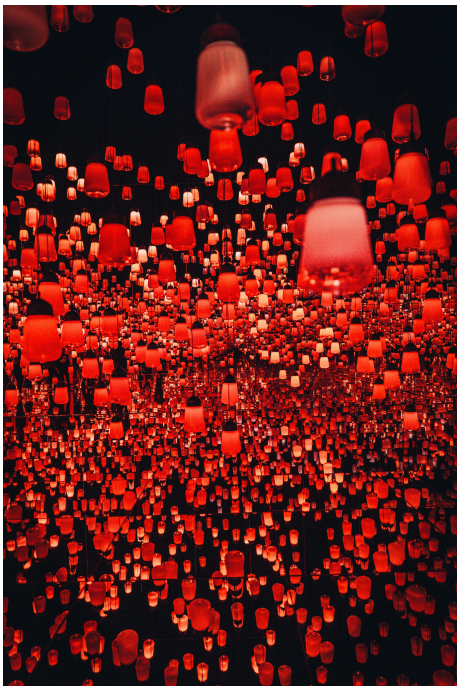


Assignment4_Group6

Chatpisut Makornkhan
Master of Business Analytics

BINGYU YANG
Master of Business Analytics

Phuong Trinh
Master of Business Analytics

Report for
Monash University

21 May 2022

Faculty of Business & Economics

(03) 9905 2478
questions@company.com

ABN: 12 377 614 630

Contents

Vietnam and New Zealand

1 Section 2

```
death <- read_csv(here::here("Data/death-rates-from-air-pollution.csv"))
```

```
## Rows: 6840 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): Entity, Code
## dbl (5): Year, Deaths - Cause: All causes - Risk: Household air pollution fr...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

2 Introduction:

In this section we will compare the different types of pollution from the period of 1990 to 2019 in two countries, Vietnam and New Zealand.

```
death_rate <- death %>%
  filter(Entity %in% c("Vietnam", "New Zealand")) %>%
  rename(Household_air_pollution = "Deaths - Cause: All causes - Risk: Household air pollution
         Ambient_particulate_matter_pollution = "Deaths - Cause: All causes - Risk: Ambient par
         Air_pollution = "Deaths - Cause: All causes - Risk: Air pollution - Sex: Both - Age: A
         Ambient_ozone_pollution = "Deaths - Cause: All causes - Risk: Ambient ozone pollution
  pivot_longer(cols = c("Household_air_pollution",
                "Ambient_particulate_matter_pollution",
                "Air_pollution",
                "Ambient_ozone_pollution"),
                names_to = "Risk_factor",
                values_to = "Rate")
```

3 Research question:

We will explore two research questions:

- Q1: Which type of pollution is the most common attribute of death over years in each of the countries?
- Q2: How did the death rates corresponding to different types of pollution change over the years and how different the trends are in both countries?

4 Analysis:

4.1 Q1:

```
death_cause <- death_rate %>%
  group_by(Entity, Risk_factor) %>%
  summarise(mean_rate = mean(Rate))
```

```
## `summarise()` has grouped output by 'Entity'. You can override using the
## `.groups` argument.
```

```
death_cause$mean_rate <- round(death_cause$mean_rate, digits = 2)
```

```
jp <- death_cause %>%
  filter(Entity == "New Zealand")


kable(jp, "latex", caption = "Death rate related to different risk factors in New Zealand", boo
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 1: *Death rate related to different risk factors in New Zealand*

Entity	Risk_factor	mean_rate
New Zealand	Air_pollution	7.11
New Zealand	Ambient_ozone_pollution	0.20
New Zealand	Ambient_particulate_matter_pollution	6.77
New Zealand	Household_air_pollution	0.14

From the table 1 we can see that air pollution contributed most to the death rates in Japan with 13.12 %, following by ambient particulate matter pollution, ambient ozone pollution and household air pollution from fossil fuels.

```
vn <- death_cause %>%
  filter(Entity == "Vietnam")


kable(vn, "latex", caption = "Death rate related to different risk factors in Vietnam", booktab
  kable_styling(latex_options = c("striped", "hold_position"))
```

Table 2: *Death rate related to different risk factors in Vietnam*

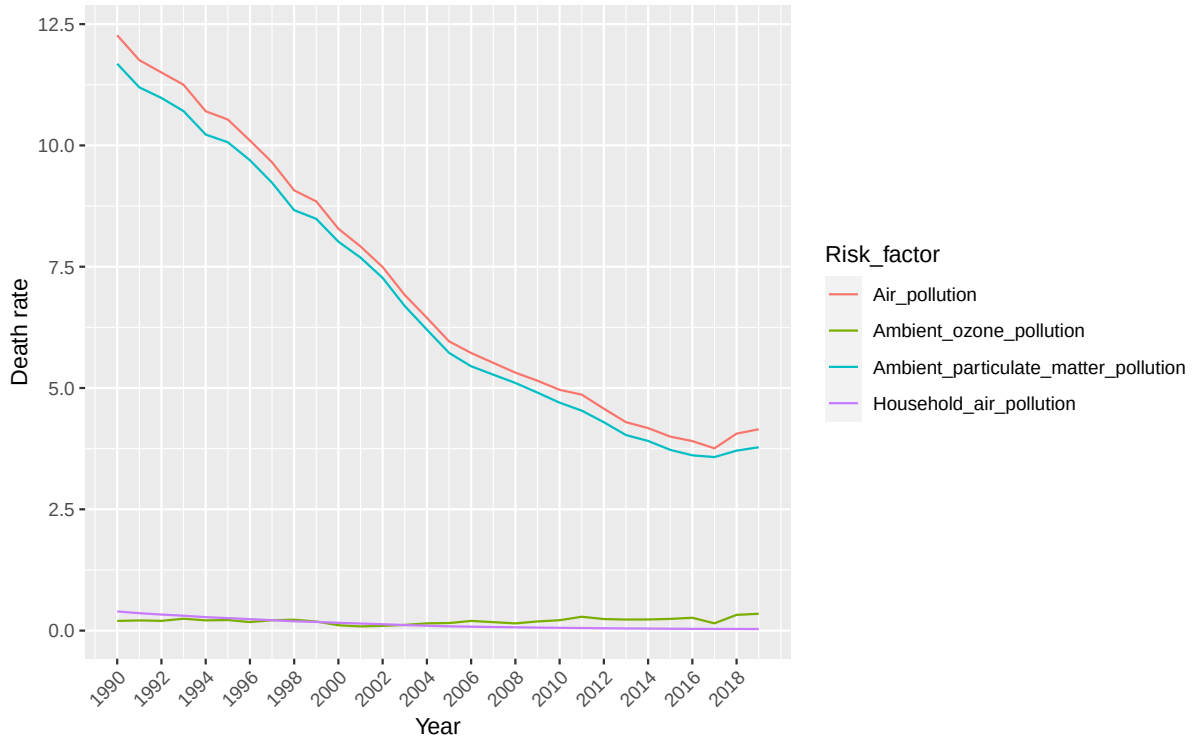
Entity	Risk_factor	mean_rate
Vietnam	Air_pollution	131.61
Vietnam	Ambient_ozone_pollution	1.74
Vietnam	Ambient_particulate_matter_pollution	38.99
Vietnam	Household_air_pollution	91.56

Similarly in table 2, we can see that air pollution is also the largest cause of pollution in Vietnam, accounting for a total of 131.61%. The second large contributor is household air pollution from fossil fuels, followed by ambient particulate matter pollution and ambient ozone pollution.

4.2 Q2:

```
jp_rate <- death_rate %>%
  filter(Entity == "New Zealand")

p1 <- ggplot(jp_rate, aes(x = Year, y = Rate, color = Risk_factor)) +
  geom_line() +
  scale_x_continuous(breaks = seq(1990, 2019, by=2)) +
  labs(y = "Death rate") +
  theme(axis.text.x = element_text(angle=45, hjust = 1))
p1
```

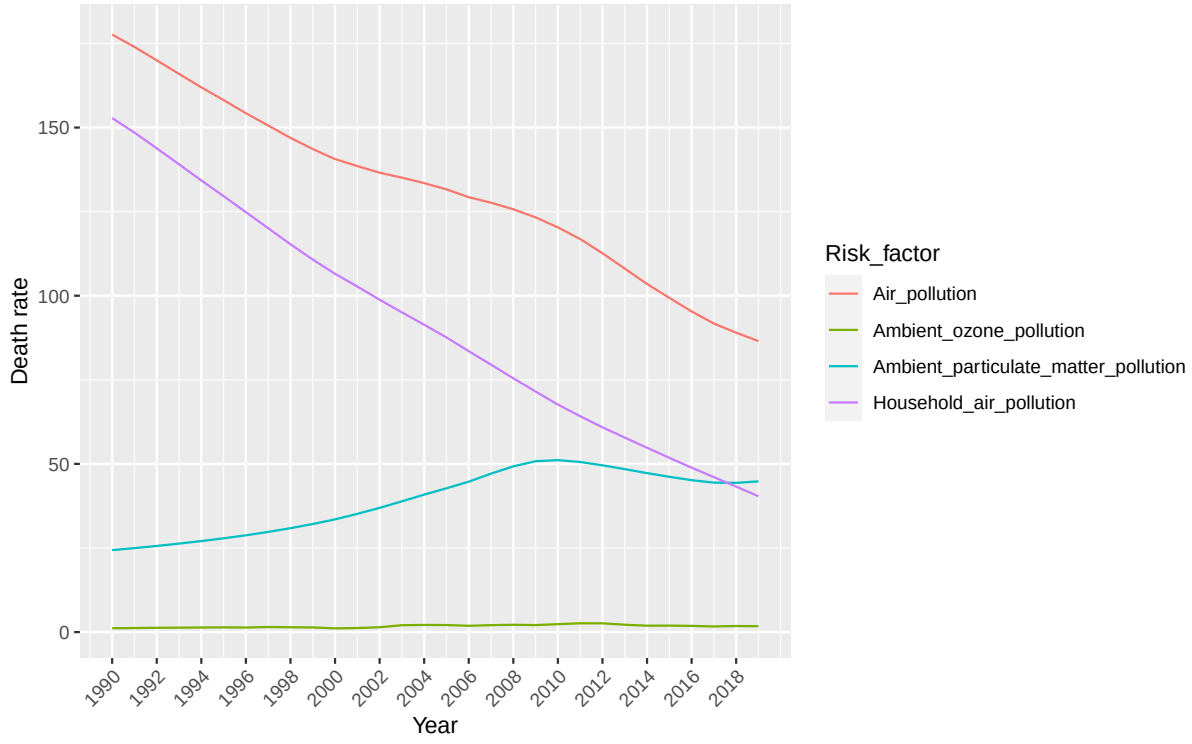


The graph **??** shows the rate at which Japanese and Vietnamese people died because of different types of pollution over a 29-year period from 1990 to 2019.

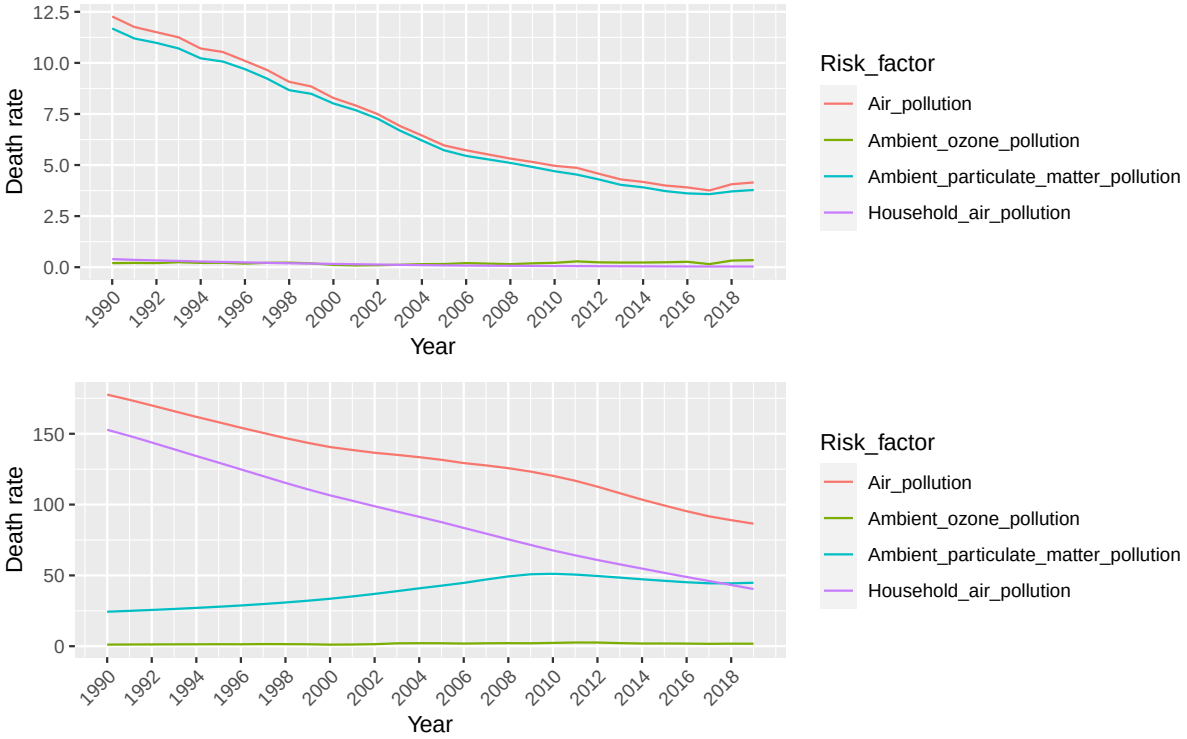
```
vn_rate <- death_rate %>%
  filter(Entity == "Vietnam")

p2 <- ggplot(vn_rate, aes(x = Year, y = Rate, color = Risk_factor)) +
```

```
  geom_line() +
  scale_x_continuous(breaks = seq(1990, 2019, by=2)) +
  labs(y = "Death rate") +
  theme(axis.text.x = element_text(angle=45, hjust = 1))
p2
```




```
gridExtra::grid.arrange(p1, p2, nrow=2)
```


Death rate
12.5
10.0
7.5
5.0
2.5
0.0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Year
Risk_factor
Air_pollution
Ambient_ozone_pollution
Ambient_particulate_matter_pollution
Household_air_pollution

Death rate
150
100
50
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Year
Risk_factor
Air_pollution
Ambient_ozone_pollution
Ambient_particulate_matter_pollution
Household_air_pollution