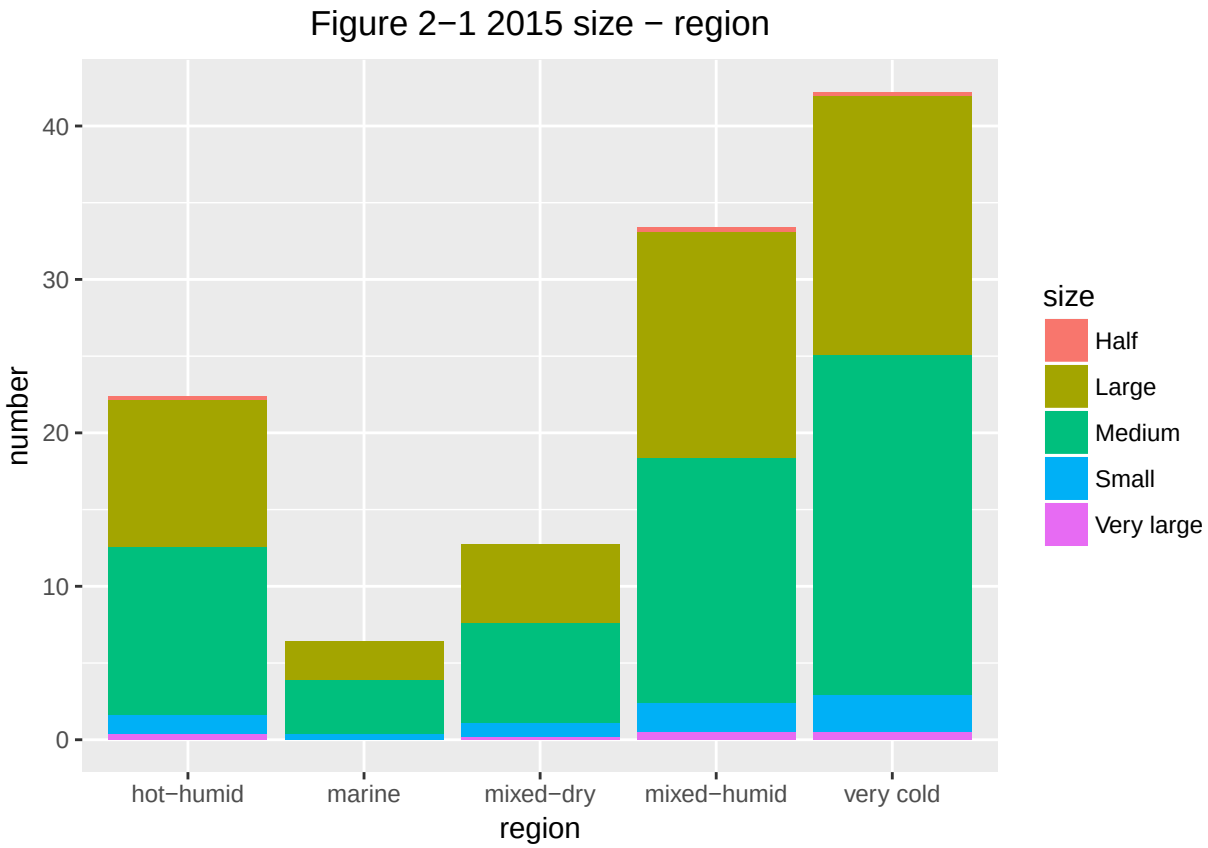
R_EDA

```
library(ggplot2)
```

Plan related to Size

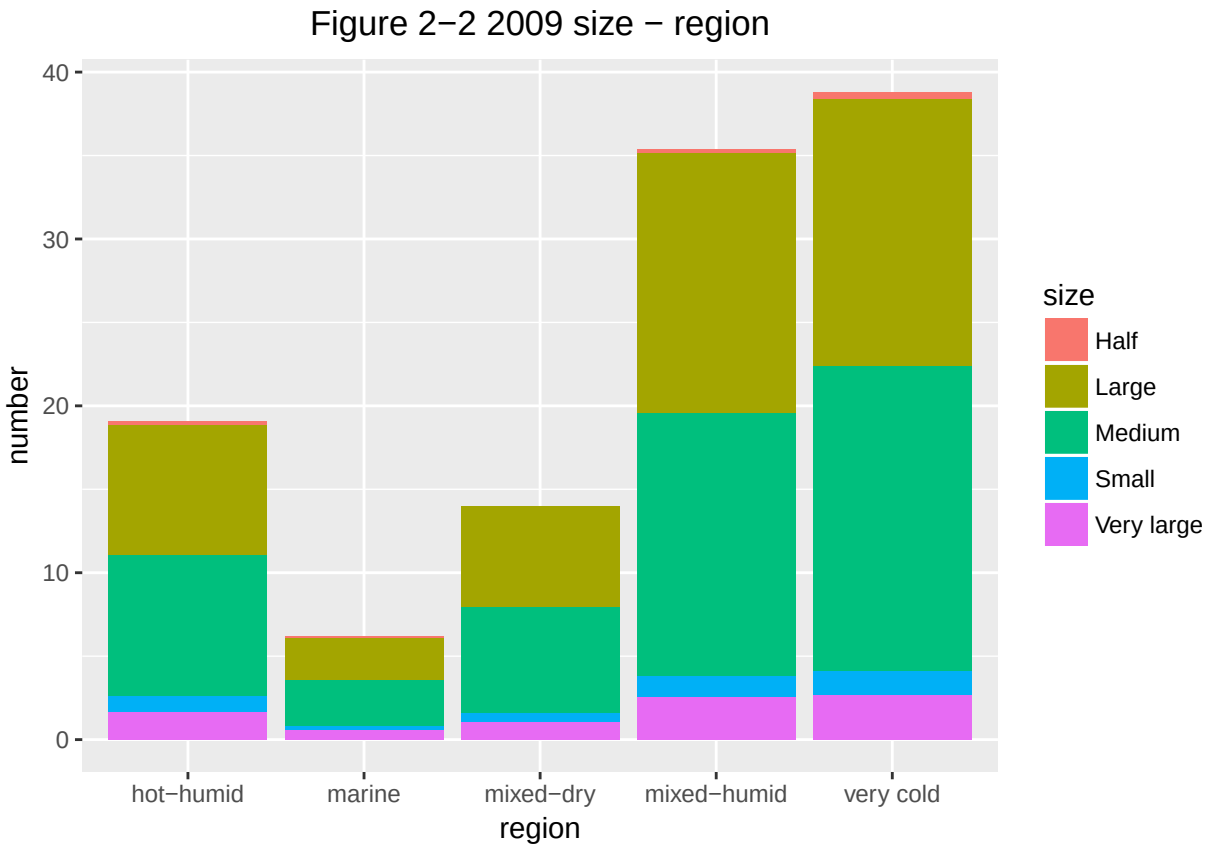
```
data <- read.table("2015.csv", header = TRUE, sep = ",")
size2015 <- ggplot(data[1:25,], aes(x = region, y = number, fill = size ))+
geom_bar(stat = "identity")+
labs(title = "Figure 2-1 2015 size - region") + theme(plot.title = element_text(hjust = 0.5))
size2009 <- ggplot(data[26:50,], aes(x = region, y = number, fill = size))+ geom_bar(stat = "identity")+
labs(title = "Figure 2-2 2009 size - region") +
theme(plot.title = element_text(hjust = 0.5))
size2015
```

```
## Warning: Removed 3 rows containing missing values (position_stack).
```



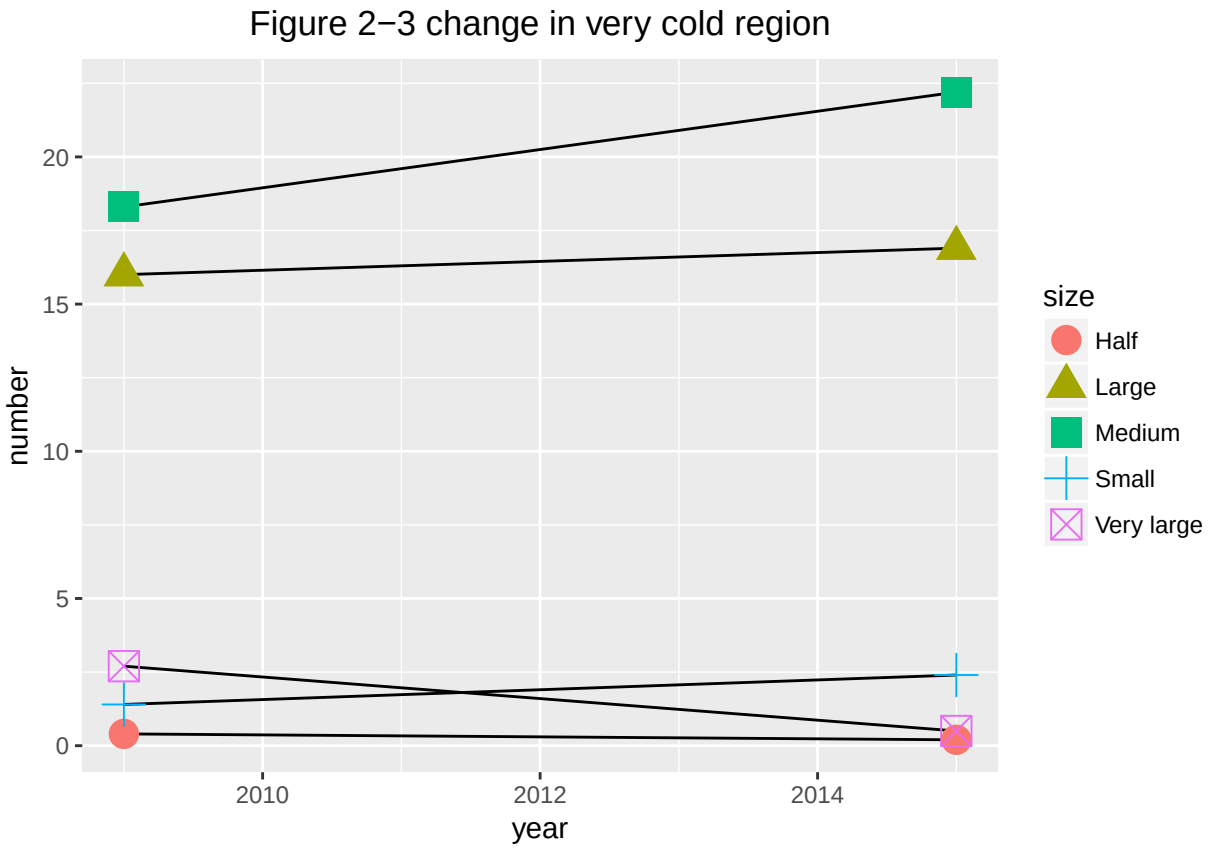

```
size2009
```

```
## Warning: Removed 1 rows containing missing values (position_stack).
```



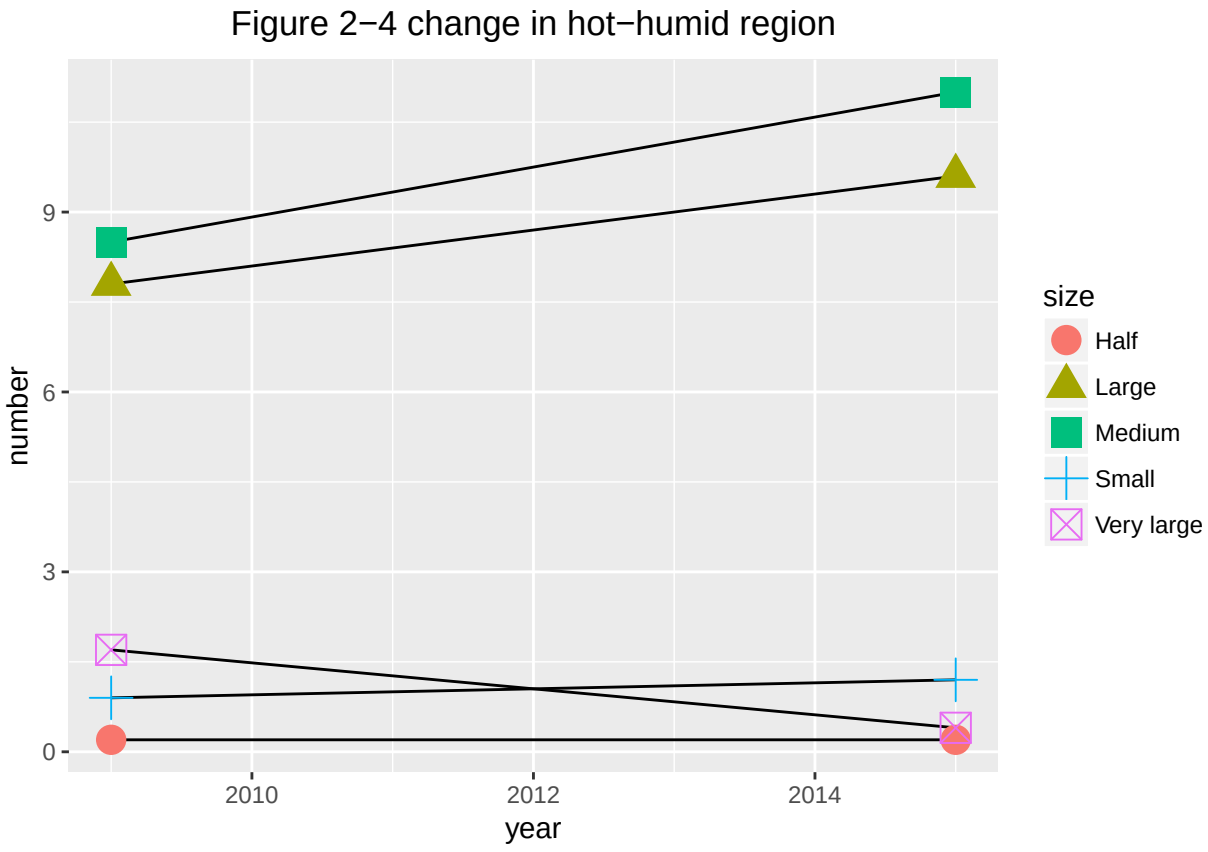
According to figure 2-1 and figure 2-2,the number of housing units that used refrigerators are different in different regions,and the refrigerartors of different size have similar proportion in different regions.

```
ggplot(data[data$regionnum == 5,], aes(x = year, y = number, group = size))+ geom_line()+
geom_point(aes(shape = size, colour = size),size = 5)+
labs(title = "Figure 2-3 change in very cold region")+
theme(plot.title = element_text(hjust = 0.5))
```



According to the figure 2-3,the number of Mediumn size changes the most,but all of the sizes have not changed much.The number of Medium,Large,Small size slightly increased,and the number of Very large,Half size slightly decreased.

```
ggplot(data[data$regionnum == 4,], aes(x = year, y = number, group = size))+ geom_line()+
geom_point(aes(shape = size, colour = size),size = 5)+
labs(title = "Figure 2-4 change in hot-humid region")+
theme(plot.title = element_text(hjust = 0.5))
```

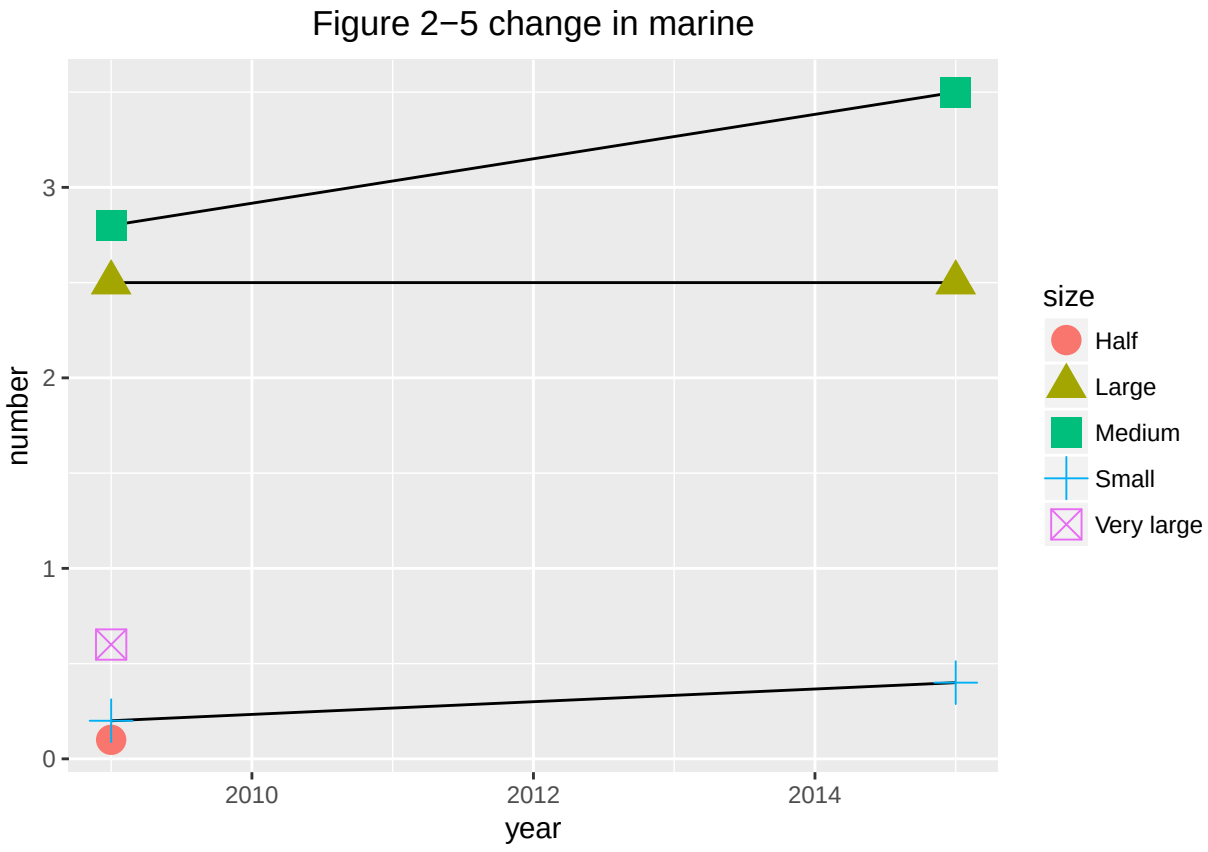


According to the figure 2-4,the number of Mediumn size changes the most,but all of the sizes have not changed much.The number of Medium,Large,Small size slightly increased,and the number of Very large,Half size slightly decreased.

```
ggplot(data[data$regionnum == 3,], aes(x = year, y = number, group = size))+ geom_line()+
geom_point(aes(shape = size, colour = size),size = 5)+
labs(title = "Figure 2-5 change in marine")+
theme(plot.title = element_text(hjust = 0.5))
```

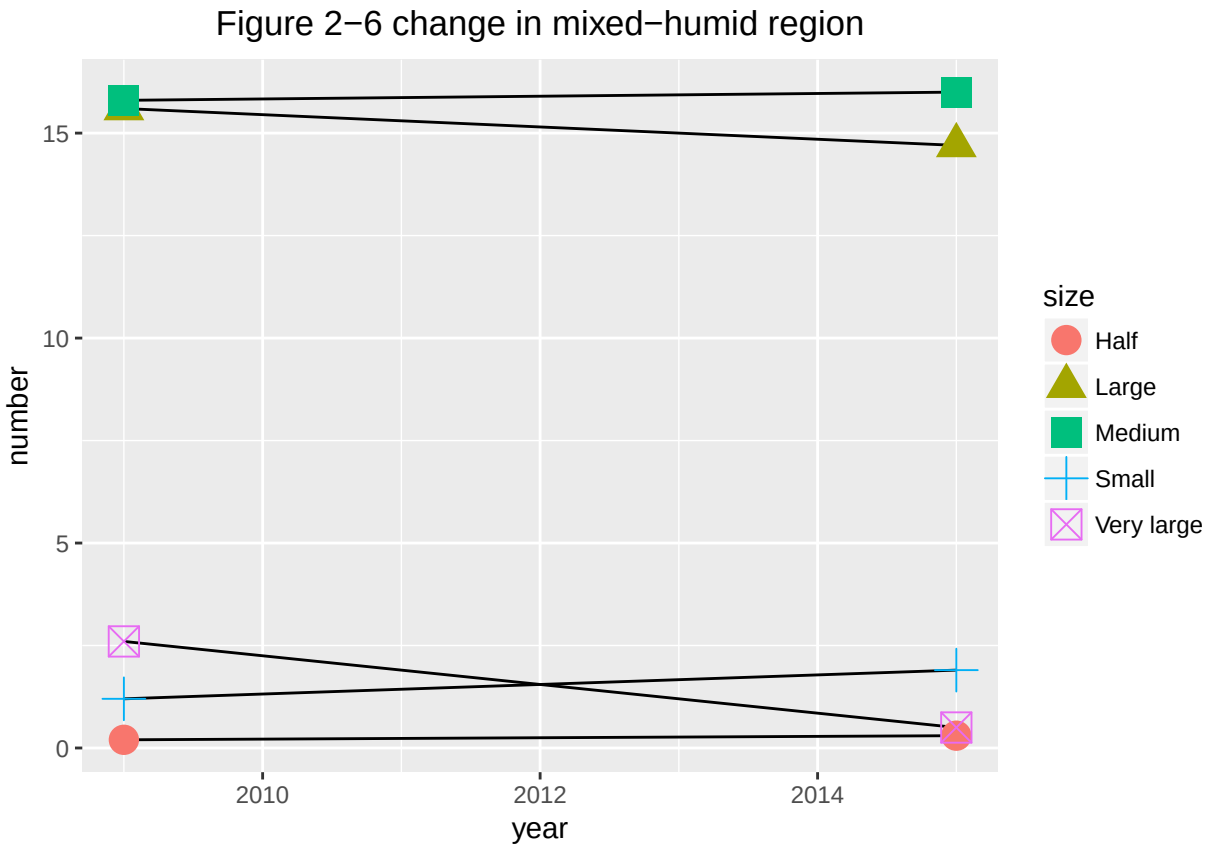
```
## Warning: Removed 2 rows containing missing values (geom_path).
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```



According to the figure 2-5,the number of Mediumn size changes the most,but all of the sizes have not changed much.The number of Medium,Large,Small size slightly increased,and the number of Very large,Half size missed.

```
ggplot(data[data$regionnum == 2,], aes(x = year, y = number, group = size))+ geom_line()+
geom_point(aes(shape = size, colour = size),size = 5)+
labs(title = "Figure 2-6 change in mixed-humid region")+
theme(plot.title = element_text(hjust = 0.5))
```

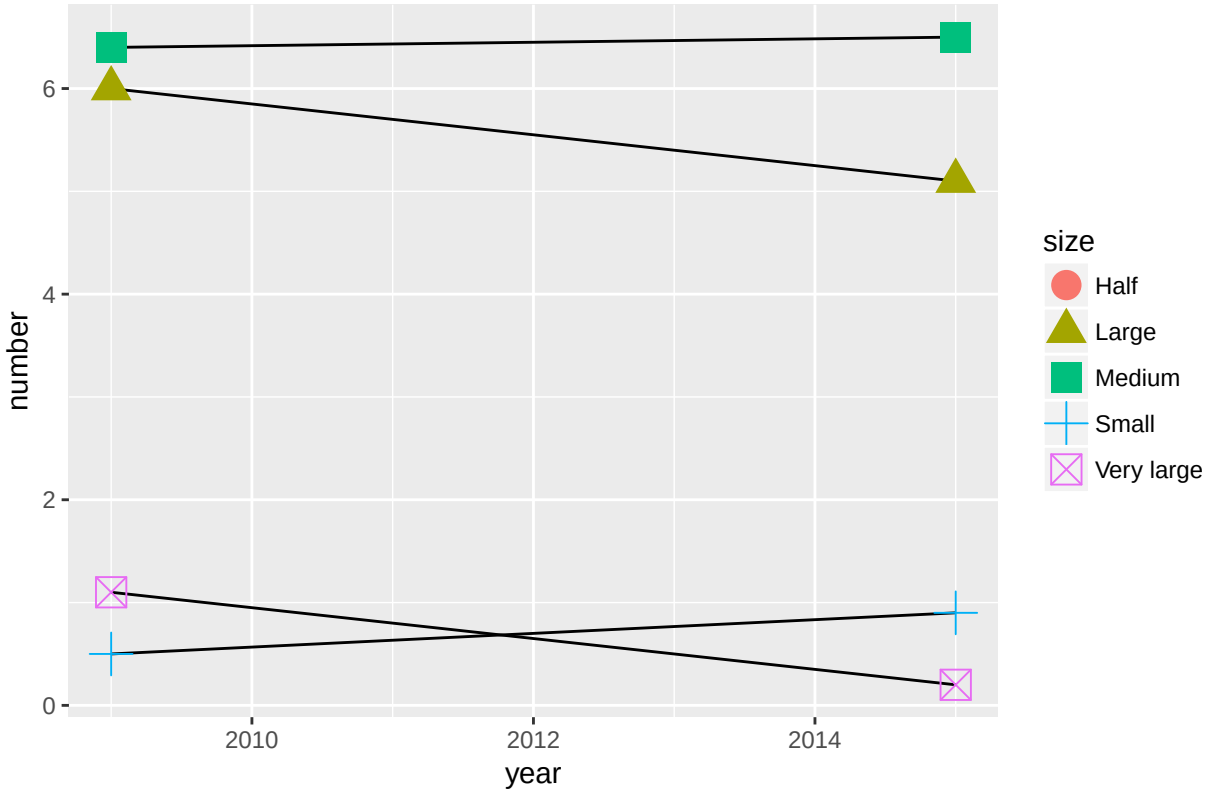


According to the figure 2-6, all of the sizes have not changed much.The number of Medium,Half,Small size slightly increased,and the number of Very large,Large size slightly decreased,which indicates the housing units in mixed-humid region tend to use smaller size.

```
ggplot(data[data$regionnum == 1,], aes(x = year, y = number, group = size))+ geom_line()+
geom_point(aes(shape = size, colour = size),size = 5)+
labs(title = "Figure 2-7 change in mixed-dry region")+
theme(plot.title = element_text(hjust = 0.5))
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```

Figure 2−7 change in mixed−dry region



According to the figure 2-7,the number of Large size changes the most,but all of the sizes have not changed much.The number of Medium and Small size slightly increased,and the number of Very large and Large size slightly decreased,which indicates the housing units in mixed-dry region tend to use smaller size.

In conclusion,the Medium and Large size are used in most housing units,and the Medium size has remained the high quantity,the Large size has decrease a little in some region.So we can reduce the number of Large size in mixed-dry and mixed-humid regions.The Half,Small and Very large size are used in much less housing units,and the number of Small size has remained the level or increased a little,the number of Very large size has decreased across the crountry,the number of Half size missed a lot but it has basically remain the level.So we should increase the number of Small size a little,and decrease the number of Very large size.