

Metodbeskrivning KKiK-rapport

Inledning

Det här dokumentet ska på en övergripande nivå beskriva de olika delarna som utgör de script som skapar de kommununika KKiK-rapporterna med AI-stöd.

Tre script

Rapporten bygger på tre olika script samt en del externa filer som läses in. De externa filerna innehåller exempelvis metadata som inte är tillgängliga via Koladas API och manuellt satta variabler som styr nyckeltalsurval och liknande.

De tre scriptens huvudfunktioner är

- Hämta och förbereda data
- Generera rapportens layout och innehåll (diagram och tabeller)
- Loop-script för att generera en rapport per kommun

Scripten är i stor utsträckning skrivna i språket R med inslag av HTML och CSS på de delar som utgör själva rapporten.

1. Datainsamling och bearbetning

Det första scriptet hämtar data och förbereder den i största möjliga utsträckning för rapportscriptet.

Förändring

Scriptet genomför alla jämförelser mellan åren. Utgångspunkten är att jämförelsen ska göras mellan år T och T-1 (där år T är det valda fokusåret för rapporten). Om denna jämförelse inte är möjlig tillämpas följande prioriteringsordning:

År T	År T-1	År T-2	År T-3	Jämförelse
Finns	Finns			T mot T-1
Finns	Finns inte	Finns		T mot T-2
Finns inte	Finns	Finns		T-1 mot T-2
Finns inte	Finns	Finns inte	Finns	T-1 mot T-3
Finns inte	Finns inte	Finns	Finns	T-2 mot T-3
Finns	Finns inte	Finns inte	Finns	Jämförs ej
Finns inte	Finns	Finns inte	Finns inte	Jämförs ej

Förändringen beräknas som procentuell förändring från föregående period till senaste period. Exempelvis:

$$\frac{'Värde \text{ år } T' - 'Värde \text{ år } T_{-1}'}{'Värde \text{ år } T_{-1}'} \cdot 100$$

Samma förändringsberäkning görs oavsett vilken enhet nyckeltalet är angivet i. I de ovanliga undantagsfall där en kommun gått från ett värde på 0 till något annat än noll innebär beräkningen ovan en nolldivision. I dessa fall exkluderas nyckeltalet från jämförelse. Fortsatt utveckling av jämförelsemetoden är planerad.

Exempelkommuner

I rapporten kan exempel på kommuner som haft fördelaktig utveckling inom ett nyckeltal listas i anslutning till nyckeltal där fokuskommunen haft ofördelaktig utveckling. Listan på dessa kommuner tas fram per nyckeltal. Listan riskerar att inte innehålla samtliga nyckeltal som ursprungligen hämtades.

Ett nyckeltal blir utan exempelkommuner om:

- Nyckeltalet inte uppfyller kravet på minimalt antal år med data. Om alla kommuner för ett specifikt nyckeltal inte har minst tre år exkluderas
- Nyckeltalet saknar värdering i om det är fördelaktigt eller inte med låga värden.
- Ingen kommun har en konsekvent fördelaktig utveckling.

Endast kommuner med liknande befolkningsstorlek får listas som 'bra exempel'. Detta för att mindre kommuner inte ska jämföras med storstäder och tvärt om. I nuläget är det fokuskommunens 75 närmsta grannar sett till befolkningsstorlek som tillåts vara med som potentiella 'bra exempel' – givet att de också uppfyller kraven för att anses vara just 'bra exempel'. De 75 närmsta grannarna kan förstås ändå vara rätt avlägsna. Speciellt om fokuskommunen är en av de, sett till befolkning, minsta eller största kommunerna. Följande kriterier måste uppfyllas för att en kommun ska kunna listas som ett 'bra exempel':

- Den ska ha data för minst tre av fyra år.
- Har den data för fyra år ska tre av dessa gå i fördelaktig riktning varav det senaste året måste vara ett av dessa.
- Har den data för tre år ska samtliga gå i fördelaktig riktning.
- Kommunen måste tillhöra de 25 % av kommunerna med mest fördelaktigt resultat det sista året.
- Kommunen ska finnas med i fokuskommunens grupp med de 75 kommuner som ligger närmast befolkningsmässigt.

Listan sorteras därefter vilka som gjort bäst fördelaktig förändring och topp tre väljs ut.

Läge

Scriptet beräknar också om en kommuns nyckeltalsvärde över flera år legat på en anmärkningsvärt fördelaktig eller ofördelaktig nivå sett till övriga kommuners värde och taggar dessa i en tabell.

För att en kommun-/nyckeltalkombination ska taggas som 'fördelaktig/ofördelaktig över tid' ska:

- Nyckeltalets värde, beroende på om höga eller låga värden är fördelaktiga, hör till de tio procenten mest fördelaktiga eller ofördelaktiga jämfört med andra kommuner. Värdet ska ha hört till de tio procenten senaste året och totalt minst tre av fyra år i tidsserien.
- Om tidsserien endast omfattar ska tre år ska nyckeltalets värde ligga bland de tio procent mest fördelaktiga eller ofördelaktiga under samtliga av dessa tre år.

2. Generering av Rapport

Rapporten genereras i RMarkdown med HTML som primär output. Efter att HTML-filen skapats tas också pdf-versionen fram.

För samtliga av rapportens AI-genererade texter har OpenAI's modell 'gpt-4o-mini' använts. Temperaturinställningen är satt mellan 0,1 och 0,5.

Nedan beskrivs rapportens enskilda delar och vad de består av:

Förord

En statisk text, identisk för alla kommuner.

Om rapporten

En statisk text, identisk för alla kommuner.

Sammanfattning

Beskriver varje nyckeltal som nämns under avsnitten 'Ofördelaktiga förändringar sedan förra året', 'Fördelaktiga förändringar sedan förra året' och 'Nyckeltal med anmärkningsvärda nivåer under längre tid'. Vissa undantag görs dock. En extern lista läses in där de nyckeltal som får vara med i sammanfattningen har vitlistats – exempelvis är kostnadsnyckeltal uteslutna från sammanfattningen.

Textstyckena är AI-genererade och beroende på varför det enskilda nyckeltalet omnämnts i sammanfattningen så skiljer sig prompten som AI-svaret utgår ifrån.

För detaljerad redogörelse av de prompter som ligger till grund för sammanfattningen se bilaga 1 (fördelaktig utveckling eller läge) och bilaga 2 (ofördelaktig utveckling eller läge).

Tabellsidorna

Tabellsidorna är skapade för att efterlikna Kolada-verktyget Jämföraren. Samma tabellordning och samma nyckeltalsordning i tabellerna eftersträvas.

En sak som skiljer sig från Koladas Jämförar-sida är förutom nyanserna p bakgrunden även kolumnen "Förändring" där förändringsvärdet redovisas. Om en asterisk ('*') står intill förändringsvärdet betyder det att förändringen är beräknad på två år före referensåret (som angivs i kolumnen år) och inte ett år som är den eftersträlvade standarden. Asterisken visas alltså när data saknas för det närmast föregående året.

Bakgrundsfärgerna i kolumnerna 'Värde' och 'Förändring' har olika betydelser:

- **Bakgrundsfärgerna för värdekolumnen** - Färgerna visar hur resultatet förhåller sig till andra kommuner. Grönt representerar de 25 % av kommunerna med mer fördelaktiga resultat, rött de 25 % med mindre fördelaktiga resultat, och grått de 50 % däremellan. Notera att färgerna endast visar relativa skillnader mellan kommuner - ett objektvt bra resultat kan få röd färg om andra kommuner har ännu mer fördelaktiga värden.
- **Bakgrundsfärgerna för förändringskolumnen** - Grön avser en förändring i fördelaktig riktning, röd bakgrund avser en förändring i ofördelaktig riktning. Bakgrunden lämnas vit när ingen förändring har skett eller om nyckeltalet inte har en värdering om höga eller låga värden. Förändringen avser procentuell förändring oavsett nyckeltalets enhet.

Fördelaktiga förändringar sedan förra året

Avsnittet listar de fem nyckeltal som gjort störst fördelaktig förändring sedan föregående tidsperiod.

Förutom att visa kommunens data i jämförelse med värdet för 'Alla kommuner (ovägt medel') finns också en AI-genererad text som kommenterar kommunens värden.

För detaljerad redogörelse av prompten som ligger bakom AI-svaret se bilaga 3.

Ofördelaktiga förändringar sedan förra året

Avsnittet följer samma struktur som föregående avsnitt men i stället listas de fem nyckeltal där kommunen gjort störst ofördelaktig utveckling sedan föregående tidsperiod.

Det här avsnittet har också AI-genererad text och använder sig av samma prompt som föregående avsnitt.

Nyckeltal med anmärkningsvärda nivåer under längre tid

Det här avsnittet finns nödvändigtvis inte med i alla rapporter. Har kommunen inga nyckeltal där den stabilt legat bland de 10 % mest fördelaktiga eller 10 % mest ofördelaktiga under en längre tidsperiod utesluts avsnittet i sin helhet.

Minst ett nyckeltal som uppfyller kriterierna behövs för att avsnittet ska inkluderas. Det finns dock ingen övre gräns hur många nyckeltal som redovisas. En AI-genererad text beskriver läget.

För detaljerad redogörelse av prompten som ligger bakom AI-svaret se bilaga 4.

Bilaga 1

```
prompt_text <- paste0(

  "Här följer information och data om en kommun och ett nyckeltal: ",

  "Kommunens namn är: ", selected_municipality, ". ",

  "Det aktuella nyckeltalet är: ", kpi_name_in_front, ". ",

  "Här är kommunens nyckeltalsvärden för de senaste åren", ": ", x, ". ",

  ifelse(u == TRUE,

    paste0(

      "Kommunens värde har under de senaste åren legat ",

      "på en fördelaktigt ",

      ifelse(z == 0, "hög ", "låg "),

      "nivå i relation till andra kommuner. Detta är viktigt att nämna."

    ),

    paste0(

      "Kommunens värde har, när vi ser på förändringen ",

      "från föregående år",

      ifelse(z == 0, "ökat ", "minskat "),

      "vilket anses fördelaktigt. Detta är viktigt att nämna."

    )

  ),

  "Det är ",

  ifelse(z == 0, "fördelaktigt", "ofördelaktigt"),

  " om värdet har ökat jämfört med föregående år. ",

  "Det är alltså ",

  ifelse(z == 0, "fördelaktigt ", "ofördelaktigt "),

  "med höga värden.",

  "Sammanfatta följande information med fokus på de viktigaste punkterna:",

  "1. Börja med att beskriva utvecklingen i kommunens nyckeltal och huruvida ",

  "uppgångar eller nedgångar är fördelaktiga eller ofördelaktiga.",

  "2. Kommentera även på utvecklingen mellan det näst sista året och det senaste, ",

  "och om den utvecklingen har varit fördelaktig eller ofördelaktig.",

  "Använd en professionell och informativ ton. Undvik kraftuttryck och ordet 'trend'. ",

  "Börja utan rubrik eller ingress, och undvik konkreta förslag till förbättring.",

  "Håll texten kort, cirka 75 ord. Varje gång nyckeltalet nämns, använd dubbla asterisker runt namnet, till exempel: **", kpi_name_in_front, "**. ",

  "Använd kommatecken som decimaltecken vid värden som innehåller decimaler.",

  "Använd mellanslag som tusenavskiljare. ")
```

Bilaga 2

```
prompt_text <- paste0(

  "Här följer information och data om en kommun och ett nyckeltal: ",

  "Kommunens namn är: ", selected_municipality, ". ",

  "Det aktuella nyckeltalet är: ", kpi_name_in_front, ". ",

  "Här är kommunens nyckeltalsvärden för de senaste åren", ": ", x, ". ",

  ifelse(u == TRUE,

    paste0(

      "Kommunens värde har under de senaste åren legat ",

      "på en ofördelaktigt ",

      ifelse(z == 0, "låg ", "hög "),

      "nivå i relation till andra kommuner. Detta är viktigt att nämna. "

    ),

    paste0(

      "Kommunens värde har, när vi ser på förändringen ",

      "från föregående år ",

      ifelse(z == 0, "minskat ", "ökat "),

      "vilket anses ofördelaktigt. Detta är viktigt att nämna. "

    )

  ),

  "Det är ",

  ifelse(z == 0, "fördelaktigt", "ofördelaktigt"),

  " om värdet har ökat jämfört med föregående år. ",

  "Det är alltså ",

  ifelse(z == 0, "fördelaktigt ", "ofördelaktigt "),

  "med höga värden.",

  "Sammanfatta informationen ovan med fokus på de viktigaste punkterna:",

  "1. Börja med att beskriva utvecklingen i kommunens nyckeltal och huruvida ",

  "uppgångar eller nedgångar är fördelaktiga eller ofördelaktiga.",

  "2. Kommentera även på utvecklingen mellan det näst sista året och det senaste, ",

  "och om den utvecklingen har varit fördelaktig eller ofördelaktig.",

  "Använd en professionell och informativ ton. Undvik kraftuttryck och ordet 'trend'. ",

  "Börja utan rubrik eller ingress, och undvik konkreta förslag till förbättring.",

  "Håll texten kort, cirka 75 ord. Varje gång nyckeltalet nämns, använd dubbla asterisker runt namnet, till exempel: **", kpi_name_in_front, "**. ",

  "Använd kommatecken som decimaltecken vid värden som innehåller decimaler.",

  "Använd mellanslag som tusenavskiljare. ")
```

Bilaga 3

```
prompt_text <- paste0(
  "Här är ett nyckeltal: '",
  kpi_name, "':", x,
  ". Medelvärde för alla kommuner var ",
  y,
  "Kommunens senaste värde är alltså ",
  ifelse(q == TRUE, "högre", "lägre"),
  " än medelvärdet för alla kommuner samma år.",
  ". Det är fördelaktigt att ha ett ",
  ifelse(z == 0, "högt", "lågt"),
  " värde. ",
  "Beskriv kommunens resultat, betona det senaste året och jämför med",
  " kommungenomsnittet. Fatta dig kort, uteslut ordet trend eller ",
  "kraftuttryck. Undvik att försöka förklara orsaker till utvecklingen. ",
  "Använd max 90 ord.",
  "Använd mellanslag som tusenavskiljare. ",
  "Innehåller ditt svar värden med decimaler ska ",
  "skiljetecknet vara ett kommatecken."
)
```

Bilaga 4

```
prompt_text <- paste0(
  "Här är värden för nyckeltalet '",
  kpi_name,
  "': ",
  x,
  " avseende kommunen ",
  selected_municipality,
  ". Det genomsnittliga värdet för Sveriges alla kommuner har under samma",
  " tidsperiod varit: ",
  y,
  ". Det är önskvärt att ha ett ,",
  ifelse(z == 0, "högt", "lågt"),
  " värde. ",
  "Skriv något om detta men ge inga konkreta förslag eller förklara orsaker",
  " till utvecklingen. Uteslut ordet trend och uteslut kraftuttryck.",
  " Använd max 100 ord. ",
  "Använd mellanslag som tusenavskiljare. ",
  "Innehåller ditt svar värden med decimaler ska ",
  "skiljetecknet vara ett kommatecken."
)
```