

# Task 1

In [2]:

```
library(tidyverse)
library(dplyr)
library(reshape2)
library(modeest)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

Attaching package: 'reshape2'

The following object is masked from 'package:tidyverse':

smiths

-- Attaching packages --

----- tidyverse 1.3.1 -----

```
v ggplot2 3.3.5      v purrr   0.3.4
v tibble   3.1.5      v stringr 1.4.0
v readr    2.1.1      vforcats 0.5.1
```

-- Conflicts --

----- tidyverse\_conflicts() --

```
x dplyr::filter() masks stats::filter()
x dplyr::lag()   masks stats::lag()
```

Registered S3 method overwritten by 'rmutil':

```
method      from
print.response httr
```

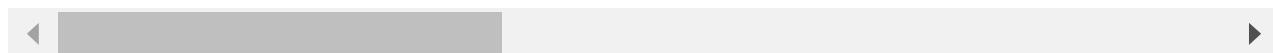
In [3]:

```
# Using tidyverse to load dataset who
who_data = tidyverse::who
who_data
```

country	iso2	iso3	year	new_sp_m014	new_sp_m1524	new_sp_m2534	new_sp_m3544	new_s
<chr>	<chr>	<chr>	<int>	<int>	<int>	<int>	<int>	<int>
Afghanistan	AF	AFG	1980	NA	NA	NA	NA	NA

country	iso2	iso3	year	new_sp_m014	new_sp_m1524	new_sp_m2534	new_sp_m3544	new_s
<chr>	<chr>	<chr>	<int>	<int>	<int>	<int>	<int>	<int>
Afghanistan	AF	AFG	1981	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1982	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1983	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1984	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1985	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1986	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1987	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1988	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1989	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1990	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1991	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1992	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1993	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1994	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1995	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1996	NA	NA	NA	NA	NA
Afghanistan	AF	AFG	1997	0	10	6	3	
Afghanistan	AF	AFG	1998	30	129	128	90	
Afghanistan	AF	AFG	1999	8	55	55	47	
Afghanistan	AF	AFG	2000	52	228	183	149	
Afghanistan	AF	AFG	2001	129	379	349	274	
Afghanistan	AF	AFG	2002	90	476	481	368	
Afghanistan	AF	AFG	2003	127	511	436	284	
Afghanistan	AF	AFG	2004	139	537	568	360	
Afghanistan	AF	AFG	2005	151	606	560	472	
Afghanistan	AF	AFG	2006	193	837	791	574	
Afghanistan	AF	AFG	2007	186	856	840	597	
Afghanistan	AF	AFG	2008	187	941	773	545	
Afghanistan	AF	AFG	2009	200	906	705	499	
...	...	...	...	...	...	...	...	...
Zimbabwe	ZW	ZWE	1984	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1985	NA	NA	NA	NA	NA

country	iso2	iso3	year	new_sp_m014	new_sp_m1524	new_sp_m2534	new_sp_m3544	new_s
<chr>	<chr>	<chr>	<int>	<int>	<int>	<int>	<int>	<int>
Zimbabwe	ZW	ZWE	1986	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1987	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1988	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1989	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1990	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1991	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1992	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1993	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1994	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1995	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1996	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1997	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1998	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	1999	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	2000	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	2001	NA	NA	NA	NA	NA
Zimbabwe	ZW	ZWE	2002	191	600	2548	1662	
Zimbabwe	ZW	ZWE	2003	133	874	3048	2228	
Zimbabwe	ZW	ZWE	2004	187	833	2908	2298	
Zimbabwe	ZW	ZWE	2005	210	837	2264	1855	
Zimbabwe	ZW	ZWE	2006	215	736	2391	1939	
Zimbabwe	ZW	ZWE	2007	138	500	3693	0	
Zimbabwe	ZW	ZWE	2008	127	614	0	3316	
Zimbabwe	ZW	ZWE	2009	125	578	NA	3471	
Zimbabwe	ZW	ZWE	2010	150	710	2208	1682	
Zimbabwe	ZW	ZWE	2011	152	784	2467	2071	
Zimbabwe	ZW	ZWE	2012	120	783	2421	2086	
Zimbabwe	ZW	ZWE	2013	NA	NA	NA	NA	



## Q1

In [4]:

```
who1 <- who_data %>%
  pivot_longer(
```

```

  cols = new_sp_m014:newrel_f65,
  names_to = "key",
  values_to = "value",
  values_drop_na = TRUE
)
who1

```

A tibble: 76046 × 6

country	iso2	iso3	year	key	value
<chr>	<chr>	<chr>	<int>	<chr>	<int>
Afghanistan	AF	AFG	1997	new_sp_m014	0
Afghanistan	AF	AFG	1997	new_sp_m1524	10
Afghanistan	AF	AFG	1997	new_sp_m2534	6
Afghanistan	AF	AFG	1997	new_sp_m3544	3
Afghanistan	AF	AFG	1997	new_sp_m4554	5
Afghanistan	AF	AFG	1997	new_sp_m5564	2
Afghanistan	AF	AFG	1997	new_sp_m65	0
Afghanistan	AF	AFG	1997	new_sp_f014	5
Afghanistan	AF	AFG	1997	new_sp_f1524	38
Afghanistan	AF	AFG	1997	new_sp_f2534	36
Afghanistan	AF	AFG	1997	new_sp_f3544	14
Afghanistan	AF	AFG	1997	new_sp_f4554	8
Afghanistan	AF	AFG	1997	new_sp_f5564	0
Afghanistan	AF	AFG	1997	new_sp_f65	1
Afghanistan	AF	AFG	1998	new_sp_m014	30
Afghanistan	AF	AFG	1998	new_sp_m1524	129
Afghanistan	AF	AFG	1998	new_sp_m2534	128
Afghanistan	AF	AFG	1998	new_sp_m3544	90
Afghanistan	AF	AFG	1998	new_sp_m4554	89
Afghanistan	AF	AFG	1998	new_sp_m5564	64
Afghanistan	AF	AFG	1998	new_sp_m65	41
Afghanistan	AF	AFG	1998	new_sp_f014	45
Afghanistan	AF	AFG	1998	new_sp_f1524	350
Afghanistan	AF	AFG	1998	new_sp_f2534	419
Afghanistan	AF	AFG	1998	new_sp_f3544	194
Afghanistan	AF	AFG	1998	new_sp_f4554	118
Afghanistan	AF	AFG	1998	new_sp_f5564	61
Afghanistan	AF	AFG	1998	new_sp_f65	20

country	iso2	iso3	year	key	value
<chr>	<chr>	<chr>	<int>	<chr>	<int>
Afghanistan	AF	AFG	1999	new_sp_m014	8
Afghanistan	AF	AFG	1999	new_sp_m1524	55
...	...	...	...	...	...
Zimbabwe	ZW	ZWE	2012	new_sn_f5564	516
Zimbabwe	ZW	ZWE	2012	new_sn_f65	432
Zimbabwe	ZW	ZWE	2012	new_ep_m014	233
Zimbabwe	ZW	ZWE	2012	new_ep_m1524	214
Zimbabwe	ZW	ZWE	2012	new_ep_m2534	658
Zimbabwe	ZW	ZWE	2012	new_ep_m3544	789
Zimbabwe	ZW	ZWE	2012	new_ep_m4554	331
Zimbabwe	ZW	ZWE	2012	new_ep_m5564	178
Zimbabwe	ZW	ZWE	2012	new_ep_m65	182
Zimbabwe	ZW	ZWE	2012	new_ep_f014	208
Zimbabwe	ZW	ZWE	2012	new_ep_f1524	319
Zimbabwe	ZW	ZWE	2012	new_ep_f2534	710
Zimbabwe	ZW	ZWE	2012	new_ep_f3544	579
Zimbabwe	ZW	ZWE	2012	new_ep_f4554	228
Zimbabwe	ZW	ZWE	2012	new_ep_f5564	140
Zimbabwe	ZW	ZWE	2012	new_ep_f65	143
Zimbabwe	ZW	ZWE	2013	newrel_m014	1315
Zimbabwe	ZW	ZWE	2013	newrel_m1524	1642
Zimbabwe	ZW	ZWE	2013	newrel_m2534	5331
Zimbabwe	ZW	ZWE	2013	newrel_m3544	5363
Zimbabwe	ZW	ZWE	2013	newrel_m4554	2349
Zimbabwe	ZW	ZWE	2013	newrel_m5564	1206
Zimbabwe	ZW	ZWE	2013	newrel_m65	1208
Zimbabwe	ZW	ZWE	2013	newrel_f014	1252
Zimbabwe	ZW	ZWE	2013	newrel_f1524	2069
Zimbabwe	ZW	ZWE	2013	newrel_f2534	4649
Zimbabwe	ZW	ZWE	2013	newrel_f3544	3526
Zimbabwe	ZW	ZWE	2013	newrel_f4554	1453
Zimbabwe	ZW	ZWE	2013	newrel_f5564	811

country	iso2	iso3	year	key	value
<chr>	<chr>	<chr>	<int>	<chr>	<int>
Zimbabwe	ZW	ZWE	2013	newrel_f65	725

## Q2

In [5]:

```
who2 %>% mutate(key = str_replace(key, "newrel","new_rel"))
who2
```

A tibble: 76046 × 6

country	iso2	iso3	year	key	value
<chr>	<chr>	<chr>	<int>	<chr>	<int>
Afghanistan	AF	AFG	1997	new_sp_m014	0
Afghanistan	AF	AFG	1997	new_sp_m1524	10
Afghanistan	AF	AFG	1997	new_sp_m2534	6
Afghanistan	AF	AFG	1997	new_sp_m3544	3
Afghanistan	AF	AFG	1997	new_sp_m4554	5
Afghanistan	AF	AFG	1997	new_sp_m5564	2
Afghanistan	AF	AFG	1997	new_sp_m65	0
Afghanistan	AF	AFG	1997	new_sp_f014	5
Afghanistan	AF	AFG	1997	new_sp_f1524	38
Afghanistan	AF	AFG	1997	new_sp_f2534	36
Afghanistan	AF	AFG	1997	new_sp_f3544	14
Afghanistan	AF	AFG	1997	new_sp_f4554	8
Afghanistan	AF	AFG	1997	new_sp_f5564	0
Afghanistan	AF	AFG	1997	new_sp_f65	1
Afghanistan	AF	AFG	1998	new_sp_m014	30
Afghanistan	AF	AFG	1998	new_sp_m1524	129
Afghanistan	AF	AFG	1998	new_sp_m2534	128
Afghanistan	AF	AFG	1998	new_sp_m3544	90
Afghanistan	AF	AFG	1998	new_sp_m4554	89
Afghanistan	AF	AFG	1998	new_sp_m5564	64
Afghanistan	AF	AFG	1998	new_sp_m65	41
Afghanistan	AF	AFG	1998	new_sp_f014	45
Afghanistan	AF	AFG	1998	new_sp_f1524	350
Afghanistan	AF	AFG	1998	new_sp_f2534	419

country	iso2	iso3	year	key	value
<chr>	<chr>	<chr>	<int>	<chr>	<int>
Afghanistan	AF	AFG	1998	new_sp_f3544	194
Afghanistan	AF	AFG	1998	new_sp_f4554	118
Afghanistan	AF	AFG	1998	new_sp_f5564	61
Afghanistan	AF	AFG	1998	new_sp_f65	20
Afghanistan	AF	AFG	1999	new_sp_m014	8
Afghanistan	AF	AFG	1999	new_sp_m1524	55
...	...	...	...	...	...
Zimbabwe	ZW	ZWE	2012	new_sn_f5564	516
Zimbabwe	ZW	ZWE	2012	new_sn_f65	432
Zimbabwe	ZW	ZWE	2012	new_ep_m014	233
Zimbabwe	ZW	ZWE	2012	new_ep_m1524	214
Zimbabwe	ZW	ZWE	2012	new_ep_m2534	658
Zimbabwe	ZW	ZWE	2012	new_ep_m3544	789
Zimbabwe	ZW	ZWE	2012	new_ep_m4554	331
Zimbabwe	ZW	ZWE	2012	new_ep_m5564	178
Zimbabwe	ZW	ZWE	2012	new_ep_m65	182
Zimbabwe	ZW	ZWE	2012	new_ep_f014	208
Zimbabwe	ZW	ZWE	2012	new_ep_f1524	319
Zimbabwe	ZW	ZWE	2012	new_ep_f2534	710
Zimbabwe	ZW	ZWE	2012	new_ep_f3544	579
Zimbabwe	ZW	ZWE	2012	new_ep_f4554	228
Zimbabwe	ZW	ZWE	2012	new_ep_f5564	140
Zimbabwe	ZW	ZWE	2012	new_ep_f65	143
Zimbabwe	ZW	ZWE	2013	new_rel_m014	1315
Zimbabwe	ZW	ZWE	2013	new_rel_m1524	1642
Zimbabwe	ZW	ZWE	2013	new_rel_m2534	5331
Zimbabwe	ZW	ZWE	2013	new_rel_m3544	5363
Zimbabwe	ZW	ZWE	2013	new_rel_m4554	2349
Zimbabwe	ZW	ZWE	2013	new_rel_m5564	1206
Zimbabwe	ZW	ZWE	2013	new_rel_m65	1208
Zimbabwe	ZW	ZWE	2013	new_rel_f014	1252
Zimbabwe	ZW	ZWE	2013	new_rel_f1524	2069

country	iso2	iso3	year	key	value
<chr>	<chr>	<chr>	<int>	<chr>	<int>
Zimbabwe	ZW	ZWE	2013	new_rel_f2534	4649
Zimbabwe	ZW	ZWE	2013	new_rel_f3544	3526
Zimbabwe	ZW	ZWE	2013	new_rel_f4554	1453
Zimbabwe	ZW	ZWE	2013	new_rel_f5564	811
Zimbabwe	ZW	ZWE	2013	new_rel_f65	725

## Q3

In [6]:

```
who3 <- who2 %>% separate(key,c("new", "type", "sexage"),sep="_")
who3
```

A tibble: 76046 × 8

country	iso2	iso3	year	new	type	sexage	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<int>
Afghanistan	AF	AFG	1997	new	sp	m014	0
Afghanistan	AF	AFG	1997	new	sp	m1524	10
Afghanistan	AF	AFG	1997	new	sp	m2534	6
Afghanistan	AF	AFG	1997	new	sp	m3544	3
Afghanistan	AF	AFG	1997	new	sp	m4554	5
Afghanistan	AF	AFG	1997	new	sp	m5564	2
Afghanistan	AF	AFG	1997	new	sp	m65	0
Afghanistan	AF	AFG	1997	new	sp	f014	5
Afghanistan	AF	AFG	1997	new	sp	f1524	38
Afghanistan	AF	AFG	1997	new	sp	f2534	36
Afghanistan	AF	AFG	1997	new	sp	f3544	14
Afghanistan	AF	AFG	1997	new	sp	f4554	8
Afghanistan	AF	AFG	1997	new	sp	f5564	0
Afghanistan	AF	AFG	1997	new	sp	f65	1
Afghanistan	AF	AFG	1998	new	sp	m014	30
Afghanistan	AF	AFG	1998	new	sp	m1524	129
Afghanistan	AF	AFG	1998	new	sp	m2534	128
Afghanistan	AF	AFG	1998	new	sp	m3544	90
Afghanistan	AF	AFG	1998	new	sp	m4554	89
Afghanistan	AF	AFG	1998	new	sp	m5564	64

country	iso2	iso3	year	new	type	sexage	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<int>
Afghanistan	AF	AFG	1998	new	sp	m65	41
Afghanistan	AF	AFG	1998	new	sp	f014	45
Afghanistan	AF	AFG	1998	new	sp	f1524	350
Afghanistan	AF	AFG	1998	new	sp	f2534	419
Afghanistan	AF	AFG	1998	new	sp	f3544	194
Afghanistan	AF	AFG	1998	new	sp	f4554	118
Afghanistan	AF	AFG	1998	new	sp	f5564	61
Afghanistan	AF	AFG	1998	new	sp	f65	20
Afghanistan	AF	AFG	1999	new	sp	m014	8
Afghanistan	AF	AFG	1999	new	sp	m1524	55
...	...	...	...	...	...	...	...
Zimbabwe	ZW	ZWE	2012	new	sn	f5564	516
Zimbabwe	ZW	ZWE	2012	new	sn	f65	432
Zimbabwe	ZW	ZWE	2012	new	ep	m014	233
Zimbabwe	ZW	ZWE	2012	new	ep	m1524	214
Zimbabwe	ZW	ZWE	2012	new	ep	m2534	658
Zimbabwe	ZW	ZWE	2012	new	ep	m3544	789
Zimbabwe	ZW	ZWE	2012	new	ep	m4554	331
Zimbabwe	ZW	ZWE	2012	new	ep	m5564	178
Zimbabwe	ZW	ZWE	2012	new	ep	m65	182
Zimbabwe	ZW	ZWE	2012	new	ep	f014	208
Zimbabwe	ZW	ZWE	2012	new	ep	f1524	319
Zimbabwe	ZW	ZWE	2012	new	ep	f2534	710
Zimbabwe	ZW	ZWE	2012	new	ep	f3544	579
Zimbabwe	ZW	ZWE	2012	new	ep	f4554	228
Zimbabwe	ZW	ZWE	2012	new	ep	f5564	140
Zimbabwe	ZW	ZWE	2012	new	ep	f65	143
Zimbabwe	ZW	ZWE	2013	new	rel	m014	1315
Zimbabwe	ZW	ZWE	2013	new	rel	m1524	1642
Zimbabwe	ZW	ZWE	2013	new	rel	m2534	5331
Zimbabwe	ZW	ZWE	2013	new	rel	m3544	5363
Zimbabwe	ZW	ZWE	2013	new	rel	m4554	2349

country	iso2	iso3	year	new	type	sexage	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<int>
Zimbabwe	ZW	ZWE	2013	new	rel	m5564	1206
Zimbabwe	ZW	ZWE	2013	new	rel	m65	1208
Zimbabwe	ZW	ZWE	2013	new	rel	f014	1252
Zimbabwe	ZW	ZWE	2013	new	rel	f1524	2069
Zimbabwe	ZW	ZWE	2013	new	rel	f2534	4649
Zimbabwe	ZW	ZWE	2013	new	rel	f3544	3526
Zimbabwe	ZW	ZWE	2013	new	rel	f4554	1453
Zimbabwe	ZW	ZWE	2013	new	rel	f5564	811
Zimbabwe	ZW	ZWE	2013	new	rel	f65	725

## Q4

In [7]:

```
who4 <- who3 %>% separate(sexage,c("sex","Age"),sep="(?<=[a-z])(?=[0-9])")
who4
```

A tibble: 76046 × 9

country	iso2	iso3	year	new	type	sex	Age	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<int>
Afghanistan	AF	AFG	1997	new	sp	m	014	0
Afghanistan	AF	AFG	1997	new	sp	m	1524	10
Afghanistan	AF	AFG	1997	new	sp	m	2534	6
Afghanistan	AF	AFG	1997	new	sp	m	3544	3
Afghanistan	AF	AFG	1997	new	sp	m	4554	5
Afghanistan	AF	AFG	1997	new	sp	m	5564	2
Afghanistan	AF	AFG	1997	new	sp	m	65	0
Afghanistan	AF	AFG	1997	new	sp	f	014	5
Afghanistan	AF	AFG	1997	new	sp	f	1524	38
Afghanistan	AF	AFG	1997	new	sp	f	2534	36
Afghanistan	AF	AFG	1997	new	sp	f	3544	14
Afghanistan	AF	AFG	1997	new	sp	f	4554	8
Afghanistan	AF	AFG	1997	new	sp	f	5564	0
Afghanistan	AF	AFG	1997	new	sp	f	65	1
Afghanistan	AF	AFG	1998	new	sp	m	014	30
Afghanistan	AF	AFG	1998	new	sp	m	1524	129

country	iso2	iso3	year	new	type	sex	Age	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<int>
Afghanistan	AF	AFG	1998	new	sp	m	2534	128
Afghanistan	AF	AFG	1998	new	sp	m	3544	90
Afghanistan	AF	AFG	1998	new	sp	m	4554	89
Afghanistan	AF	AFG	1998	new	sp	m	5564	64
Afghanistan	AF	AFG	1998	new	sp	m	65	41
Afghanistan	AF	AFG	1998	new	sp	f	014	45
Afghanistan	AF	AFG	1998	new	sp	f	1524	350
Afghanistan	AF	AFG	1998	new	sp	f	2534	419
Afghanistan	AF	AFG	1998	new	sp	f	3544	194
Afghanistan	AF	AFG	1998	new	sp	f	4554	118
Afghanistan	AF	AFG	1998	new	sp	f	5564	61
Afghanistan	AF	AFG	1998	new	sp	f	65	20
Afghanistan	AF	AFG	1999	new	sp	m	014	8
Afghanistan	AF	AFG	1999	new	sp	m	1524	55
...	...	...	...	...	...	...	...	...
Zimbabwe	ZW	ZWE	2012	new	sn	f	5564	516
Zimbabwe	ZW	ZWE	2012	new	sn	f	65	432
Zimbabwe	ZW	ZWE	2012	new	ep	m	014	233
Zimbabwe	ZW	ZWE	2012	new	ep	m	1524	214
Zimbabwe	ZW	ZWE	2012	new	ep	m	2534	658
Zimbabwe	ZW	ZWE	2012	new	ep	m	3544	789
Zimbabwe	ZW	ZWE	2012	new	ep	m	4554	331
Zimbabwe	ZW	ZWE	2012	new	ep	m	5564	178
Zimbabwe	ZW	ZWE	2012	new	ep	m	65	182
Zimbabwe	ZW	ZWE	2012	new	ep	f	014	208
Zimbabwe	ZW	ZWE	2012	new	ep	f	1524	319
Zimbabwe	ZW	ZWE	2012	new	ep	f	2534	710
Zimbabwe	ZW	ZWE	2012	new	ep	f	3544	579
Zimbabwe	ZW	ZWE	2012	new	ep	f	4554	228
Zimbabwe	ZW	ZWE	2012	new	ep	f	5564	140
Zimbabwe	ZW	ZWE	2012	new	ep	f	65	143
Zimbabwe	ZW	ZWE	2013	new	rel	m	014	1315

country	iso2	iso3	year	new	type	sex	Age	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<int>
Zimbabwe	ZW	ZWE	2013	new	rel	m	1524	1642
Zimbabwe	ZW	ZWE	2013	new	rel	m	2534	5331
Zimbabwe	ZW	ZWE	2013	new	rel	m	3544	5363
Zimbabwe	ZW	ZWE	2013	new	rel	m	4554	2349
Zimbabwe	ZW	ZWE	2013	new	rel	m	5564	1206
Zimbabwe	ZW	ZWE	2013	new	rel	m	65	1208
Zimbabwe	ZW	ZWE	2013	new	rel	f	014	1252
Zimbabwe	ZW	ZWE	2013	new	rel	f	1524	2069
Zimbabwe	ZW	ZWE	2013	new	rel	f	2534	4649
Zimbabwe	ZW	ZWE	2013	new	rel	f	3544	3526
Zimbabwe	ZW	ZWE	2013	new	rel	f	4554	1453
Zimbabwe	ZW	ZWE	2013	new	rel	f	5564	811
Zimbabwe	ZW	ZWE	2013	new	rel	f	65	725

## Q5

In [8]:

```
head(who4)
tail(who4)
```

A tibble: 6 × 9

country	iso2	iso3	year	new	type	sex	Age	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<int>
Afghanistan	AF	AFG	1997	new	sp	m	014	0
Afghanistan	AF	AFG	1997	new	sp	m	1524	10
Afghanistan	AF	AFG	1997	new	sp	m	2534	6
Afghanistan	AF	AFG	1997	new	sp	m	3544	3
Afghanistan	AF	AFG	1997	new	sp	m	4554	5
Afghanistan	AF	AFG	1997	new	sp	m	5564	2

A tibble: 6 × 9

country	iso2	iso3	year	new	type	sex	Age	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<int>
Zimbabwe	ZW	ZWE	2013	new	rel	f	1524	2069
Zimbabwe	ZW	ZWE	2013	new	rel	f	2534	4649
Zimbabwe	ZW	ZWE	2013	new	rel	f	3544	3526

country	iso2	iso3	year	new	type	sex	Age	value
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<chr>	<int>
Zimbabwe	ZW	ZWE	2013	new	rel	f	4554	1453
Zimbabwe	ZW	ZWE	2013	new	rel	f	5564	811
Zimbabwe	ZW	ZWE	2013	new	rel	f	65	725

## Q6

```
In [9]: write.csv(who4, "who.csv")
```

## Task 2

```
In [10]: data(iris)
head(iris, 6)
```

A data.frame: 6 × 5

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
	<dbl>	<dbl>	<dbl>	<dbl>	<fct>
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa

## Q1 Calculating the mean, median and mode of sepal length

```
In [11]: mean(iris$Sepal.Length, na.rm = TRUE)
median(iris$Sepal.Length, na.rm = TRUE)
mfv(iris$Sepal.Length, na_rm=TRUE)
```

5.84333333333333

5.8

5

## Q2

```
In [12]: min(iris$Sepal.Length,na_rm=TRUE)
max(iris$Sepal.Length,na_rm=TRUE)
```

```
range(iris$Sepal.Length, na.rm=TRUE)
```

1  
7.9  
1 · 7.9

## Q3

```
In [13]: IQR(iris$Sepal.Length, na.rm=TRUE)
quantile(iris$Sepal.Length, na.rm=TRUE)
```

1.3  
**0%: 4.3 25%: 5.1 50%: 5.8 75%: 6.4 100%: 7.9**

## Q4

```
In [14]: var(iris$Sepal.Length, na.rm = TRUE)
sd(iris$Sepal.Length, na.rm = TRUE)
```

0.685693512304251  
0.828066127977863

## Q5

```
In [15]: summary(iris$Sepal.Length)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
4.300	5.100	5.800	5.843	6.400	7.900

## Q6

```
In [16]: sapply(select(iris, -c('Species')), mean)
sapply(select(iris, -c('Species')), quantile, na.rm =TRUE)
```

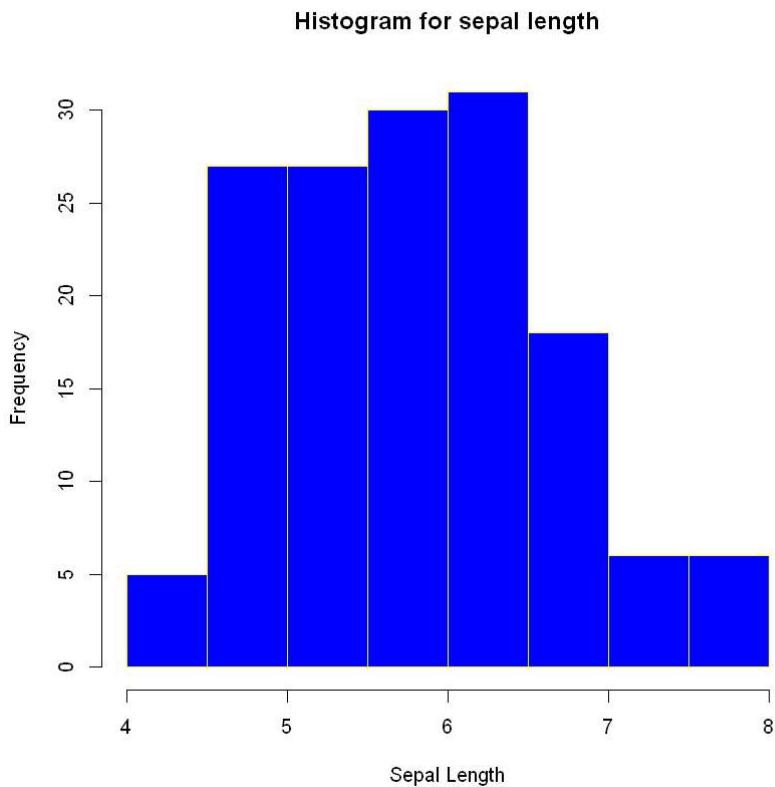
**Sepal.Length:** 5.84333333333333 **Sepal.Width:** 3.05733333333333 **Petal.Length:** 3.758  
**Petal.Width:** 1.19933333333333

A matrix: 5 × 4 of type dbl

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
<b>0%</b>	4.3	2.0	1.00	0.1
<b>25%</b>	5.1	2.8	1.60	0.3
<b>50%</b>	5.8	3.0	4.35	1.3
<b>75%</b>	6.4	3.3	5.10	1.8
<b>100%</b>	7.9	4.4	6.90	2.5

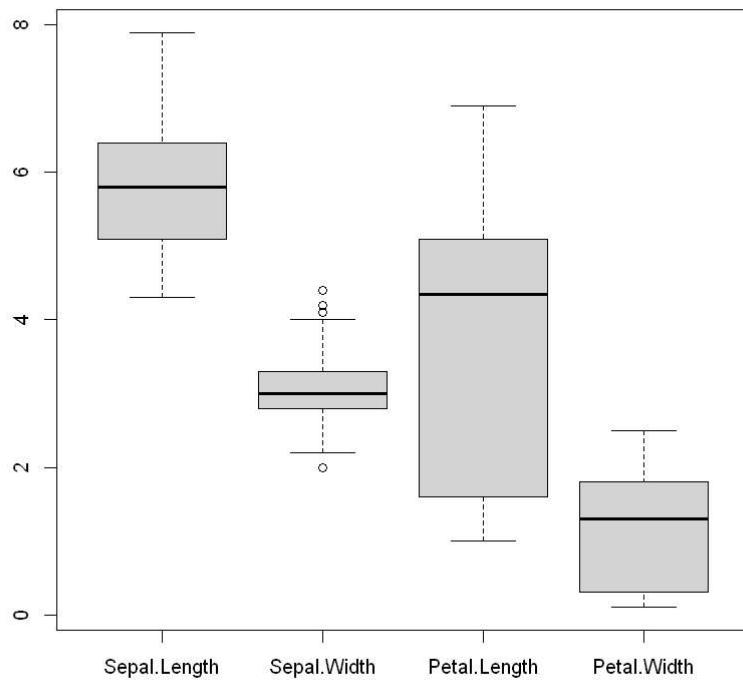
## Q7

```
In [17]: hist(iris$Sepal.Length,main='Histogram for sepal length', xlab = 'Sepal Length', col =
```



## Q8

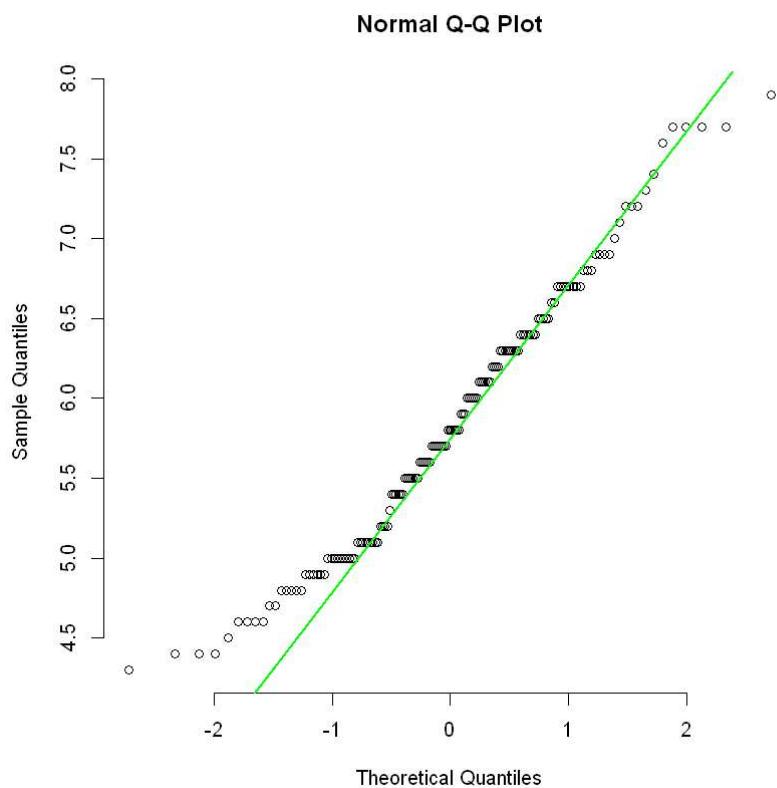
```
In [18]: boxplot(iris[,-c('Species')])
```



## Q9

In [19]:

```
qqnorm(iris$Sepal.Length, pch = 1, frame = FALSE)
qqline(iris$Sepal.Length, col = "green", lwd = 2)
```



# Task-3

In [20]:

```
head(mpg)
```

A tibble: 6 × 11

manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
<chr>	<chr>	<dbl>	<int>	<int>	<chr>	<chr>	<int>	<int>	<chr>	<chr>
audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact

## Q1

In [21]:

```
unique(mpg$manufacturer)

df_grp_manufacturer = mpg %>% group_by(manufacturer) %>%
  summarise(city_mpg = mean(cty),
            high_waympb = mean(hwy),
            .groups = 'drop')

df_grp_manufacturer
df_manufacturer <- melt(df_grp_manufacturer[,c('manufacturer','city_mpg','high_waympb')]
df_manufacturer
```

'audi' · 'chevrolet' · 'dodge' · 'ford' · 'honda' · 'hyundai' · 'jeep' · 'land rover' · 'lincoln' · 'mercury' · 'nissan' · 'pontiac' · 'subaru' · 'toyota' · 'volkswagen'

A tibble: 15 × 3

manufacturer	city_mpg	high_waympb
--------------	----------	-------------

<chr>	<dbl>	<dbl>
audi	17.61111	26.44444
chevrolet	15.00000	21.89474
dodge	13.13514	17.94595
ford	14.00000	19.36000
honda	24.44444	32.55556
hyundai	18.64286	26.85714
jeep	13.50000	17.62500
land rover	11.50000	16.50000
lincoln	11.33333	17.00000

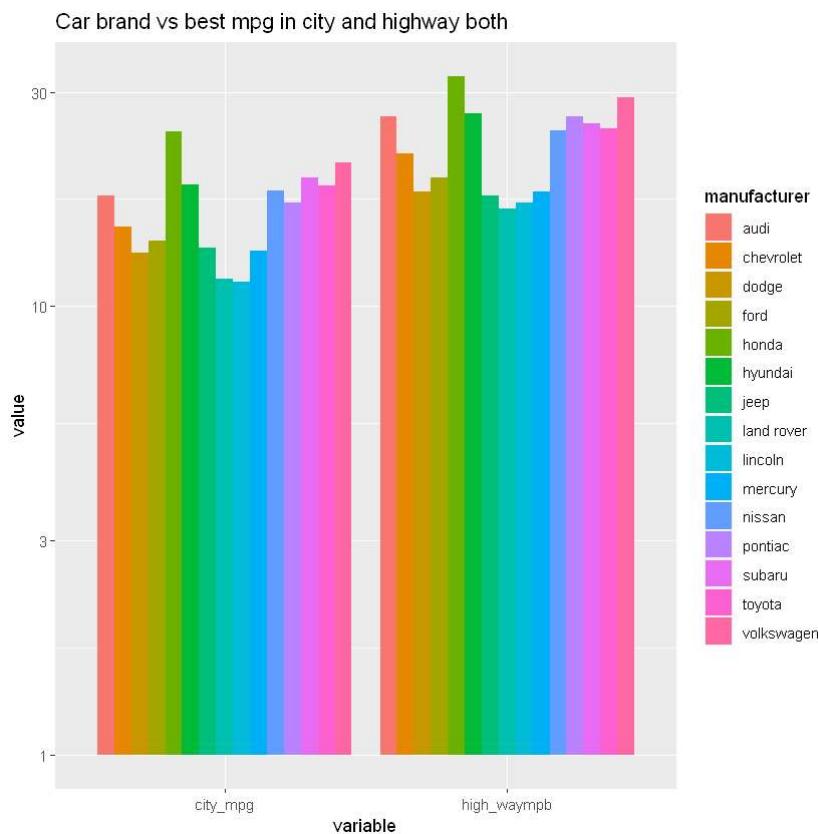
manufacturer	city_mpg	high_waympb
	<chr>	<dbl>
mercury	13.25000	18.00000
nissan	18.07692	24.61538
pontiac	17.00000	26.40000
subaru	19.28571	25.57143
toyota	18.52941	24.91176
volkswagen	20.92593	29.22222

A data.frame: 30 × 3

manufacturer	variable	value
	<chr>	<fct>
audi	city_mpg	17.61111
chevrolet	city_mpg	15.00000
dodge	city_mpg	13.13514
ford	city_mpg	14.00000
honda	city_mpg	24.44444
hyundai	city_mpg	18.64286
jeep	city_mpg	13.50000
land rover	city_mpg	11.50000
lincoln	city_mpg	11.33333
mercury	city_mpg	13.25000
nissan	city_mpg	18.07692
pontiac	city_mpg	17.00000
subaru	city_mpg	19.28571
toyota	city_mpg	18.52941
volkswagen	city_mpg	20.92593
audi	high_waympb	26.44444
chevrolet	high_waympb	21.89474
dodge	high_waympb	17.94595
ford	high_waympb	19.36000
honda	high_waympb	32.55556
hyundai	high_waympb	26.85714
jeep	high_waympb	17.62500
land rover	high_waympb	16.50000
lincoln	high_waympb	17.00000

manufacturer	variable	value
<chr>	<fct>	<dbl>
mercury	high_waympb	18.00000
nissan	high_waympb	24.61538
pontiac	high_waympb	26.40000
subaru	high_waympb	25.57143
toyota	high_waympb	24.91176
volkswagen	high_waympb	29.22222

```
In [22]: ggplot(df_manufacturer,aes(x = variable,y = value)) +
  geom_bar(aes(fill = manufacturer),stat = "identity",position = "dodge") +
  ggtitle("Car brand vs best mpg in city and highway both") + scale_y_log10()
```



## Q2

```
In [23]: type_car <- mpg %>%group_by(displ, class) %>% summarise(city_mpg = min(cty), .groups = df_type_car)
```

Error in eval(expr, envir, enclos): object 'df\_type\_car' not found  
Traceback:

```
In [24]: ggplot(df_type_car, aes(x=class, y=city_mpg, fill = factor(displ))) +
  geom_bar(stat = 'identity', position = 'dodge') +ggtitle("Lowest mpg in city regarding
  geom_text(aes(label = city_mpg), vjust = -0.2, size = 5, position = position_dodge(0.
```

```
Error in ggplot(df_type_car, aes(x = class, y = city_mpg, fill = factor(displ))): object
'df_type_car' not found
Traceback:
```

```
1. ggplot(df_type_car, aes(x = class, y = city_mpg, fill = factor(displ)))
```

## Q3

In [25]:

```
df_best_performance_car <- mpg %>% group_by(displ) %>% summarise(city_mpg = max(cty), hwy)
```

A tibble: 35 × 3

displ	city_mpg	hwy_mpg
<dbl>	<int>	<int>
1.6	28	33
1.8	28	37
1.9	35	44
2.0	22	31
2.2	21	29
2.4	22	31
2.5	23	32
2.7	17	24
2.8	18	26
3.0	19	26
3.1	18	27
3.3	19	28
3.4	15	19
3.5	19	29
3.6	17	26
3.7	15	19
3.8	18	28
3.9	14	17
4.0	17	26
4.2	16	23
4.4	12	18
4.6	15	23
4.7	14	19
5.0	13	17

displ	city_mpg	hwy_mpg
<dbl>	<int>	<int>
5.2	11	17
5.3	16	25
5.4	14	20
5.6	12	18
5.7	16	26
5.9	11	15
6.0	12	17
6.1	11	14
6.2	16	26
6.5	14	17
7.0	15	24

```
In [26]: melt_car_type <- melt(df_best_performance_car[,c('displ','city_mpg','hwy_mpg')],id.vars=melt_car_type)
```

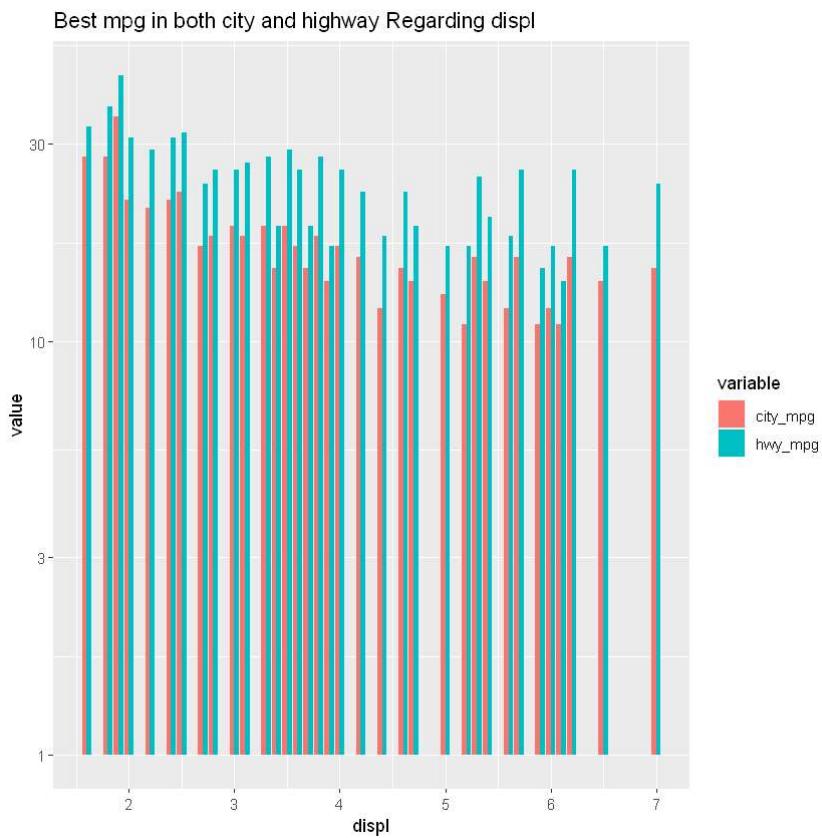
A data.frame: 70 × 3

displ	variable	value
<dbl>	<fct>	<int>
1.6	city_mpg	28
1.8	city_mpg	28
1.9	city_mpg	35
2.0	city_mpg	22
2.2	city_mpg	21
2.4	city_mpg	22
2.5	city_mpg	23
2.7	city_mpg	17
2.8	city_mpg	18
3.0	city_mpg	19
3.1	city_mpg	18
3.3	city_mpg	19
3.4	city_mpg	15
3.5	city_mpg	19
3.6	city_mpg	17
3.7	city_mpg	15

displ	variable	value
<dbl>	<fct>	<int>
3.8	city_mpg	18
3.9	city_mpg	14
4.0	city_mpg	17
4.2	city_mpg	16
4.4	city_mpg	12
4.6	city_mpg	15
4.7	city_mpg	14
5.0	city_mpg	13
5.2	city_mpg	11
5.3	city_mpg	16
5.4	city_mpg	14
5.6	city_mpg	12
5.7	city_mpg	16
5.9	city_mpg	11
...	...	...
2.4	hwy_mpg	31
2.5	hwy_mpg	32
2.7	hwy_mpg	24
2.8	hwy_mpg	26
3.0	hwy_mpg	26
3.1	hwy_mpg	27
3.3	hwy_mpg	28
3.4	hwy_mpg	19
3.5	hwy_mpg	29
3.6	hwy_mpg	26
3.7	hwy_mpg	19
3.8	hwy_mpg	28
3.9	hwy_mpg	17
4.0	hwy_mpg	26
4.2	hwy_mpg	23
4.4	hwy_mpg	18
4.6	hwy_mpg	23

displ	variable	value
<dbl>	<fct>	<int>
4.7	hwy_mpg	19
5.0	hwy_mpg	17
5.2	hwy_mpg	17
5.3	hwy_mpg	25
5.4	hwy_mpg	20
5.6	hwy_mpg	18
5.7	hwy_mpg	26
5.9	hwy_mpg	15
6.0	hwy_mpg	17
6.1	hwy_mpg	14
6.2	hwy_mpg	26
6.5	hwy_mpg	17
7.0	hwy_mpg	24

```
In [27]: ggplot(melt_car_type,aes(x = displ,y = value)) +
  geom_bar(aes(fill = variable),stat = "identity",position = "dodge") +ggtitle("Best mpg
  scale_y_log10()
```



```
In [28]:
```

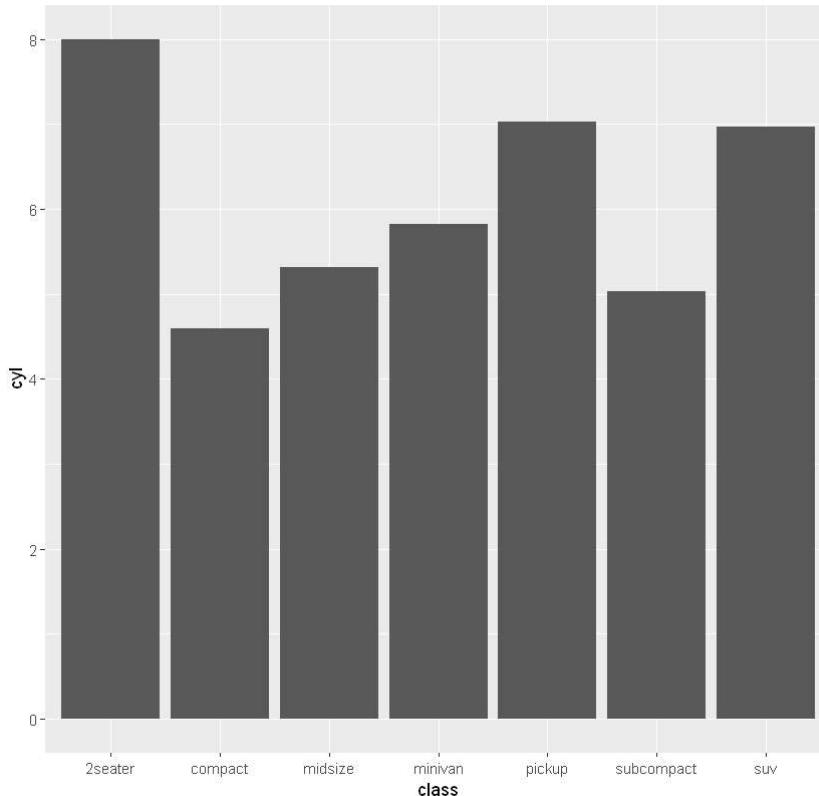
```
df_cat_by_cyl_typ = mpg %>% group_by(class) %>%
  summarise(cyl = mean(cyl),
            .groups = 'drop')

df_cat_by_cyl_typ
ggplot(df_cat_by_cyl_typ, aes(class, cyl)) +
  ggtitle("Car class categorised by Cylinder and type") + geom_col()
```

A tibble: 7 × 2

class	cyl
<chr>	<dbl>
2seater	8.000000
compact	4.595745
midsize	5.317073
minivan	5.818182
pickup	7.030303
subcompact	5.028571
suv	6.967742

Car class categorised by Cylinder and type



In [29]:

```
df_grp_class = mpg %>% group_by(class) %>%
  summarise(displ_mpg = mean(displ),
            hwy_mpg = mean(hwy),
            .groups = 'drop')

df_grp_class
```

```
melt_df_grp_class<- melt(df_grp_class[,c('class','displ_mpg', 'hwy_mpg')],id.vars = 1)
melt_df_grp_class
```

A tibble: 7 × 3

class	displ_mpg	hwy_mpg
<chr>	<dbl>	<dbl>
2seater	6.160000	24.80000
compact	2.325532	28.29787
midsize	2.921951	27.29268
minivan	3.390909	22.36364
pickup	4.418182	16.87879
subcompact	2.660000	28.14286
suv	4.456452	18.12903

A data.frame: 14 × 3

class	variable	value
<chr>	<fct>	<dbl>
2seater	displ_mpg	6.160000
compact	displ_mpg	2.325532
midsize	displ_mpg	2.921951
minivan	displ_mpg	3.390909
pickup	displ_mpg	4.418182
subcompact	displ_mpg	2.660000
suv	displ_mpg	4.456452
2seater	hwy_mpg	24.800000
compact	hwy_mpg	28.297872
midsize	hwy_mpg	27.292683
minivan	hwy_mpg	22.363636
pickup	hwy_mpg	16.878788
subcompact	hwy_mpg	28.142857
suv	hwy_mpg	18.129032

In [30]:

```
ggplot(melt_df_grp_class,aes(x = class,y = value)) +ggtitle( 'car who have High liter v
geom_bar(aes(fill = variable),stat = "identity",position = "dodge") +
scale_y_log10()
```

car who have High liter vehical and drive mostly in highway

