

Longitudinal approach

재앙의 종류 Group, subGroup, Type, subType별로 시간에 따른 빈도수의 추이를 확인해보자.

각 재앙별로 자주 발생하는 시기(계절 혹은 월) 이 있는가? 혹은 년도별로 추이는 어떻게 되는가?

1.1 어떤 시간 변수를 활용함이 적절한가?

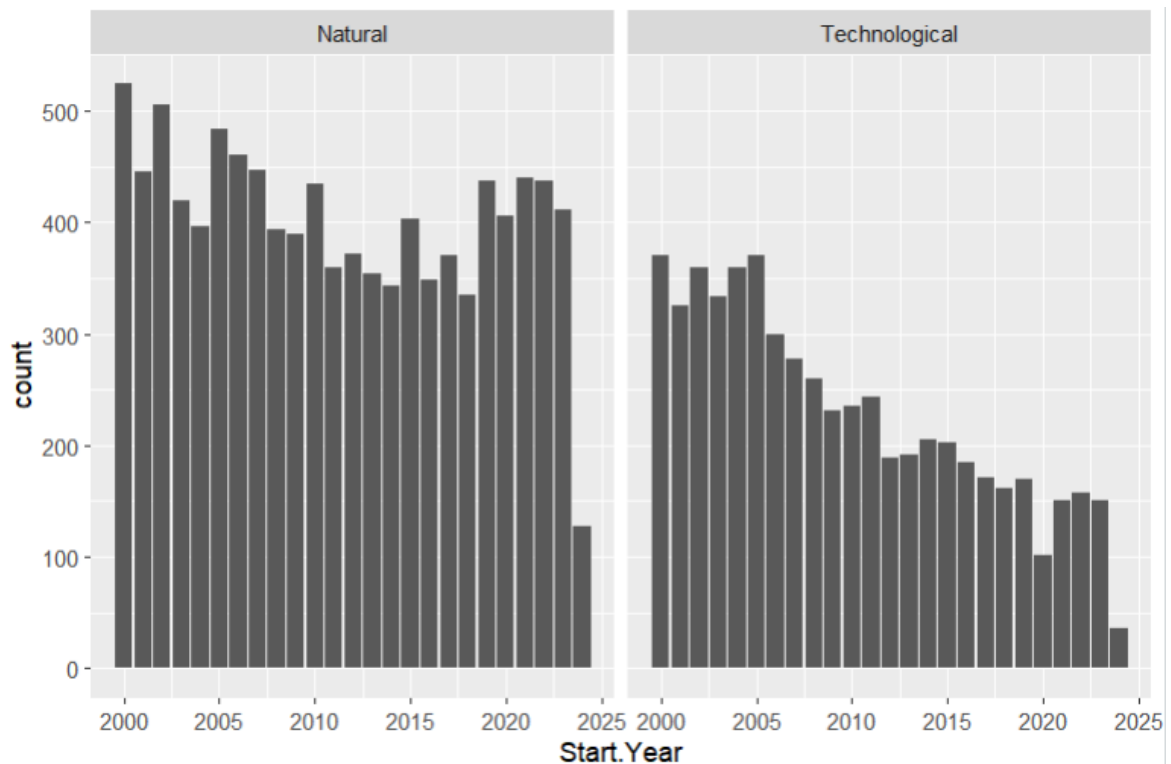
시간에 따른 빈도수를 disaster 별로 시각화해보자.

→ 재앙의 종류, 분류 등등에 따라 비교

Disaster.group → 2가지 종류가 있으며, 자연재해와 기술적 사고가 있다.

기술적 재해는 관련한 제도 및 법규에 따라 그 빈도수가 년도별로 감소할 것이라고 생각하며, 월별로 일정한 편균 빈도를 가질 것이라 예상한다.

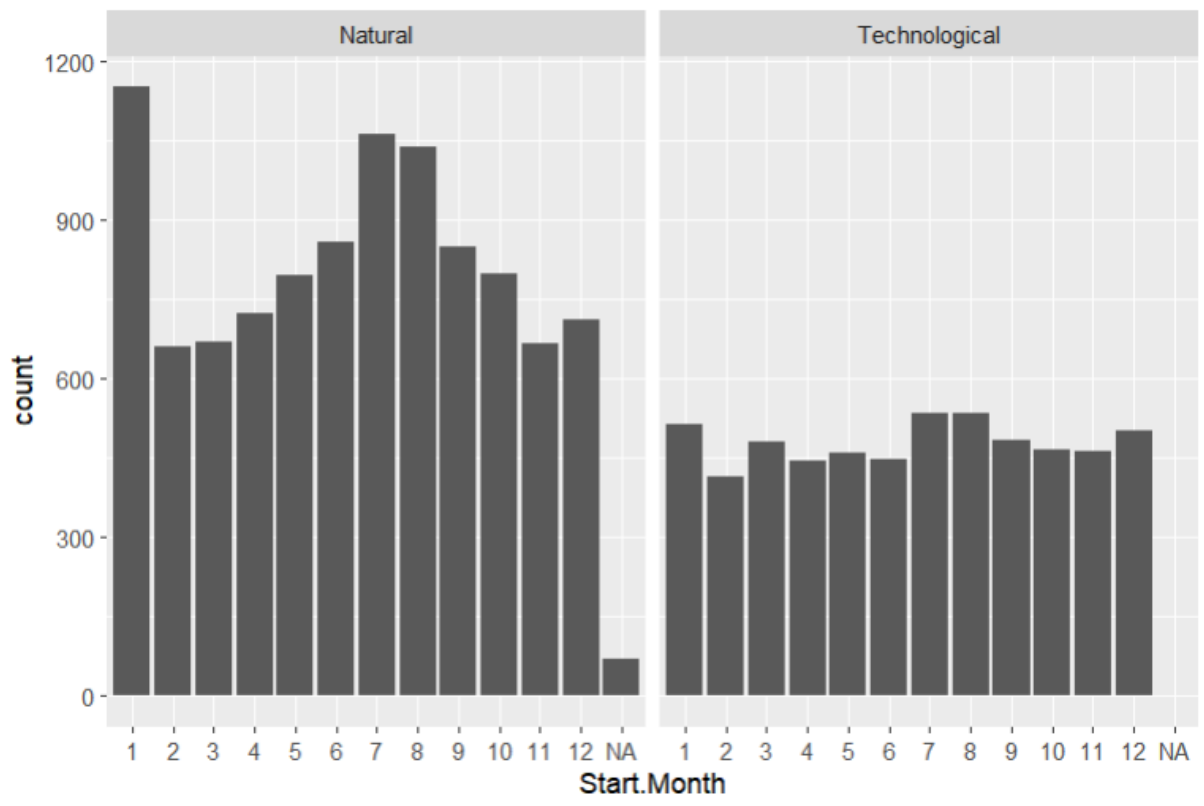
```
data%>%  
  group_by(Disaster.Group)%>%  
  ggplot(aes(x = Start.Year))+  
  geom_bar()+  
  facet_wrap(~Disaster.Group)
```



년도별로 빈도수를 시각화함. 기본적으로 자연재해가 더 높게 나타남을 확인.

자연재해의 경우 대부분 비슷하게 나타나지만 기술재해의 경우 점점 감소하는 추세를 보임

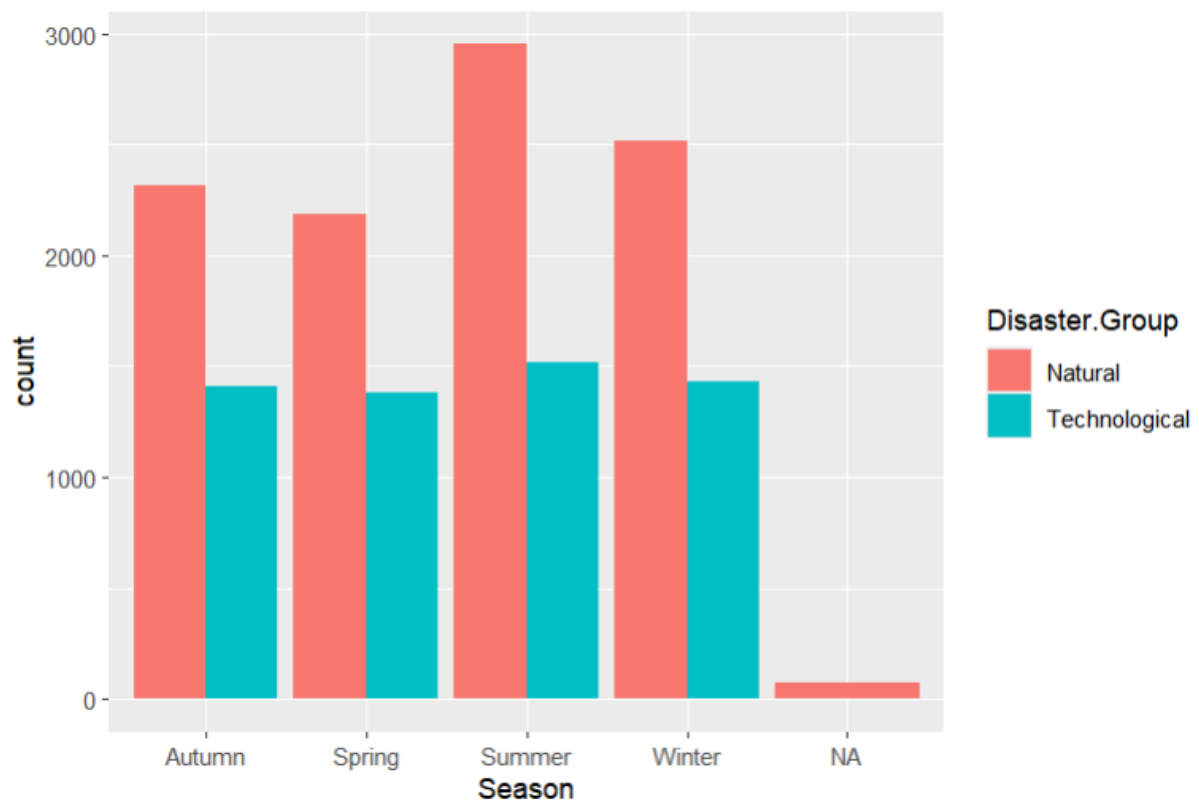
```
data%>%
  group_by(Disaster.Group)%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Group)
```



이번에는 월별로 자연재해 및 기술재해의 빈도를 시각화함.

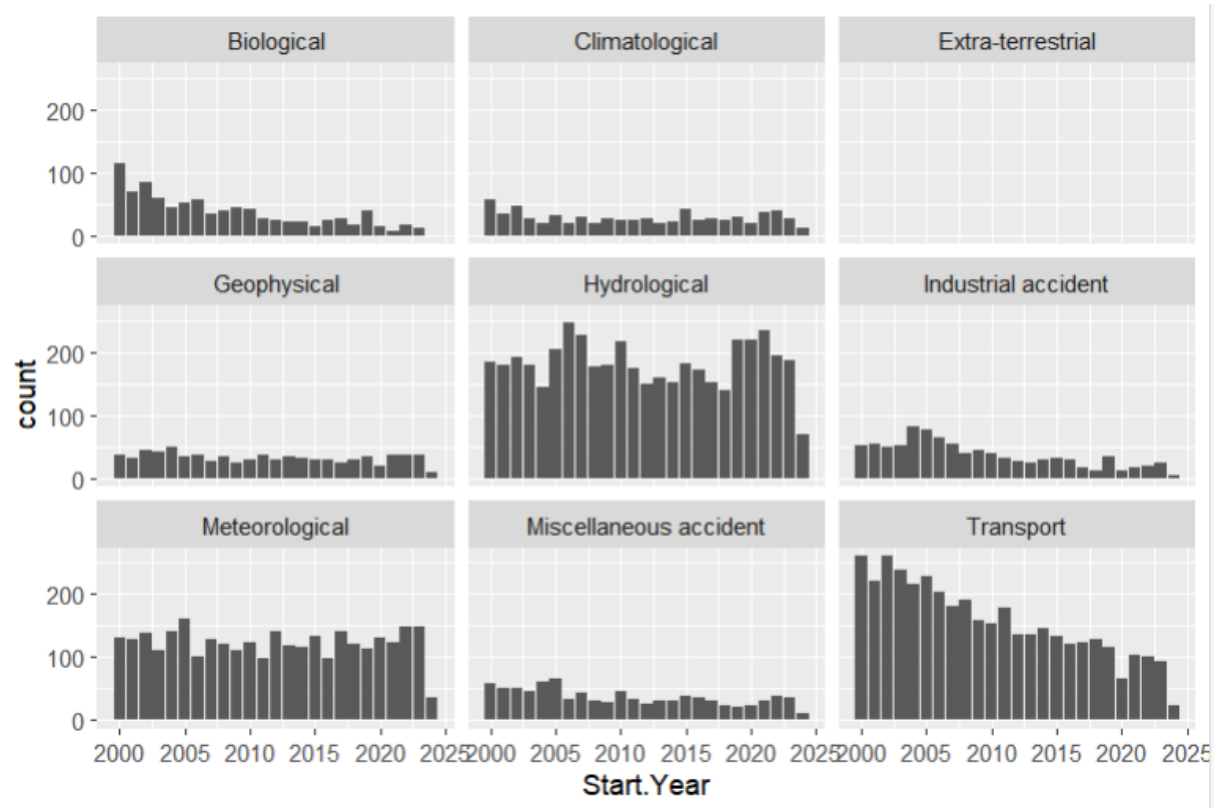
기술적 재해의 경우 자연환경의 영향을 받지 않으므로 거의 모든 달이 일정하게 나왔으나, 자연재해의 경우 특정 월에 더 치우쳐져 있음.

```
data %>%
  mutate(Season = case_when(
    Start.Month %in% c(12, 1, 2) ~ "Winter",
    Start.Month %in% c(3, 4, 5) ~ "Spring",
    Start.Month %in% c(6, 7, 8) ~ "Summer",
    Start.Month %in% c(9, 10, 11) ~ "Autumn"
  )) %>%
  ggplot(aes(x = Season, fill = Disaster.Group)) +
  geom_bar(position = "dodge")
```



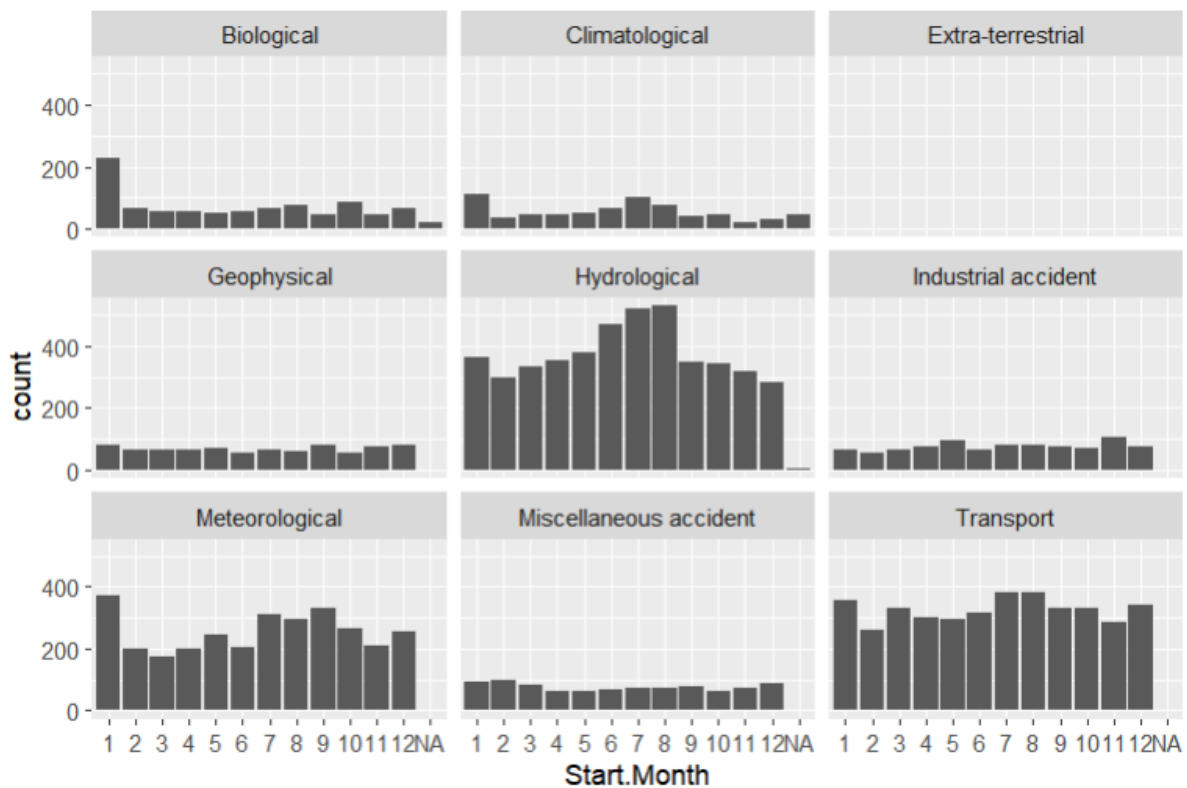
조금더 보기 쉽도록 계절별로 그룹화를 한 후 비교함. 여름과 겨울에 비교적 자연재해가 많이 발생함

```
data%>%
  group_by(Disaster.Subgroup)%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subgroup)
```



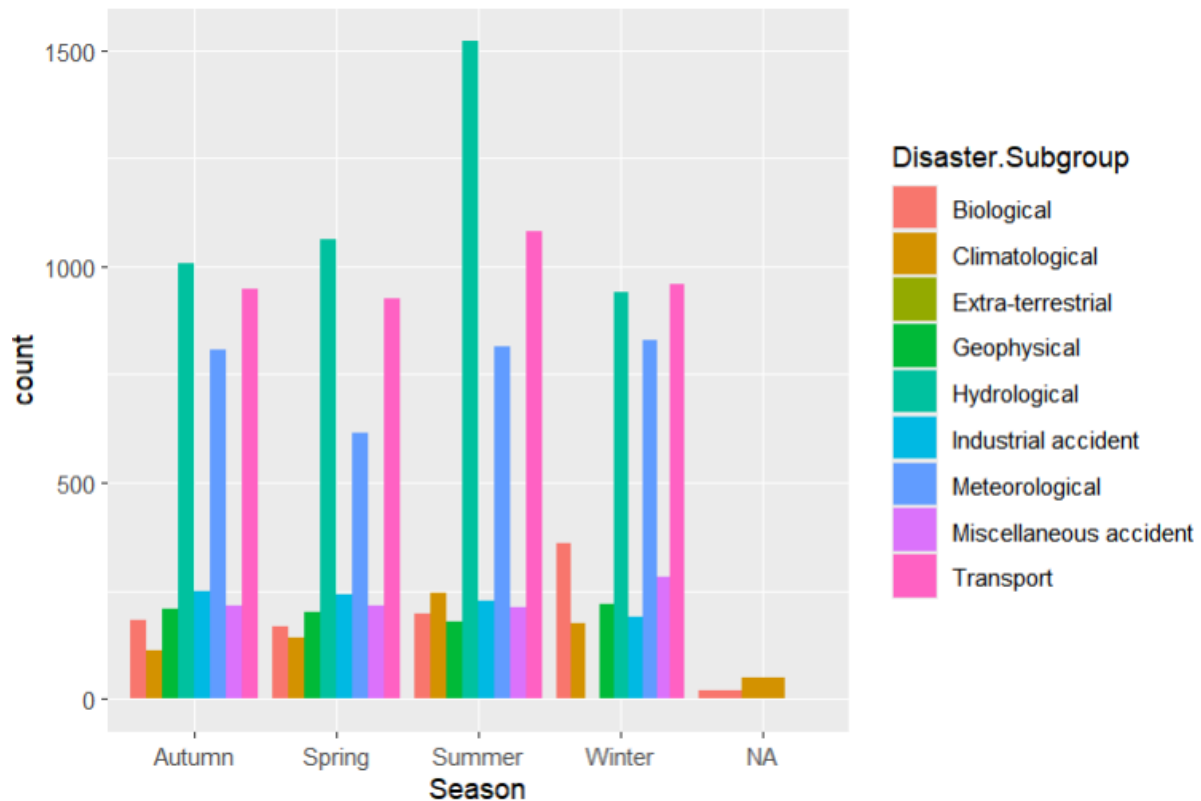
이번에는 년도별 재해 종류를 시각화함. 가장 빈도가 높아보이는 재해는 Hydrological과 Meteorological, Transport이다. 이중 Meteorological과 Hydrological은 년도별로 고른 빈도를 보였고, Transport는 2000년대 초반에는 높았으나 점점 해가 지날수록 감소하고 있음.

```
data%>%
  group_by(Disaster.Subgroup)%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Subgroup)
```



동일하게 월별로 다시 시각화를 진행함, 홍수의 경우 여름철 장마와 함께 높은 빈도를 보이며, 지진의 경우에는 그래도 각 월별로 고르긴 하지만 1월에 조금 더 많은 평균 빈도를 보임

```
data %>%
  mutate(Season = case_when(
    Start.Month %in% c(12, 1, 2) ~ "Winter",
    Start.Month %in% c(3, 4, 5) ~ "Spring",
    Start.Month %in% c(6, 7, 8) ~ "Summer",
    Start.Month %in% c(9, 10, 11) ~ "Autumn"
  )) %>%
  ggplot(aes(x = Season, fill = Disaster.Subgroup)) +
  geom_bar(position = "dodge")
```



계절별로 나누어 표현함

이번에는 disaster subGroup과 그 하위 분류인 Type, subType에 대해 시각화를 진행함

▼ Climatological기후 관련 재해

기후와 관련된 재해는 년도별 빈도수에서는 지구 온난화와 연관지어 그 빈도수가 늘어날 것이라고 예상한다.

```
data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
  filter(Disaster.Subgroup == "Climatological") %>%
  distinct()
```

	Disaster.Type	Disaster.Subgroup
1	Drought	Climatological
2	wildfire	Climatological
3	Glacial lake outburst flood	Climatological

> |

Drought (가뭄)

Wildfire (산불)

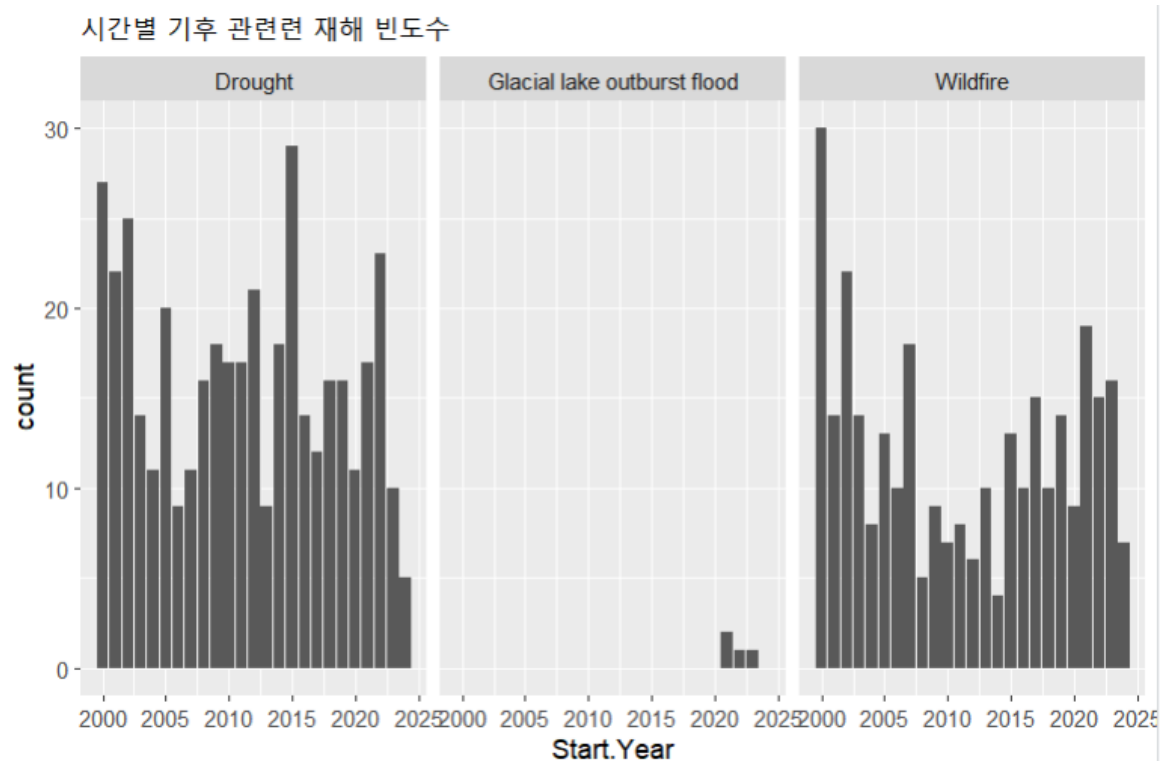
Glacial lake outburst flood (빙하호 폭발 홍수)

subGroup인 climatological에는 3개의 Type이 있음

시간별 기후 관련 재해 정리

년도-빈도

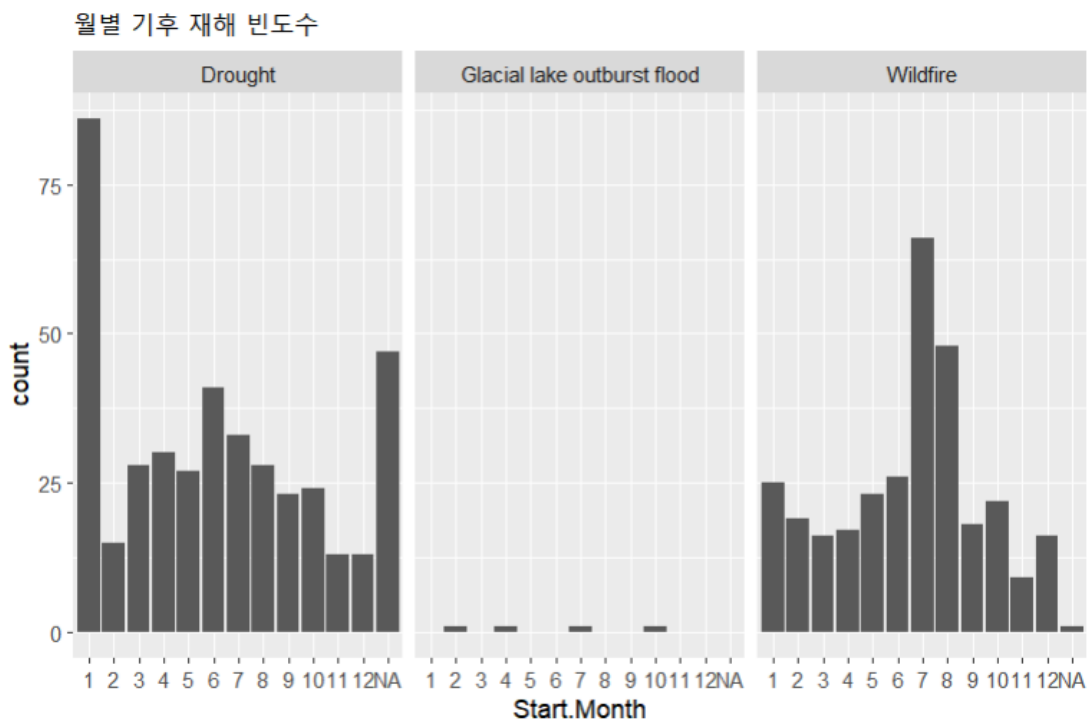
```
data%>%  
  group_by(Disaster.Group)%>%  
  filter(Disaster.Subgroup == "Climatological")%>%  
  ggplot(aes(x = Start.Year))+  
  geom_bar()+  
  facet_wrap(~Disaster.Type)+  
  labs(title = "시간별 기후 관련 재해 빈도수")
```



년도별 기후 관련 재해 빈도를 보면, 가뭄과 산불의 빈도수가 높음을 알 수 있음. 또한 빙하호 폭발 호수는 근래 2020년대 이후로 등장하여 기후변화와 연관성을 찾을 필요가 있어보임.

이번에는 월별로 특이사항을 확인.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Climatological")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "월별 기후 재해 빈도수")
```



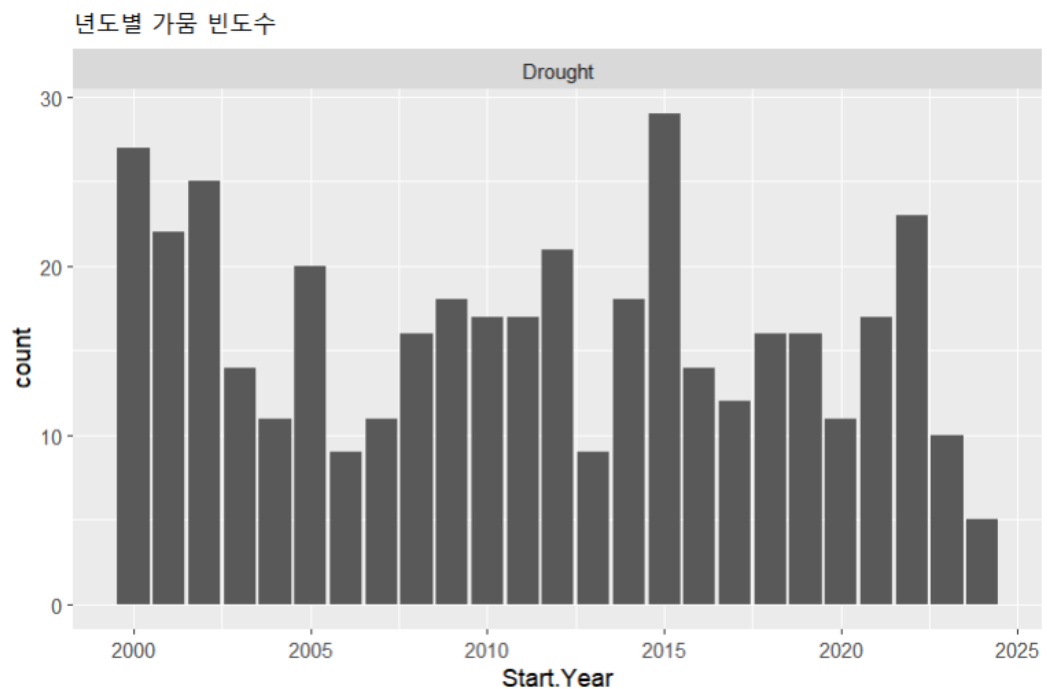
특이한 점은 산불에 대한 재해가 건조한 가을, 겨울이 아닌 여름에 많이 나타났고, 가뭄 피해가 1월에 많이 나타남.

이번에는 Type안에서도 세부 Type별 빈도수를 체크함

▼ Drought

먼저 가뭄의 경우에는 세부 항목 역시 가뭄으로 하나뿐이라 위에서 진행한 시각화 자료와 동일하다.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Drought")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "년도별 가뭄 빈도수")
```



▼ Wildfire

다음으로 산불과 관련된 세부 항목과 빈도수를 확인

산불의 세부 항목으로는

Disaster.Type	Disaster.Subtype
1 Wildfire	Land fire (Brush, Bush, Pasture)
2 Wildfire	Forest fire
3 Wildfire	Wildfire (General)

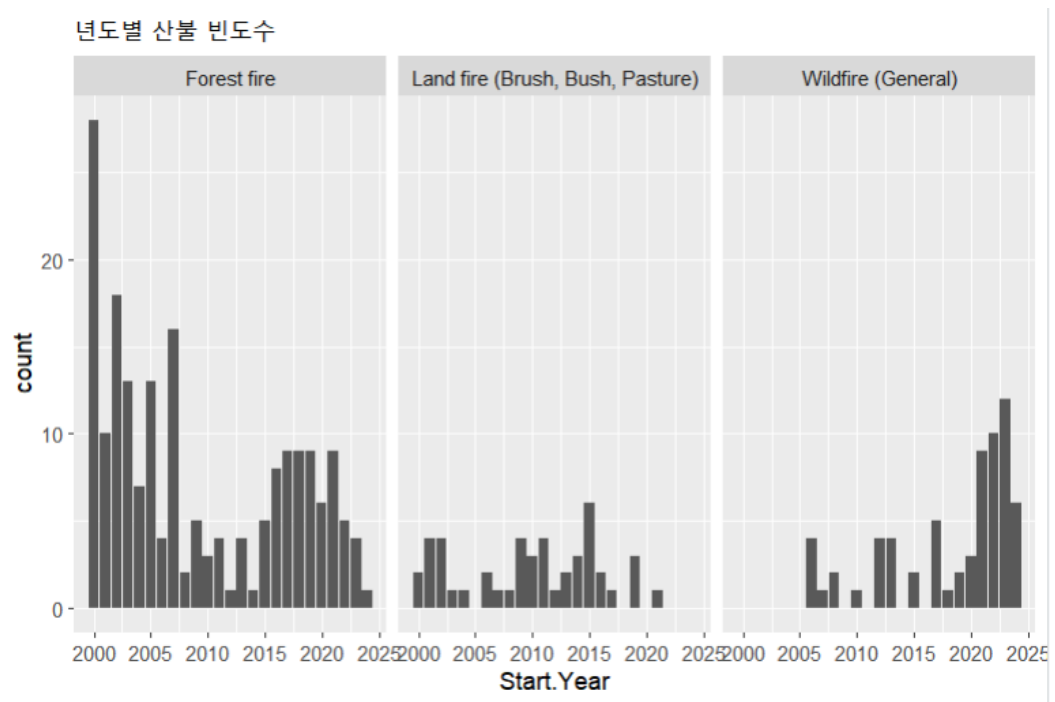
Land fire (Brush, Bush, Pasture) - 덩불, 관목지, 초지 화재

Forest fire - 산림 화재

Wildfire (General) - 일반적인 산불

이 있으며 이를 년도별로 시각화하면

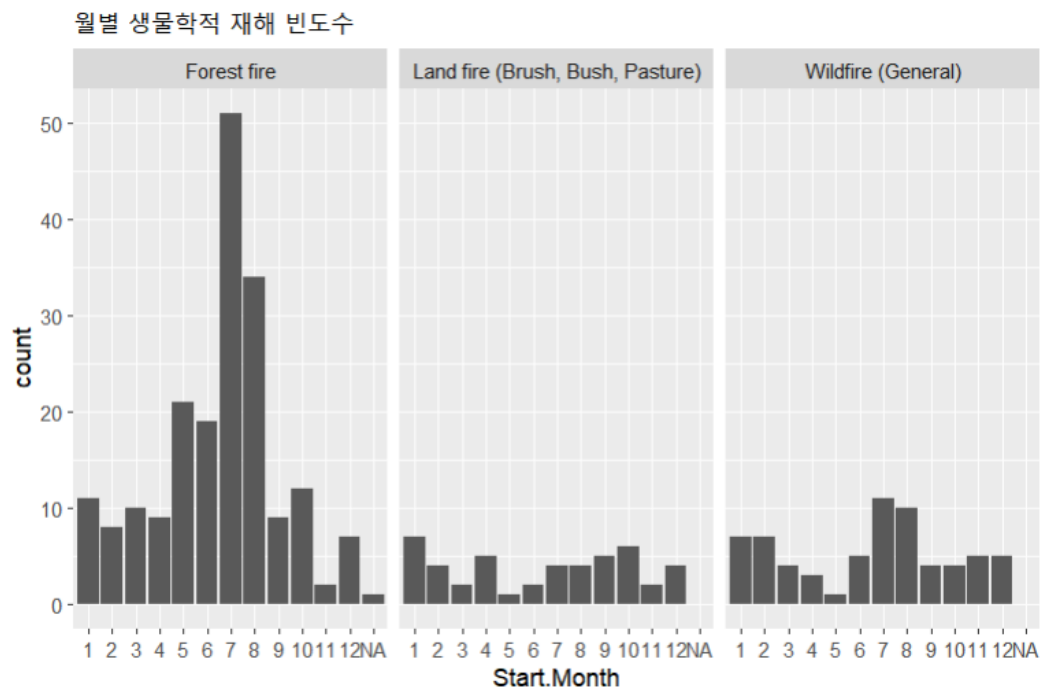
```
data%>%  
  group_by(Disaster.Group)%>%  
  filter(Disaster.Type == "Wildfire")%>%  
  ggplot(aes(x = Start.Year))+  
  geom_bar()+  
  facet_wrap(~Disaster.Subtype)+  
  labs(title = "년도별 산불 빈도수")
```



산림 산불의 경우 가장 빈도수가 많았지만 최근들어 빈도가 줄었고, 일반적인 산불이 오히려 빈도가 증가함을 확인할 수 있음.

월별로도 체크해본다면

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Wildfire")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "월별 산불 재해 빈도수")
```



다음과 같이 산림 산불의 경우 여름철에 주로 몰려있음을 확인할 수 있음.

▼ Transport 교통 관련 재해

교통 관련 재해에서는

```
data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
```

```
filter(Disaster.Subgroup == "Transport") %>%
distinct()
```

	Disaster.Type	Disaster.Subgroup
1	Road	Transport
2	Rail	Transport
3	Air	Transport
4	Water	Transport

Road (도로 사고)

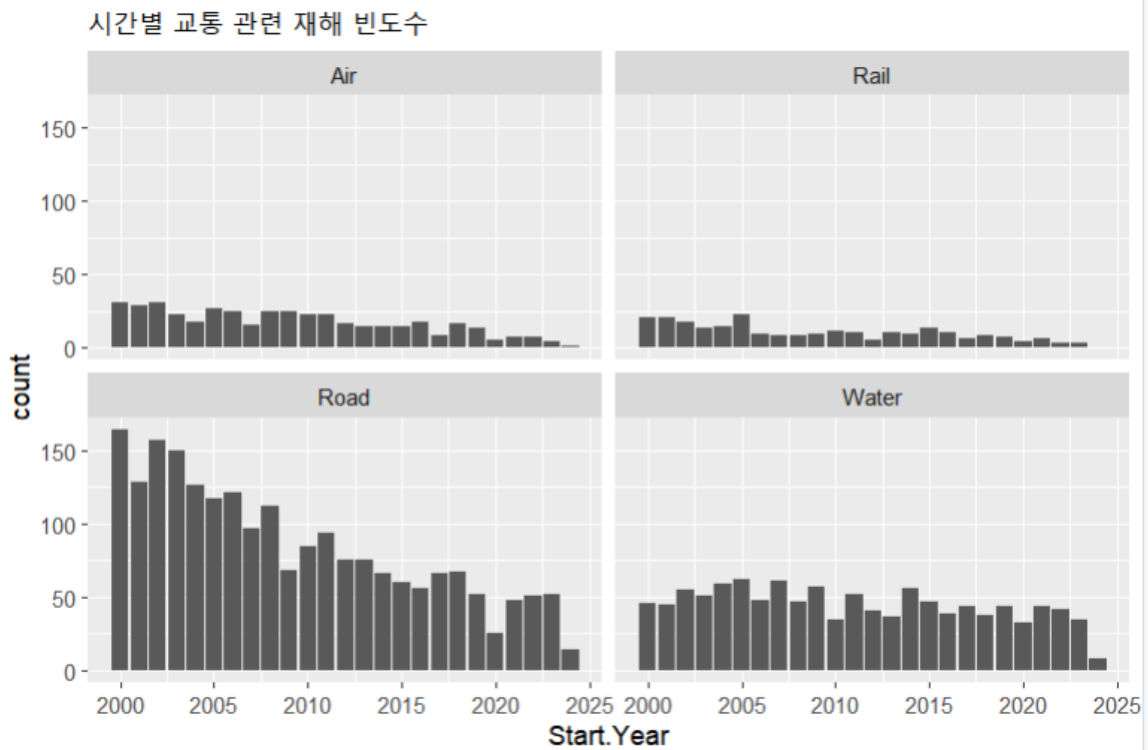
Rail (철도 사고)

Air (항공 사고)

Water (수상 사고)

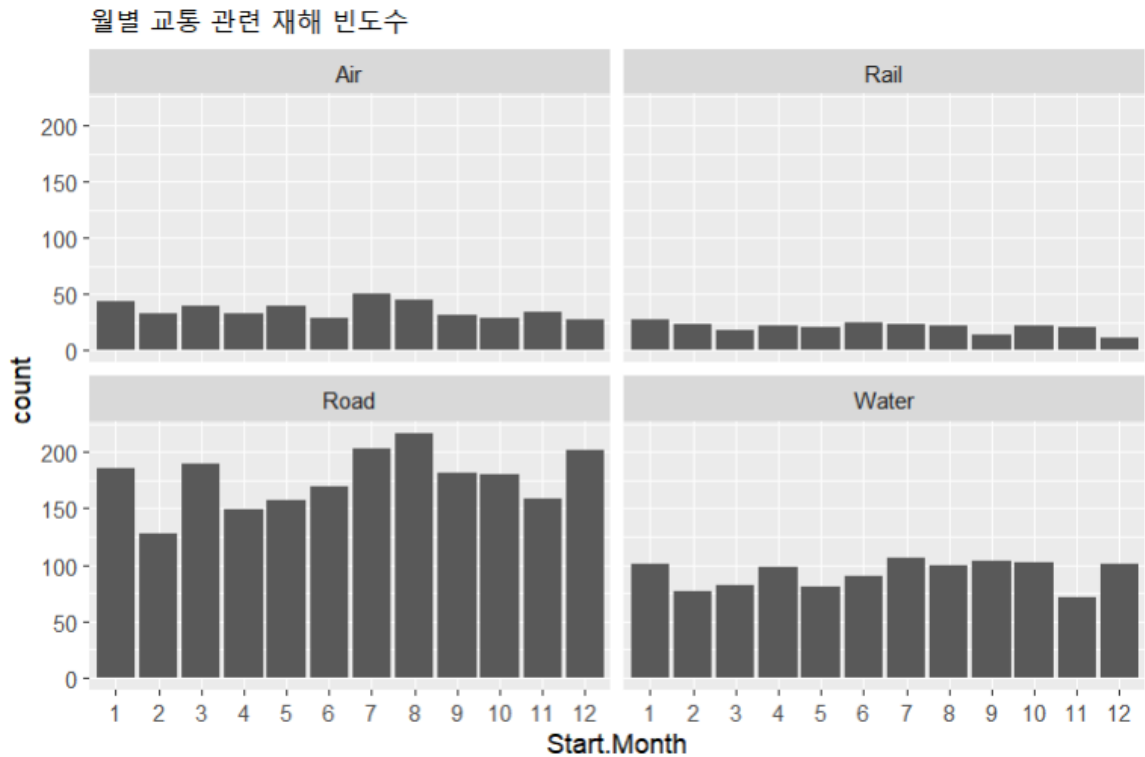
이렇게 4가지 유형이 있음

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Transport")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "시간별 교통 관련 재해 빈도수")
```



교통 관련 재해의 경우 자연재해와는 달리 환경과 기후 등의 조건에 따라 빈도수가 물리고 하지는 않음. 그러나 4가지 유형 모두 전반적으로 빈도수가 갈수록 감소하는 것에는 관련된 법규나 제약이 생기며 빈도수를 줄인다고 할 수 있음

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Transport")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "월별 교통 관련 재해 빈도수")
```



월별로 빈도수를 보았을때는 그 빈도에서 큰 차이를 보이지는 않음 대신 도로사고와 수상사고의 빈도가 항공이나 철도 사고에 비해 큰 것으로 보아 이용자 수와 연관시키는 것이 가능함

이번에는 subType별 빈도수를 확인

```
Disaster.Type Disaster.Subtype
1           Air           Air
```

```
Disaster.Type Disaster.Subtype
1           Rail          Rail
```

```
Disaster.Type Disaster.Subtype
1           Road          Road
```

	Disaster.Type	Disaster.Subtype
1	Water	Water

교통사고의 경우에는 세부 Type들이 기존 Type과 동일하므로 그 빈도수는 Type의 빈도와 같음.

▼ Hydrological (수문학적 재해)

```
data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
  filter(Disaster.Subgroup == "Hydrological") %>%
  distinct()
```

	Disaster.Type	Disaster.Subgroup
1	Flood	Hydrological
2	Mass movement (wet)	Hydrological

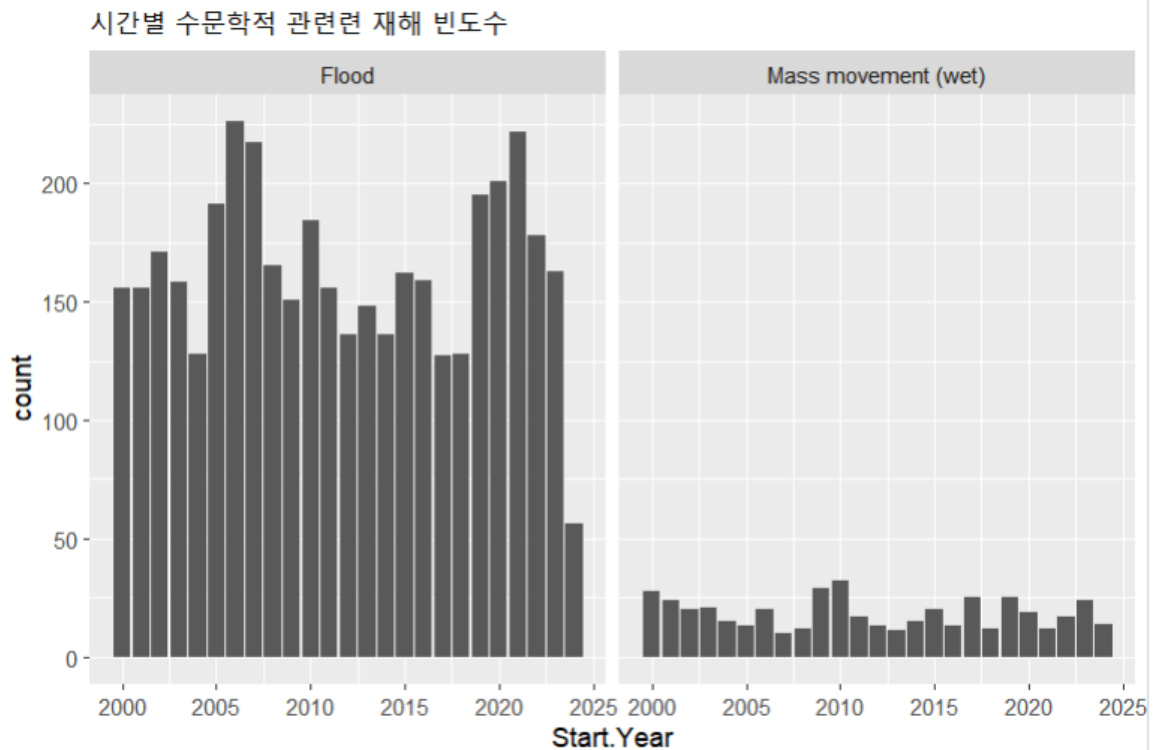
> |

Flood (홍수)

Mass movement (wet) (대규모 이동 - 습한 조건)

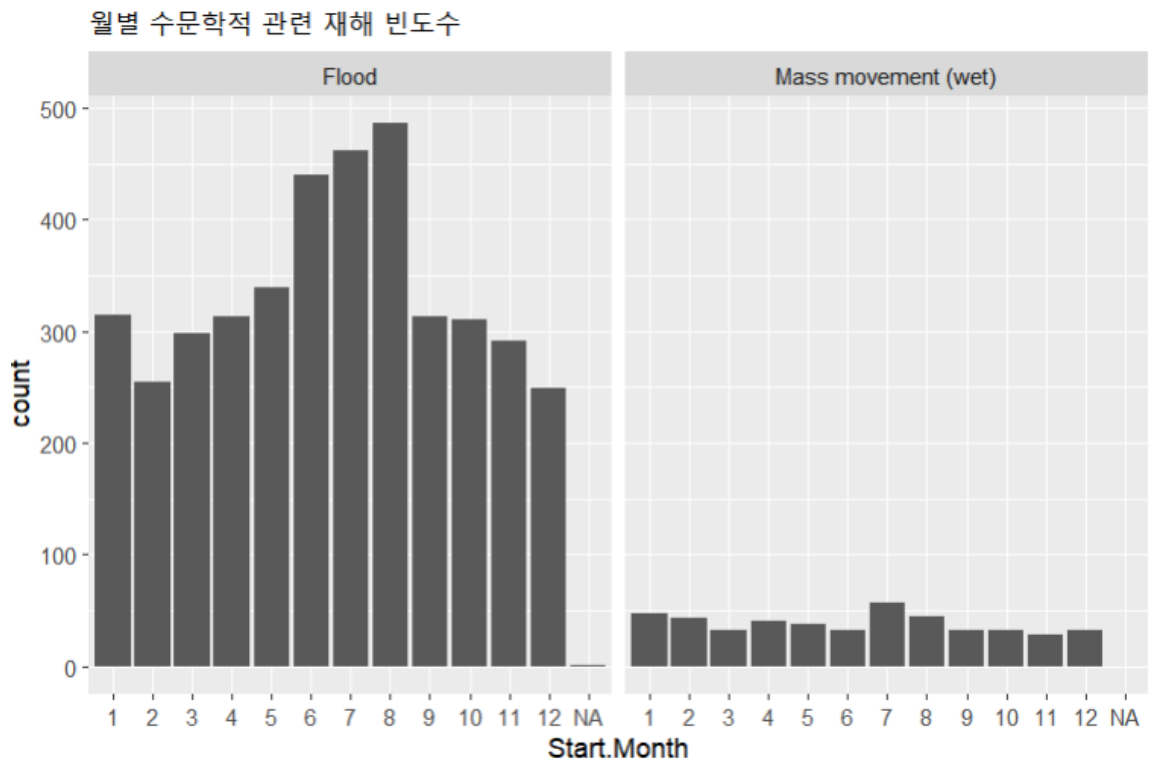
수해에는 2가지 유형이 있음.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Hydrological")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "시간별 수문학적 관련 재해 빈도수")
```

년도별로 빈도를 표시했을때 홍수의 빈도가 굉장히 많음을 확인할 수 있음

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Hydrological")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "월별 수문학적 관련 재해 빈도수")
```



월별 수문학적 재해를 확인했을 때, 주로 6,7,8월 장마철과 함께 빈도가 올라감을 확인할 수 있음

수문학적 재해 역시 세부Type별로 종류를 찾고 빈도를 조사

▼ Flood

	Disaster.Type	Disaster.Subtype
1	Flood	Riverine flood
2	Flood	Flash flood
3	Flood	Coastal flood
4	Flood	Flood (General)

가장 먼저 홍수에는 4 종류가 있으며,

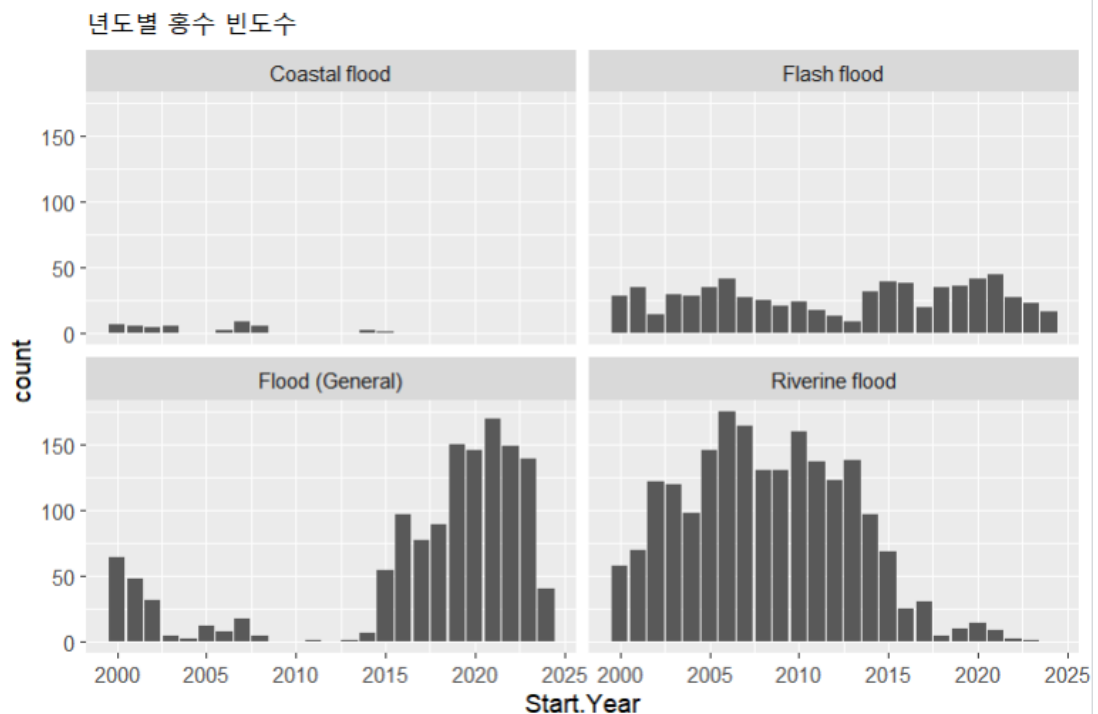
Riverine flood-하천 홍수

Flash flood- 돌발 홍수

Coastal flood-해안 홍수

Flood (General)-일반적인 홍수 로 구분됨

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Flood")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "년도별 홍수 빈도수")
```

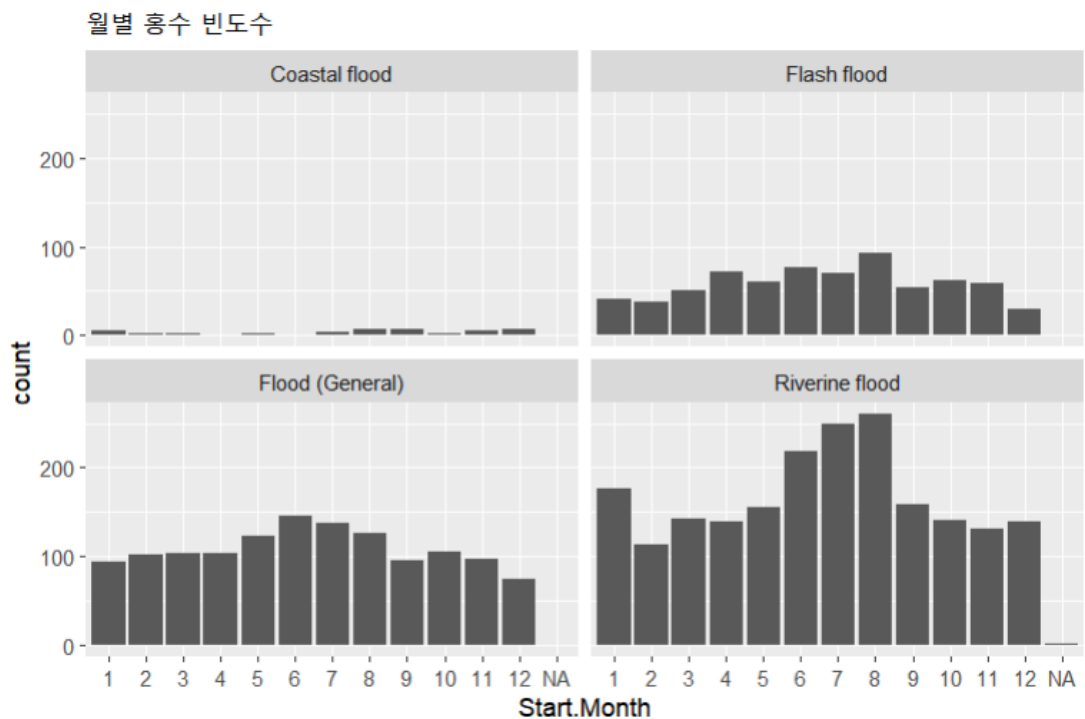


년도별 홍수 빈도수를 비교해 보았을때 하천홍수의 경우 2000년대 초반부터 2010년도 중반까지 가장 많은 빈도를 가졌으며 최근에는 그 수가 줄었지만 일반적인 홍수의 경우 반대로 최근들어 급격히 그 빈도가 증가함

이번에는 월별로 확인

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Flood")%>%
  mutate(Start.Month = factor(Start.Month))%>%
```

```
ggplot(aes(x = Start.Month)) +  
  geom_bar() +  
  facet_wrap(~Disaster.Subtype) +  
  labs(title = "월별 홍수 빈도수")
```



▼ Mass movement (wet)

다음으로는 Mass movement (wet) 산사에 대한 종류로는

Disaster.Type	Disaster.Subtype
1 Mass movement (wet)	Landslide (wet)
2 Mass movement (wet)	Mudslide
3 Mass movement (wet)	Avalanche (wet)
4 Mass movement (wet)	Sudden Subsidence (wet)
5 Mass movement (wet)	Rockfall (wet)

Landslide (wet) - 습한 조건에서의 산사태

Mudslide - 진흙사태

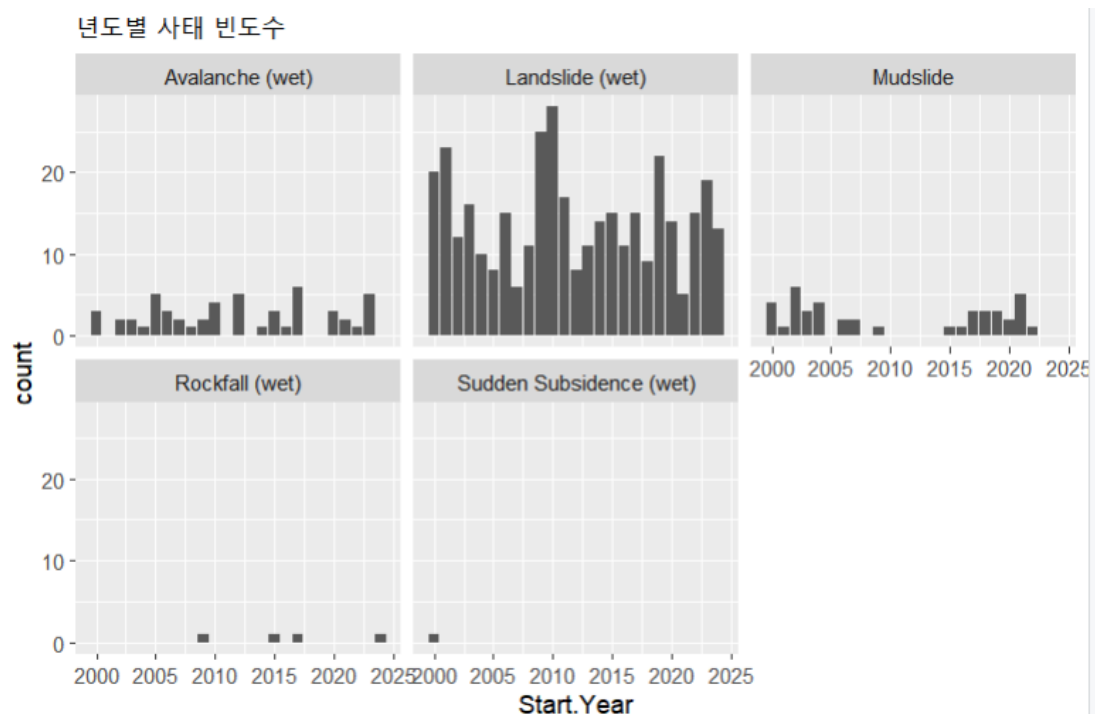
Avalanche (wet) - 습한 조건에서의 눈사태

Sudden Subsidence (wet) - 습한 조건에서의 급격한 침하

Rockfall (wet) - 습한 조건에서의 낙석

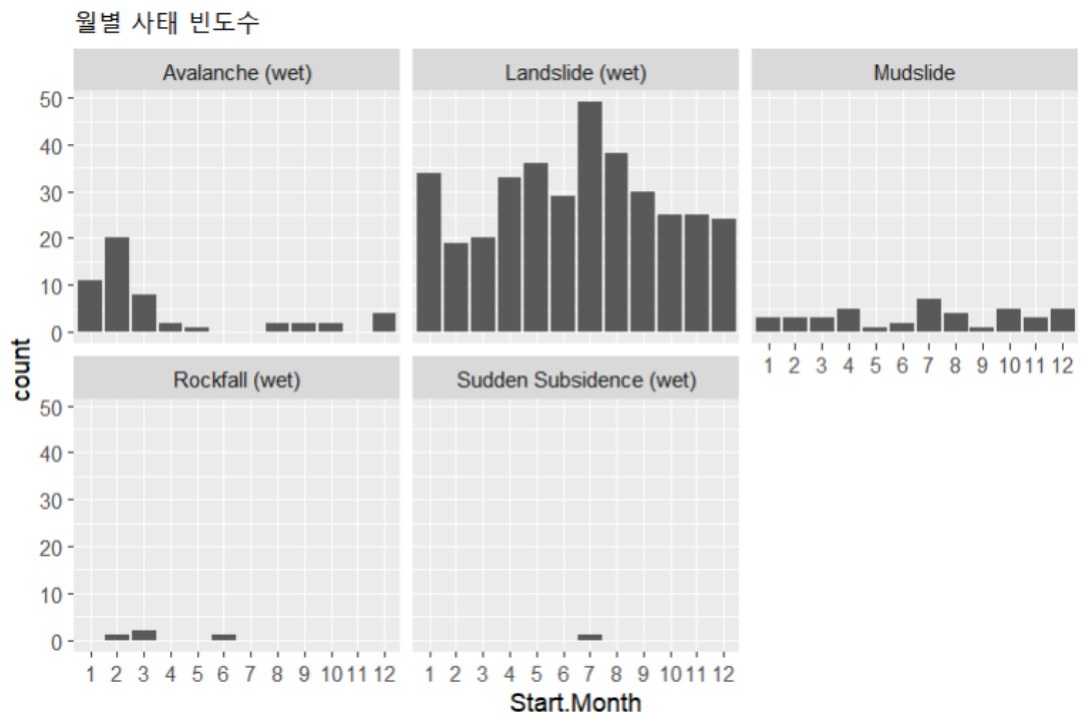
가 있으며 이번에도 년도별, 월별 빈도수를 시각화

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Mass movement (wet)")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "년도별 사태 빈도수")
```



년도별 사태 종류에서는 산사태의 빈도가 가장 많았음을 확인 가능함.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Mass movement (wet)")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "월별 사태 빈도수")
```



▼ Meteorological (기상 관련 재해)

```
data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
  filter(Disaster.Subgroup == "Meteorological") %>%
  distinct()
```

	Disaster.Type	Disaster.Subgroup
1	Extreme temperature	Meteorological
2	Storm	Meteorological

> |

Storm (폭풍)

Extreme temperature (극한 온도)

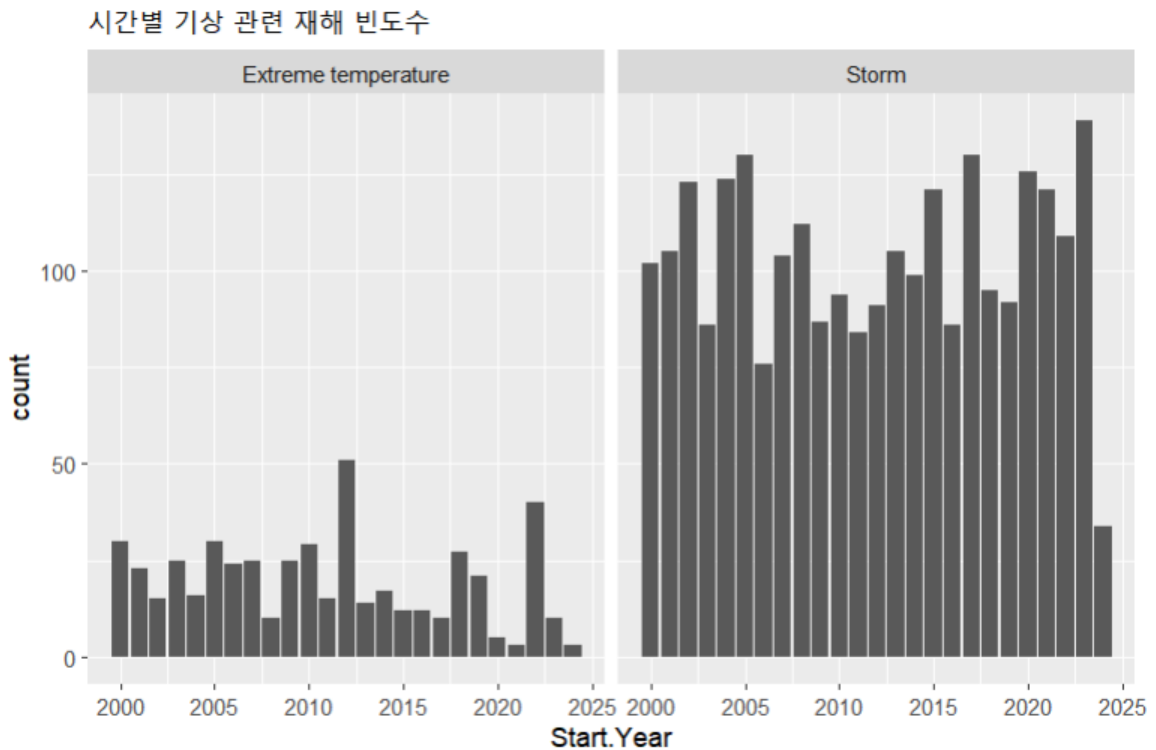
기상 관련 재해로는 폭풍과 극한 온도가 있음

```
data%>%
  group_by(Disaster.Group)%>%
```

```

filter(Disaster.Subgroup == "Meteorological")%>%
ggplot(aes(x = Start.Year))+
geom_bar()+
facet_wrap(~Disaster.Type)+
labs(title = "시간별 기상 관련 재해 빈도수")

```

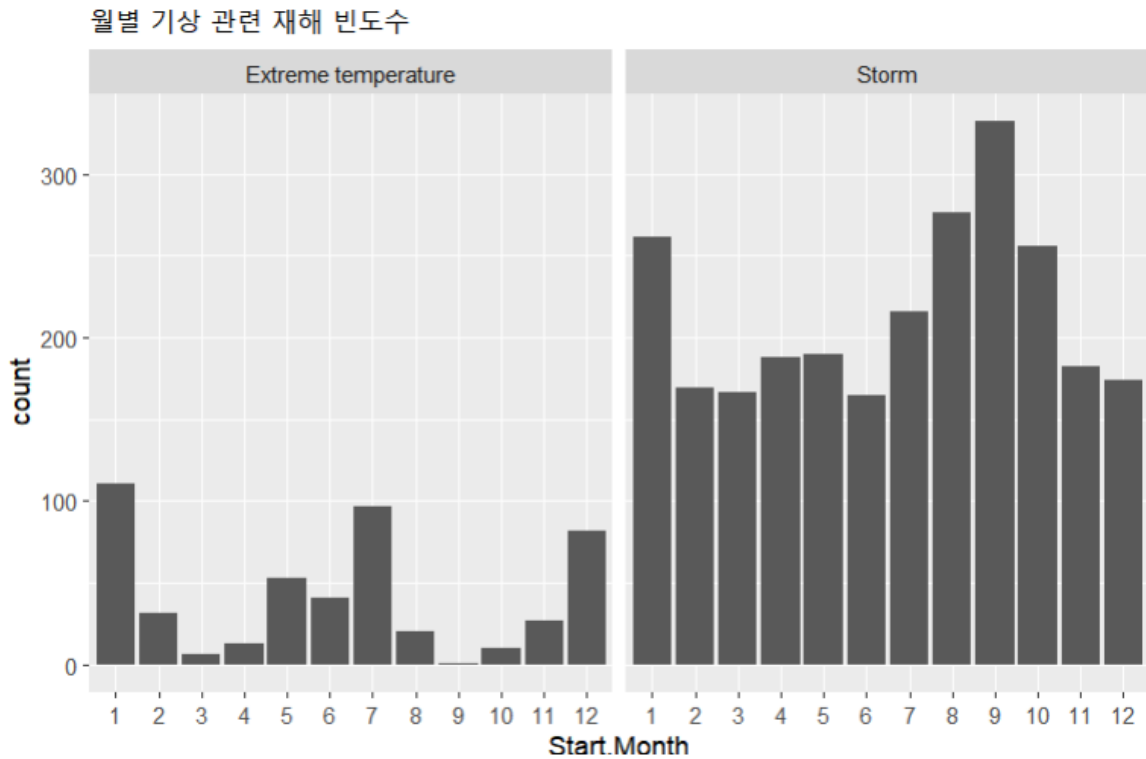


각각의 빈도수는 폭풍의 빈도가 굉장히 높았고, 폭풍의 경우 조금씩 증가하는 추세를 보임

```

data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Meteorological")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "월별 기상 관련 재해 빈도수")

```



극한 온도의 경우 겨울과 여름에 많이 발생하며 12월~1월에 발생하는 극한 온도와 7월에 많이 발생하는 극한 온도의 조건이 다를 것임.

이번에도 Type별 세부 항목을 시각화

▼ Extreme temperature

Disaster.Type	Disaster.Subtype
1 Extreme temperature	Cold wave
2 Extreme temperature	Heat wave
3 Extreme temperature	Severe winter conditions

Cold wave - 한파

Heat wave - 폭염

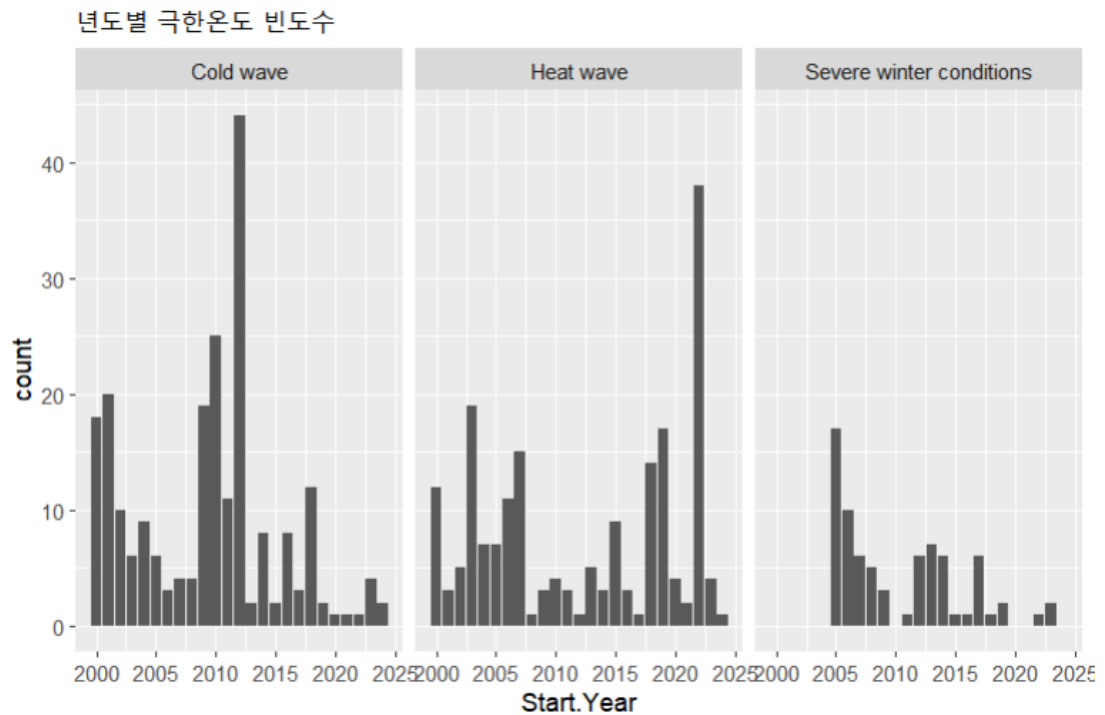
Severe winter conditions - 혹독한 겨울 조건

이상 세 개의 세부Type이 있음

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Extreme temperature")%>%
```

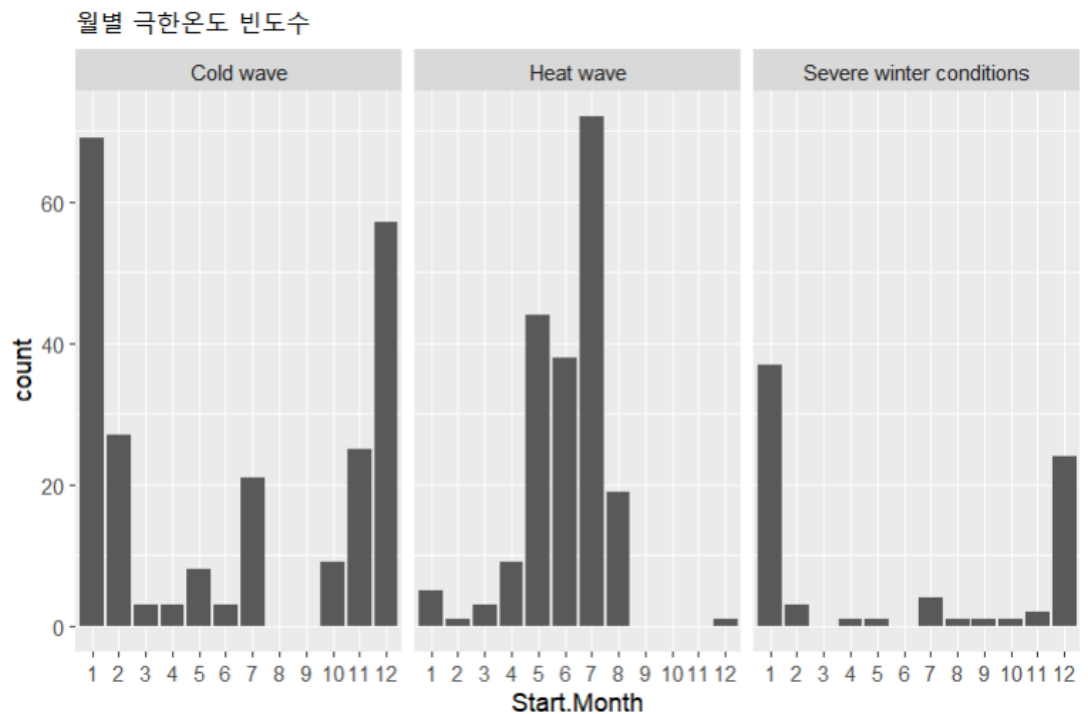


```
ggplot(aes(x = Start.Year)) +
  geom_bar() +
  facet_wrap(~Disaster.Subtype) +
  labs(title = "년도별 극한온도 빈도수")
```



한파의 경우 2010년도, 2011년도에 매우 그 빈도가 잦았고, 그 외에는 감소하는 추세를 보이는 반면, 폭염은 최근들어 점점 그 빈도가 증가하여 기후변화와 연관성을 찾을 수 있음.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Extreme temperature")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month)) +
  geom_bar() +
  facet_wrap(~Disaster.Subtype) +
  labs(title = "월별 극한온도 빈도수")
```



▼ Storm

이번에는 폭풍과 관련하여 세부 type을 찾아보자

	Disaster.Type	Disaster.Subtype
1	Storm	Storm (General)
2	Storm	Tornado
3	Storm	Blizzard/Winter storm
4	Storm	Tropical cyclone
5	Storm	Hail
6	Storm	Severe weather
7	Storm	Lightning/Thunderstorms
8	Storm	Sand/Dust storm
9	Storm	Extra-tropical storm
10	Storm	Storm surge
11	Storm	Derecho

11종류의 폭풍 종류가 있으며,

Storm (General) - 일반적인 폭풍

Tornado - 토네이도

Blizzard/Winter storm - 눈보라/겨울 폭풍

Tropical cyclone - 열대성 저기압

Hail - 우박

Severe weather - 악천후

Lightning/Thunderstorms - 번개/뇌우

Sand/Dust storm - 모래/먼지 폭풍

Extra-tropical storm - 온대성 저기압

Storm surge - 폭풍 해일

Derecho - 데레초

종류가 많아 하나의 그래프로 시각화를 하기에는 굉장히 복잡하고 효율적이지 못하기에 조금 그룹화를 진행함.

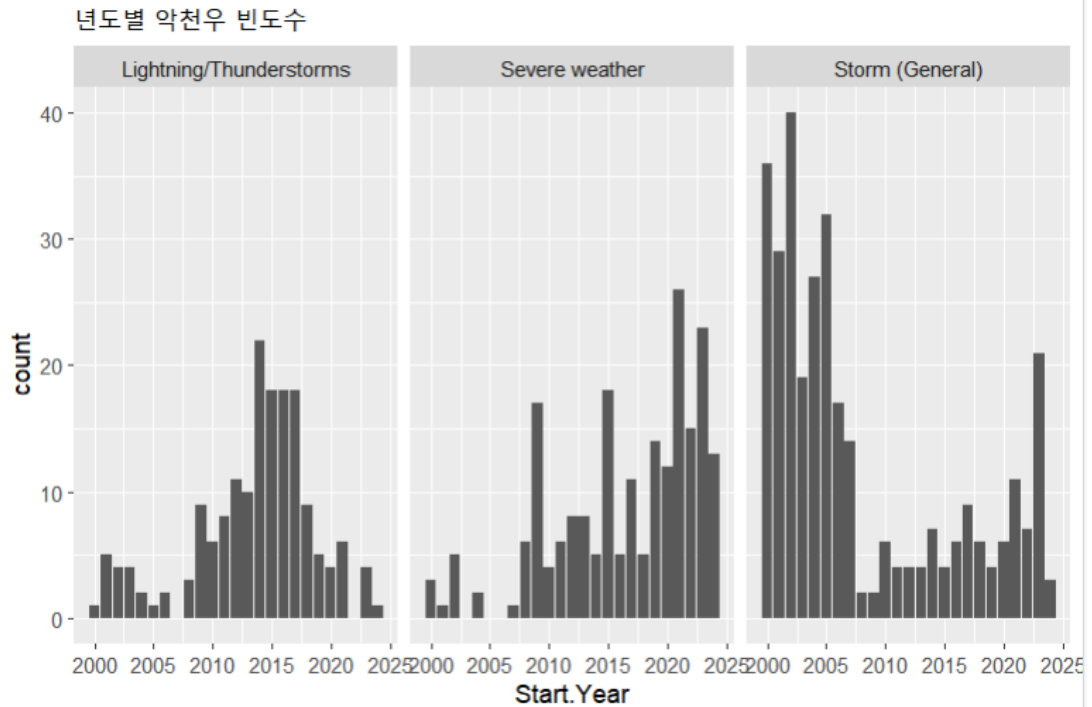
Severe Weather (악천후)

→ **Storm (General), Severe weather, Lightning/Thunderstorms**

일반적인 폭풍, 악천후, 번개/뇌우

```
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm", "Severe winter conditions") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
    Disaster.Subtype %in% c("Storm surge", "Hail") ~ "Water-Related Events",
    TRUE ~ "Other"
  ))%>%
  filter(Storm.Group == "Severe Weather") %>%
  ggplot(aes(x = Start.Year)) +
```

```
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "년도별 악천우 빈도수")
```



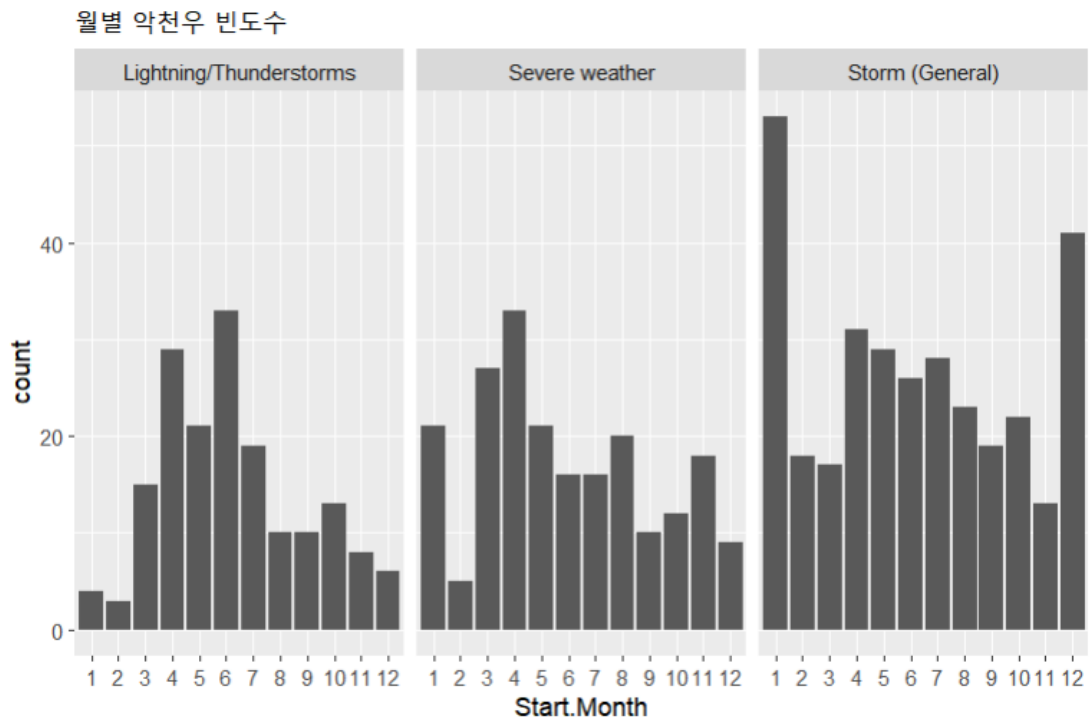
악천후를 확인해볼때, 번개와 뇌우는 2015년도에 자주 발생했고, 악천후의 경우 점
점 증가하는 추세를 보이며, 단순 폭풍의 경우 2000년대 초반에 굉장히 빈도가 많
았던 것 이후로 확 꺾였다가 최근 다시 증가하고 있음

```
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Sever
e weather", "Lightning/Thunderstorms") ~ "Severe Weat
her",
    Disaster.Subtype %in% c("Tropical cyclone", "Extr
a-tropical storm") ~ "Tropical and Extra-tropical Cyc
lones",
    Disaster.Subtype %in% c("Blizzard/Winter storm",
"Severe winter conditions") ~ "Winter and Cold Weathe
r",
```

```

Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
Disaster.Subtype %in% c("Storm surge", "Hail") ~ "Water-Related Events",
TRUE ~ "Other"
))%>%
filter(Storm.Group == "Severe Weather") %>%
mutate(Start.Month = factor(Start.Month))%>%
ggplot(aes(x = Start.Month)) +
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "월별 악천우 빈도수")

```



월별 빈도를 보았을때, 뇌우를 동반한 악천후는 주로 3,4,5,6월 에 많이 발생하고, 일반적인 폭풍은 12,1월에 많이 발생하고 있다.

Tropical and Extra-tropical Cyclones (열대성 및 온대성 저기압)

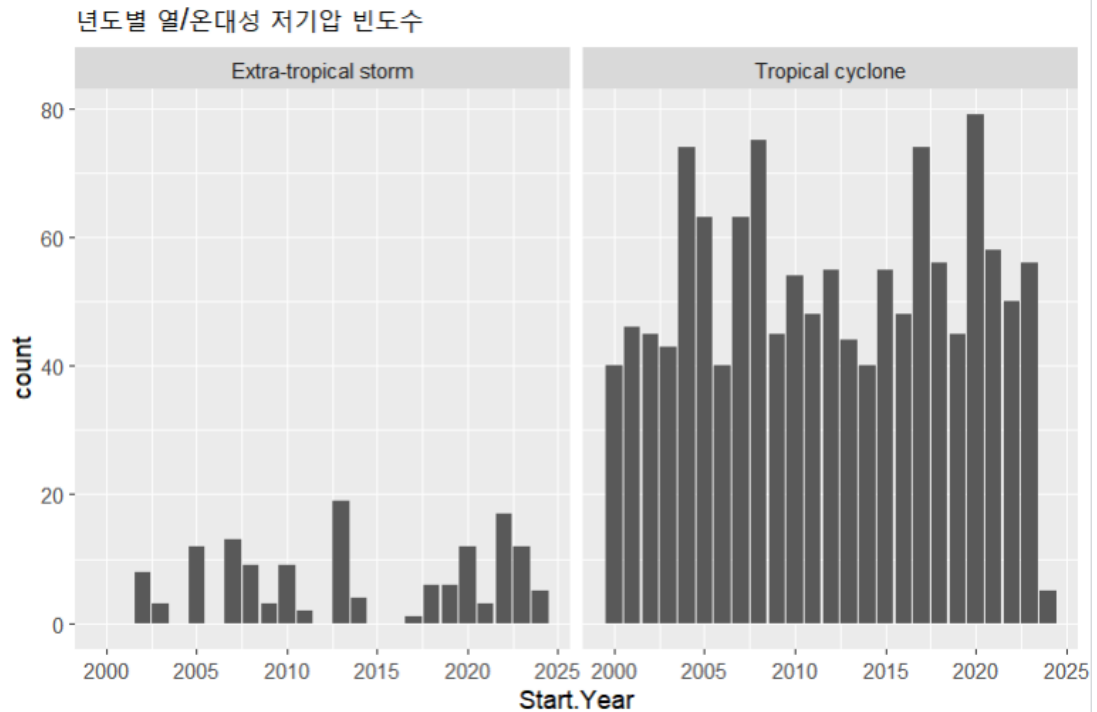
→ **Tropical cyclone, Extra-tropical storm**

열대성 저기압, 온대성 저기압

```

data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm", "Severe winter conditions") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
    Disaster.Subtype %in% c("Storm surge", "Hail") ~ "Water-Related Events",
    TRUE ~ "Other"
  ))%>%
  filter(Storm.Group == "Tropical and Extra-tropical Cyclones") %>%
  ggplot(aes(x = Start.Year)) +
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "년도별 열/온대성 저기압 빈도수")

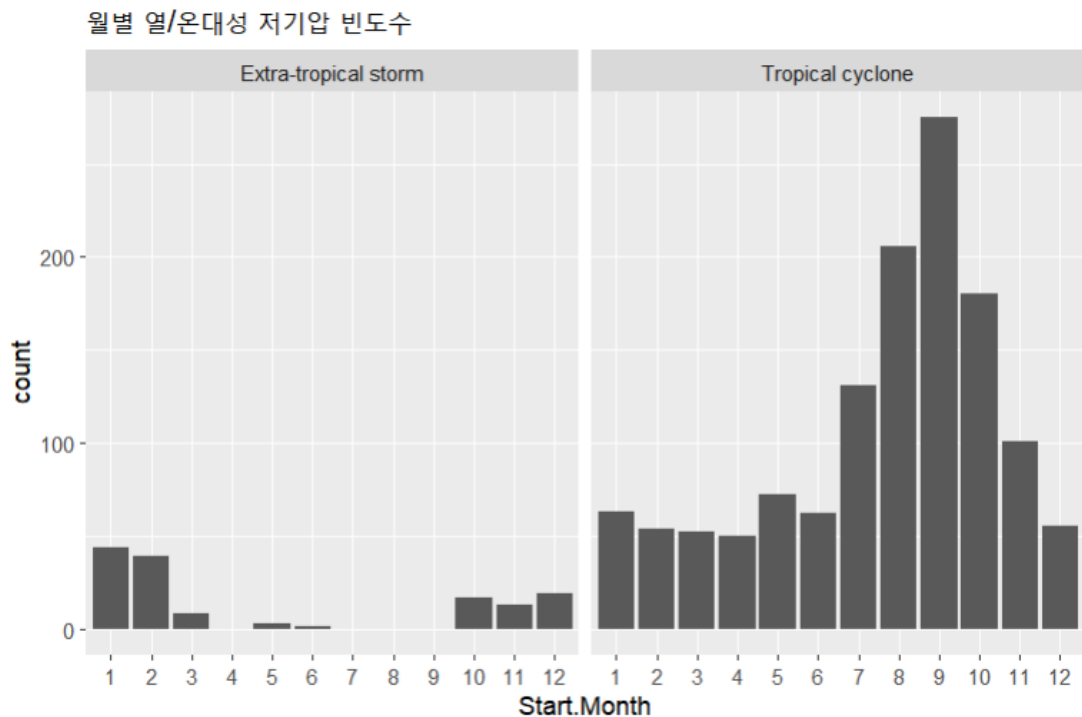
```



년도별 데이터상으로는 온대성 저기압보다 열대성 저기압의 비중이 더 높았으며

```
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm", "Severe winter conditions") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
    Disaster.Subtype %in% c("Storm surge", "Hail") ~ "Water-Related Events",
    TRUE ~ "Other"
  ))%>%
  filter(Storm.Group == "Tropical and Extra-tropical Cyclones") %>%
  mutate(Start.Month = factor(Start.Month))%>%
```

```
ggplot(aes(x = Start.Month)) +
  geom_bar() +
  facet_wrap(~Disaster.Subtype) +
  labs(title = "월별 열/온대성 저기압 빈도수")
```



월별 데이터상에서 열대성 저기압의 경우 9월에 가장 많이 나타났음

Winter and Cold Weather (겨울 및 한파)

→ Blizzard/Winter storm, Severe winter conditions

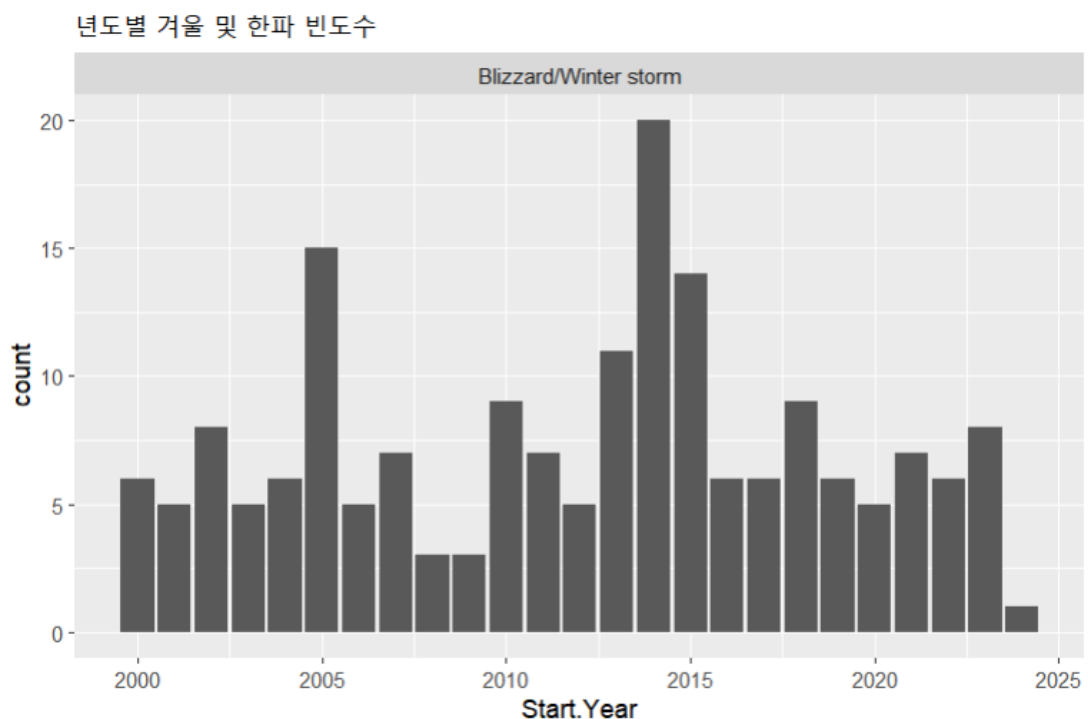
```
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sa
```



```

nd/Dust storm") ~ "Wind and Dust",
  Disaster.Subtype %in% c("Storm surge", "Hail") ~
"Water-Related Events",
  TRUE ~ "Other"
))%>%
filter(Storm.Group == "Winter and Cold Weather") %
>%
ggplot(aes(x = Start.Year)) +
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "년도별 겨울 및 한파 빈도수")

```



겨울 폭풍의 경우 2015년 전후로 빈도가 높았음

```

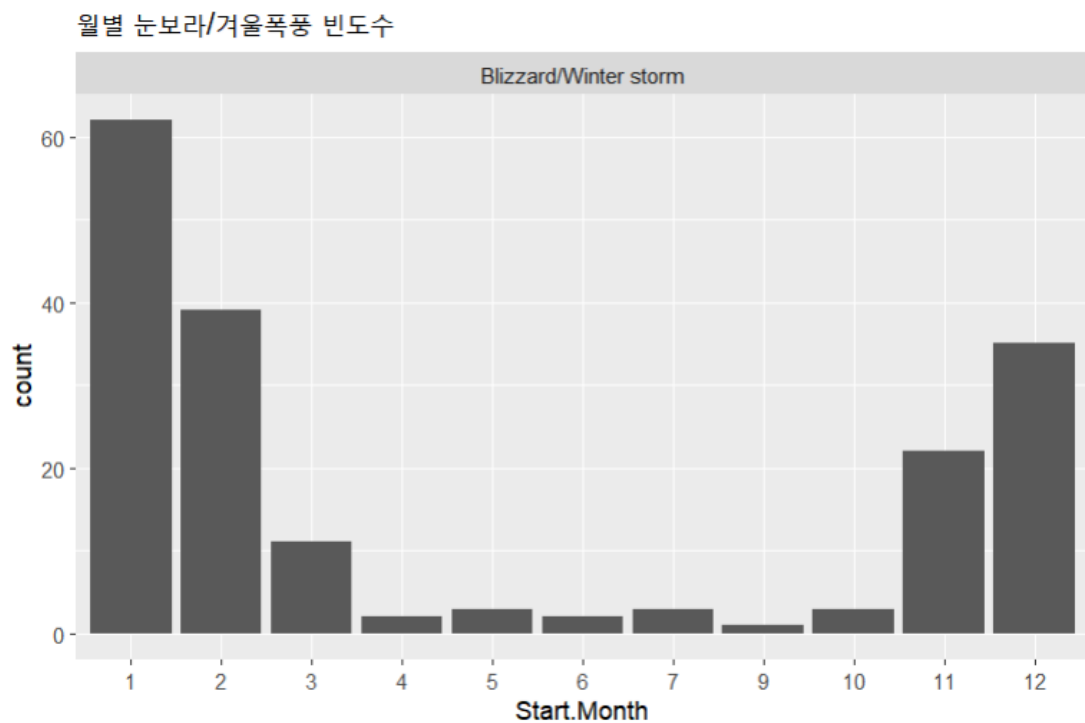
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe
e weather", "Lightning/Thunderstorms") ~ "Severe Weat
her",
    Disaster.Subtype %in% c("Tropical cyclone", "Extr
a-tropical storm") ~ "Tropical and Extra-tropical Cyc
lones",

```

```

Disaster.Subtype %in% c("Blizzard/Winter storm")
~ "Winter and Cold Weather",
Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
Disaster.Subtype %in% c("Storm surge", "Hail") ~ "Water-Related Events",
TRUE ~ "Other"
))%>%
filter(Storm.Group == "Winter and Cold Weather") %>%
mutate(Start.Month = factor(Start.Month))%>%
ggplot(aes(x = Start.Month)) +
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "월별 눈보라/겨울폭풍 빈도수")

```



월별 빈도 역시 겨울에 집중되어있음

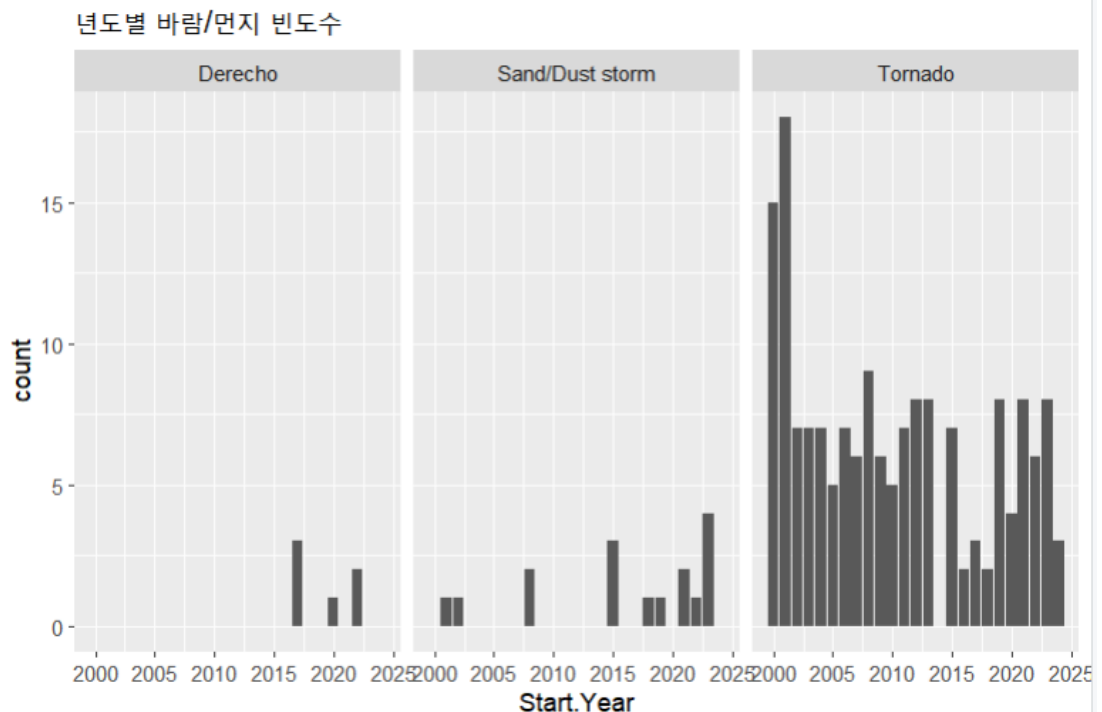
Wind and Dust (바람과 먼지 관련 재해)

→ Tornado, Derecho, Sand/Dust storm

```

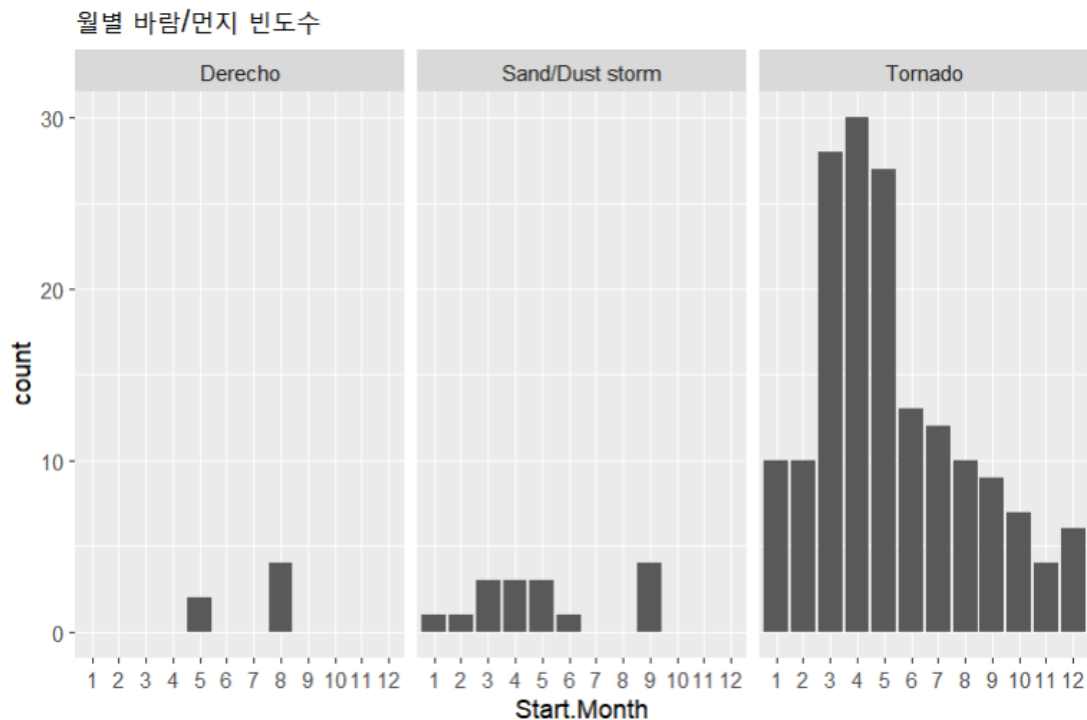
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm", "Severe winter conditions") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
    Disaster.Subtype %in% c("Storm surge", "Hail") ~ "Water-Related Events",
    TRUE ~ "Other"
  ))%>%
  filter(Storm.Group == "Wind and Dust") %>%
  ggplot(aes(x = Start.Year)) +
  geom_bar() +
  facet_wrap(~Disaster.Subtype) +
  labs(title = "년도별 바람/먼지 빈도수")

```



```
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm", "Severe winter conditions") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
    Disaster.Subtype %in% c("Storm surge", "Hail") ~ "Water-Related Events",
    TRUE ~ "Other"
  ))%>%
  filter(Storm.Group == "Wind and Dust") %>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month)) +
  geom_bar()+
```

```
facet_wrap(~Disaster.Subtype)+
labs(title = "월별 바람/먼지 빈도수")
```



Water-Related Events (물 관련 현상)

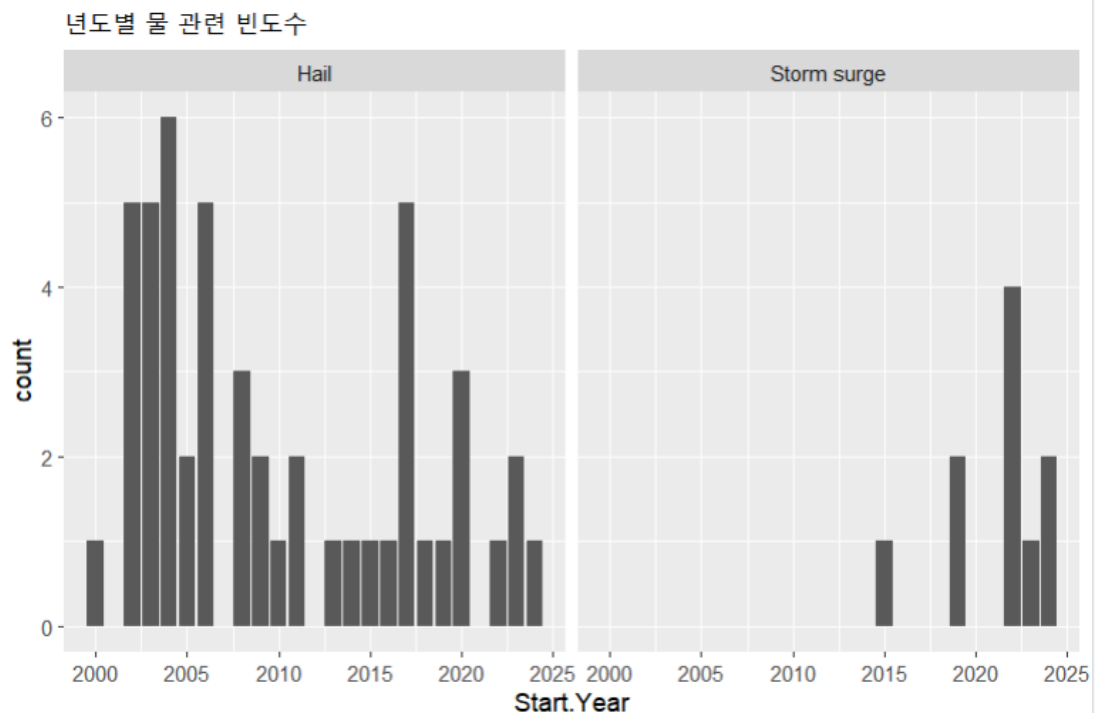
→ Storm surge, Hail

```
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm", "Severe winter conditions") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sand/Dust storm") ~ "Wind and Dust",
    Disaster.Subtype %in% c("Storm surge", "Hail") ~
```

```

"Water-Related Events",
  TRUE ~ "Other"
))%>%
filter(Storm.Group == "Water-Related Events") %>%
ggplot(aes(x = Start.Year)) +
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "년도별 물 관련 빈도수")

```



```

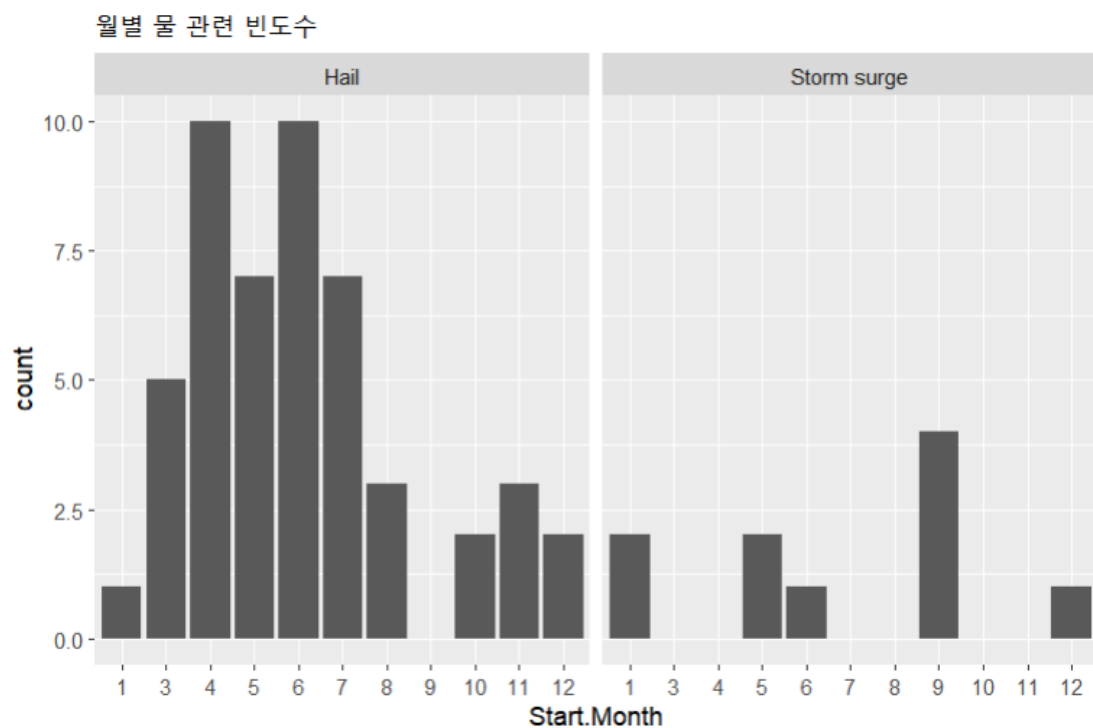
data %>%
  mutate(Storm.Group = case_when(
    Disaster.Subtype %in% c("Storm (General)", "Severe weather", "Lightning/Thunderstorms") ~ "Severe Weather",
    Disaster.Subtype %in% c("Tropical cyclone", "Extra-tropical storm") ~ "Tropical and Extra-tropical Cyclones",
    Disaster.Subtype %in% c("Blizzard/Winter storm", "Severe winter conditions") ~ "Winter and Cold Weather",
    Disaster.Subtype %in% c("Tornado", "Derecho", "Sa

```

```

nd/Dust storm") ~ "Wind and Dust",
  Disaster.Subtype %in% c("Storm surge", "Hail") ~
"Water-Related Events",
  TRUE ~ "Other"
))%>%
filter(Storm.Group == "Water-Related Events") %>%
mutate(Start.Month = factor(Start.Month))%>%
ggplot(aes(x = Start.Month)) +
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "월별 물 관련 빈도수")

```



▼ Miscellaneous accident (기타 사고)

```

data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
  filter(Disaster.Subgroup == "Miscellaneous accident")
%>%
  distinct()

```

	Disaster.Type	Disaster.Subgroup
1	Fire (Miscellaneous)	Miscellaneous accident
2	Collapse (Miscellaneous)	Miscellaneous accident
3	Explosion (Miscellaneous)	Miscellaneous accident
4	Miscellaneous accident (General)	Miscellaneous accident

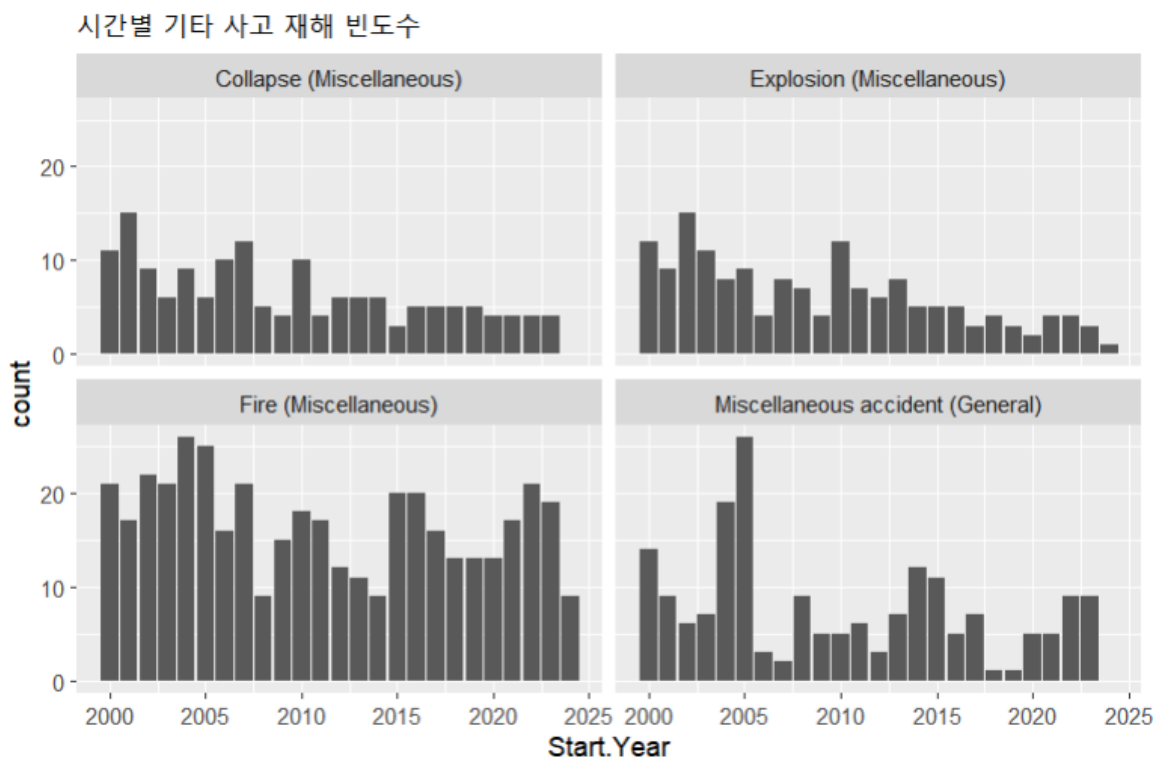
Collapse (Miscellaneous) (붕괴 - 기타)

Fire (Miscellaneous) (화재-기타)

Explosion (Miscellaneous) (폭발 - 기타)

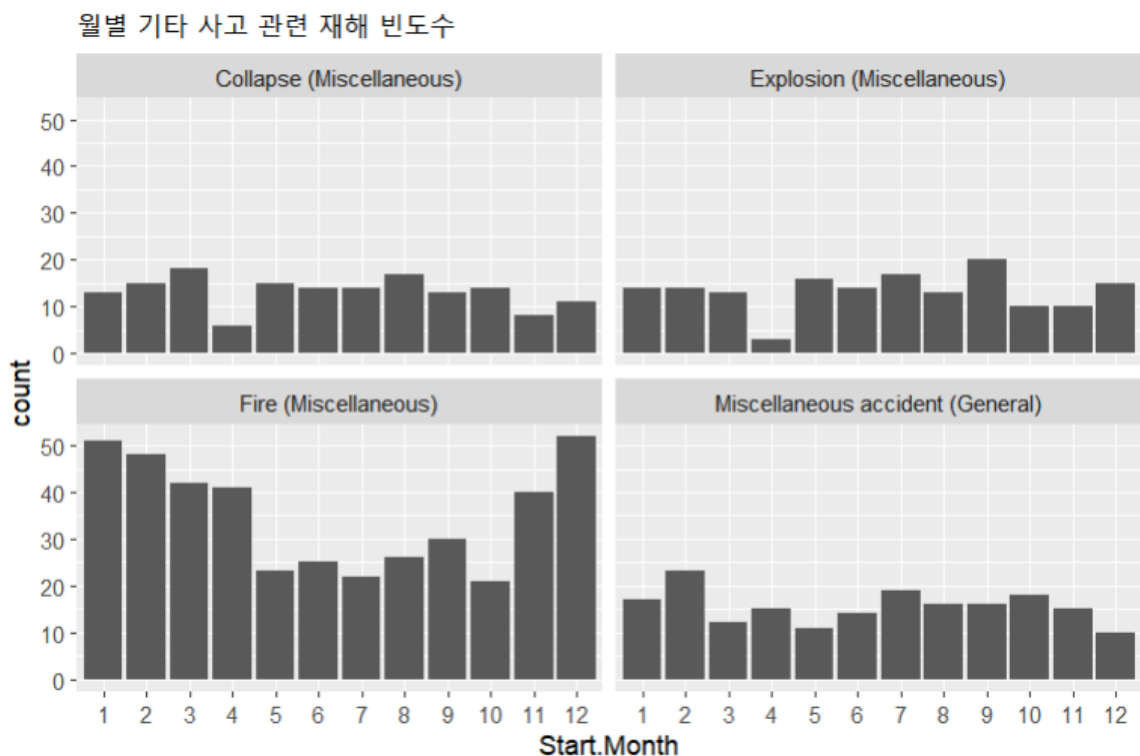
Miscellaneous accident (General) (기타 사고 - 일반)

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Miscellaneous accident")%
  >%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "시간별 기타 사고 재해 빈도수")
```



기타 사고중 폭발과 붕괴는 해마다 수가 감소하는 추세로 관련 법률이나 제도의 변경으로 사고가 줄어들었다고 생각할 수 있음

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Miscellaneous accident")%
  >%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "월별 기타 사고 관련 재해 빈도수")
```



기타 사고 역시 세부 type별 분류를 하여 시각화를 시작한다

Disaster.Type	Disaster.Subtype
1 Collapse (Miscellaneous)	Collapse (Miscellaneous)

Disaster.Type	Disaster.Subtype
1 Fire (Miscellaneous)	Fire (Miscellaneous)

	Disaster.Type	Disaster.Subtype
1	Explosion (Miscellaneous)	Explosion (Miscellaneous)

	Disaster.Type	Disaster.Subtype
1	Miscellaneous accident (General)	Miscellaneous accident (General)

기타 사고의 세부 type은 모두 하나씩 동일하므로 Type별 시각화 자료와 동일하다

▼ Geophysical (지질학적 재해)

```
data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
  filter(Disaster.Subgroup == "Geophysical") %>%
  distinct()
```

	Disaster.Type	Disaster.Subgroup
1	Volcanic activity	Geophysical
2	Earthquake	Geophysical
3	Mass movement (dry)	Geophysical

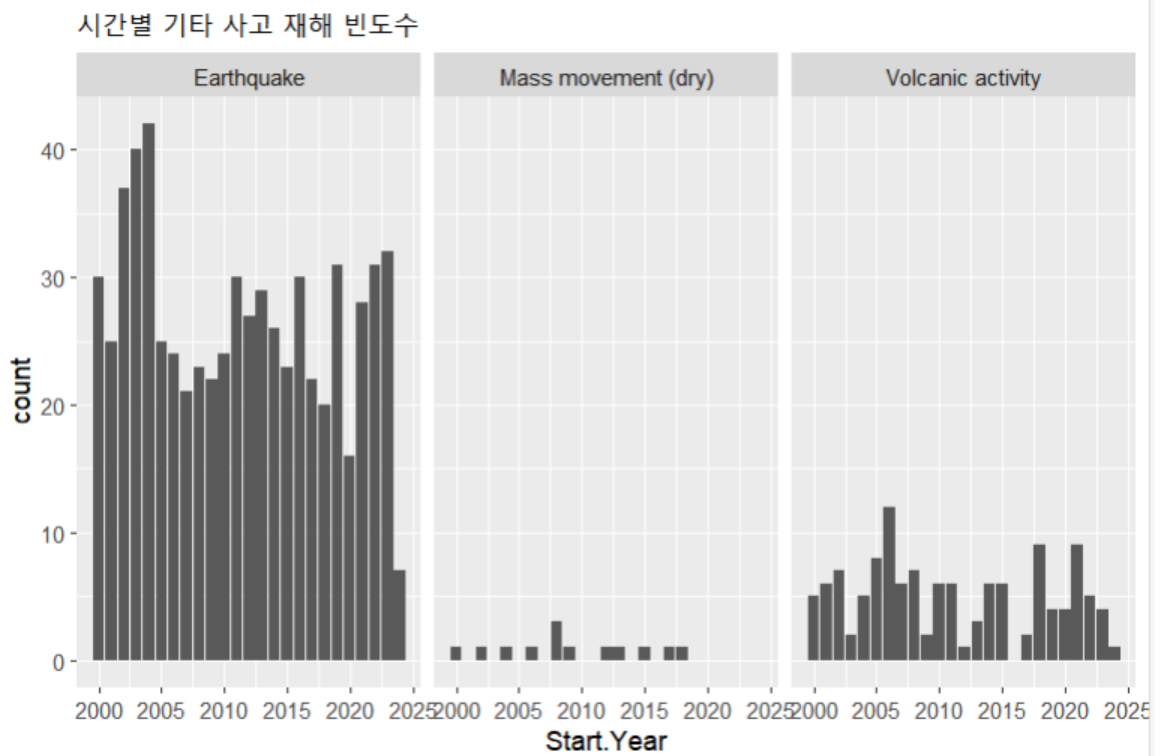
Earthquake (지진)

Volcanic activity (화산 활동)

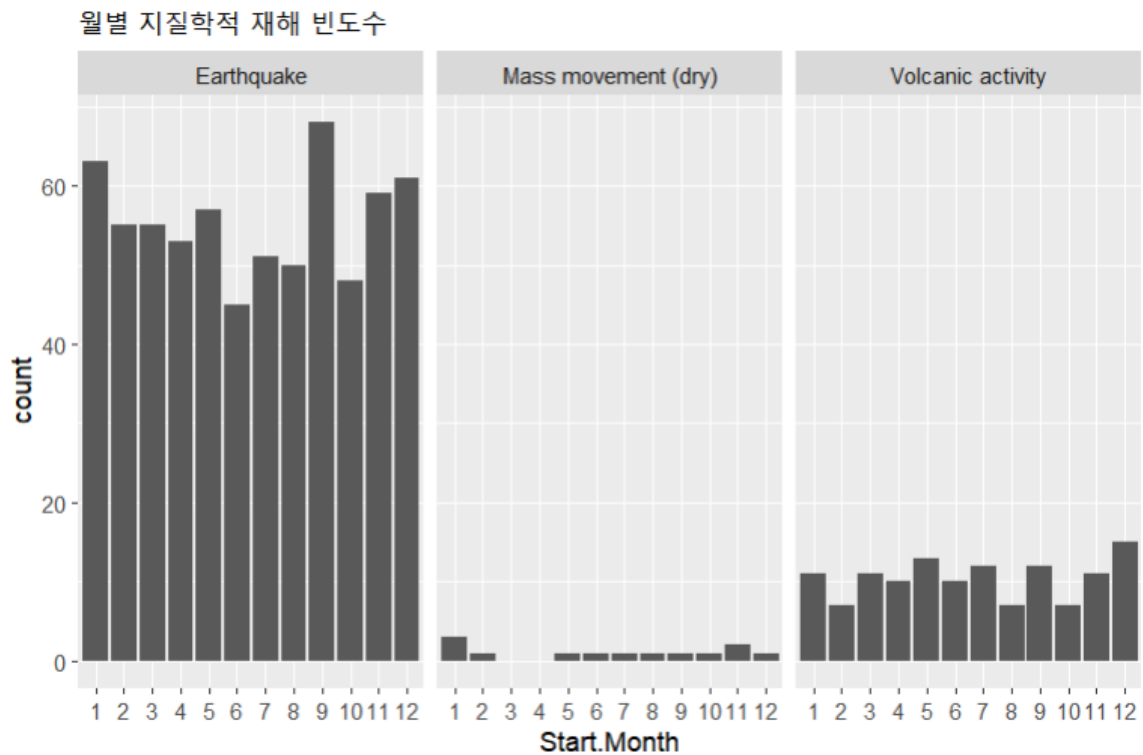
Mass movement (dry) (대규모 이동 - 건조한 조건)

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Geophysical")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
```

```
facet_wrap(~Disaster.Type)+
labs(title = "시간별 지질학적 재해 빈도수")
```



```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Geophysical")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "월별 지질학적 재해 빈도수")
```



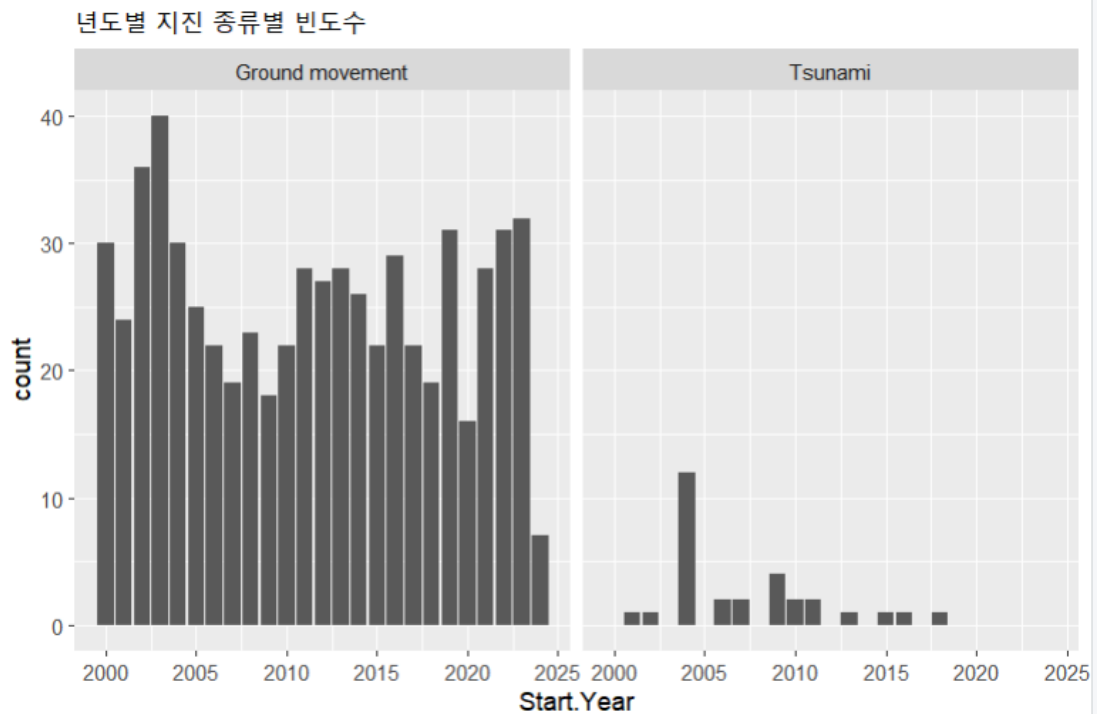
▼ Earthquake

	Disaster.Type	Disaster.Subtype
1	Earthquake	Ground movement
2	Earthquake	Tsunami

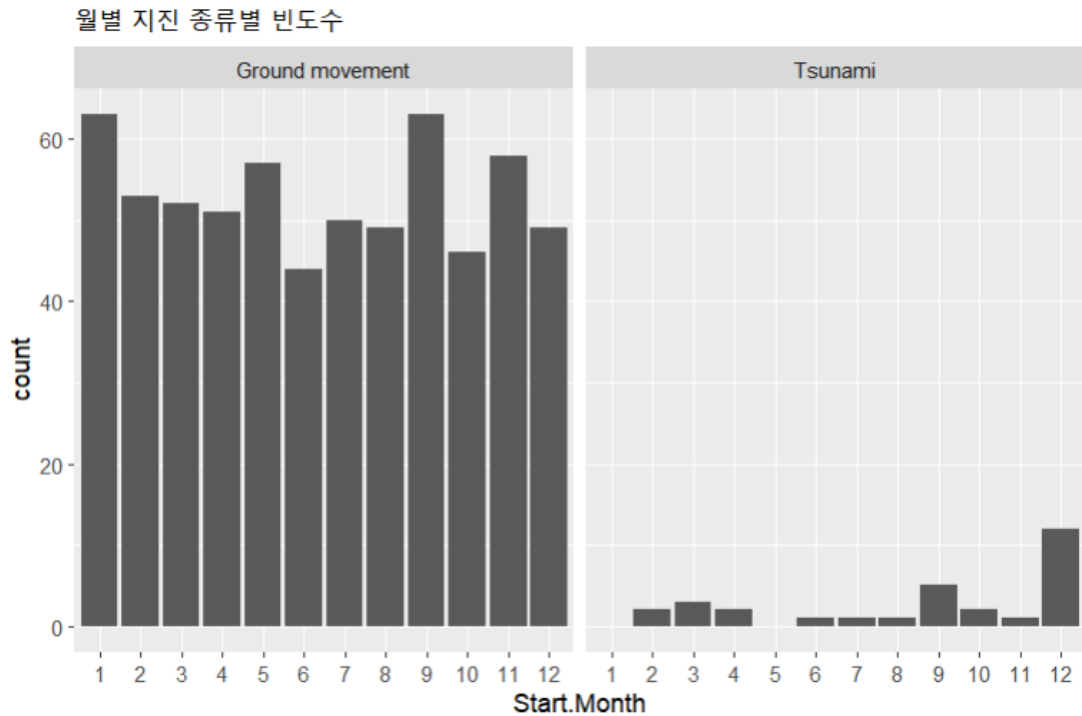
Earthquake (Ground movement) - 지진 (지반 이동)

Earthquake (Tsunami) - 지진 (쓰나미)

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Earthquake")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "년도별 지진 종류별 빈도수")
```



```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Earthquake")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "월별 지진 종류별 빈도수")
```



▼ Volcanic activity

화산 활동과 관련해서는

Disaster.Type	Disaster.Subtype
1 Volcanic activity	Ash fall
2 Volcanic activity	Volcanic activity (General)
3 Volcanic activity	Lava flow
4 Volcanic activity	Pyroclastic flow

4개의 세부 type이 있음

Volcanic activity (Ash fall) - 화산 활동 (화산재 낙하)

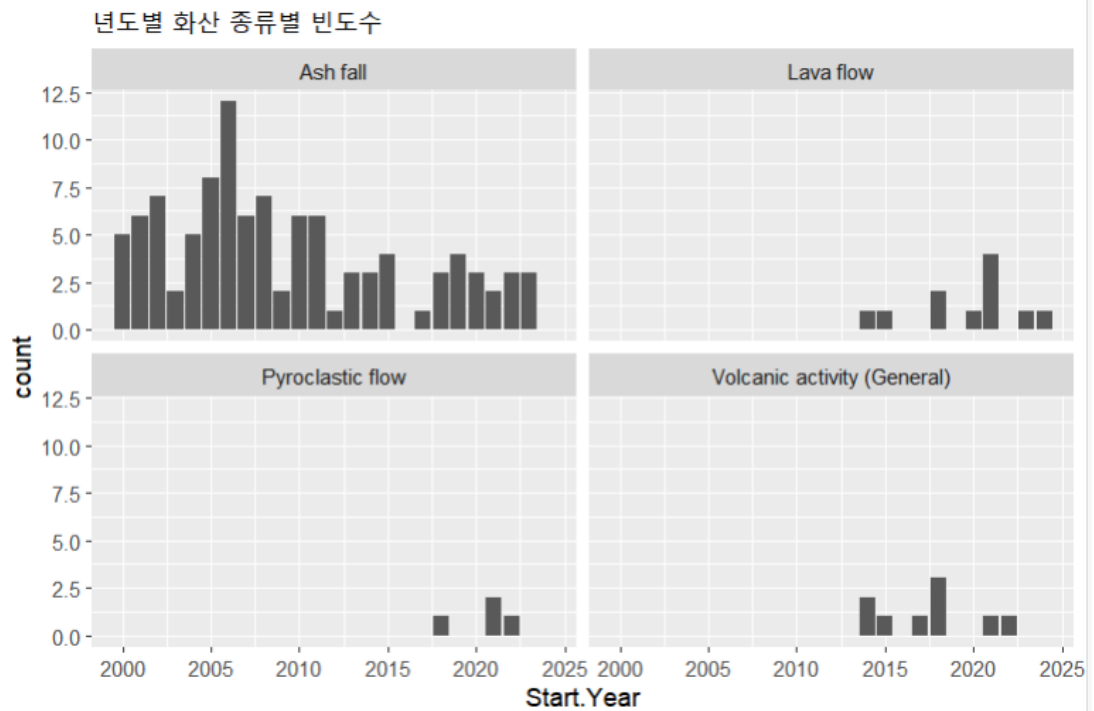
Volcanic activity (Volcanic activity - General) (일반)

Volcanic activity (Lava flow) - 화산 활동 (용암류)

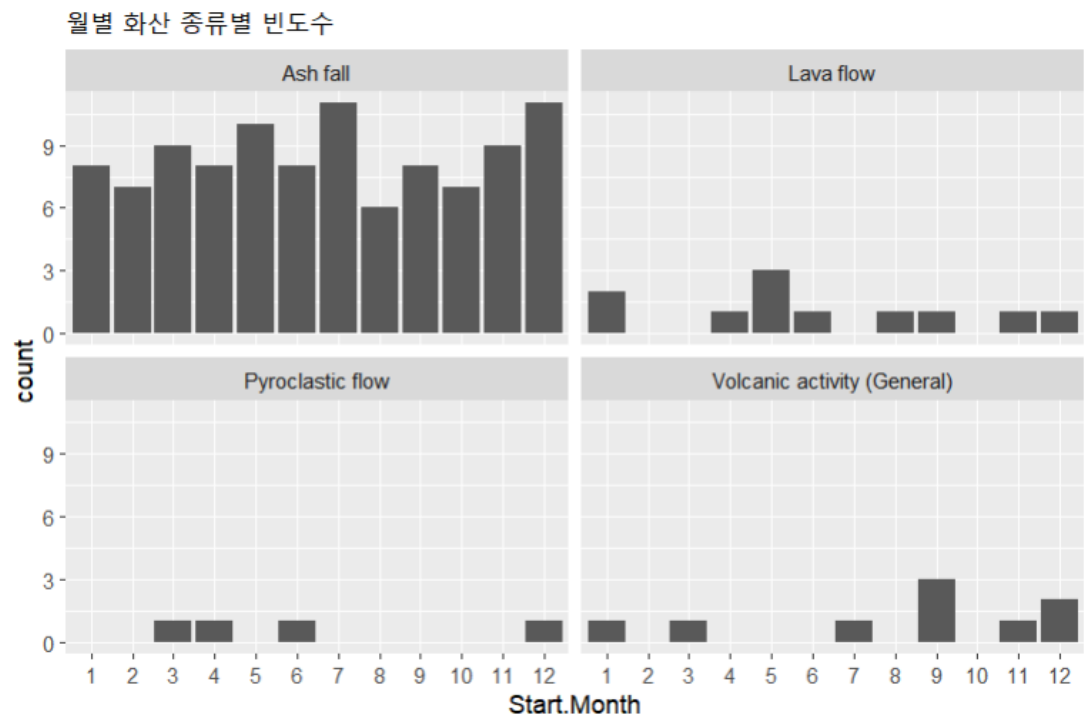
Volcanic activity (Pyroclastic flow) - 화산 활동 (화쇄류)

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Volcanic activity")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
```

```
facet_wrap(~Disaster.Subtype)+
labs(title = "년도별 화산 종류별 빈도수")
```



```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Volcanic activity")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "월별 화산 종류별 빈도수")
```



▼ Mass movement (dry)

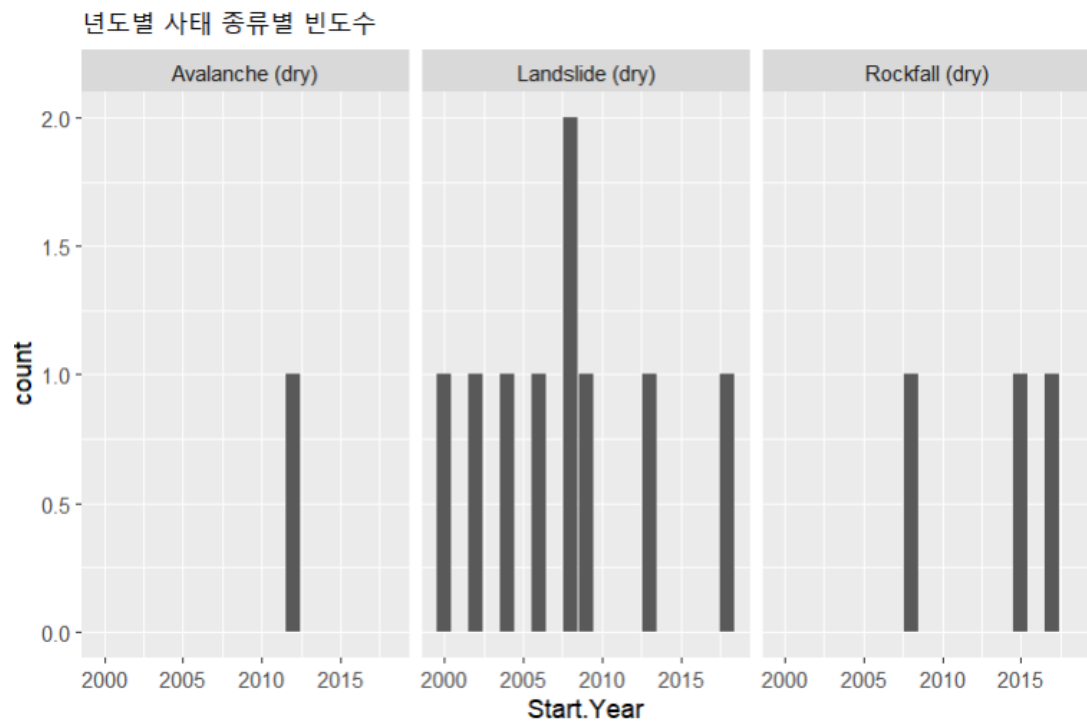
	Disaster.Type	Disaster.Subtype
1	Mass movement (dry)	Landslide (dry)
2	Mass movement (dry)	Rockfall (dry)
3	Mass movement (dry)	Avalanche (dry)

Mass movement (dry) (Landslide - dry) 산사태

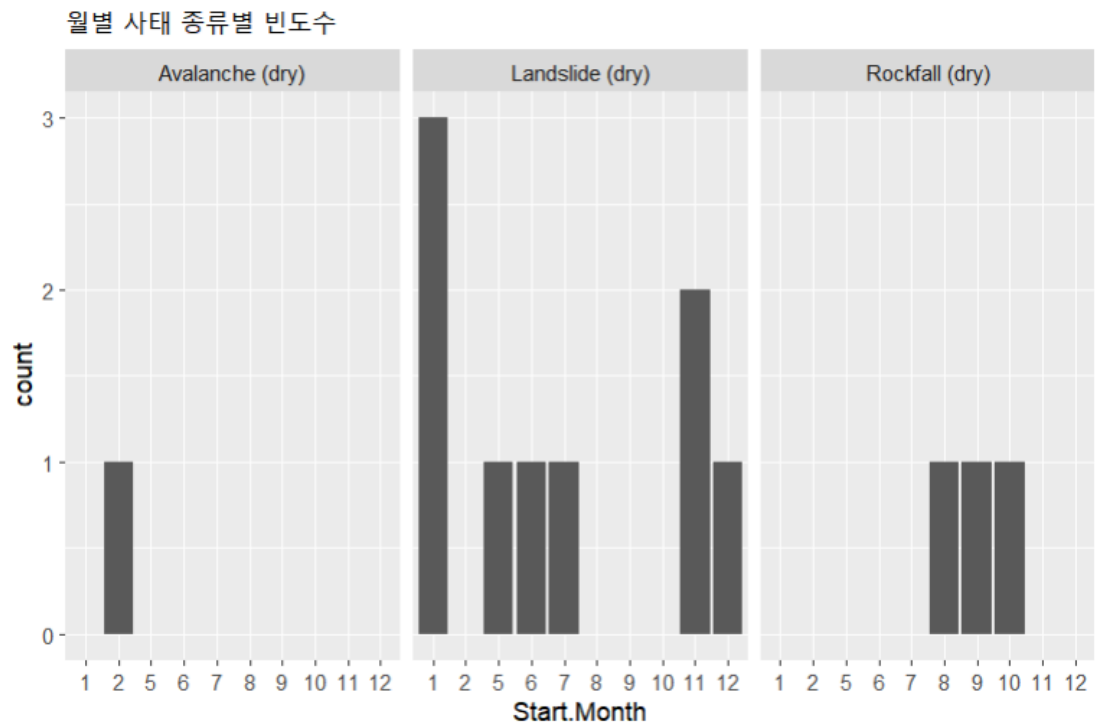
Mass movement (dry) (Rockfall - dry) 낙석

Mass movement (dry) (Avalanche - dry) 눈사태

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Mass movement (dry)")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "년도별 사태 종류별 빈도수")
```

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Mass movement (dry)")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "월별 사태 종류별 빈도수")
```



▼ Industrial accident (산업 재해)

```
data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
  filter(Disaster.Subgroup == "Industrial accident") %>%
  distinct()
```

	Disaster.Type	Disaster.Subgroup
1	Collapse (Industrial)	Industrial accident
2	Fire (Industrial)	Industrial accident
3	Explosion (Industrial)	Industrial accident
4	Chemical spill	Industrial accident
5	Gas leak	Industrial accident
6	Poisoning	Industrial accident
7	Industrial accident (General)	Industrial accident
8	Radiation	Industrial accident
9	Oil spill	Industrial accident

> |

Fire (Industrial) (화재 - 산업)

Explosion (Industrial) (폭발 - 산업): 산업 현장에서 발생한 폭발.

Chemical spill (화학물질 유출): 화학 물질이 유출되어 발생하는 사고.

Gas leak (가스 누출): 산업 현장에서 가스가 누출되어 발생하는 재해.

Radiation (방사능 누출): 방사능이 누출되어 환경이나 인체에 영향을 미치는 재해.

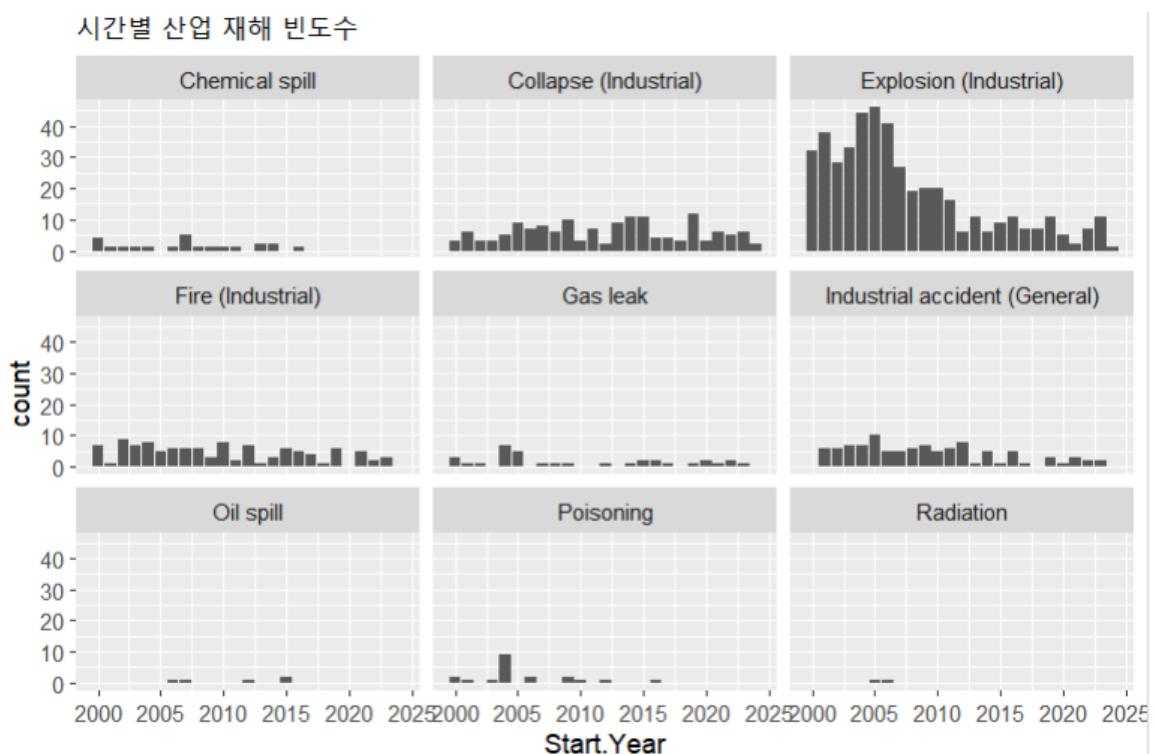
Industrial accident (General) (산업 사고 - 일반): 산업 전반에서 발생하는 기타 사고.

Oil spill (기름 유출)

Poisoning

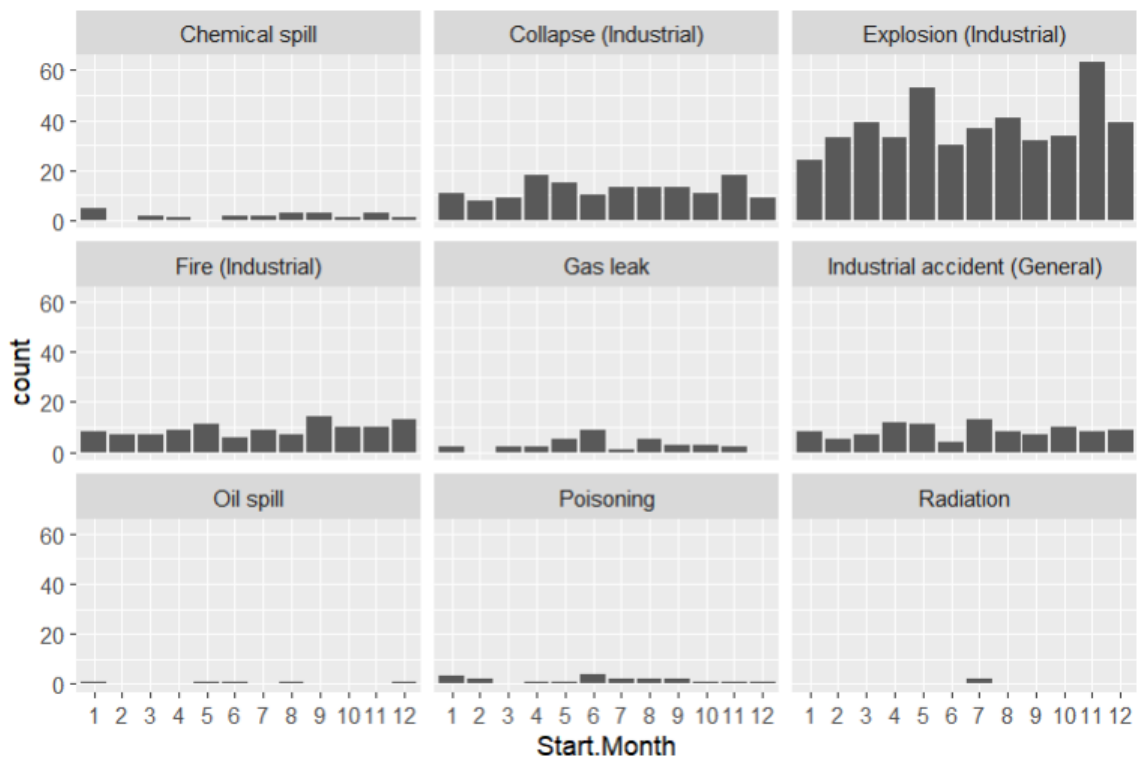
Collapse (Industrial)

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "시간별 산업 재해 빈도수")
```



월별 빈도수 추이

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Type)
```



type별 세부 유형 확인

```
Disaster.Type  Disaster.Subtype
1 Fire (Industrial) Fire (Industrial)
```

```
Disaster.Type  Disaster.Subtype
1 Explosion (Industrial) Explosion (Industrial)
```

```
Disaster.Type  Disaster.Subtype
1 Chemical spill Chemical spill
```

	Disaster.Type	Disaster.Subtype
1	Gas leak	Gas leak

	Disaster.Type	Disaster.Subtype
1	Radiation	Radiation

	Disaster.Type	Disaster.Subtype
1	Industrial accident (General)	Industrial accident (General)

	Disaster.Type	Disaster.Subtype
1	Oil spill	Oil spill

	Disaster.Type	Disaster.Subtype
1	Poisoning	Poisoning

	Disaster.Type	Disaster.Subtype
1	Collapse (Industrial)	Collapse (Industrial)

산업재해 역시 type별 세부 type이 동일해 년도, 월별 빈도는 동일하다

▼ Biological (생물학적 재해)

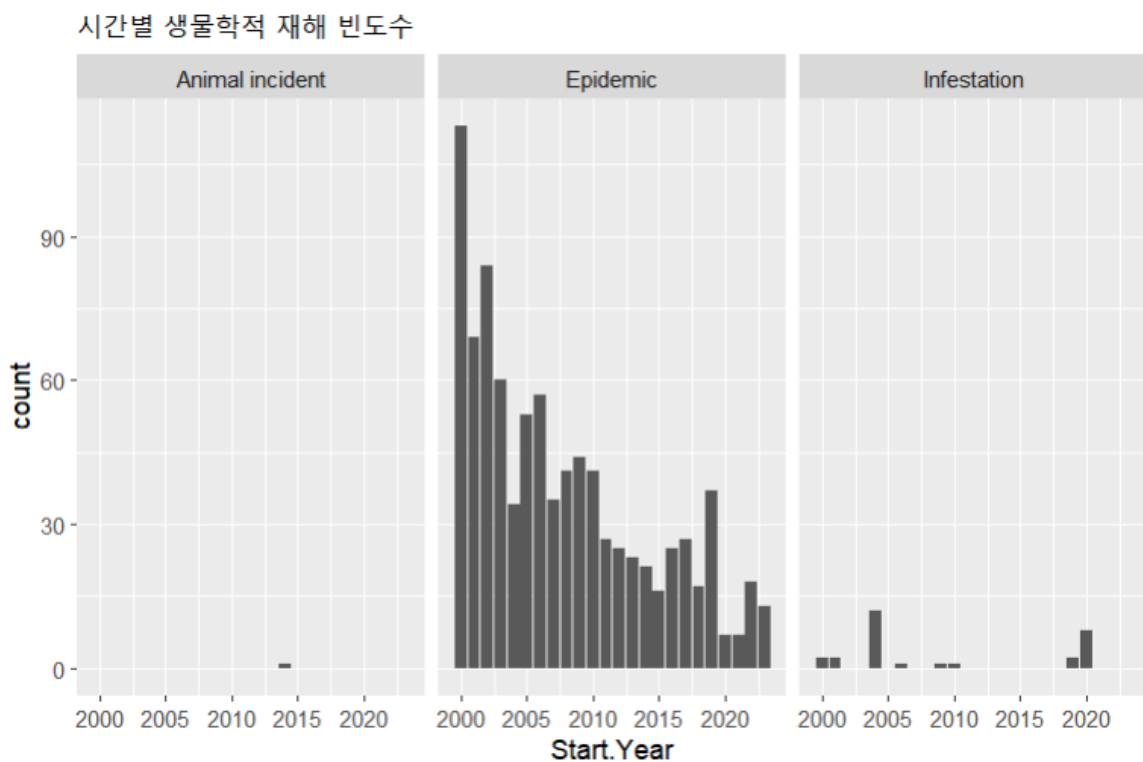
```
data %>%
  select(Disaster.Type, Disaster.Subgroup) %>%
  filter(Disaster.Subgroup == "Biological") %>%
  distinct()
```

	Disaster.Type	Disaster.Subgroup
1	Epidemic	Biological
2	Infestation	Biological
3	Animal incident	Biological

• Epidemic (전염병)

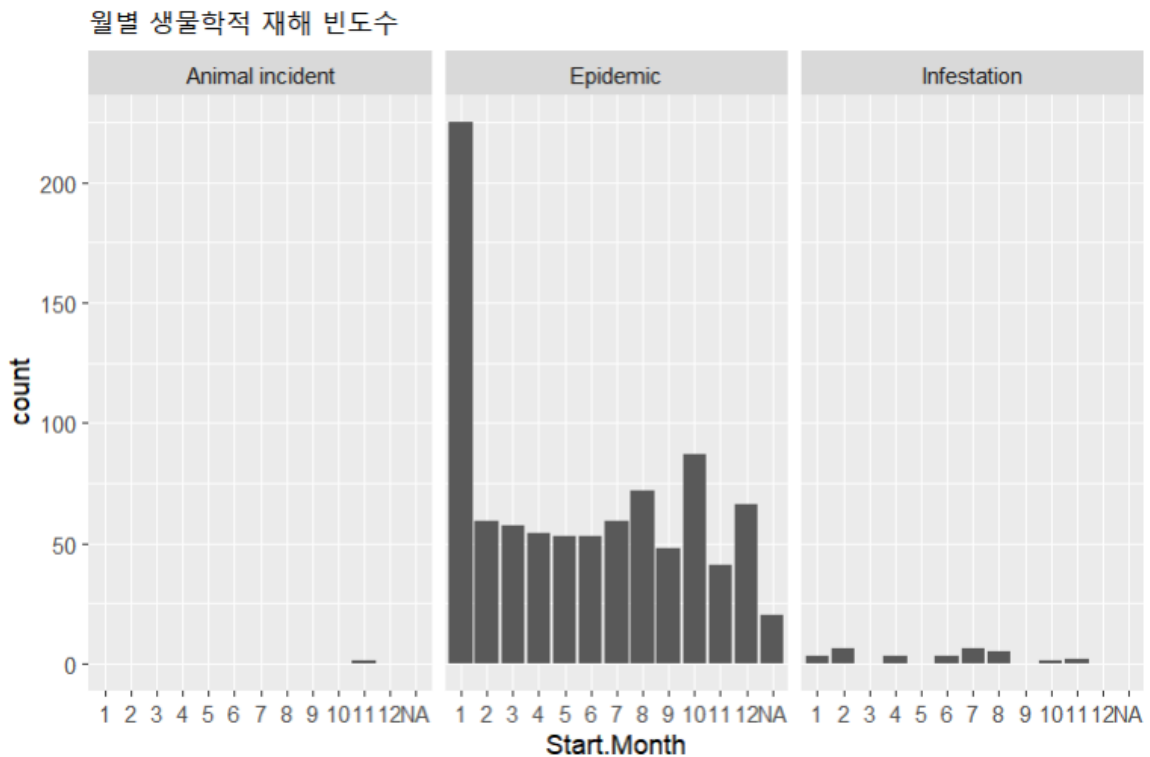
- Infestation (해충 침입)
- Animal incident (동물 사고)

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Biological")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Type)+
  labs(title = "시간별 생물학적 재해 빈도수")
```



```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Biological")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
```

```
facet_wrap(~Disaster.Type)+
labs(title = "월별 생물학적 재해 빈도수")
```



생물학 재해에서 유형별 세부 type

▼ Epidemic

	Disaster.Type	Disaster.Subtype
1	Epidemic	Viral disease
2	Epidemic	Infectious disease (General)
3	Epidemic	Bacterial disease
4	Epidemic	Parasitic disease

Epidemic (Viral disease) - 전염병 (바이러스 질병)

Epidemic (Infectious disease - General) - 전염병 (일반적인 감염병)

Epidemic (Bacterial disease) - 전염병 (세균성 질병)

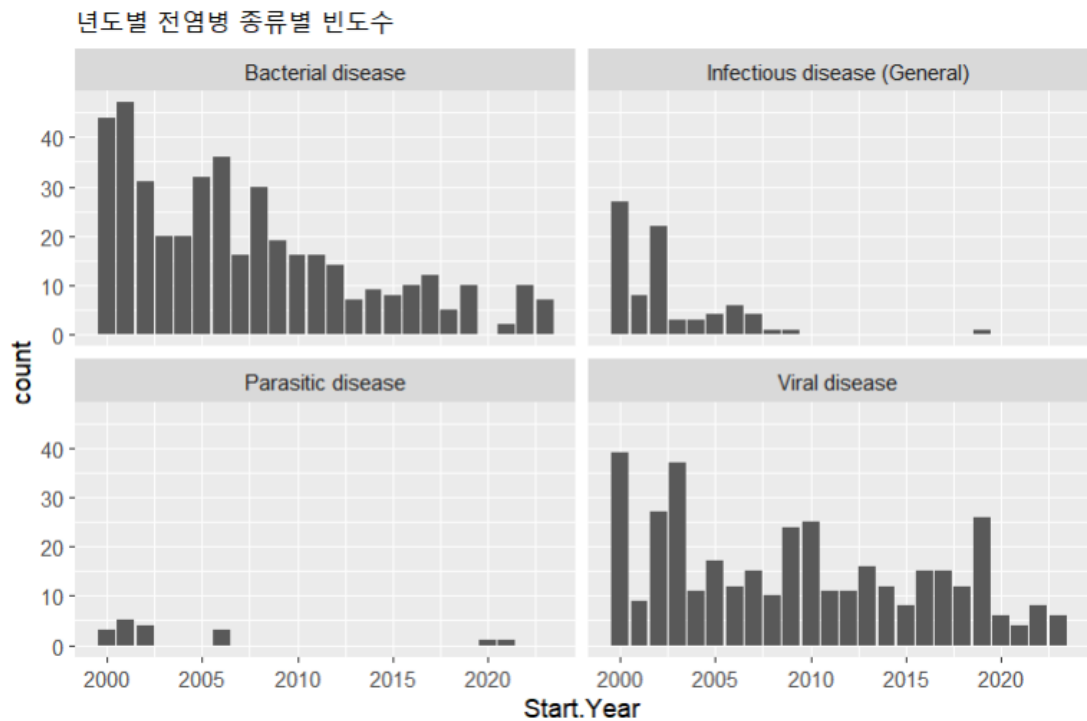
Epidemic (Parasitic disease) - 전염병 (기생충 질병)

```
data%>%
  group_by(Disaster.Group)%>%
```

```

filter(Disaster.Type == "Epidemic")%>%
ggplot(aes(x = Start.Year))+
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "년도별 전염병 종류별 빈도수")

```

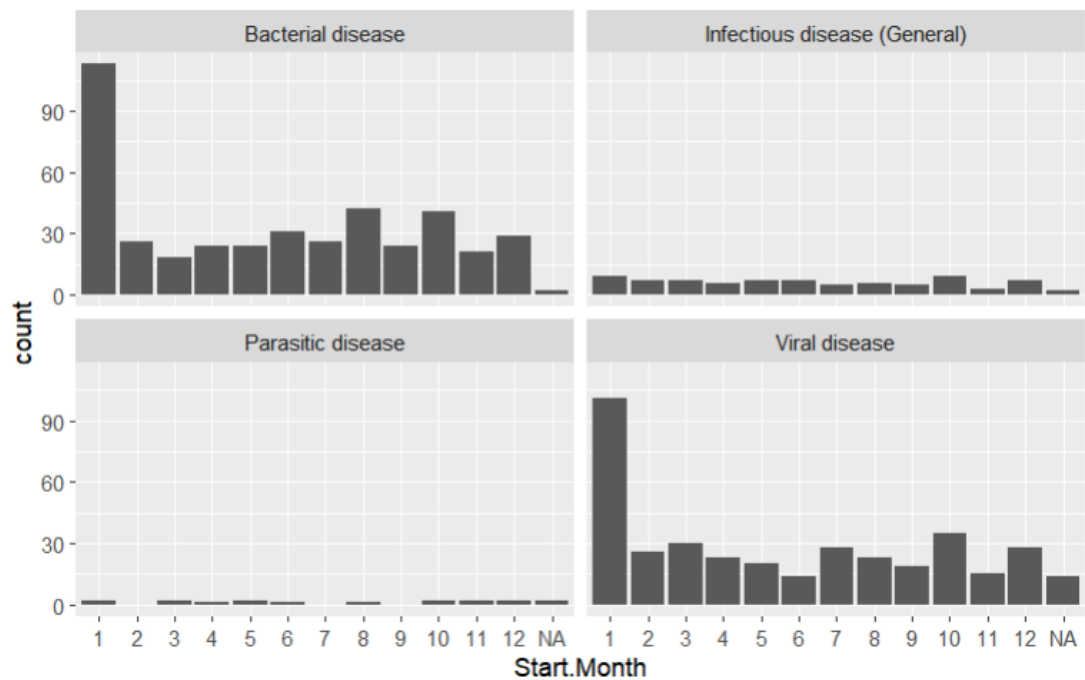


```

data%>%
group_by(Disaster.Group)%>%
filter(Disaster.Type == "Epidemic")%>%
mutate(Start.Month = factor(Start.Month))%>%
ggplot(aes(x = Start.Month))+
geom_bar()+
facet_wrap(~Disaster.Subtype)+
labs(title = "월별 전염병 종류별 빈도수")

```


월별 전염병 종류별 빈도수



▼ Infestation

	Disaster.Type	Disaster.Subtype
1	Infestation	Locust infestation
2	Infestation	Worms infestation

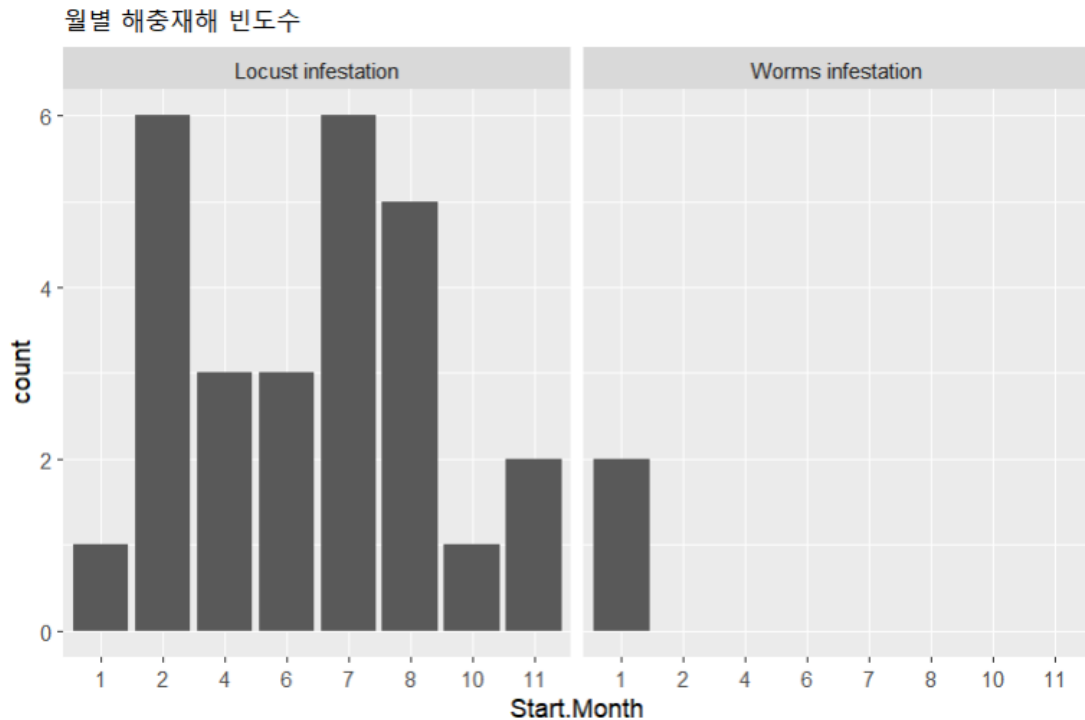
Infestation (Locust infestation) - 해충 침입 (메뚜기 침입)

Infestation (Worms infestation) - 해충 침입 (벌레 침입)

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Infestation")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "년도별 해충재해 빈도수")
```



```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Infestation")%>%
  mutate(Start.Month = factor(Start.Month))%>%
  ggplot(aes(x = Start.Month))+
  geom_bar()+
  facet_wrap(~Disaster.Subtype)+
  labs(title = "월별 해충재해 빈도수")
```



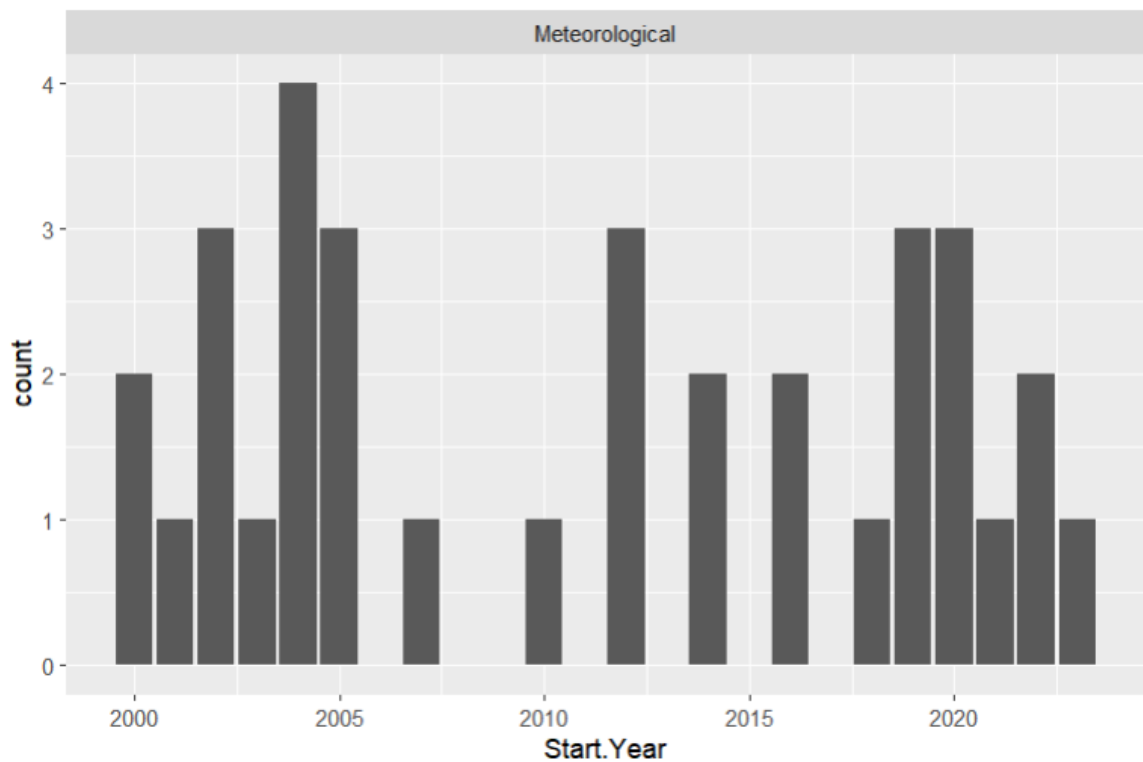
```
Disaster.Type Disaster.Subtype
1 Animal incident Animal incident
```

동물 사고의 경우 그 유형이 한가지로, type과 빈도수가 같다

▼ 한국의 년도별 기상관련 재해와 대응

초기 가설: 한국에서 발생하는 기상 재해는 주로 폭풍이 있다. 그렇다면 이 폭풍과 관련해서 우리나라의 피해는 어떤 추세를 보일까? 기술이 발전하면서 폭풍에 대한 대처가 잘 되고 있을까? 피해 규모가 점점 해마다 줄어드는지 확인해보자

```
data%>%
  group_by(Disaster.Subgroup)%>%
  filter(Disaster.Subgroup == "Meteorological")%>%
  filter(Country == "Republic of Korea") %>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  facet_wrap(~Disaster.Subgroup)
```



2000년도부터 한국의 기상관련 재해 빈도수를 체크

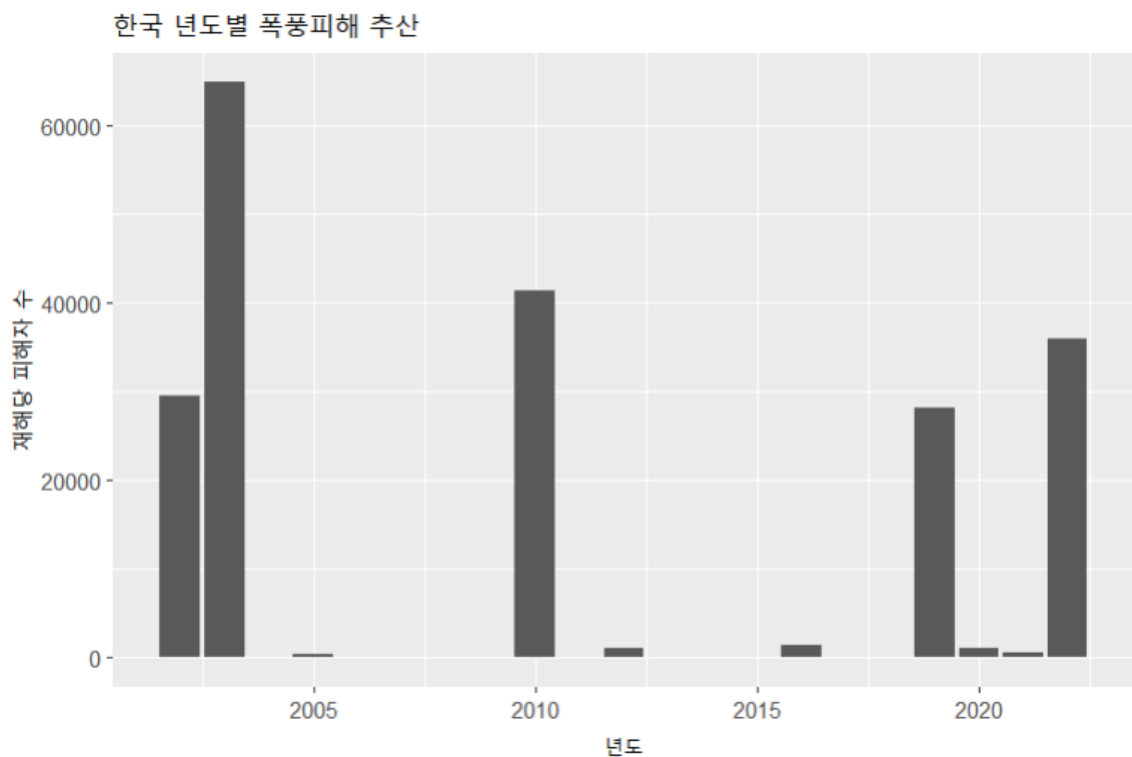
빈도는 년 평균 2회 정도 되는 것 같다. 그렇다면 년도별 회당 피해상황을 확인해보고, 기상관련 재해에 관하여 특별한 점이 있는지 체크

```
data%>%
  filter(Disaster.Type == "Extreme temperature")%>%
  filter(Country == "Republic of Korea") %>%
  group_by(Start.Year) %>%
  summarise(
    Total_Affected = sum(No..Affected, na.rm = TRUE),
    # 피해자 수 총합
    Frequency = n()
    # 재해 발생 빈도 (건수)
  )
```

	Start.Year	Total_Affected	Frequency
	<int>	<int>	<int>
1	2016	0	1
2	2018	0	1
3	2022	0	1

우선 한국의 극한 온도에 대해서는 총 3건이 존재했으며 각각에 대한 피해자는 0명이었고, 이 외 기상재해는 전부 폭풍으로 인한 피해라고 볼 수 있다.

```
data%>%
  filter(Disaster.Type == "Storm")%>%
  filter(Country == "Republic of Korea") %>%
  group_by(Start.Year) %>%
  summarise(
    Total_Affected = sum(No..Affected, na.rm = TRUE),
    # 피해자 수 총합
    Frequency = n()
    # 재해 발생 빈도 (건수)
  ) %>%
  mutate(Affected_per_Event = Total_Affected / Frequency) %>%
  filter(Affected_per_Event > 1)%>%
  ggplot(aes(x = Start.Year, y = Affected_per_Event))+
  geom_col()+
  labs(title = "한국 년도별 폭풍피해 추산", x = "년도", y = "재해당 피해자 수")
```



같은 폭풍이라고 해도 그 규모나 발생한 위치, 시간대에 따라 피해의 규모가 천차만별이므로 단순히 재해당 피해자 수로는 대처를 잘 했는지 아닌지 확인하기가 어렵다. 이는 폭풍 뿐 아니라 자연재해의 특징이 아닐까 한다.

▼ 아프리카의 전염병 발생과 전 세계 발생의 추이

아프리카의 경우 가장 낙후된 지역으로, 위생과 점염병에 취약하며 그로 인해 가장 전염병의 발생 빈도가 높을 것이라 예상된다. 또한, 최근 뉴스에 나왔던 엠폭스의 경우 그 발원지가 아프리카인것으로 밝혀져 아프리카의 세균 혹은 바이러스성 전염병이 아프리카에서 많이 발생하고 세계적으로 확산되는지 확인이 필요하다

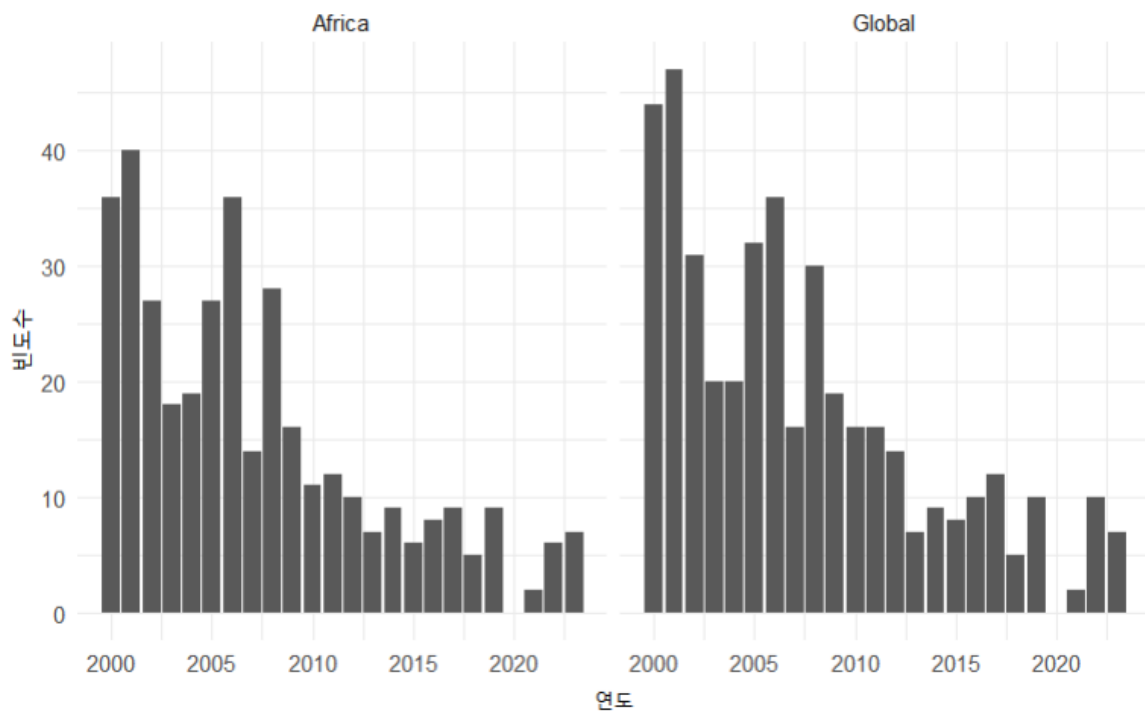
```
data_africa <- data %>%
  filter(Disaster.Subtype == "Bacterial disease") %>%
  filter(Region == "Africa") %>%
  mutate(Region = "Africa") # 지역을 "Africa"로 명시

# 두 번째 데이터에 "Global" 명시
data_global <- data %>%
  filter(Disaster.Subtype == "Bacterial disease") %>%
  mutate(Region = "Global") # 전 세계 데이터를 "Global"로
  명시

# 두 데이터를 결합
combined_data <- bind_rows(data_africa, data_global)

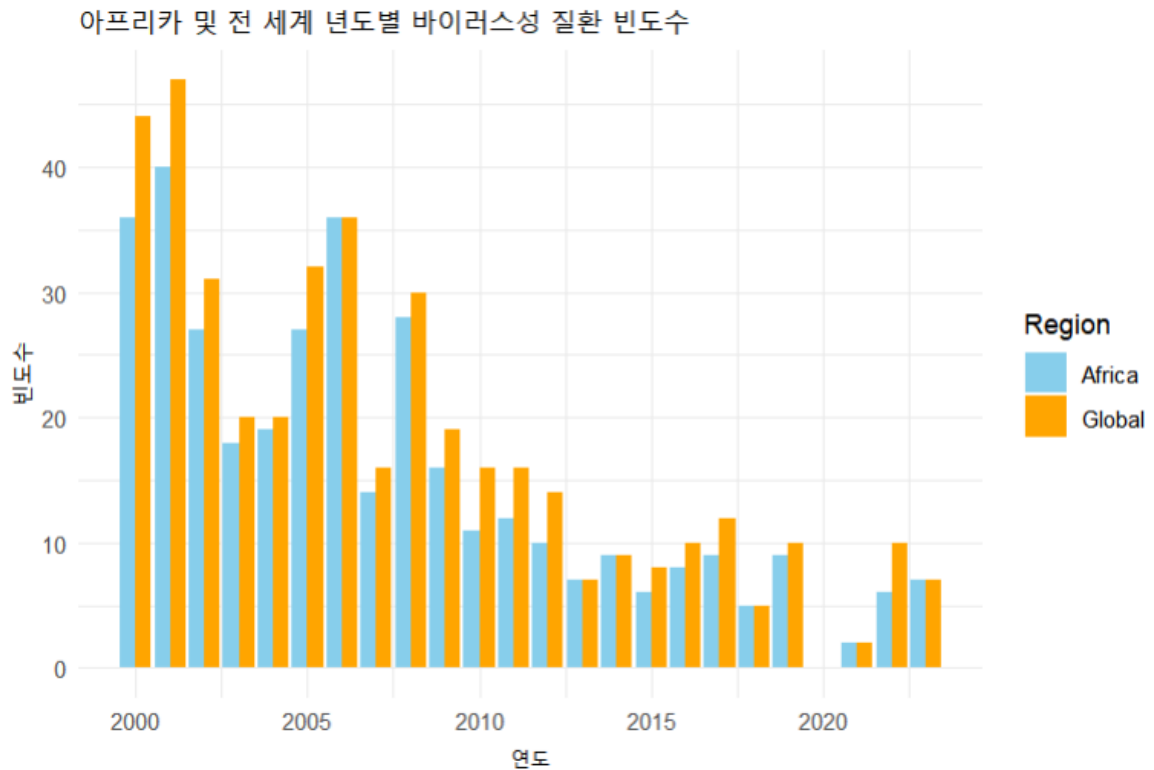
# 결합된 데이터로 시각화
combined_data %>%
  ggplot(aes(x = Start.Year)) +
  geom_bar() +
  facet_wrap(~Region) + # Region별로 시각화
  labs(title = "아프리카 및 전 세계 연도별 박테리아 질환 빈도수",
        x = "연도",
        y = "빈도수") +
  theme_minimal()
```

아프리카 및 전 세계 년도별 박테리아 질환 빈도수



조금 더 보기 편하도록 두 시각화 그래프를 합쳐보면

```
combined_data %>%
  ggplot(aes(x = Start.Year, fill = Region)) +
  geom_bar(position = "dodge") + # Region별로 나란히 표시
  labs(title = "아프리카 및 전 세계 년도별 바이러스성 질환 빈도수",
        x = "연도",
        y = "빈도수") +
  scale_fill_manual(values = c("Africa" = "skyblue", "Global" = "orange")) + # 색상 설정
  theme_minimal()
```



아프리카의 박테리아 질환 빈도수는 전 세계 빈도수의 추이와 동일하다는 것을 확인할 수 있으며 대부분의 박테리아 질환이 아프리카에서 많이 발생함을 확인할 수 있다.

그렇다면 세균성 질환은 어떨까?

```
data_africa1 <- data %>%
  filter(Disaster.Subtype == "Viral disease") %>%
  filter(Region == "Africa") %>%
  mutate(Region = "Africa") # 지역을 "Africa"로 명시

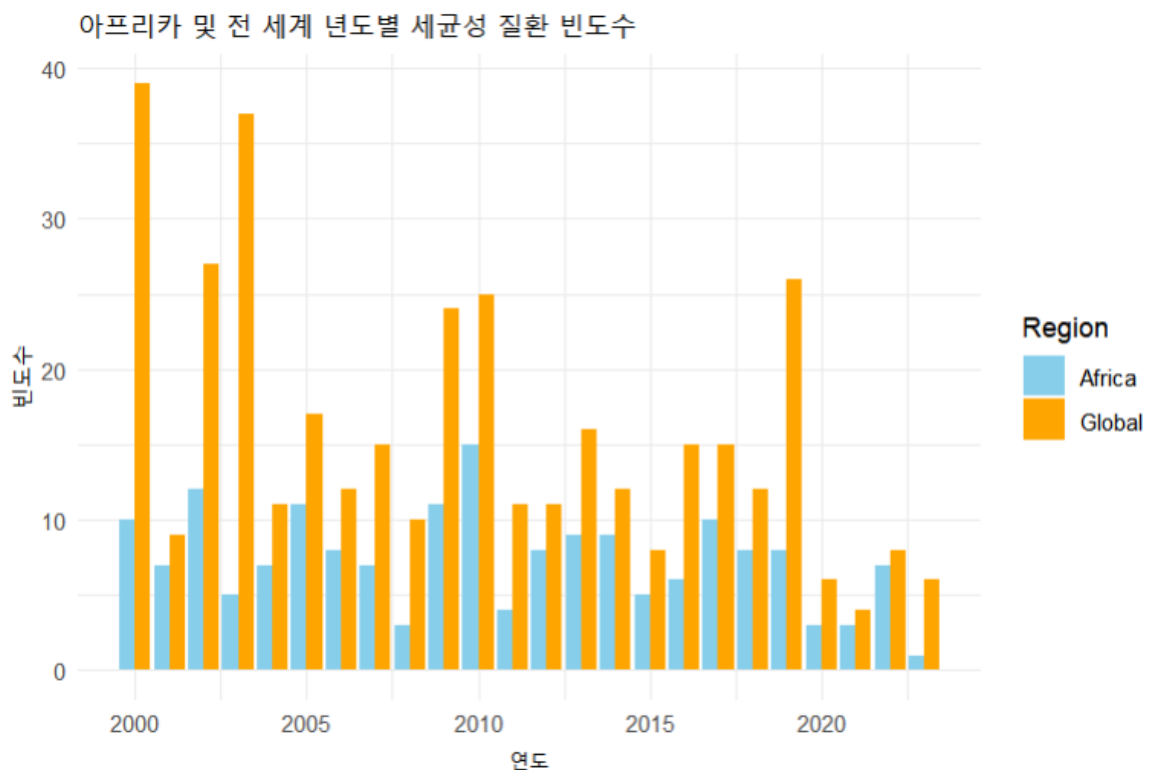
# 두 번째 데이터에 "Global" 명시
data_global1 <- data %>%
  filter(Disaster.Subtype == "Viral disease") %>%
  mutate(Region = "Global") # 전 세계 데이터를 "Global"로 명시

# 두 데이터를 결합
combined_data1 <- bind_rows(data_africa1, data_global1)

# 결합된 데이터로 시각화
combined_data1 %>%
```



```
ggplot(aes(x = Start.Year, fill = Region)) +
  geom_bar(position = "dodge") + # Region별로 나란히 표시
  labs(title = "아프리카 및 전 세계 연도별 세균성 질환 빈도수",
        x = "연도",
        y = "빈도수") +
  scale_fill_manual(values = c("Africa" = "skyblue", "Global" = "orange")) + # 색상 설정
  theme_minimal()
```



세균성 질환은 바이러스성 질환에 비해서는 전 세계적 비중보다는 적으나 그래도 큰 부분을 차지하고 있다.

다음으로는 피해자 수를 비교해보자

```
data_africa2 <- data %>%
  filter(Disaster.Subtype == "Bacterial disease") %>%
  filter(Region == "Africa") %>%
  group_by(Start.Year) %>%
  summarise(
    Total_Affected = sum(No..Affected, na.rm = TRUE)
  ) %>%
```

```

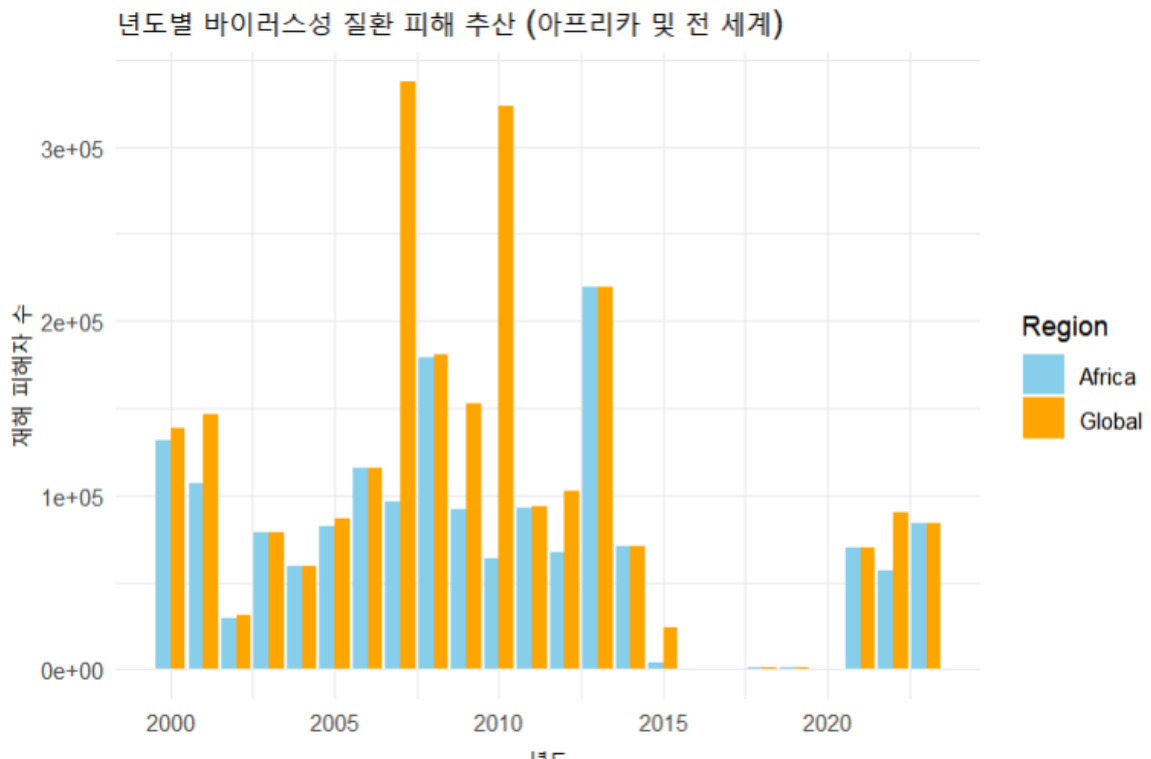
filter(Total_Affected > 1) %>%
mutate(Region = "Africa") # 지역을 "Africa"로 명시

# 두 번째 데이터(Global) 처리
data_global2 <- data %>%
  filter(Disaster.Subtype == "Bacterial disease") %>%
  group_by(Start.Year) %>%
  summarise(
    Total_Affected = sum(No..Affected, na.rm = TRUE)
  ) %>%
  filter(Total_Affected > 1) %>%
  mutate(Region = "Global") # 지역을 "Global"로 명시

# 두 데이터를 결합
combined_data2 <- bind_rows(data_africa2, data_global2)

# 결합된 데이터로 시각화
combined_data2 %>%
  ggplot(aes(x = Start.Year, y = Total_Affected, fill =
Region)) +
  geom_col(position = "dodge") + # 두 데이터를 나란히 비교
  labs(title = "년도별 세균성 질환 피해 추산 (아프리카 및 전 세
계)",
        x = "년도",
        y = "재해 피해자 수") +
  scale_fill_manual(values = c("Africa" = "skyblue", "Gl
obal" = "orange")) + # 색상 설정
  theme_minimal()

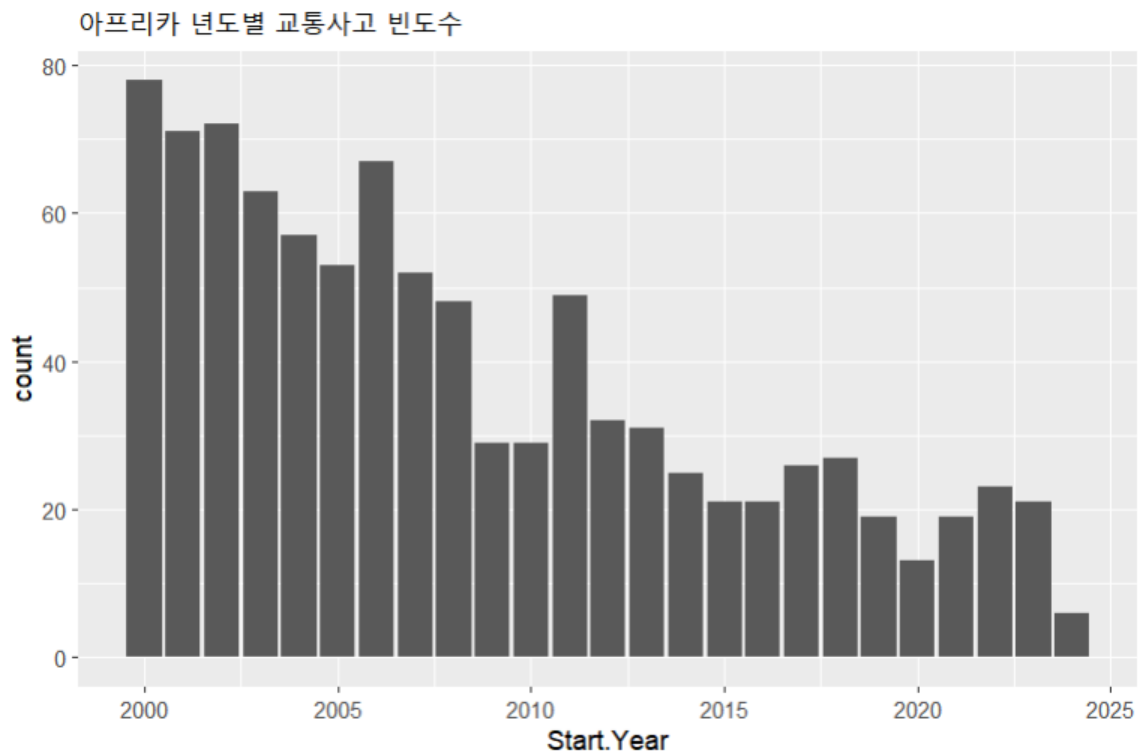
```



▼ 유럽과 아프리카 교통사고 빈도수 및 피해자 추산

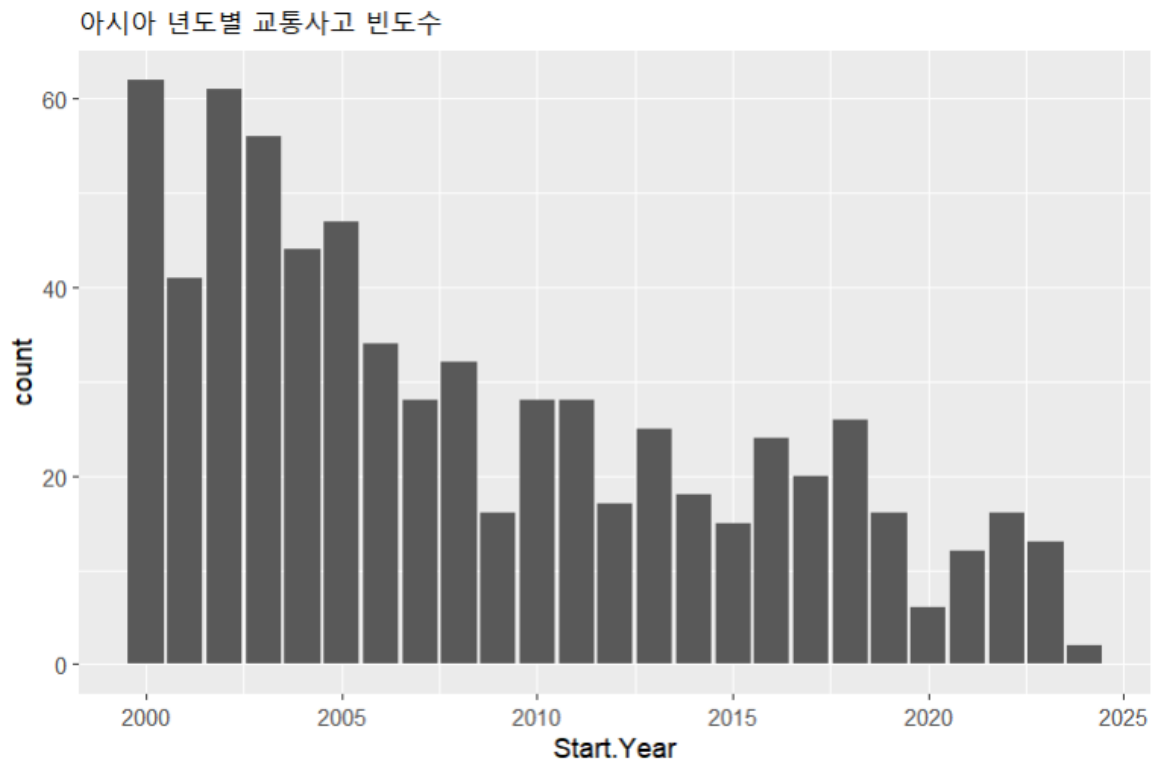
교통사고 road의 경우 유럽과 같이 선진화된 대륙은 인프라가 잘 갖추어져 사고의 빈도가 적을 것이고, 반대로 아프리카는 도로가 깔리지 않아 교통사고의 빈도가 높을 것이라 예상한다.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Road")%>%
  filter(Region == "Africa") %>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "아프리카 년도별 교통사고 빈도수")
```



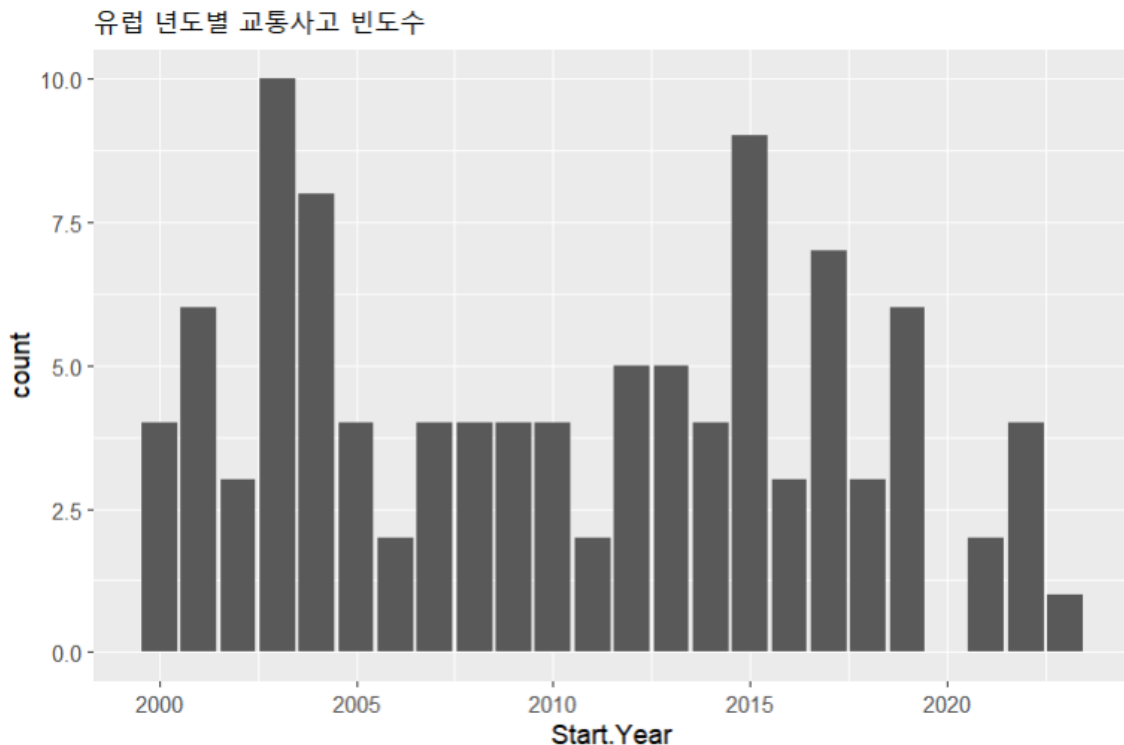
아프리카의 년도별 교통사고 빈도수는 해마다 감소하는 추세로 볼 수 있다.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Road")%>%
  filter(Region == "Asia") %>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "아시아 년도별 교통사고 빈도수")
```



아시아의 경우에도 마찬가지로 줄어드는 추세이며 아프리카와 비교했을때에도 조금 더 빈도가 적다 그렇다면 유럽은 어떨까?

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Road")%>%
  filter(Region == "Europe") %>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "유럽 년도별 교통사고 빈도수")
```



유럽은 이미 교통 인프라가 갖춰졌기에 안정화 단계에 왔다고 볼 수 있을 것 같다.
그런데 이후 피해자를 추산하는 과정에서 이상한 점을 발견했다.

```
data%>%
  filter(Disaster.Type == "Road")%>%
  filter(Region == "Asia")%>%
  select(Start.Year, Country, Disaster.Type, No..Affected)%
>%
  filter(!is.na(No..Affected))%>%
  group_by(Start.Year)
```

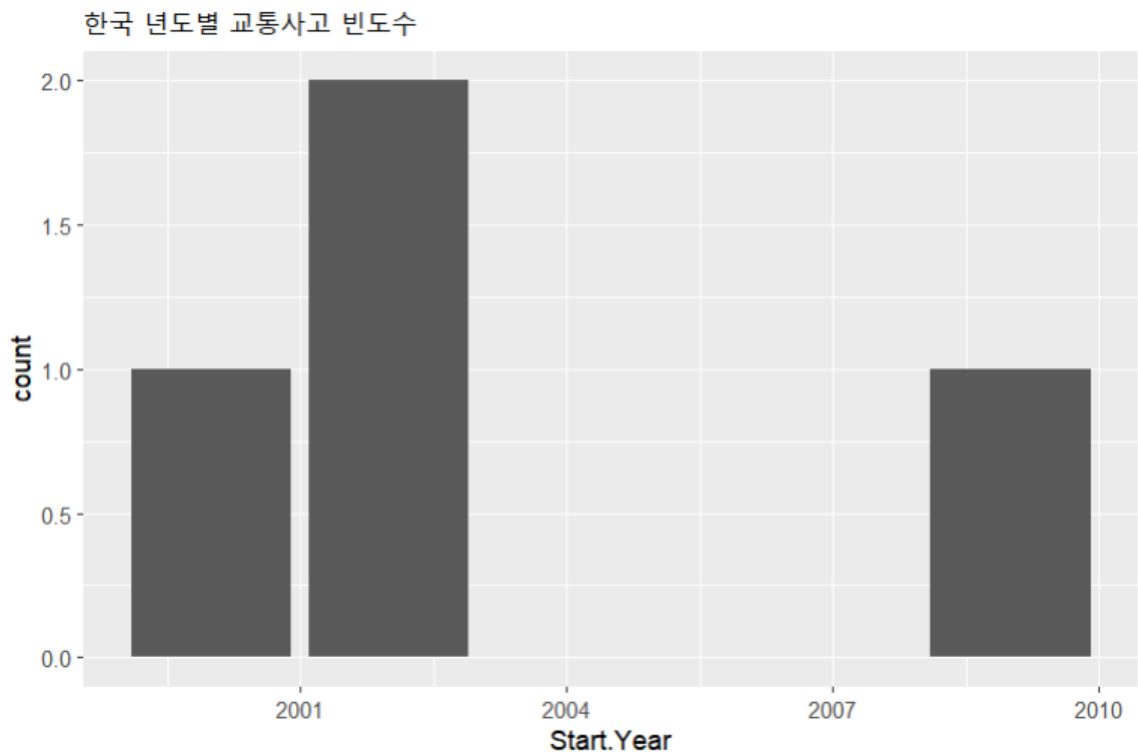
	Start.Year	Country	Disaster.Type	No..Affected
	<int>	<chr>	<chr>	<int>
1	2003	India	Road	7
2	2005	Viet Nam	Road	1
3	2005	India	Road	42
4	2005	India	Road	27
5	2006	Bangladesh	Road	10
6	2006	Nepal	Road	9
7	2006	India	Road	5

8	2006	India	Road	4
9	2010	Pakistan	Road	10
10	2010	India	Road	40
11	2010	Pakistan	Road	5
12	2011	India	Road	10
13	2011	Philippines	Road	1
14	2020	Pakistan	Road	7
15	2021	India	Road	7

아시아의 국가별 교통사고 피해자 수를 확인한 것인데, NA값을 제외하고 보면 피해자 확인이 된 것이 15건밖에 되지 않는다. 우리가 생각하는 일반적인 차량 교통사고가 아닌 것 같다.

이후 다시 우리나라의 교통사고 빈도수를 확인하니

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Type == "Road")%>%
  filter(Region == "Asia") %>%
  filter(Country == "Republic of Korea")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "한국 년도별 교통사고 빈도수")
```

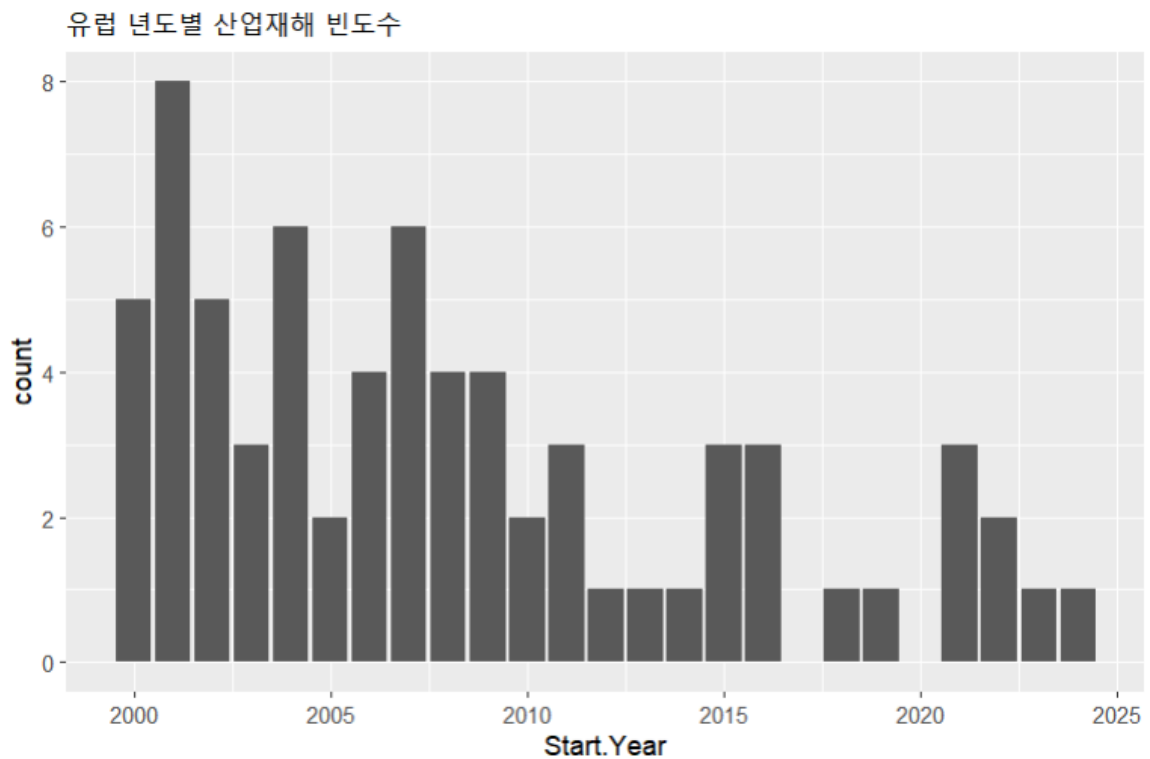


확인된 것이 총 4건으로 데이터가 담기지 않았거나, 우리가 생각하는 일반적인 교통사고가 아닌 것이라고 생각한다.

▼ 산업 재해의 빈도와 피해자

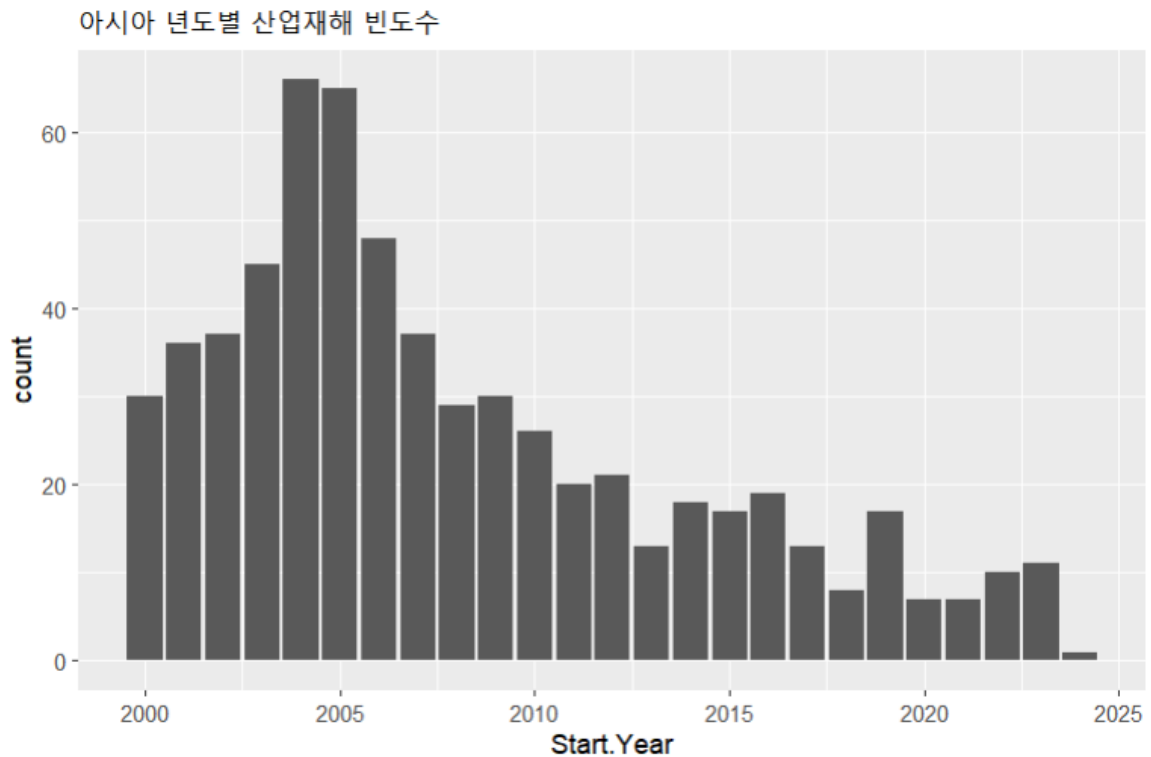
가장 선진화된 유럽의 년도별 산업재해 빈도와 피해자를 확인한 뒤, 개발 도상국이라고 할 수 있는 아시아 지역과 후진국이라고 할 수 있는 아프리카의 산업재해 빈도와 피해자를 확인해 보겠다.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  filter(Region == "Europe") %>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "유럽 년도별 산업재해 빈도수")
```

가장 먼저 유럽의 산업재해 역시 점점 감소하며 2000년대에 이미 산업재해가 상당히 적음을 확인할 수 있다.

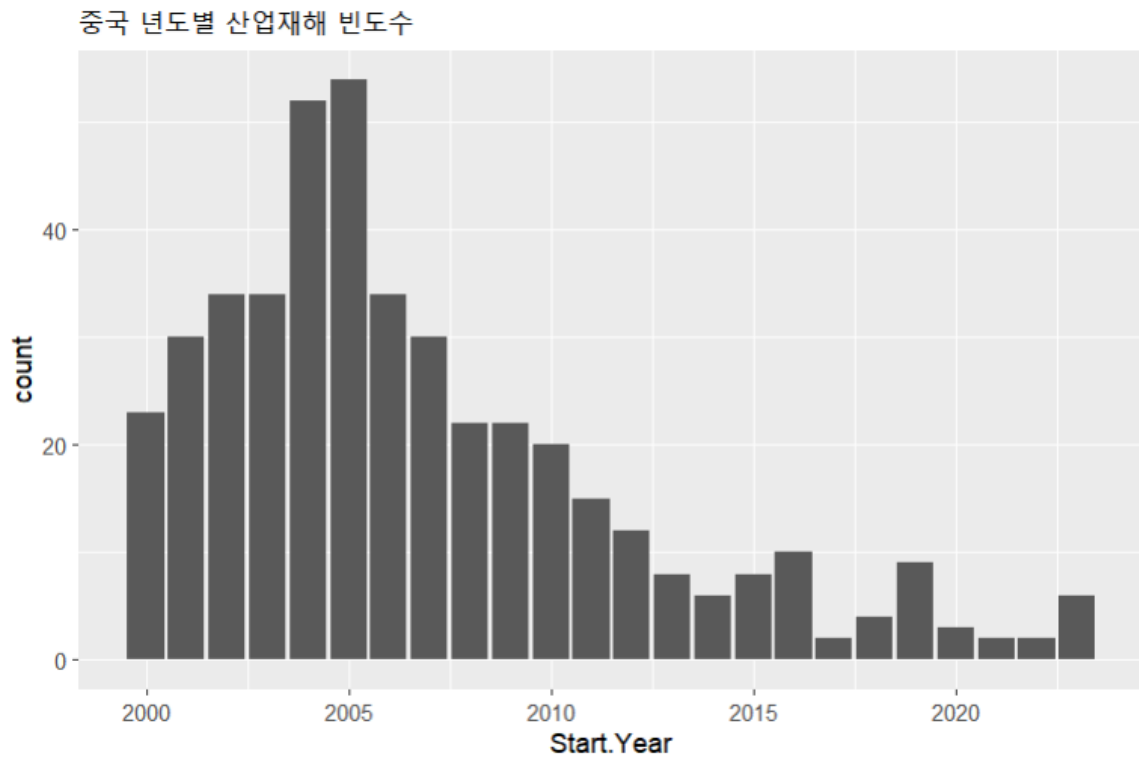
```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  filter(Region == "Asia") %>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "아시아 년도별 산업재해 빈도수")
```



다음으로 아시아의 경우 역시 점점 감소하고 있지만 2005년도에 급격히 산업재해가 올라감을 확인할 수 있다.

그중에서 중국에 대한 내용을 확인

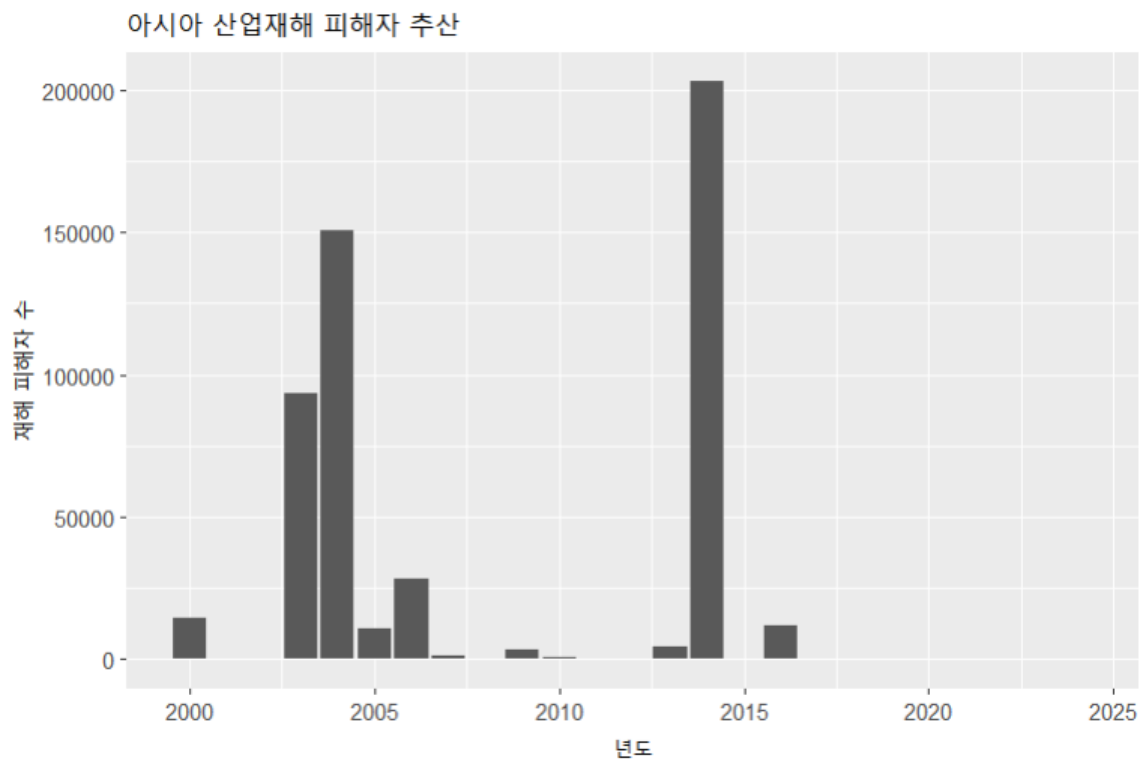
```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  filter(Region == "Asia") %>%
  filter(Country == "China")%>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "중국 년도별 산업재해 빈도수")
```



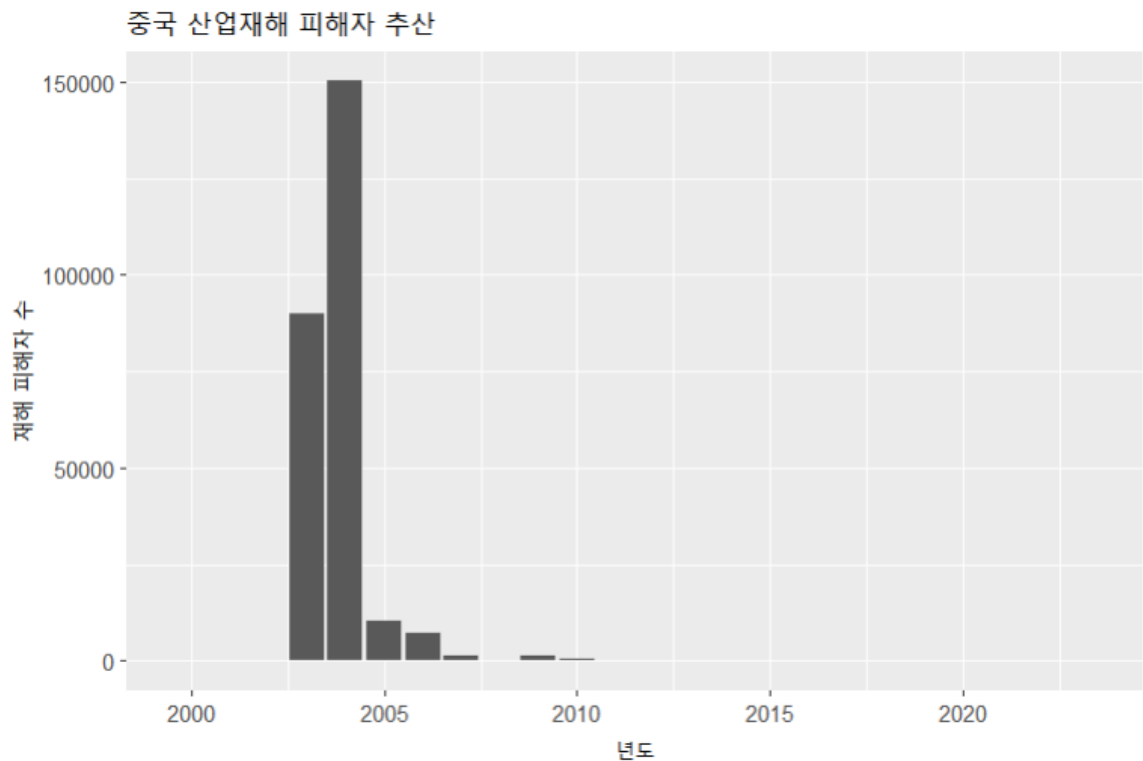
중국의 산업재해는 아시아 전체의 산업재해와 매우 비슷한 비율이다.

이번엔 산업피해자 수를 확인해본다

```
data%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  filter(Region == "Asia")%>%
  group_by(Start.Year) %>%
  summarise(
    Total_Affected = sum(No..Affected, na.rm = TRUE),
    # 피해자 수 총합
    # 재해 발생 빈도 (건수)
  ) %>%
  ggplot(aes(x = Start.Year, y = Total_Affected))+
  geom_col()+
  labs(title = "아시아 산업재해 피해자 추산", x = "년도", y =
    "재해 피해자 수")
```



```
data%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  filter(Country == "China")%>%
  group_by(Start.Year) %>%
  summarise(
    Total_Affected = sum(No..Affected, na.rm = TRUE),
    # 피해자 수 총합
    # 재해 발생 빈도 (건수)
  ) %>%
  ggplot(aes(x = Start.Year, y = Total_Affected))+
  geom_col()+
  labs(title = "중국 산업재해 피해자 추산", x = "년도", y = "재
해 피해자 수")
```



아시아와 중국의 산업재해 피해자 수를 확인해본 결과 2004년 많이 발생했던 산업재해에서는 중국의 피해자 수와 일치하고, 2013년에도 많은 피해자가 발생한 사고가 있었는데,

```
data%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  filter(Region == "Asia")%>%
  filter(Start.Year == 2014)%>%
  select(Country, No..Affected)%>%
  filter(!is.na(No..Affected))
```


	Country	No..Affected
1	T\xfcrkiye	1
2	Maldives	203000

몰디브에서 20만명이 피해를 입은 산업재해가 있었는데, 하나는 하수 처리 회사의 해수 담수화 기술에 문제가 있어 20만명 이상이 물 부족 사태를 겪은 사건이 있다.

<https://thediomat.com/2014/12/maldives-faces-drinking-water-crisis/>

Maldives Faces Drinking Water Crisis

A fire at the water sewage treatment plant in the country's capital has left at least 100,000 without drinking water.

 <https://thedi diplomat.com/2014/12/maldives-faces-drinking-water-crisis/>



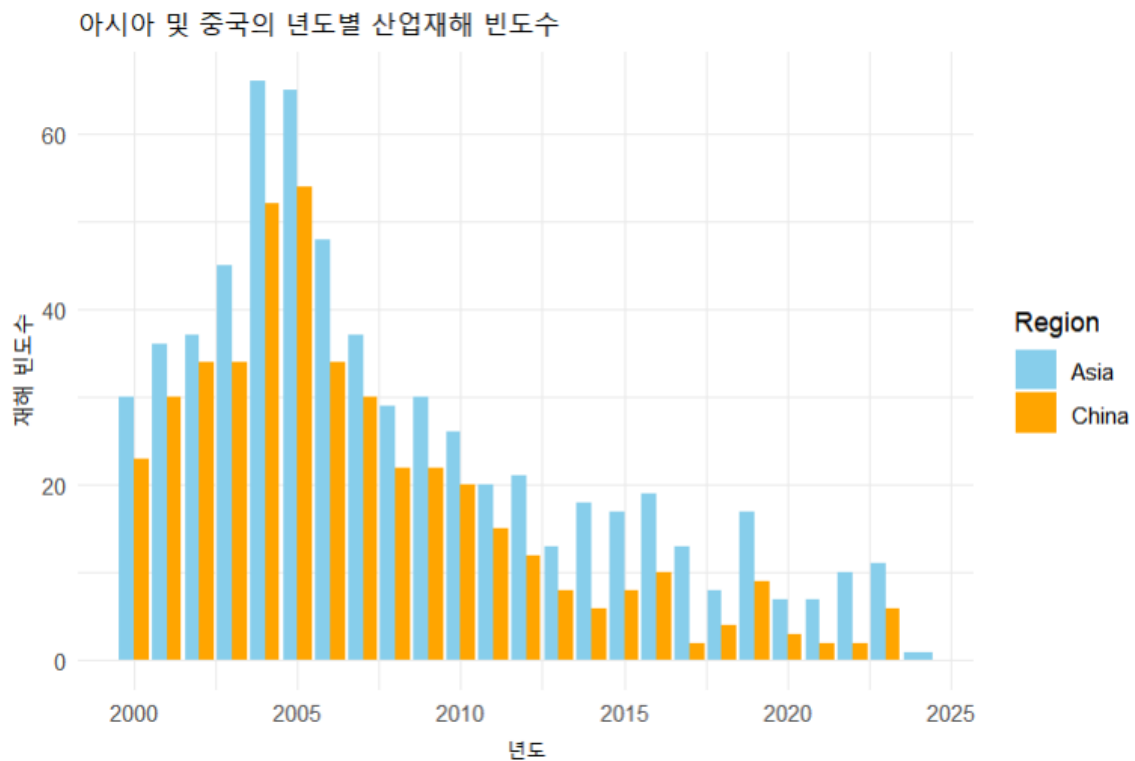
```
# 첫 번째 데이터 (Asia) 처리
data_asia4 <- data %>%
  filter(Disaster.Subgroup == "Industrial accident") %>%
  filter(Region == "Asia") %>%
  group_by(Start.Year) %>%
  summarise(Frequency = n()) %>%
  mutate(Region = "Asia") # 지역을 "Asia"로 명시

# 두 번째 데이터 (China) 처리
data_china4 <- data %>%
  filter(Disaster.Subgroup == "Industrial accident") %>%
  filter(Region == "Asia") %>%
  filter(Country == "China") %>%
  group_by(Start.Year) %>%
  summarise(Frequency = n()) %>%
  mutate(Region = "China") # 지역을 "China"로 명시

# 두 데이터를 결합
combined_data4 <- bind_rows(data_asia4, data_china4)

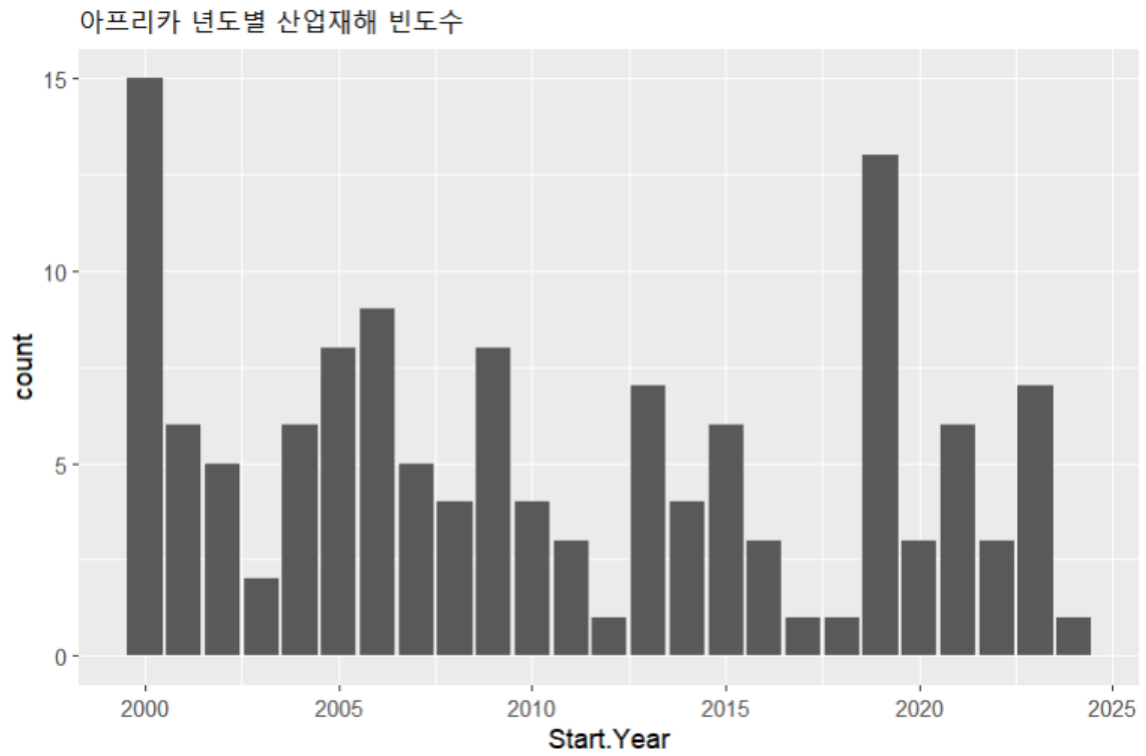
# 결합된 데이터로 시각화
combined_data4 %>%
  ggplot(aes(x = Start.Year, y = Frequency, fill = Region)) +
  geom_bar(stat = "identity", position = "dodge") + #
  # 막대 나란히 표시
  labs(title = "아시아 및 중국의 연도별 산업재해 빈도수",
        x = "연도",
        y = "재해 빈도수") +
```

```
scale_fill_manual(values = c("Asia" = "skyblue", "China" = "orange")) + # 색상 설정
theme_minimal()
```



중국의 산업재해가 아시아 전체의 산업재해의 많은 비중을 가짐을 알 수 있다.

```
data%>%
  group_by(Disaster.Group)%>%
  filter(Disaster.Subgroup == "Industrial accident")%>%
  filter(Region == "Africa") %>%
  ggplot(aes(x = Start.Year))+
  geom_bar()+
  labs(title = "아프리카 연도별 산업재해 빈도수")
```



아프리카의 경우 오히려 산업재해가 아시아보다 적음을 확인할 수 있음.