



# Metodikk for beregning av byindeks

Et mål på endring i trafikkmengde for et byområde

# Metodikk for beregning av byindeks

---

## Innhold

Metodebeskrivelse for byindeksen .....	2
Datagrunnlag.....	2
Beregningsmetode.....	5
Usikkerhet i indekstallene .....	6

## Metodebeskrivelse for byindeksen

Byindeksen er en metode for å estimere endringen i trafikkmengden på vegnettet i byområder. Beregningsmetoden er den samme som for vegtrafikkindeksens fylkesindeks, men det geografiske området avgrenses av de utvalgte trafikkregistreringspunktene i byområdet.

Byindeksen brukes til oppfølging av byvekstavtaler og belønningsavtaler.

### Datagrunnlag

Byindeksen bygger på datagrunnlaget som samles inn fra utvalgte trafikkregistreringsstasjoner, der utstyr kontinuerlig registrerer trafikkdata i hvert kjørefelt. En trafikkregistreringsstasjon er geografisk tilknyttet et punkt på vegens senterlinje, og registrerer all trafikk som passerer gjennom en tenkt linje som dekker alle kjørefelt på den aktuelle vegen. Dette kaller vi et trafikkregistreringsnitt. Vanligvis brukes begrepet *registreringspunkt* om stedet der det registreres trafikk i alle kjørefelt gjennom trafikkregistreringsnittet. I forbindelse med byindeks omtales punktene gjerne også som indekspunkt.

Standarden for utstyr i dag er at data om kjøretøy registreres og lagres enkeltvis. Motorsykler registreres på lik linje med andre motorkjøretøy. For indekser som beregnes til og med 2019 holdes motorsykler utenfor datagrunnlaget. Kriteriet for å filtrere ut motorsykler (o.l.) er at de klassifiseres som motorsykler (klasse 1) eller har kjøretøylengde  $\leq 1,8$  m. Fra og med 2020 vil motorsykler inngå i datagrunnlaget, både i basis- og beregningsår.

Videre består datagrunnlaget for indeksen av trafikkvolumet aggregert til timenivå. Alle data som inngår i beregningen av byindeksen skal bruke trafikkvolumet  $Q$ , fra trafikkregistreringspunkt  $j$ , aggregert til timenivå og summert over alle kjørefelt for begge kjøreretninger:

$$Q_{j,h,d,s,t}$$

for time  $h$ , dag  $d$ , måned  $s$  og år  $t$ . Indeksen sammenligner data fra beregningsåret  $t$  og basisåret  $t - 1$ .

Alle data som skal brukes i indeksen skal være faktisk registrerte data for trafikkvolum som er registrert på et godkjent indekspunkt.

### Ekskludering og matching

Metoden anvender time for time- og dato for dato-prinsippet. Ulike ukedager kan variere med hensyn på trafikkvolum, men i løpet av et år vil dette utjevne seg.

Prosessen for å lage en indeks er:

1. Velge ut godkjente trafikkregistreringspunkt.

2. Velge ut kvalitetssikrede data fra de godkjente trafikkregistreringspunktene.
3. Alle registrerte data for måned  $s$  og trafikkregistreringspunkt  $j$  sorteres etter dato og time for år  $t$ .
4. Alle registrerte data for måned  $s$  og trafikkregistreringspunkt  $j$  sorteres etter dato og time for år  $t - 1$ .
5. Alle tilsvarende timer på samme dato som har registreringer i de to årene ( $t, t - 1$ ) velges. Dette betyr at timer som ikke har verdi i begge de to årene utelates.

Data sammenlignes kjørefelt for kjørefelt, og det kreves at registreringene har vært utført på alle kjørefelt samtidig. Etter denne prosessen sitter en igjen med godkjente punkt som inneholder kvalitetssikrede data, med registrerte timer i begge årene på alle kjørefelt.

### *Kriterier for ekskludering av glisne døgn og måneder*

Når bortfall av data forekommer, stilles følgende krav:

1. Et døgn må inneholde minimum 16 matchende, godkjente timeverdier.
2. En måned må inneholde minimum 16 matchende, godkjente døgn per punkt.

### *Ekskludering av lengdeklasser*

Trafikkregistreringene omfatter lengdeklassifisering av alle kjøretøy, og tabellen nedenfor viser inndelingen i lengdeklasser. I byindeksen benyttes dette til å beregne indekser for lette og tunge kjøretøy, hvor skillet mellom lette og tunge er satt til kjøretøylengde på 5,6 m.

Lengdeklassifisering i byindeks.

Lengdegruppe	Beskrivelse
L20	Alle lengder unntatt motorsykler
L21	Lengde under 5,6 m, unntatt motorsykler
L22	Lengde mellom 5,6 m og 7,6 m
L23	Lengde mellom 7,6 m og 12,5 m
L24	Lengde mellom 12,5 m og 16 m
L25	Lengde over 16 m

Erfaring har vist at lengdeklassifiseringen ikke alltid er komplett i rådatamaterialet. Derfor må det kontrolleres om lengdeklassifiseringen er tilnærmet komplett. Alle avvik per time mellom summen L21+L22+L23+L24+L25 og L20 summeres for hvert døgn. Dersom avviket er mer enn 5 % av totalsum for L20 i hele døgnet, strykes lengdeklassifiseringen for hele dette døgnet. Dersom avviket er 5 % eller mindre beholdes lengdeklassifiseringen som den er.

Fra nytt registreringsutstyr med enkeltkjøretøy, filtreres dårlige lengdemålinger ut av grunnlagsdata. Følgende kriterier benyttes:

- Kjøretøyregistreringer med lengde under 1,0 m filtreres helt bort. Motorsykler utgjør en stor andel av disse, og de skal ikke være med i grunnlaget for byindeksen.
- Kjøretøyregistreringer med lengde over 27 m tas med i totaltrafikken, men ikke i lengdeklasseinndelingen.
- Kjøretøyregistreringer med fart under 7 km/h tas med i totaltrafikken, men ikke i lengdeklasseinndelingen.
- Kjøretøyregistreringer med godkjent nøyaktighet på fartsmåling, det vil si at fartsmålingen ikke skal variere med mer enn 3 % over sensorene.

Den samme kompletthetskontrollen som nevnt over, gjelder for disse registreringene. Kravet om minst 16 døgn i måneden med godkjent lengdeklassifisering gjelder for utregning av lengdeklasseindeks.

### Indeksdata

Etter at ekskludering og matching er gjennomført, gjenstår de dataene som danner grunnlaget for selve byindeksberegningene. Disse dataene kalles indeksdata  $D^i$ , og utgjør et utvalg, hele eller deler av, de opprinnelige rådataene  $D$ , dvs. at  $D^i \subseteq D$ . For byindeksen har vi at

$$Q_{j,s,t} = \sum_{h,d} Q_{j,h,d,s,t} \quad \exists Q_{j,h,d,s,t-1}.$$

Dette danner utgangspunktet for videre beregning av indeksene.

### Gjeldende sifre

I byindeksen benyttes trafikkvolum som er heltall for antall kjøretøy per time, summert over alle kjørefelt. Antall sifre vil her variere fra ett til fire, oftest to eller tre. Månedsvolum vil typisk summere seg til 5–7 sifre. Kravet til registreringsutstyret er at minst 99 % av trafikken skal registreres, noe som gir usikkerhet i tredje gjeldende siffer for trafikkvolumet. Når indekstall skal presenteres, skal det derfor benyttes to gjeldende sifre. Dersom indeksen blir 0,xx, må den oppgis med to desimaler, ellers med kun én desimal. I mellomregninger beholdes alle sifre som er tilgjengelig, og avrunding skjer til slutt ved å benytte Bankers metode (runde f.eks. xxx5 opp eller ned til nærmeste partall).

## Beregningsmetode

Det vaskede datagrunnlaget utgjør nå en database med verdier fra matchede timer. Først beregnes trafikkvolum og indeks for enkeltpunkt, som senere benyttes for å finne verdier for et byområde.

### Indeks per trafikkregistreringspunkt

Månedsindeksen benytter indeksdata og er definert som

$$Q_{j,s,t}^i = \frac{Q_{j,s,t}}{Q_{j,s,t-1}}.$$

Her er indeksen gitt på forholdsform. Når indeksen publiseres skal den gis på prosentform,

$$Q^{\text{ip}} = 100 \cdot (Q^i - 1),$$

med to gjeldende sifre.

En kan også regne ut indeks for en periode på flere måneder, hvor

$$Q_{j,\Sigma s,t}^i = \sum_{h,d,s} Q_{j,h,d,s,t} \quad \exists Q_{j,h,d,s,t-1},$$

som gir indeksen

$$Q_{j,\Sigma s,t}^i = \frac{Q_{j,\Sigma s,t}}{Q_{j,\Sigma s,t-1}} = \frac{\sum_s Q_{j,s,t}}{\sum_s Q_{j,s,t-1}}.$$

### Indeks for et byområde

En indeks for et byområde  $b$ , som inneholder registreringspunktene  $j$ , baseres på indeksdata

$$Q_{b,s,t} = \sum_{j,h,d} Q_{j,h,d,s,t} \quad \exists Q_{j,h,d,s,t-1}.$$

Månedsindeks for by  $b$  er da

$$Q_{b,s,t}^i = \frac{Q_{b,s,t}}{Q_{b,s,t-1}}.$$

Alternativt kan byindeksen ses på som et vektet gjennomsnitt av enkeltpunktindeksene, der de normaliserte vektene dannes av trafikkvolumene fra basisåret:

$$Q_{b,s,t}^i = \frac{1}{\sum_j Q_{j,s,t-1}} \sum_j Q_{j,s,t}^i \cdot Q_{j,s,t-1} = \frac{1}{Q_{b,s,t-1}} \sum_j \frac{Q_{j,s,t}}{Q_{j,s,t-1}} \cdot Q_{j,s,t-1} = \frac{Q_{b,s,t}}{Q_{b,s,t-1}}$$

Byindeks for flere måneder er basert på

$$Q_{b,\Sigma s,t} = \sum_{j,h,d,s} Q_{j,h,d,s,t} \quad \exists Q_{j,h,d,s,t-1},$$

som gir

$$Q_{b,\Sigma,t}^i = \frac{Q_{b,\Sigma,t}}{Q_{b,\Sigma,t-1}}$$

## Indeks over flere år

For å beregne et estimat på en indeks over flere år, kan trafikkmengden fra de aktuelle årene benyttes eller alle mellomliggende års indekser multipliseres:

$$\begin{aligned} Q_{j,s,t,t-n}^i &= \frac{Q_{j,s,t}}{Q_{j,s,t-n}} = \frac{Q_{j,s,t}}{Q_{j,s,t-1}} \cdot \frac{Q_{j,s,t-1}}{Q_{j,s,t-2}} \cdots \frac{Q_{j,s,t-1}}{Q_{j,s,t-n}} \\ &= Q_{j,s,t,t-1}^i \cdot Q_{j,s,t-1,t-2}^i \cdots Q_{j,s,t-n+1,t-n}^i \end{aligned}$$

Dette gjelder også for aggregerte indekser.

## Usikkerhet i indekstallene

Det er heftet usikkerhet ved flere ledd i kjeden fra datainnsamlingen til en aggregert indeks. Registreringsutstyret skal ifølge kravspesifikasjonen få med seg minst 99 % av alle passerende kjøretøy når alt utstyr er i normal drift. Men av og til vil vi få nedetid, og dermed mister vi deler av en måleserie.

## Fulltallighet

Registreringsutstyret legger på et løpenummer for hvert kjøretøy som registreres. Løpenummeret skal øke med 1 for hvert nye kjøretøy, og dersom noe går galt i lagring av registreringen, vil vi oppdage at enkeltregistreringer mangler når det mangler løpenumre. Vi kaller «fulltallighet» et mål på andel lagrede løpenumre, og beregner en prosentverdi for alle timer med data. For byindeksen er det krav til større enn 99 % fulltallighet per time.

## Dekningsgrad

For indeksen som baserer seg på timeverdier, vil enkelte timer mangle på grunn av nedetid eller for lav fulltallighet, og dette gjenspeiler vi i en størrelse i vi kaller dekningsgrad.

Dekningsgraden for en punktindeks  $Q_{j,h,d,s,t}^i$  er definert som

$$G_j(Q_j) = \frac{\text{antall timer med indeksdata}}{\text{antall timer i perioden}} \cdot 100 \% = \frac{h_{j,s,t}^i}{h_{j,s,t}} \cdot 100 \%$$

Dekningsgraden er altså den relative andelen i prosent av antall timer med indeksdata delt på antallet timer en skulle hatt dersom det er 100 % oppetid på registreringsutstyret. Dekningsgraden blir lavere enn 100 % dersom enten fulltalligheten er lav (under 99 %), stasjonen har vært ute av drift eller at data er manuelt merket med en hendelse som gir dårlig datakvalitet eller skjevhet i datagrunnlaget for en indeks. Skjevheter kan oppstå som følge av stengte veier og omkjøringer i forbindelse med vegarbeid.

For aggregerte indekser er dekningsgraden den sammensatte dekningsgraden på alle punktindeksene:

$$G_b(Q_b) = \frac{\sum_j h_{j,s,t}^i}{h_{j,s,t} \sum_j 1} \cdot 100\%.$$

### Statistisk variasjon

Trafikkregistreringspunktene som bidrar med data til indeksen utgjør et utvalg av alle veglenker i et område. Selv om plasseringen av registreringspunktene er valgt med omhu med tanke på datakvalitet og spredning på vegnettet, antar vi at de likevel representerer et tilnærmet tilfeldig og representativt utvalg veglenker. Punktindeksene antar vi er normalfordelte, men standardavviket er ikke kjent, og må estimeres for hvert tilfelle.

### Standardavvik

For byindeksen  $Q^i$  er et forventningsrett, vektet standardavvik

$$\sigma_{b,s,t} = \sqrt{\frac{1}{1 - \sum_j \left(\frac{Q_{j,s,t-1}}{Q_{b,s,t-1}}\right)^2} \sum_j \left(\frac{Q_{j,s,t-1}}{Q_{b,s,t-1}} (Q_{j,s,t}^{\text{ip}} - Q_{b,s,t}^{\text{ip}})^2\right)}$$

Her er det brukt at trafikkvolumvektene er normaliserte, men ikke-stokastiske.

### Konfidensintervall

Alle indekstall skal oppgis med et tosidig konfidensintervall på 95 %-nivå. Da populasjonens standardavvik er ukjent, benyttes  $t$ -fordeling. Med konfidensfaktor  $\tau$  gitt av  $t$ -fordeling, er konfidensintervallet

$$K_{b,s,t} = \tau \frac{\sigma_{b,s,t}}{\sqrt{n_{b,s,t}}},$$

hvor  $n$  er antall trafikkregistreringspunkter.

### Variansen til indeks over to år

Indeks over to år beregnes ved å multiplisere de to indeksverdiene  $Q^i$ . Variansen er da gitt ved

$$\text{Var}(Q_1^i \cdot Q_2^i) = \frac{\text{Var}(Q_1^i) \cdot \text{Var}(Q_2^i)}{n_1 n_2} + \frac{\text{Var}(Q_1^i)}{n_1} \cdot Q_2^{i2} + \frac{\text{Var}(Q_2^i)}{n_2} \cdot Q_1^{i2}.$$





Statens vegvesen  
Abels gate 5  
7030 Trondheim

Tlf: (+47)22073000  
trafikdata@vegvesen.no

vegvesen.no

**Trygt fram sammen**