

# BioQuery 4

## Vejledning i brug af BioQuery4

11. december 2006

## Indhold

1. INDLEDNING .....	3
2. OPBYGNING AF DATAUDTRÆK.....	3
2.1 Tabelnavigatoren .....	3
2.2 Feltlisten .....	3
2.3 Opstille kriterier .....	5
2.3.1 Overordnet opbygning.....	5
2.3.2 Kriteriecellen .....	6
2.3.3 Eksistens-kriterier .....	7
2.3.4 Kriterier på felter tilknyttet opslagsliste.....	7
2.4 Kør udtrækket .....	8
3. FUNKTIONER .....	10
3.1 Funktioner på tabelniveau .....	10
3.1.1 Udtræk i kolonner (fælleslister).....	10
3.1.2 Udtræk i kolonner (specifikke).....	11
3.1.3 Udtræk nyeste .....	11
3.1.4 Data skal eksistere.....	11
3.1.5 Fasthold stationsoplysninger .....	12
3.1.6 Afgræns til Fauna og/eller Fiskeundersøgelser .....	12
3.1.7 0-observationer ved udtræk af fiske-data .....	12
3.2 Unikke rækker .....	12
3.3 Grupperede udtræk.....	13

### Historik

Dokumentid.	Dato	Udarb. af	Ændringer
A346.82.1	28.11.2006	Stmik	Initial version
A346.82.2	11.12.2006	Mipdr	Rettelser/Layout ændringer efter review

## 1. Indledning

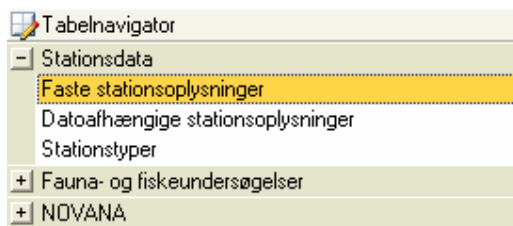
Formålet med modulet er at sætte slutbrugere i stand til at danne avancerede dataudtræk fra WinBio4-databasen (inkl. NOVANA), uden forudgående kendskab til SQL-spørgesproget, samt den tekniske opbygning af selve tabelstrukturen i WinBio4.

## 2. Opbygning af dataudtræk

Opbygning af et dataudtræk gennemløber følgende etaper (alle punkter beskrives i detaljer senere):

- Via tabelnavigatoren vælges en tabel.
- Fra feltlisten vælges én eller flere felter
- Ovenstående gentages indtil alle ønskede felter fra alle ønskede tabeller er valgt
- Opstil evt. kriterier på et eller flere af felterne
- Kør dataudtrækket
- Udfør evt. supplerende manipuleringer på det færdige dataudtræk
- Eksporter evt. data til Excel eller kolonne-fil
- Gem evt. den opbyggede forespørgsel som en rapportdefinition

### 2.1 Tabelnavigatoren




Navigatoren giver en hierarkisk oversigt over de mulige tabeller. Oversigten svarer i sin opbygning til de skærbilleder og begrebsmæssige strukturer, der er anvendt i WinBio4. Du åbner/lukker dele af navigator-træet ved at klikke på hhv. +/- knapperne i træet.

### 2.2 Feltlisten

Felter
Ejer
Stationsnr
Navn
Lokalitet
Oprettet den
Vandområde,nr
Vandområde,navn
Kortnummer
Kommune,nr
Kommune,navn

Når en tabel er valgt, vises de mulige kolonner/felter, der kan vælges fra tabellen. Du kan udvælge felter på følgende måder:


- dobbeltklik
- klik på et felt og træk det med musen

- trykke på knappen 

Medmindre du vælger at dobbeltklikke, har du mulighed for at vælge mere end ét felt af gangen:

- Hold Shift-tasten nedtrykket samtidigt med at du klikker på start/slut i et interval
- Hold Ctrl-tasten nedtrykket samtidig med at du klikker på de ønskede felter (kan også benyttes selvom der først er udvalgt et sammenhængende interval med Shift-tasten).

Herefter kan du

- klikke på et udvalgt felt og trække alle felter med musen
- trykke på knappen 


Når et eller flere felter er valgt, vil de stå i udtræks-listen:

Udvalgte kolonner								Suppl.
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier	eller	eller	Sorter
Undersøgelse	Ejer		<input checked="" type="checkbox"/>					
Undersøgelse	Stationarit		<input checked="" type="checkbox"/>					
Undersøgelse	Plævedato		<input checked="" type="checkbox"/>	01-01-1999, 31-12-1999				
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>					
Bund	Ø	Blad bund	<input checked="" type="checkbox"/>					
Bund	Ø	Grøt sten	<input checked="" type="checkbox"/>					

De første fire kolonner er:

- Navnet på den tabel informationen stammer fra
- Navnet på kolonnen
- Evt. identifikation på specifikke kolonner (se senere)
- Afkrydsning af om kolonnen skal udskrives (vises) i udtrækket eller ej.

Herefter følger kriteriecellerne og til slut Sorterings-kolonnen, hvori evt. sorteringsorden kan angives (enten stigende eller aftagende).

Du kan fjerne valgte kolonner ved at udvælge en eller flere, og herefter klikke på .


Udover de valgte kolonner rummer skærbilledet også en sektion, hvori de valgte tabeller er vist:

Øversigt over tabeller i forespørgslen	Evt. id på specifikke kolonner	Udtræk i kolonnen	Udtræk nyeste	Data skal ekstraheres
Undersøgelse			<input type="checkbox"/>	<input type="checkbox"/>
Bund	Blad bund			<input type="checkbox"/>
Bund	Grøt sten			<input type="checkbox"/>

De første to kolonner viser:

- Navnet på tabellen
- Evt. identifikation på specifikke kolonner

Herefter følger tre kolonner, der kan benyttes ved forskellige manipuleringer af udtrækket. Disse beskrives i afsnit 3.

Du kan fjerne alle kolonner hørende til en eller flere tabeller ved at udvælge disse og herefter klikke på .

## 2.3 Opstille kriterier

### 2.3.1 Overordnet opbygning

I langt de fleste udtræk ønskes datamængden afgrænset på en eller anden måde: data for et bestemt år, prøver hvor en bestemt art er fundet, stationer med en bestemt målsætning osv.

I de simple tilfælde ønsker man måske kun simple kriterier som f.eks.: find alle prøver i 2004, hvor målsætningen er opfyldt.

I andre tilfælde er det måske mere avanceret: Find alle prøver i 2004, hvor Gammarus eller Asellus er fundet i sparkeprøven (Asellus med 5 eller flere individer).

BioQuery4 giver mulighed for at opstille endog meget avancerede kriterier for udvælgelse.

Kriterier skrives i nedenstående sektion.

Faste kriterier	Variable kriterier	..eller	..eller

I kolonnen 'Faste kriterier' skrives de kriterier, der skal gælde uanset hvad der ellers måtte blive anført af kriterier.

Kolonnerne 'Variable kriterier' og alle '..eller' kolonner benyttes, når data skal medtages, hvis enten det ene eller det andet er opfyldt.

Samtidig gælder det, at de kriterier, der er opstillet i samme kolonne også skal være opfyldt samtidigt.

### Eksempel

Spørgsmål: Find prøver i 2004, der har DVFI=7 **eller** Fysisk index=I.

Udvalgte kolonner				Faste kriterier	Variable kriterier	..eller	..eller
Tabel	Felt	Kolonneid	Udskriv				
Undersøgelse	Ejer		<input checked="" type="checkbox"/>				
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>	01-01-2004..31-12-2004			
Undersøgelse	Prøvedato		<input checked="" type="checkbox"/>				
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>		{?}		
Undersøgelse	Fysisk index		<input checked="" type="checkbox"/>				{I}

Kriteriet på perioden er anført i 'Faste kolonner', idet dette altid skal være opfyldt. Om en prøve i 2004 nu skal med, varierer altså, afhængigt af om DVFI=7 **ELLER** Fysisk index=I. Ovenstående er altså den korrekte opbygning til besvarelse af det ønskede spørgsmål.

Lad os prøve at ændre lidt på ”fordelingen” af kriterier for at skærpe forståelsen af vigtigheden af den korrekte opbygning:

Udvalgte kolonner				Faste kriterier	Variable kriterier	..eller	..e
Tabel	Felt	Kolonneid	Udskriv				
Undersøgelse	Ejer		<input checked="" type="checkbox"/>				
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>				
Undersøgelse	Prøvedato		<input checked="" type="checkbox"/>		01-01-2004..31-12-2004		
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>		7		
Undersøgelse	Fysisk index		<input checked="" type="checkbox"/>				{'1'}

Kriteriet for perioden er nu flyttet ud i første variable kolonne. Det er nu en helt anden forespørgsel. Ovenstående finder de prøver, der i 2004 har DVFI=7, eller de prøver, der har Fysisk index=1 uanset hvornår den er taget. Dvs. en prøve fra f.eks. 1999 kommer med i udtrækket blot Fysisk index=1. For 2004-prøverne er det kun dem med DVFI=7, der kommer med.

Udvalgte kolonner				Faste kriterier	Variable kriterier	..
Tabel	Felt	Kolonneid	Udskriv			
Undersøgelse	Ejer		<input checked="" type="checkbox"/>			
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>			
Undersøgelse	Prøvedato		<input checked="" type="checkbox"/>		01-01-2004..31-12-2004	
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>			{'7'}
Undersøgelse	Fysisk index		<input checked="" type="checkbox"/>			{'1'}

I ovenstående har vi nu både DVFI- og Fysisk index-kriteriet stående i samme kolonne. Kun prøver fra 2004 medtages. Men nu skal prøven både have DVFI=7 og Fysisk index=1 for at blive medtaget. Alle kriterier skal altså nu være opfyldt samtidigt!

Ovenstående illustrerer, hvilke faldgruber man kan falde i, i formuleringen af kriterier. Udtrækkene er selvfølgelig gode nok, hvis det var det, vi ville. De besvarer bare ikke det stillede spørgsmål.

### 2.3.2 Kriteriecellen

I det foregående beskrev vi, hvorledes den overordnede logiske opbygning fungerer. Lad os nu se på, hvorledes den enkelte kriteriecelle kan udfyldes.

Når der stilles kriterier på datofelter, skal formatet være DD-MM-YYYY. For decimalfelter skal der benyttes punktum (.) som decimaltegn.

For datoer og numeriske felter kan der anvendes intervaller og logiske operatoren:

- Interval: Start..Slut (bemærk 2 punktummer imellem Start og slut (..))
  - o F.eks. 01-01-2004..31-12-2004 for afgrænsning af perioden 2004, eller
  - o 2.4..3.5 for afgrænsning af intervallet 2.4 til 3.5 (begge inklusive).
- Logiske operatoren: >,<=,>=,<=>
  - o F.eks. >=01-01-2004 for alle datoer fra og med 01-01-2004, eller
  - o <2.4 for alle værdier skarpt mindre end 2.4.

For tekst-felter gælder det, at hvis der blot skrives tekst, søges efter denne eksakte tekst:

- Angives f.eks. kriteriet 'Mølle' i feltet stations-navn, findes de stationer der præcist har navnet 'Mølle'.

Der kan dog anvendes %-tegn som "wildcards".

- Angives f.eks. kriteriet 'Mølle%', findes alle de stationer, hvis navn starter med Mølle. Således vil både 'Møllebæk' og 'Mølleå' blive medtaget.
- Angives kriteriet '%mølle', findes alle de stationer, hvis navn slutter med 'mølle'. 'Storemølle' og 'Lillemølle' vil altså begge kvalificere sig, hvorimod 'Lillemølleå' ikke vil.
- Angives kriteriet '%mølle%', findes alle stationer, hvis navn rummer teksten 'mølle'.

Det er tilladt at angive mere end ét kriterium i samme celle. De enkelte kriterier skal blot adskilles med semi-kolon (;). Hvis vi f.eks. ønsker at finde de prøver, der har middelbredde 1, 2, 3 eller 4m, kan dette gøres med kriteriet: 1;2;3;4. (Bemærk, det ikke er nødvendigt at skrive =1;=2;=3;=4. Såfremt der ikke er angivet en operator benyttes '=' som default.) Hvis vi ønsker prøver med middelbredde mindre end 2 meter eller større end 4 meter, kan dette gøres med kriteriet: <2;>4.



Dette gælder også tekster og datoer. F.eks. kunne vi finde alle stationer, der rummer teksten 'mølle' eller teksten 'lille' i navnet med flg. kriterium: %mølle%;%lille%. Eller vi kunne finde alle prøver i 1999 eller 2004 med kriteriet: 01-01-1999..31-12-1999;01-01-2004..31-12-2004.

I øvrigt er det også tilladt at anvende et mængde kriterium. Dette skrives som (k1,k2,...,kn). I ovenstående eksempel med middelbredder på 1, 2, 3 eller 4m kunne man altså skrive (1,2,3,4) (bemærk at det er komma, der adskiller de enkelte elementer i mængden). Begrebsmæssigt er det præcist det samme, det giver blot en mere effektiv (hurtigere svartid), hvis der kan benyttes mængder.

### 2.3.3 Eksistens-kriterier

Når cursoren placeres i en kriteriecelle, vil den have nedenstående udseende (hvis feltet er koblet til en opslagsliste vil der være en knap mere, se afsnit 2.3.4):



Hvis der trykkes på knappen , skrives 'Udfyldt' i cellen. Vi er altså ligeglade med hvad der står, feltet skal bare være udfyldt. Trykkes der derimod på , skrives 'Ikke udfyldt'. Vi er altså kun interesserede i de rækker, hvor feltet ikke er udfyldt (tomt), f.eks. alle de stationer, som ikke har fået indtastet en målsætning.

Man behøver ikke trykke på knapperne, men må gerne selv skrive 'Udfyldt' eller 'Ikke udfyldt'. I øvrigt er det også tilladt at skrive 'Not null' (svarer til udfyldt) eller Null (svarer til ikke udfyldt).



### 2.3.4 Kriterier på felter tilknyttet opslagsliste

En lang række af de kolonner, der udtrækkes data fra, er tilknyttet diverse opslagslister, f.eks. DVFI, målsætninger, prøvetagere, osv. Når cursoren placeres i kriterium-cellen for et sådant felt, har den nedenstående udseende:




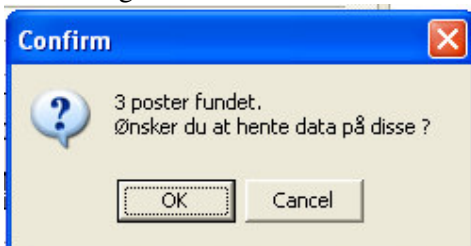
Når der trykkes på knappen , vises opslagslisten for den pågældende kolonne. Var kolonnen f.eks. DVFI ville følgende fremkomme:



Der kan nu vælges ét eller flere elementer fra listen. Trykkes nu på knappen , opstilles kriteriet, at alle de valgte elementer skal være tilstede. Trykkes i stedet på knappen , opstilles kriteriet, at ALT ANDET end de valgte elementer skal være tilstede (komplementærmængden).

## 2.4 Kør udtrækket

Når alle ønskede kolonner er valgt og kriterier opstillet, er det tid til at få kørt udtrækket. Dette gøres med et tryk på knappen . I første omgang vises en meddelelse om, hvor mange rækker udtrækket returnerer, f.eks.



Hvis der trykkes på Cancel, foretager programmet sig ikke yderligere. Trykkes der på Ok, hentes data fra serveren.

Fanebladet 'Data' bliver synligt, værktøjslinien skifter udseende og data vises:

Forespørgsel	
Undersøgel	Undersøgel
Antal: Stationsnr	DVFI faunaklasse
261	
115	4
83	5

Der returneres til selve forespørgsels-definitionen ved at klikke på fanebladet 'Forespørgsel'.

Data kan nu udtrækkes til enten Excel  eller som Tabulator-adskilt kolonne-fil  (tekst-format).

Endvidere kan data-udtrækket manipuleres yderligere:



Slår øverste kolonneskrift (tabel) til og fra.



Giver mulighed for at sortere i data ved et klik på kolonneoverskriften.



Åbner mulighed for yderligere afgrænsning ved at vælge fra kombo-boks i kolonneoverskriften.



Åbner mulighed for at gruppere data efter indholdet i én eller flere kolonner.



Medfører at kolonne-bredder tilpasses til indhold af data og kolonneoverskrifter.

### 3. Funktioner

#### 3.1 Funktioner på tabelniveau

##### 3.1.1 Udtræk i kolonner (fælleslister)

I et traditionelt udtræk, vil data altid blive udtrukket i rækker. Antag f.eks. at vi ønsker at udtrække alle faunaprøver for et bestemt år og dette resulterer i 500 rækker (dvs. 500 prøver). Hvis vi supplerer udtrækket med også at rumme information om bundforhold, f.eks. intensiteten, resultater udtrækket pludselig i langt flere rækker. Nemlig én række for hver registrering af bundforhold på de 500 prøver. For nemheds skyld kan vi antage, at der i snit er 3 registreringer pr. prøve, så udtrækket resulterer nu i 1500 rækker.

Data, der kun hører til faunaprøven, f.eks. DVFI, duplikeres/gentages altså for hver registrering af bundforhold, der er på prøven.

Oversigten over valgte tabeller vil have følgende udseende:

Oversigt over tabeller i forespørgslen	Evt. id på specifikke kolonner	Udtræk i kolonner	Udtræk nyeste	Data skal eksistere
Undersøgelse			<input type="checkbox"/>	<input type="checkbox"/>
Bund		<input type="checkbox"/>		<input type="checkbox"/>

Ovenstående udtræk er for så vidt helt tilfredsstillende, men i nogen sammenhænge ville det være rart at kunne fastholde, at der kun udtrækkes én række pr. prøve, mens oplysningerne om bundforhold udtrækkes i separate kolonner. Dette opnås ganske enkelt ved at sætte et flueben i feltet 'Udtræk i kolonner' ud for tabellen 'Bund'. I så fald oprettes der en fælles liste af kolonner for de bundforhold, der er benyttet for de 500 faunaprøver. For den enkelte række udfyldes kolonnerne med værdierne for den konkrete prøve.

Udover at give et overskueligt layout, der ligeledes er velegnet til videre databehandling i f.eks. regneark, har denne funktionalitet en yderligere fordel:

Antag vi i samme udtræk også ønsker at medtage intensiteten for oplysninger om vandet. Tabeloversigten vil nu få følgende udseende:

Oversigt over tabeller i forespørgslen	Evt. id på specifikke kolonner	Udtræk i kolonner	Udtræk nyeste	Data skal eksistere
Undersøgelse			<input type="checkbox"/>	<input type="checkbox"/>
Bund		<input type="checkbox"/>		<input type="checkbox"/>
Vand		<input type="checkbox"/>		<input type="checkbox"/>

Dette udtræk giver faktisk ingen mening (i den række-baserede version), idet der jo ikke eksisterer en relation imellem en bund-registrering og en vand-registrering. Forespørgslen vil derfor danne et krydsprodukt (kartesisk produkt), som sjældent er det ønskede resultat. Hvis du på en prøve har to bundregistreringer, f.eks. 'Blød' og 'Slam' og to vand-registreringer f.eks. 'Uklart' og 'Jævn', så er krydsproduktet fire rækker:

Blød Uklart  
 Slam Uklart  
 Blød Jævn  
 Slam Jævn


Altså alle de mulige kombinationer.

Hvis vi derimod sætter flueben i 'Udtræk i kolonner' for både 'Bund' og 'Vand', udtrækkes der fortsat kun én række pr. faunaprøve, og der dannes nu fællesliste af kolonner for både bund- og vand-registreringer.

### 3.1.2 Udtræk i kolonner (specifikke)

I ovenstående eksempler blev der dannet de kolonner for bund og vand, som var repræsenteret i de valgte prøver. Nyttigt, da vi jo ikke på forhånd ved hvilke bund- og vandforhold der er registreret.

Hvis man derimod i et udtræk kun går efter ganske bestemte fysiske forhold, f.eks. intensiteten på 'Blød bund' og 'Gråt slam' i bundforhold kan disse meget simpelt udtrækkes i kolonner.

Når tabellen 'Bundforhold' er valgt i tabel-navigatoren, vil knappen  blive aktiv. Trykkes på den, vises et separat afsnit med mulige valg af bundforhold:

Felter	Udvælg specifikke kolonner. Vælg herefter de felter du ønsker udtrukket på disse kolonner
Ejer	
Stationsnr	
Prøvedato	Grene, sten (skjul for fisk)
Bund,kode	Groft sand (1-3 mm)
Bund,navn	Gruset
Intensitet (X,1,2,3)	Gråt slam

- find 'Gråt slam' i listen og klik på den
- hold Ctrl-tasten nede, find 'Blød bund' og klik på den
- dobbeltklik på 'Intensitet (X,1,2,3)

Når udtrækket køres, oprettes der en kolonne for 'Blød bund' og en for 'Gråt slam'. Kolonnerne udfyldes med intensiteten.

### 3.1.3 Udtræk nyeste

I udtræk, der involverer faunaprøver eller fiskeundersøgelser, er det muligt, kun at udtrække den nyeste prøve (der opfylder de opstillede kriterier). Hvis man f.eks. interesserer sig for udbredelsen af en specifik art, kan man naturligvis danne et udtræk, der finder alle de prøver, hvor arten er fundet. Ved at sætte flueben i 'Udtræk nyeste', finder man kun de seneste bedømmelser, hvor arten er fundet.

### 3.1.4 Data skal eksistere

Antag, at vi danner et udtræk der involverer felter, dels fra de faste stationsoplysninger, dels fra fauna- og fiskeundersøgelser i f.eks. 2004.

Hvis der på en givet station ikke er foretaget undersøgelser i 2004, hvad så: Ønsker vi alligevel stationen med i udtrækket, eller vil vi kun have stationen med, hvis der også er taget en prøve på den.


Hvis man insisterer på, at der skal være en prøve, sættes flueben i 'Data skal eksistere' ud for tabellen 'Undersøgelse'.

Oversigt over tabeller i forespørgslen	Evt. id på specifikke kolonner	Udtræk i kolonner	Udtræk nyeste	Data skal eksistere
Stationfast				<input type="checkbox"/>
Undersøgelse			<input type="checkbox"/>	<input checked="" type="checkbox"/>

### 3.1.5 Fasthold stationsoplysninger

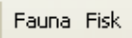
Antag, at vi danner et udtræk, der involverer felter, dels fra de datoafhængige stationsoplysninger (f.eks. målsætningen), dels fra fauna- og fiskeundersøgelser.

Idet målsætningen kan have ændret sig over tid, er det altså ikke sikkert, at målsætningsfeltet er udfyldt ens for alle prøver, selvom der er tale om den samme station. Dette vil ikke altid være ønskeligt, idet man i nogen sammenhænge er interesseret i altid at vise den nugældende målsætning, selvom det ikke var den, der gjaldt da prøven blev taget.

Styringen af dette foretages med knappen . Hvis den er trykket ned, vises altid de nugældende stationsoplysninger, ellers vises de stationsoplysninger der var gældende da prøven blev taget.

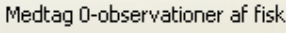
### 3.1.6 Afgræns til Fauna og/eller Fiskeundersøgelser

Da man ofte ønsker at afgrænse sit udtræk til enten fauna eller fiskeundersøgelser, er der oprettet særlige knapper til dette formål. Anvendelse af knapperne giver naturligvis kun mening, hvis udtrækket rummer undersøgelses-data.

Knapperne  kan være trykket ind eller ud. Hvis de begge er trykket ind eller begge er ude, opereres der på alle data. Hvis kun én er trykket, opereres der kun på data af den pågældende type.

### 3.1.7 0-observationer ved udtræk af fiske-data

Ved udtræk af f.eks. bestandstætheder for fisk, vil sådanne underkastes en efterfølgende statistisk bearbejdning. I sådanne tilfælde er det nødvendigt, at udtrækket kan medtage 0-observationer, dvs. medtage 0 (nul)-data for de arter, der IKKE er fundet ved befiskningen. Er der f.eks. foretaget befiskning på fem stationer, men kun fundet Ørred på de to, er det afgørende at de tre øvrige medtages i den statistiske beregning med værdien 0.


Ønsker man at medtage 0-observationer af fisk i et udtræk, trykkes knappen  ned. I modsat fald medtages kun de faktiske observationer.

## 3.2 Unikke rækker

Hvilke arter er der fundet ved undersøgelser i 1999? Dette spørgsmål lader sig nemt besvare ved brug af unikke rækker. Opbyg først en forespørgsel der finder alle fauna-registreringer i 1999:

Udvalgte kolonner					
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier
Undersøgelse	Prøvedato		<input type="checkbox"/>	01-01-1999..31-12-1999	
Fauna	Art_navn		<input checked="" type="checkbox"/>		

Fjern flueben i Udskriv-kolonnen for prøvedato feltet. Hvis forespørgslen køres nu, returneres der én række med artsnavnet for hver fauna-registrering i 1999. Karakterisk for dette udtræk er jo, at det samme artsnavn optræder en masse gange. Det vi ønskede, var at finde ud af, hvilke arter vi har fundet, så vi vil have fjernet dubletterne.

Dette løses af funktionen Unikke rækker, der aktiveres ved at trykke på knappen .

### 3.3 Grupperede udtræk

Grupperede er et stærkt element i SQL-spørgesproget, som også kan anvendes i Bio-Query4. Der er dog tale om mere avancerede udtræk, hvorfor den utrænede bruger bør være forsigtig! Simpelthen fordi resultatet af en forespørgsel måske danner grundlag for videre beslutninger, hvorfor man bør være helt sikker på, at dataudtrækket faktisk illustrerer det, man tror det gør!

Grupperede udtræk har det karakteristika, at man på baggrund af nogle fællestræk i data samler data i grupper og herefter udfører funktioner (tælle, sum, gennemsnit o.a.) på disse grupper.

Antag f.eks., at vi ønsker et udtræk der viser antal prøver i 1999 fordelt på DVFI.

Vi vælger først de relevante felter og opstiller et periode-kriterie på prøvedatoen:

Udvalgte kolonner					
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier
Undersøgelse	Ejer		<input checked="" type="checkbox"/>		
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>		
Undersøgelse	Prøvedato		<input checked="" type="checkbox"/>	01-01-1999..31-12-1999	
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>		

Foreløbig er der altså tale om et ganske simpelt udtræk, som jo altså netop er de data, vi ønsker at beskæftige os med. Dette er altså kandidat-mængden til det endelige resultat.

Tryk nu på knappen  i værktøjslinien.

Oversigten over felter tilføjes nu kolonnen 'Grupper' (efter sorterings-kolonnen).

Udvalgte kolonner						Suppl.	
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier	..eller	..eller
Undersøgelse	Ejer		<input checked="" type="checkbox"/>				
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>				
Undersøgelse	Prøvedato		<input checked="" type="checkbox"/>	01-01-1999..31-12-1999			
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>				

Hvis du kører udtrækket nu, er beskeden, at der ikke er udvalgt noget. Det skyldes, at vi har valgt grupperingsfunktionen, men har ikke benyttet grupperinger på nogle af felterne.

Vi ønskede en opgørelse fordelt på DVFI, så vi skal altså samle data i de forskellige DVFI-grupper. Vælg derfor 'Gruppe' i dette felt som vist i nedenstående:

Udvalgte kolonner							Suppl.	
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier	..eller	Sorter	Grupper
Undersøgelse	Ejer		<input checked="" type="checkbox"/>			..eller		
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>			..eller		
Undersøgelse	Prevedato		<input checked="" type="checkbox"/>	01-01-1999..31-12-1999		..eller		
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>			..eller		Gruppe

For hver DVFI skal vi tælle antal prøver. Når vi skal tælle noget, skal vi sikre os, at vi tæller på et felt, der altid er udfyldt (ellers tælles det ikke med). Det kan vi sikre os ved altid at tælle på felter, der er en del af den primære identifikation, f.eks. stationsnummret.

Vælg derfor antal 'Antal' for feltet 'Stationsnr'. Udtrækket har nu nedenstående opbygning:

Udvalgte kolonner							Suppl.	
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier	..eller	Sorter	Grupper
Undersøgelse	Ejer		<input checked="" type="checkbox"/>			..eller		
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>			..eller		Antal
Undersøgelse	Prevedato		<input checked="" type="checkbox"/>	01-01-1999..31-12-1999		..eller		
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>			..eller		Gruppe

Dette er det færdige udtræk, der leverer en optælling af antal prøver fordelt efter DVFI på 1999-data.

De felter, der ikke er anvendt grupperings-direktiver på, ignoreres ganske enkelt i udtrækket. Dog vil evt. kriterier på disse felter naturligvis gælde (til afgrænsning af kandidatmængden).

Vi kunne yderligere komplicere forespørgslen ved f.eks. at kræve, at vi kun ønsker en DVFI repræsenteret i udtrækket, hvis der er 50 eller flere prøver, der har opnået den pågældende DVFI.

Hvorledes får vi opstillet dette kriterium. Det er jo ikke kandidatmængden, der skal afgrænses, det er stadigvæk alle 1999-prøver. Vi kan jo først afgøre, om der er 50 eller flere i en DVFI-gruppe, efter at de er samlet i grupperne.

Kriteriet sættes på 'Stationsnr' feltet, da det var dette felt, vi brugte til at tælle antal prøver med. Vi får altså:

Udvalgte kolonner							Suppl.	
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier	..eller	Sorter	Grupper
Undersøgelse	Ejer		<input checked="" type="checkbox"/>			..eller		
Undersøgelse	Stationsnr		<input checked="" type="checkbox"/>	>=50		..eller		Antal
Undersøgelse	Prevedato		<input checked="" type="checkbox"/>	01-01-1999..31-12-1999		..eller		
Undersøgelse	DVFI faunaklasse		<input checked="" type="checkbox"/>			..eller		Gruppe

Reglen er altså: Kriterier, der sættes på felter, hvorpå der er angivet en grupperingsfunktion (antal, sum, gennemsnit osv.) er kriterier, der stilles på de færdige grupper. Alle

andre kriterier (inkl. dem, der måtte være anført for felter med direktivet 'Gruppe') afgrænser kandidatmængden **INDEN** data samles i grupper.

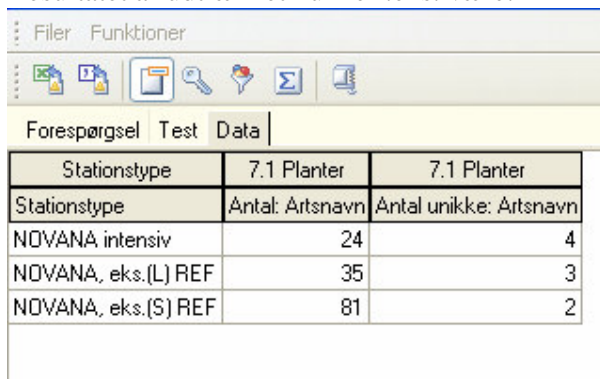
Grupperingsfunktionerne er altså operationer, der udføres på de grupper, der er dannet. Antal, sum, gennemsnit, min, max og spredning giver sig selv, hvorimod 'Unikke' er en speciel men meget nyttig funktion.

Antag f.eks. at vi ønsker en opgørelse af antal planter fra NOVANA, skema7, fordelt efter en række stationstyper. Nedenstående viser dette udtræk.

Udvalgte kolonner								Suppl.	
Tabel	Felt	Kolonneid	Udskriv	Faste kriterier	Variable kriterier	..eller	..eller	Sorter	Grupper
Stationstype	Stationstype		<input checked="" type="checkbox"/>	NOVANA intensiv, NOVANA al					Gruppe
7.1 Planter	Artsnavn		<input checked="" type="checkbox"/>						Antal
7.1 Planter	Artsnavn		<input checked="" type="checkbox"/>						Unikke

Udtrækket grupperer data efter stationstype. Dernæst tælles der på arter. Anvendelsen af funktionen tæller antal artsregistreringer. I en sådan optælling kan den samme art altså godt optræde mere end én gang. 'Unikke' derimod, tæller antal unikke forekomster af det valgte felt (arten) indenfor gruppen (stationstypen), og giver altså dermed artsantallet.

Resultatet af udtrækket kunne f.eks. være:



Stationstype	7.1 Planter	7.1 Planter
Stationstype	Antal: Artsnavn	Antal unikke: Artsnavn
NOVANA intensiv	24	4
NOVANA, eks.(L) REF	35	3
NOVANA, eks.(S) REF	81	2

Vi kan altså f.eks. se, at der er indtastet 24 artsregistreringer for stationer med typen 'NOVANA, intensiv'. Af 'Antal unikke' rækken fremgår det, at der kun er tale om fire forskellige arter.