Canada_leading_causes_of_death

Scarlet Ruoxian Wu

Introduction

Methods

The study utilized a cleaned dataset, which included data on the leading causes of death in Canada. The data was sourced from a comprehensive database containing records from 2000 to 2022. The causes of death were categorized according to the ICD-10 classification system. The dataset was filtered to focus on the data from the year 2022, and the top nine causes of death by rank were selected for in-depth analysis.

Result

```
top_10 <-
  data |>
    filter(
      ref_date == 2022,
      rank <= 10)


top_10 |>
  kable(
      col.names = c("Year", "Cause", "Deaths", "Rank", "Years"),
      align = c("l", "r", "r", "r", "r"),
      digits = 0, booktabs = TRUE, linesep = ""
    )
```

Table 1: Top-ten causes of death in Canada in 2022

Year	Cause	Deaths	Rank	Years
2022	Malignant neoplasms [C00-C97]	82412	1	23
2022	Major cardiovascular diseas...	76639	2	23
2022	Diseases of heart [I00-I09,...	57357	3	23
2022	Ischaemic heart diseases [I...	34830	4	23
2022	Dementia [F010-F019, F03]	25994	5	6
2022	Unspecified dementia [F03]	23896	6	6
2022	Other forms of chronic isch...	20126	7	23
2022	COVID-19 [U07.1, U07.2, U10.9]	19716	8	3
2022	Malignant neoplasms of trac...	19151	9	23
2022	Other heart diseases [I26-I51]	18913	10	23

```r
top_nine <-
  data |>
  filter(
    ref_date == 2022,
        ) |>
      slice_max(order_by = desc(rank), n = 9) |>
      pull(cause_of_death_icd_10)


top_9 <-
  data |>
  filter(cause_of_death_icd_10 %in% top_nine)

short_names <-
  c("Malignant neoplasms [C00-C97]" = "Malignant Neoplasms",
    "Diseases of heart [I00-I09,..." = "Diseases of Heart",
    "Malignant neoplasms of trac..." = "Respiratory Malignant Neoplasms",
    "Dementia [F010-F019, F03]" = "Dementia",
    "COVID-19 [U07.1, U07.2, U10.9]" = "COVID-19",
    "Major cardiovascular diseas..." = "Major Cardiovascular Diseases",
    "Ischaemic heart diseases [I..." = "Ischaemic Heart Diseases",
    "Unspecified dementia [F03]" = "Unspecified Dementia",
    "Other forms of chronic isch..." = "Other Chronic Ischaemic Heart Diseases"
  )

top_9 <- top_9 %>%
  mutate(cause_of_death_icd_10 = recode(cause_of_death_icd_10, !!!short_names))
```

Table 2: Summary statistics of the number of yearly deaths, by cause, in Canada

	Min	Mean	Max	SD	Var	N
value	14 466	42 831	82 822	22 480	505 344 387	153

```
top_9 |>
  ggplot(aes(x = ref_date, y = value, color = cause_of_death_icd_10)) +
  geom_line() +
  theme_minimal() +
  scale_color_brewer(palette = "Set1") +
  labs(x = "Year", y = "Annual number of deaths in Canada") +
  facet_wrap(~cause_of_death_icd_10, dir = "v", ncol = 1, scales = "free_y") +
  theme(legend.position = "none") +
  scale_y_continuous(
    labels = scales::comma,  # This formats the labels with commas for thousands
    breaks = scales::pretty_breaks(n = 5)  # This creates 5 evenly spaced breaks
  )

ggsave("leading_ning_plot.png", plot = last_plot(), width = 10, height = 8, units = "in")
```

```
poisson_summary <- summary(cause_of_death_poisson)
neg_binomial_summary <- summary(cause_of_death_neg_binomial)


tidy_poisson <- tidy(cause_of_death_poisson)
tidy_neg_binomial <- tidy(cause_of_death_neg_binomial)

combined_summary <- bind_rows(
  mutate(tidy_poisson, model = "Poisson"),
  mutate(tidy_neg_binomial, model = "Negative Binomial")
)

coef_short_names <-
  c("cause_of_death_icd_10Malignant Neoplasms"
    = "Malignant Neoplasms",
    "cause_of_death_icd_10Diseases of Heart"
    = "Diseases of Heart",
    "cause_of_death_icd_10Respiratory Malignant Neoplasms"
    = "Respiratory Malignant Neoplasms",
    "cause_of_death_icd_10Dementia [F010-F019, F03]"
```
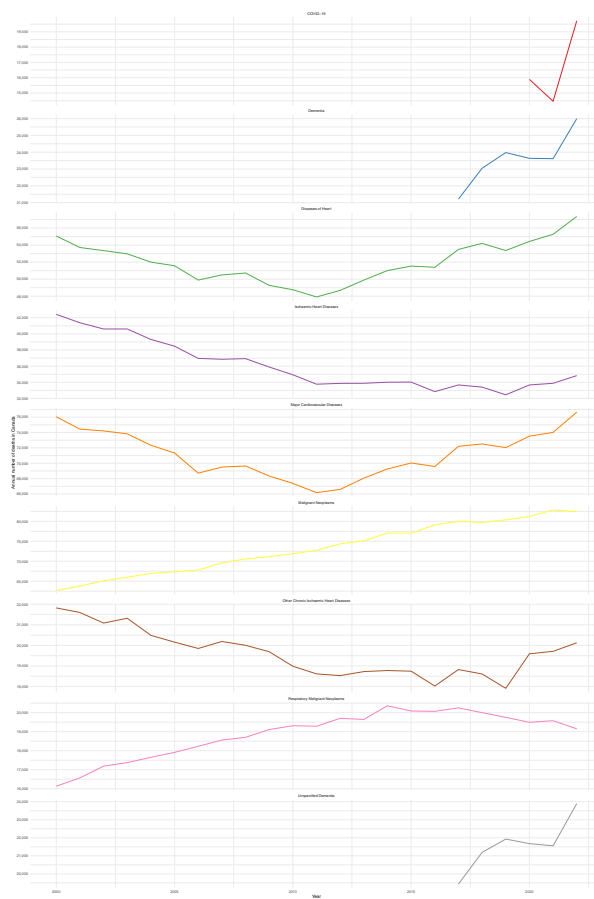
Figure 1

```
      = "Dementia",
      "cause_of_death_icd_10COVID-19"
      = "COVID-19",
      "cause_of_death_icd_10Major Cardiovascular Diseases"
      = "Major Cardiovascular Diseases",
      "cause_of_death_icd_10Ischaemic Heart Diseases"
      = "Ischaemic Heart Diseases",
      "cause_of_death_icd_10Unspecified Dementia"
      = "Unspecified Dementia",
      "cause_of_death_icd_10Other Chronic Ischchaemic Heart Diseases"
      = "Other Chronic Ischchaemic Heart Diseases"
      )

combined_summary$term <-
  ifelse(combined_summary$term %in% names(coef_short_names),
    coef_short_names[combined_summary$term],
    combined_summary$term)

models_list <- list(
  Poisson = cause_of_death_poisson,
  `Negative Binomial` = cause_of_death_neg_binomial
)


modelsummary(models_list, coef_map = coef_short_names)


pp_check(cause_of_death_poisson) +
  theme(legend.position = "bottom")

pp_check(cause_of_death_neg_binomial) +
  theme(legend.position = "bottom")


poisson <- loo(cause_of_death_poisson, cores = 2)
neg_binomial <- loo(cause_of_death_neg_binomial, cores = 2)

loo_compare(poisson, neg_binomial)


                          elpd_diff se_diff
cause_of_death_neg_binomial      0.0       0.0
cause_of_death_poisson      -14125.2    1748.0
```

Table 3: Modeling the most prevalent cause of deaths in Canada, 2001-2020

	Poisson	Negative Binomial
Malignant Neoplasms	1.479	1.472
		(0.082)
Diseases of Heart	1.137	1.129
		(0.083)
Respiratory Malignant Neoplasms	0.123	0.116
		(0.082)
Major Cardiovascular Diseases	1.450	1.443
		(0.085)
Ischaemic Heart Diseases	0.770	0.761
		(0.083)
Unspecified Dementia	0.259	0.253
		(0.096)
Num.Obs.	153	153
Log.Lik.	$-14\,947.833$	-1463.624
ELPD	$-15\,593.5$	-1468.3
ELPD s.e.	1752.0	7.1
LOOIC	$31\,186.9$	2936.6
LOOIC s.e.	3504.0	14.3
WAIC	$31\,838.1$	2936.3
RMSE	3153.30	3153.84

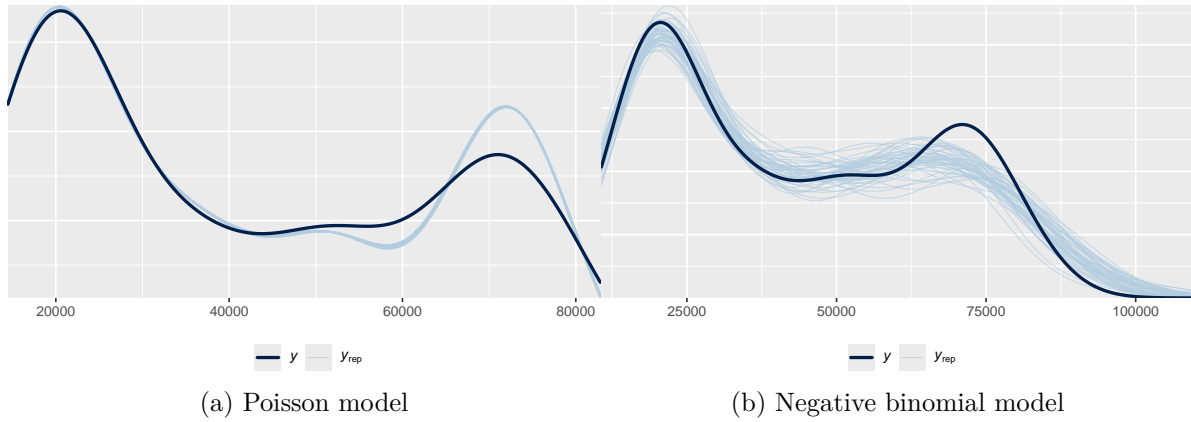(a) Poisson model (b) Negative binomial model

Figure 2: Comparing posterior prediction checks for Poisson and negative binomial models