



**Wprowadzenie do**



# Czym jest R?



- Rozbudowane narzędzie statystyczne i graficzne
- Specyficzny język programowania:
  - Język skupiony raczej wokół obiektów niż akcji
  - Obiekty złożone z danych oraz funkcji
- Prosty sposób importu/eksportu danych
- Darmowy i otwarty
- Dostępny dla wielu systemów operacyjnych (Windows, Mac, Linux, FreeBSD itd.)

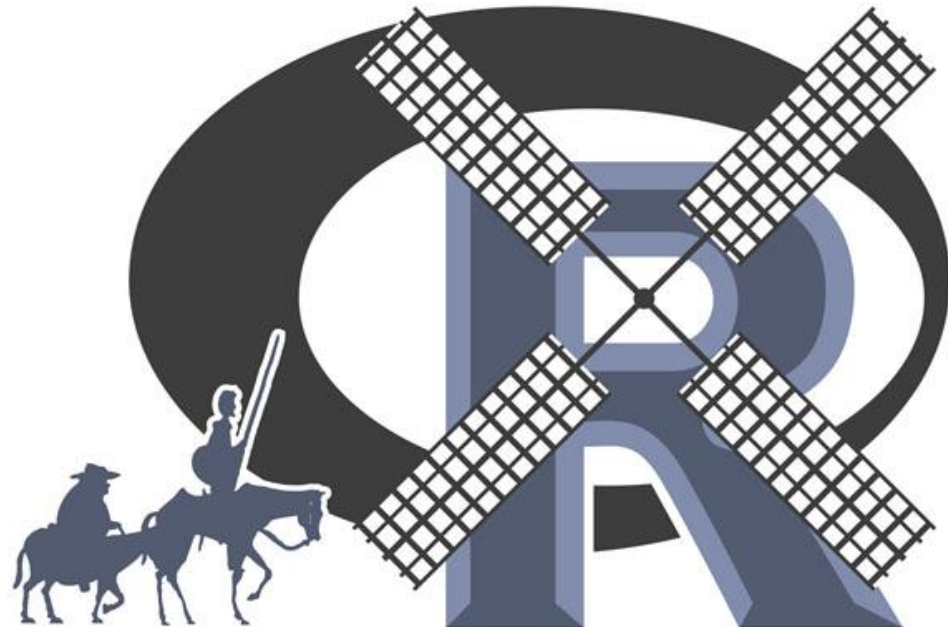
# Zalety

- Praca z wieloma zestawami danych o różnej strukturze jednocześnie:
  - listy, ramki danych, wektory, macierze itp.
- Pojawianie się nowych funkcji i bibliotek, przez co R ciągle poszerzany
- Budowa pętli, własnych funkcji w zależności od potrzeb
- Doświadczalna interaktywność – programujesz tak długo, aż otrzymasz to, co potrzebujesz
- Dostępne prawie każde narzędzie statystyczne
- Szeroko rozbudowana pomoc



# Wady

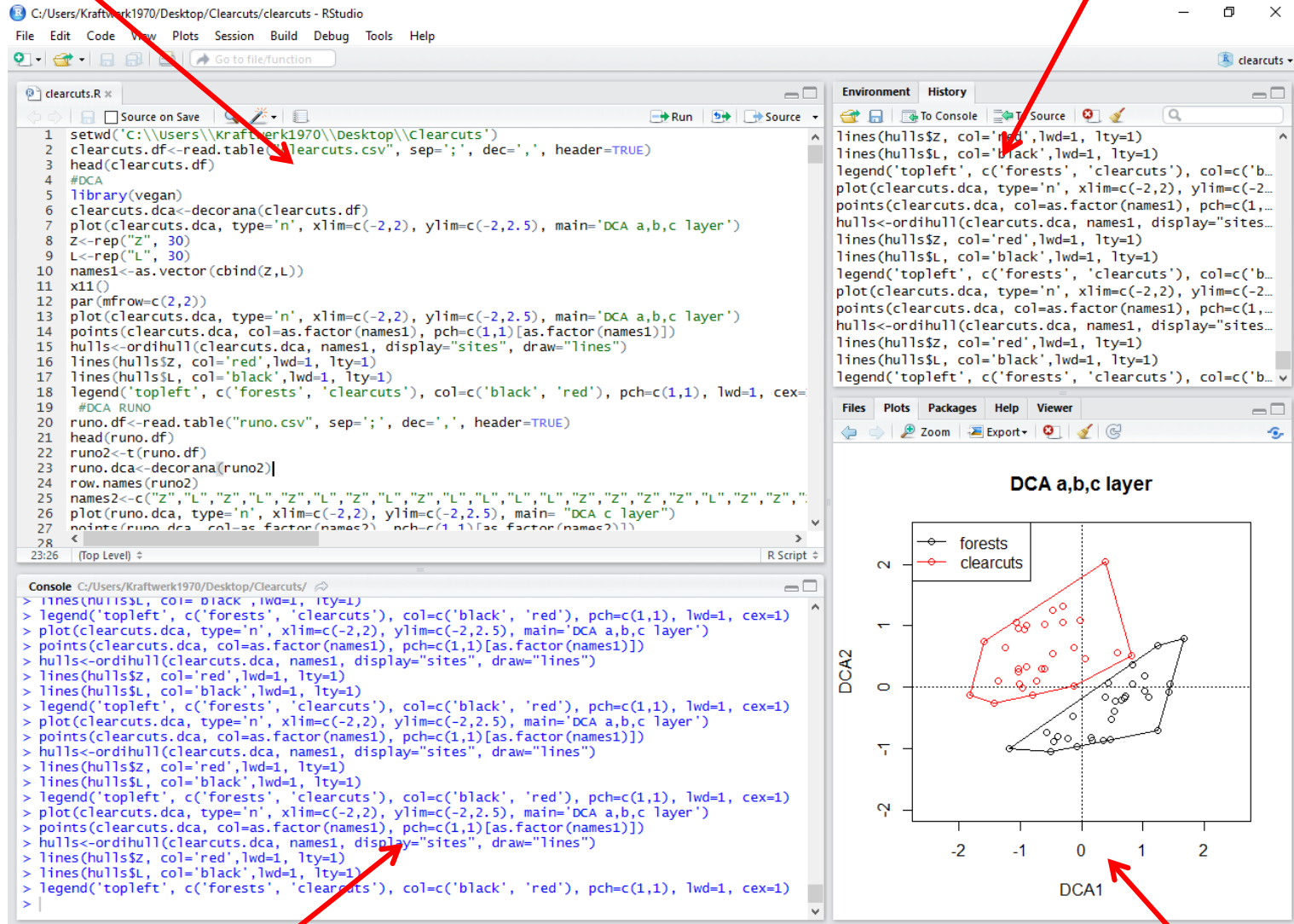
- Bardzo stroma krzywa uczenia się
- Nie jest programem typu „wskaż i kliknij”
- Łatwo o pomyłki, które są trudne do zlokalizowania
- Program metody „prób i błędów”



# Jak wygląda?

Okno historii

Edytor tekstu (skrypt)



konsola

Okno grafiki

# Pierwsze kroki

R jest typem języka obiektowego

```
x<-33
```

x jest obiektem, strzałka to symbol przypisania liczby 33 do obiektu x

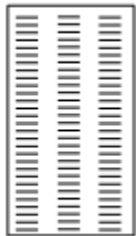
Obiekt może zawierać więcej liczb lub znaków

```
x<-c(33,34,35,36)
```

```
x<-c('R','jest','super')
```

# Typy obiektów

- **Integer** – wartość liczbowa
- **Character** – wartość złożona z jednej lub więcej znaków literowych
- **Vector** – seria liczb lub liter. Jeżeli zawiera liczby i litery, odczytywany jest jako obiekt typu „**character**”. Jeżeli nie znamy jakiejś wartości, wpisujemy „**NA – not available**”
- **Matrix** – seria wektorów w systemie kolumn i wierszy. Matrix może mieć różną liczbę kolumn i wierszy, ale każdy wiersz musi mieć taką samą liczbę kolumn, a kolumna taką samą liczbę wierszy





- **Data frame** – podstawowy typ obiektów w modelowaniu i przedstawianiu graficznemu danych

response	predictor	colour	response
1.23	A	black	1.23
2.45	A	black	2.45
1.11	A	black	NA
2.35	A	black	2.35
5.78	B	blue	NA
4.32	B	blue	4.32
3.31	B	blue	3.31
6.98	B	blue	6.98

# Podstawowe funkcje

Funkcja	Działanie
<code>read.table</code>	Wczytuje dane
<code>colnames/rownames</code>	Nazwy kolumn/wierszy
<code>colSums/rowSums</code>	Sumy wartości w kolumnach/wierszach
<code>colMeans/rowMeans</code>	Średnie wartości z kolumn/wierszy
<code>c</code>	Łączy wartości w wektor
<code>cbind/rbind</code>	Skleja obiekty kolumnami/wierszami
<code>plot</code>	Graficzne przedstawienie danych
<code>abline</code>	Dodaje linię trendu
<code>mean/median</code>	Średnia/mediana
<code>sqrt</code>	Pierwiastek kwadratowy
<code>sd</code>	Odchylenie standardowe

# Tabele

Funkcja	Działanie
<code>[ , 1]</code>	Pierwsza kolumna z tabeli
<code>[ , 1:3]</code>	Kolumny od 1 do 3
<code>[c(1,3)]</code>	Kolumna 1 i 3
<code>[- , 1]</code>	Usuwanie pierwszej kolumny
<code>[-c(1,3)]</code>	Usuwanie kolumny pierwszej i trzeciej
<code>[1, ]</code>	Pierwszy wiersz
<code>[1:3, ]</code>	Wiersze od 1 do 3
<code>[c(1,3), ]</code>	Wiersz 1 i 3
<code>[-1, ]</code>	Usuwanie pierwszego wiersza
<code>[-c(1,3), ]</code>	Usuwanie pierwszego i trzeciego wiersza

# Łączenie tabel

epiphytes.hist

	species	site	freq
1	Alyxoria varia s.l.	A1h	1
2	Arthonia arthonioides	A1h	1
3	Arthonia spadicea	A1h	1
4	Arthonia vinosa	A1h	1
5	Calicium adpersum	A1h	1
6	Chaenotheca chrysocephala	A1h	1
7	Chaenotheca furfuracea	A1h	1
8	Chaenotheca trichialis	A1h	1
9	Chrysothrix candelaris	A1h	1
10	Cladonia fimbriata	A1h	1
11	Evernia prunastri	A1h	1
12	Lecanora argentata s.l.	A1h	1
13	Lecidella elaeochroma	A1h	1
14	Loxospora elatina	A1h	1
15	Melanelixia glabratula	A1h	1
16	Opegrapha niveoatra	A1h	1
17	Opegrapha vermicellifera	A1h	1
18	Parmelia sulcata s.l.	A1h	1
19	Pertusaria amara	A1h	1
20	Pertusaria coccodes	A1h	1
21	Porina aenea	A1h	1
22	Usnea dasypoga	A1h	1
23	Cladonia coniocraea	A1h	2
24	Graphis scripta s.l.	A1h	2
25	Hypogymnia physodes	A1h	2
26	Lepraria incana s.l.	A1h	3
27	Pertusaria leioplaca	A1h	2
28	Phlyctis argena	A1h	2
29	Pyrenula nitida	A1h	2
30	Pyrenula nitidella	A1h	2
31	Thelotrema lepadinum	A1h	2
32	Zwackhia viridis	A1h	2
33	Alyxoria varia s.l.	A10h	1
34	Arthonia arthonioides	A10h	1
35	Arthonia spadicea	A10h	1
36	Chaenotheca furfuracea	A10h	1
37	Evernia prunastri	A10h	1

epiphytes.new

	species	site	freq
1	Acrocordia gemmata	A1	1
2	Alyxoria varia s.l.	A1	1
3	Arthonia arthonioides	A1	1
4	Arthonia radiata	A1	1
5	Arthothelium ruanum	A1	1
6	Chaenotheca brachypoda	A1	1
7	Chaenotheca chrysocephala	A1	1
8	Chaenotheca ferruginea	A1	1
9	Chaenotheca trichialis	A1	1
10	Cladonia digitata	A1	1
11	Inoderma byssaceum	A1	1
12	Lecanora carpinea s.l.	A1	1
13	Lecanora chlarotera	A1	1
14	Melanelixia glabratula	A1	1
15	Micarea prasina s.l.	A1	1
16	Microcalicium disseminatum	A1	1
17	Opegrapha vermicellifera	A1	1
18	Pertusaria coccodes	A1	1
19	Pertusaria ophthalmiza	A1	1
20	Pertusaria pertusa	A1	1
21	Pyrenula nitidella	A1	1
22	Ramalina farinacea	A1	1
23	Reichlingia leopoldii	A1	1
24	Arthonia spadicea	A1	2
25	Buellia griseovirens	A1	2
26	Chrysothrix candelaris	A1	2
27	Cladonia coniocraea	A1	2
28	Coenogonium pineti	A1	2
29	Fellhanera gyrophorica	A1	3
30	Graphis scripta s.l.	A1	3
31	Lecanora argentata s.l.	A1	3
32	Lecanora thysanophora	A1	3
33	Lecidella elaeochroma	A1	2
34	Lepraria finkii s.l.	A1	3
35	Lepraria incana s.l.	A1	3
36	Loxospora elatina	A1	2
37	Parmelia sulcata s.l.	A1	2
38	Pertusaria amara	A1	3
39	Phlyctis argena	A1	3
40	Porina aenea	A1	2

```
#import danych
epiphytes.hist<-read.csv('epiphytes.hist.csv', sep=';')
epiphytes.new<-read.csv('epiphytes.new.csv', sep=';')

#sprawdzenie, czy dobrze wczytaliśmy dane
head(epiphytes.hist)
head(epiphytes.new)

#sklejenie tabel wierszami
epi.all<-rbind(epiphytes.hist,epiphytes.new)
```

```
#załadowujemy bibliotekę  
library(reshape2)
```

```
#przekształcamy tabelę wąską do tabeli szerokiej  
epi.cast<-as.data.frame(dcast(epi.all, species~site,  
value.var='freq', drop=FALSE))
```

```
#eksportujemy tabelę z R do pliku .csv  
write.table(epi.cast, 'epi.casting.csv', sep=';')
```

```
#nową tabelę można znowu wgrać do R  
epi<-read.csv('epi.cast.csv', sep=';')
```

## #tabela szeroka

[illegible]

```
#transpozycja kolumn z wierszami
epi.t<-t(epi)
```

	Acr.gem	Aly.var	Ani.bif	Art.art	Art.atr	Art.did	Art.rad	Art.spa	Art.vin	Ath.rua
A10h	0	1	0	1	0	1	0	0	0	0
A11h	2	0	0	1	0	0	0	1	0	0
A1h	0	1	0	1	0	1	1	0	0	0
A2h	1	2	0	2	0	1	0	1	1	0
A3h	0	2	0	0	1	0	0	0	0	0
A4h	1	2	0	2	0	1	0	1	0	0
	Ath.spe	Bac.arc	Bac.bec	Bac.lau	Bac.rub	Bct.dry	Bia.glo	Bry.fus	Bry.imp	Bue.dis
A10h	0	0	0	0	0	0	0	0	0	0
A11h	0	0	1	0	0	0	0	0	1	0
A1h	0	0	0	0	0	0	0	0	0	1
A2h	0	0	0	0	0	0	0	0	0	0
A3h	0	0	1	0	0	0	0	0	0	0
A4h	0	0	0	1	0	0	0	0	1	0
	Bue.eru	Bue.gri	Bue.sch	Cal.ads	Cal.gla	Cal.sal	Cal.vir	Clp.pyr	Can.xan	Crb.ant
A10h	0	0	0	0	0	0	0	0	2	0
A11h	0	0	0	0	0	1	0	0	1	0
A1h	0	0	0	0	0	0	0	0	1	0
A2h	0	0	0	0	0	0	0	0	1	0
A3h	0	0	1	0	0	0	0	0	2	0
A4h	0	1	1	0	0	0	0	0	1	0



```
#przekształcenie tabeli szerokiej do wąskiej
epi.melt<-melt(epi.t.id, value.name='freq')
```

	Var1	Var2	freq
1	A10h	Acr.gem	0
2	A11h	Acr.gem	2
3	A1h	Acr.gem	0
4	A2h	Acr.gem	1
5	A3h	Acr.gem	0
6	A4h	Acr.gem	1
7	A5h	Acr.gem	0
8	A6h	Acr.gem	0
9	A7h	Acr.gem	0
10	A8h	Acr.gem	0
11	A9h	Acr.gem	1
12	B10h	Acr.gem	0
13	B11h	Acr.gem	0
14	B1h	Acr.gem	1
15	B2h	Acr.gem	0
16	B3h	Acr.gem	0
17	B4h	Acr.gem	0
18	B5h	Acr.gem	1
19	B6h	Acr.gem	0
20	B7h	Acr.gem	0

```
#zmiana nazw kolumn
colnames(epi.melt)<-c("site", "species", "frequency")
```

site	species	frequency
A10h	Acr.gem	0
A11h	Acr.gem	2
A1h	Acr.gem	0
A2h	Acr.gem	1
A3h	Acr.gem	0
A4h	Acr.gem	1
A5h	Acr.gem	0
A6h	Acr.gem	0
A7h	Acr.gem	0
A8h	Acr.gem	0
A9h	Acr.gem	1
B10h	Acr.gem	0
B11h	Acr.gem	0
B1h	Acr.gem	1
B2h	Acr.gem	0
B3h	Acr.gem	0
B4h	Acr.gem	0
B5h	Acr.gem	1
B6h	Acr.gem	0
B7h	Acr.gem	0
B8h	Acr.gem	1
B9h	Acr.gem	1
C10h	Acr.gem	2

```
#tworzenie obiektu z danymi przynależności gatunku do siedliska
forest<-rep(c("dec.forest", "con.forest"), 74)
species<-colnames(epi.t.id)
habitat<-as.data.frame(cbind(species, forest))
```

```
species      forest
Acr.gem dec.forest
Aly.var  con.forest
Ani.bif mix.forest
Art.art  dec.forest
Art.atr  con.forest
Art.did  mix.forest
Art.rad  dec.forest
Art.spa  con.forest
Art.vin  mix.forest
Ath.rua  dec.forest
Ath.spe  con.forest
Bac.arc  mix.forest
Bac.bec  dec.forest
Bac.lau  con.forest
Bac.rub  mix.forest
Bct.dry  dec.forest
Bia.glo  con.forest
Bry.fus  mix.forest
Bry.imp  dec.forest
Bue.dis  con.forest
```

```
#dołączanie do tabeli wąskiej kolumny z tabeli szerokiej
epi.join<-left_join(epi.melt, habitat, by='species')
```

	site	species	frequency	forest
1	A10h	Acr.gem	0	dec.forest
2	A11h	Acr.gem	2	dec.forest
3	A1h	Acr.gem	0	dec.forest
4	A2h	Acr.gem	1	dec.forest
5	A3h	Acr.gem	0	dec.forest
6	A4h	Acr.gem	1	dec.forest
7	A5h	Acr.gem	0	dec.forest
8	A6h	Acr.gem	0	dec.forest
9	A7h	Acr.gem	0	dec.forest
10	A8h	Acr.gem	0	dec.forest
11	A9h	Acr.gem	1	dec.forest
12	B10h	Acr.gem	0	dec.forest
13	B11h	Acr.gem	0	dec.forest
14	B1h	Acr.gem	1	dec.forest
15	B2h	Acr.gem	0	dec.forest
16	B3h	Acr.gem	0	dec.forest
17	B4h	Acr.gem	0	dec.forest
18	B5h	Acr.gem	1	dec.forest
19	B6h	Acr.gem	0	dec.forest

# Działania na tabeli przy zastosowaniu funkcji z pakietu dplyr

change.per.summit.merged

	Access	Tourism		BCdist	LoId	Lnew	ToId	Tnew	MoId	Mnew	NoId	Nnew	RoId	Rnew
1		2	1	0.3658537	4.632653	4.742424	1.714286	1.575758	3.295918	3.196970	2.553191	2.363636	3.683673	3.227273
2		3	1	0.5303030	4.589286	4.657895	1.937500	1.907895	3.098214	3.065789	2.764151	2.687500	4.172727	4.085526
3		4	2	0.6821192	4.218750	3.862069	2.148438	2.465517	3.320312	3.304598	2.629032	2.988235	3.571429	3.252874
4		2	2	0.5405405	4.632653	4.920000	1.775510	1.500000	3.285714	3.160000	2.622222	2.229167	3.530612	2.920000
5		4	2	0.6725146	4.220000	3.842975	2.500000	2.834711	2.960000	3.148760	2.867347	3.150000	4.430000	4.024793
6		4	1	0.5260116	4.587302	4.518182	1.968254	1.950000	3.055556	3.090909	2.762295	2.740566	4.137097	3.972727
7		4	1	0.9629630	4.652174	4.725806	1.739130	1.709677	3.326087	3.048387	2.195652	2.266667	2.869565	3.161290
8		3	1	0.4301676	4.301205	4.359375	2.259036	2.234375	3.259036	3.130208	2.814815	2.803191	3.855422	3.911458
9		3	2	0.3517241	4.441176	4.446602	2.029412	1.864078	3.251337	3.174757	2.795455	2.773196	3.856757	3.828431
10		1	3	0.5229358	4.476190	4.739130	1.809524	1.500000	3.277778	3.217391	2.557377	2.125000	3.317460	3.065217
11		1	1	0.2285714	4.650602	4.745614	1.656627	1.508772	3.307229	3.289474	2.416667	2.277778	3.343373	3.254386
12		4	2	0.4032258	4.416129	4.338710	1.954839	2.016129	3.241935	3.284946	2.791946	2.877778	3.810458	3.875000
13		3	1	0.5000000	4.329670	4.315603	2.170330	2.148936	3.159341	3.198582	2.857143	2.912409	4.027473	3.946429
14		2	3	0.5238095	4.379562	4.541096	2.135036	2.095890	3.175182	3.130137	2.766917	2.794521	4.103704	4.075342

#dodawanie, odejmowanie, mnożenie i dzielenie wartości pomiędzy kolumnami

```
library(dplyr)
submitsummary<-change.per.submit.merged%>%
  group_by(Access)%>%
  mutate(changeL=Lnew-Lo1d, changeT=Tnew+To1d,
changeM=Mnew*Mo1d, changeN=Nnew/No1d, changer=Rnew+Ro1d)%>%
  select(-(Lo1d:Rnew))
```

```
# A tibble: 14 x 8
# Groups:   Access [4]
   Access Tourism BCdist changeL changeT changeM changeN changer
  <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
1      2.      1.    0.366    0.110     3.29    10.5     0.926     6.91
2      3.      1.    0.530    0.0686     3.85     9.50     0.972     8.26
3      4.      2.    0.682   -0.357     4.61    11.0     1.14     6.82
4      2.      2.    0.541    0.287     3.28    10.4     0.850     6.45
5      4.      2.    0.673   -0.377     5.33     9.32     1.10     8.45
6      4.      1.    0.526   -0.0691     3.92     9.44     0.992     8.11
7      4.      1.    0.963    0.0736     3.45    10.1     1.03     6.03
8      3.      1.    0.430    0.0582     4.49    10.2     0.996     7.77
9      3.      2.    0.352    0.00543     3.89    10.3     0.992     7.69
10     1.      3.    0.523    0.263     3.31    10.5     0.831     6.38
11     1.      1.    0.229    0.0950     3.17    10.9     0.943     6.60
12     4.      2.    0.403   -0.0774     3.97    10.6     1.03     7.69
13     3.      1.    0.500   -0.0141     4.32    10.1     1.02     7.97
14     2.      3.    0.524    0.162     4.23     9.94     1.01     8.18
```

#wyłanianie z tabeli konkretnych wartości, grupowanie i  
podsumowywanie danych

meteo

	station	date	temp	day	month	year
1	Kasprowy wierch	1966-01-01	-7.0	1	1	1966
2	Kasprowy wierch	1966-01-02	-6.1	2	1	1966
3	Kasprowy wierch	1966-01-03	-7.4	3	1	1966
4	Kasprowy wierch	1966-01-04	-12.0	4	1	1966
5	Kasprowy wierch	1966-01-05	-19.0	5	1	1966
6	Kasprowy wierch	1966-01-06	-20.2	6	1	1966
7	Kasprowy wierch	1966-01-07	-17.2	7	1	1966
8	Kasprowy wierch	1966-01-08	-13.9	8	1	1966
9	Kasprowy wierch	1966-01-09	-15.9	9	1	1966
10	Kasprowy wierch	1966-01-10	-16.4	10	1	1966
11	Kasprowy wierch	1966-01-11	-17.3	11	1	1966
12	Kasprowy wierch	1966-01-12	-12.0	12	1	1966
13	Kasprowy wierch	1966-01-13	-15.8	13	1	1966
14	Kasprowy wierch	1966-01-14	-17.2	14	1	1966
15	Kasprowy wierch	1966-01-15	-15.3	15	1	1966
16	Kasprowy wierch	1966-01-16	-13.3	16	1	1966
17	Kasprowy wierch	1966-01-17	-11.2	17	1	1966
18	Kasprowy wierch	1966-01-18	-11.5	18	1	1966
19	Kasprowy wierch	1966-01-19	-7.6	19	1	1966
20	Kasprowy wierch	1966-01-20	-8.6	20	1	1966
21	Kasprowy wierch	1966-01-21	-7.6	21	1	1966
22	Kasprowy wierch	1966-01-22	-6.8	22	1	1966
23	Kasprowy wierch	1966-01-23	-5.9	23	1	1966
24	Kasprowy wierch	1966-01-24	-8.1	24	1	1966

```
meteo2<-meteo%>%filter(temp>5)%>%group_by(year,  
station)%>%summarise(length=n(), mean.temp=mean(temp))
```

	year <dbl>	station <fct>	length <int>	mean.temp <dbl>
1	1951.	Kasprowy wierch	89	9.17
2	1951.	Zakopane	189	12.1
3	1952.	Kasprowy wierch	87	8.85
4	1952.	Zakopane	171	12.2
5	1953.	Kasprowy wierch	114	8.33
6	1953.	Zakopane	186	12.2
7	1954.	Kasprowy wierch	92	8.41
8	1954.	Zakopane	180	12.1
9	1955.	Kasprowy wierch	93	7.79
10	1955.	Zakopane	173	11.5
#	... with 122 more rows			



