statistikproduktionen, SCB Dnr 2008/1743, Fi2008/6314.
- SCB (2009), Checklista för outputgranskning, Intern checklista, Stockholm/Örebro: Statistiska centralbyrån.
- F. Yu (2003), Making data quality visible in practice, Proc. Statistics Canada Symposium on Challenges in Survey Taking.