



DATA MANIPULATION WITH DPLYR

select

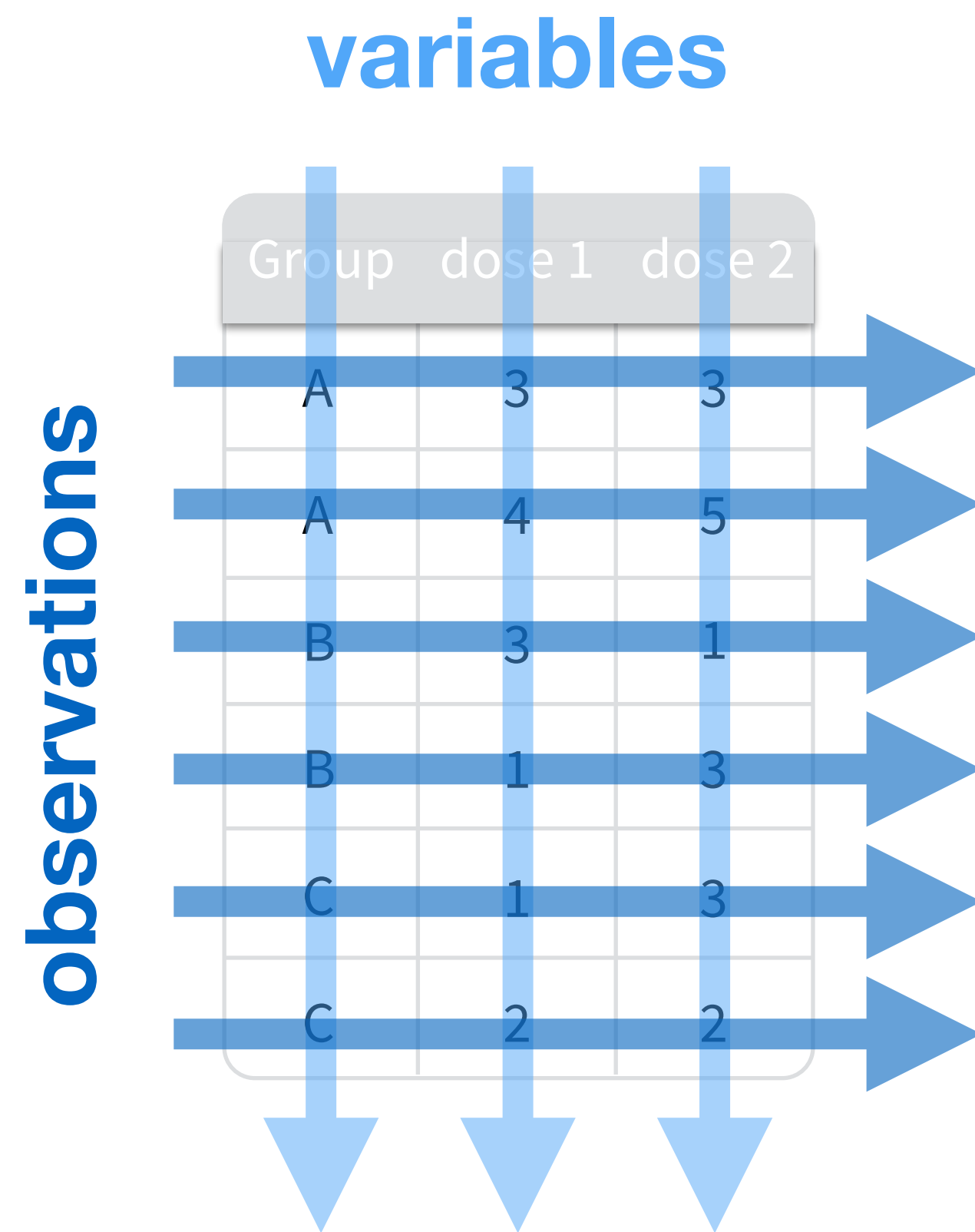
select
filter
arrange
mutate
summarize

Group	dose 1	dose 2	Sum
A	3	3	6
A	4	5	9
B	3	1	4
B	2	3	4
C	1	3	4
C	2	2	4

massage

n	min	mean	max
6	4	5.2	9

summarise



select**filter**

variables

arrange**mutate**

observations

summarize

groups

tidyr

tidy data

Year	ActualElapsedTime
Month	AirTime
DayofMonth	ArrDelay
DayofWeek	DepDelay
DepTime	Origin
ArrTime	Dest
UniqueCarrier	Distance
FlightNum	TaxiIn
TailNum	TaxiOut
	Cancelled
	CancellationCode
	Diverted

tbl

columns to
select

```
select(df, Group, Sum)
```

Group	dose 1	dose 2	Sum
A	3	3	6
A	4	5	9
B	3	1	4
B	1	3	4
C	1	3	4
C	2	2	4



DATA MANIPULATION WITH DPLYR

Let's practice!



DATA MANIPULATION WITH DPLYR

mutate

mutate

Group	dose 1	dose 2	Sum
A	3	3	6
A	4	5	9
B	3	1	4
B	1	3	4
C	1	3	4
C	2	2	4

mutate

length x width x height = volume

mass / volume = density

length	width	height	volume
	mass	volume	density
2	50	40	40
3	3	3	9
1	33	10	335



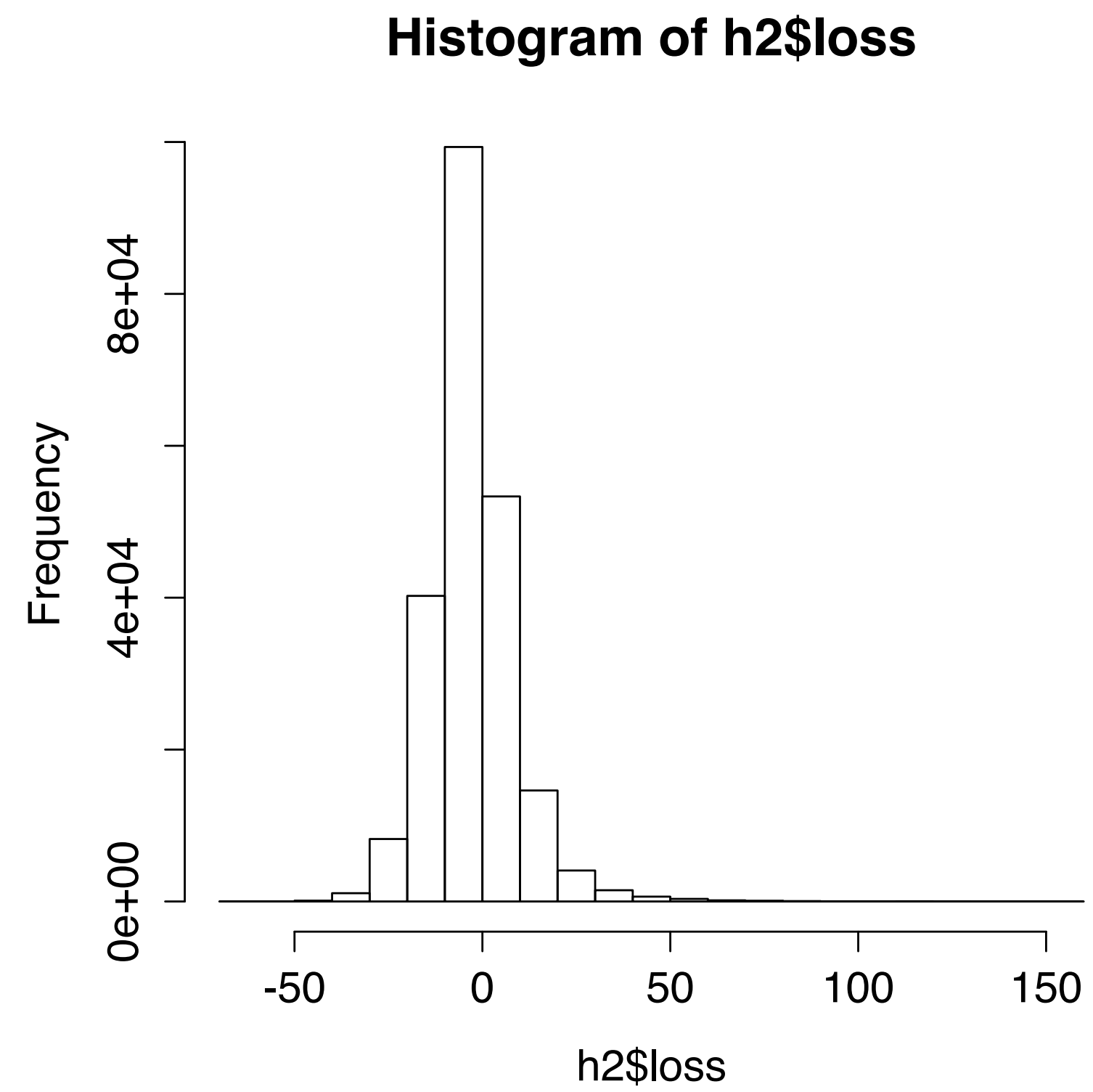
```
mutate(h1, loss = ArrDelay - DepDelay)
```

Year	ArrDelay	DepDelay	loss
2011	-10	0	-10
2011	-9	1	-10
2011	-8	-8	0
2011	3	3	0
2011	-3	5	-8
2011	-7	-1	-6

```
> mean(h2$loss, na.rm = TRUE)  
[1] -2.32
```



```
> hist(h2$loss)
```





DATA MANIPULATION WITH DPLYR

Let's practice!