

Canada_leading_causes_of_death

Scarlet Ruoxian Wu

```
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 diseases...	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

```
top_nine <-  
  data |>
```

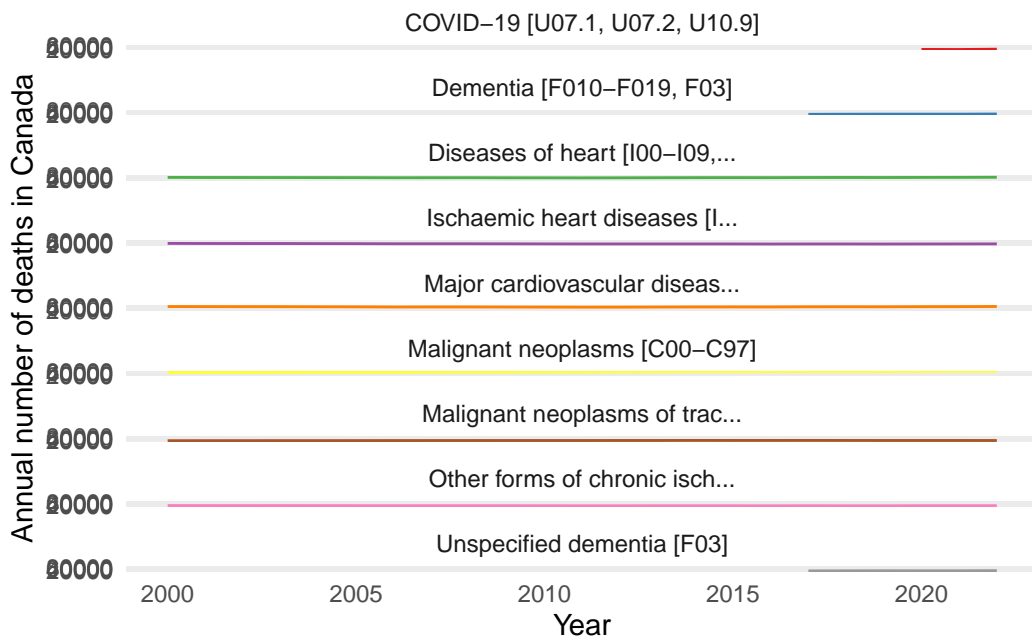
```

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)

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) +
  theme(legend.position = "none")

```



```

cause_of_death_poisson <-
  stan_glm(

```

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

```

value ~ cause_of_death_icd_10,
data = top_9,
family = poisson(link = "log"),
seed = 853
)

```

SAMPLING FOR MODEL 'count' NOW (CHAIN 1).

Chain 1:

Chain 1: Gradient evaluation took 0.017548 seconds

Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 175.48 seconds.

Chain 1: Adjust your expectations accordingly!

Chain 1:

Chain 1:

Chain 1: Iteration: 1 / 2000 [0%] (Warmup)

Chain 1: Iteration: 200 / 2000 [10%] (Warmup)

Chain 1: Iteration: 400 / 2000 [20%] (Warmup)

Chain 1: Iteration: 600 / 2000 [30%] (Warmup)

Chain 1: Iteration: 800 / 2000 [40%] (Warmup)

Chain 1: Iteration: 1000 / 2000 [50%] (Warmup)

Chain 1: Iteration: 1001 / 2000 [50%] (Sampling)

Chain 1: Iteration: 1200 / 2000 [60%] (Sampling)

Chain 1: Iteration: 1400 / 2000 [70%] (Sampling)

Chain 1: Iteration: 1600 / 2000 [80%] (Sampling)

Chain 1: Iteration: 1800 / 2000 [90%] (Sampling)

Chain 1: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 1:

Chain 1: Elapsed Time: 0.209 seconds (Warm-up)

Chain 1: 0.235 seconds (Sampling)

Chain 1: 0.444 seconds (Total)

Chain 1:

SAMPLING FOR MODEL 'count' NOW (CHAIN 2).

Chain 2:

Chain 2: Gradient evaluation took 1.4e-05 seconds

Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.14 seconds.

Chain 2: Adjust your expectations accordingly!

Chain 2:

Chain 2:

Chain 2: Iteration: 1 / 2000 [0%] (Warmup)

Chain 2: Iteration: 200 / 2000 [10%] (Warmup)

Chain 2: Iteration: 400 / 2000 [20%] (Warmup)

Chain 2: Iteration: 600 / 2000 [30%] (Warmup)

Chain 2: Iteration: 800 / 2000 [40%] (Warmup)

Chain 2: Iteration: 1000 / 2000 [50%] (Warmup)

Chain 2: Iteration: 1001 / 2000 [50%] (Sampling)

Chain 2: Iteration: 1200 / 2000 [60%] (Sampling)

Chain 2: Iteration: 1400 / 2000 [70%] (Sampling)

Chain 2: Iteration: 1600 / 2000 [80%] (Sampling)

Chain 2: Iteration: 1800 / 2000 [90%] (Sampling)

Chain 2: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 2:

Chain 2: Elapsed Time: 0.18 seconds (Warm-up)

Chain 2: 0.19 seconds (Sampling)

Chain 2: 0.37 seconds (Total)

Chain 2:

SAMPLING FOR MODEL 'count' NOW (CHAIN 3).

Chain 3:

Chain 3: Gradient evaluation took 1.3e-05 seconds

Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.13 seconds.

Chain 3: Adjust your expectations accordingly!

Chain 3:

Chain 3:

Chain 3: Iteration: 1 / 2000 [0%] (Warmup)

Chain 3: Iteration: 200 / 2000 [10%] (Warmup)

Chain 3: Iteration: 400 / 2000 [20%] (Warmup)

Chain 3: Iteration: 600 / 2000 [30%] (Warmup)

Chain 3: Iteration: 800 / 2000 [40%] (Warmup)

Chain 3: Iteration: 1000 / 2000 [50%] (Warmup)

Chain 3: Iteration: 1001 / 2000 [50%] (Sampling)

Chain 3: Iteration: 1200 / 2000 [60%] (Sampling)

Chain 3: Iteration: 1400 / 2000 [70%] (Sampling)

Chain 3: Iteration: 1600 / 2000 [80%] (Sampling)

Chain 3: Iteration: 1800 / 2000 [90%] (Sampling)

Chain 3: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 3:

Chain 3: Elapsed Time: 0.189 seconds (Warm-up)

Chain 3: 0.199 seconds (Sampling)

Chain 3: 0.388 seconds (Total)

Chain 3:

SAMPLING FOR MODEL 'count' NOW (CHAIN 4).

Chain 4:

Chain 4: Gradient evaluation took 1.5e-05 seconds

Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.15 seconds.

Chain 4: Adjust your expectations accordingly!

Chain 4:

Chain 4:

Chain 4: Iteration: 1 / 2000 [0%] (Warmup)

Chain 4: Iteration: 200 / 2000 [10%] (Warmup)

Chain 4: Iteration: 400 / 2000 [20%] (Warmup)

Chain 4: Iteration: 600 / 2000 [30%] (Warmup)

Chain 4: Iteration: 800 / 2000 [40%] (Warmup)

Chain 4: Iteration: 1000 / 2000 [50%] (Warmup)

Chain 4: Iteration: 1001 / 2000 [50%] (Sampling)

Chain 4: Iteration: 1200 / 2000 [60%] (Sampling)

Chain 4: Iteration: 1400 / 2000 [70%] (Sampling)

Chain 4: Iteration: 1600 / 2000 [80%] (Sampling)

Chain 4: Iteration: 1800 / 2000 [90%] (Sampling)

Chain 4: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 4:

Chain 4: Elapsed Time: 0.194 seconds (Warm-up)

Chain 4: 0.226 seconds (Sampling)

Chain 4: 0.42 seconds (Total)

Chain 4:

```
cause_of_death_neg_binomial <-  
  stan_glm(  
    value ~ cause_of_death_icd_10,  
    data = top_9,  
    family = neg_binomial_2(link = "log"),  
    seed = 853  
  )
```

SAMPLING FOR MODEL 'count' NOW (CHAIN 1).

Chain 1:

Chain 1: Gradient evaluation took 4e-05 seconds

Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.4 seconds.

Chain 1: Adjust your expectations accordingly!

```

Chain 1:
Chain 1:
Chain 1: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 1: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 1: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 1: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 1: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 1: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 1: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 1: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 1: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 1: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 1: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 1: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 1:
Chain 1: Elapsed Time: 0.312 seconds (Warm-up)
Chain 1:                  0.292 seconds (Sampling)
Chain 1:                  0.604 seconds (Total)
Chain 1:

```

SAMPLING FOR MODEL 'count' NOW (CHAIN 2).

```

Chain 2:
Chain 2: Gradient evaluation took 2e-05 seconds
Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.2 seconds.
Chain 2: Adjust your expectations accordingly!
Chain 2:
Chain 2:
Chain 2: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 2: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 2: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 2: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 2: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 2: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 2: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 2: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 2: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 2: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 2: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 2: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 2:
Chain 2: Elapsed Time: 0.255 seconds (Warm-up)
Chain 2:                  0.306 seconds (Sampling)
Chain 2:                  0.561 seconds (Total)

```

Chain 2:

SAMPLING FOR MODEL 'count' NOW (CHAIN 3).

Chain 3:

Chain 3: Gradient evaluation took 2e-05 seconds

Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.2 seconds.

Chain 3: Adjust your expectations accordingly!

Chain 3:

Chain 3:

Chain 3: Iteration: 1 / 2000 [0%] (Warmup)

Chain 3: Iteration: 200 / 2000 [10%] (Warmup)

Chain 3: Iteration: 400 / 2000 [20%] (Warmup)

Chain 3: Iteration: 600 / 2000 [30%] (Warmup)

Chain 3: Iteration: 800 / 2000 [40%] (Warmup)

Chain 3: Iteration: 1000 / 2000 [50%] (Warmup)

Chain 3: Iteration: 1001 / 2000 [50%] (Sampling)

Chain 3: Iteration: 1200 / 2000 [60%] (Sampling)

Chain 3: Iteration: 1400 / 2000 [70%] (Sampling)

Chain 3: Iteration: 1600 / 2000 [80%] (Sampling)

Chain 3: Iteration: 1800 / 2000 [90%] (Sampling)

Chain 3: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 3:

Chain 3: Elapsed Time: 0.274 seconds (Warm-up)

Chain 3: 0.289 seconds (Sampling)

Chain 3: 0.563 seconds (Total)

Chain 3:

SAMPLING FOR MODEL 'count' NOW (CHAIN 4).

Chain 4:

Chain 4: Gradient evaluation took 2e-05 seconds

Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.2 seconds.

Chain 4: Adjust your expectations accordingly!

Chain 4:

Chain 4:

Chain 4: Iteration: 1 / 2000 [0%] (Warmup)

Chain 4: Iteration: 200 / 2000 [10%] (Warmup)

Chain 4: Iteration: 400 / 2000 [20%] (Warmup)

Chain 4: Iteration: 600 / 2000 [30%] (Warmup)

Chain 4: Iteration: 800 / 2000 [40%] (Warmup)

Chain 4: Iteration: 1000 / 2000 [50%] (Warmup)

Chain 4: Iteration: 1001 / 2000 [50%] (Sampling)

Chain 4: Iteration: 1200 / 2000 [60%] (Sampling)

Chain 4: Iteration: 1400 / 2000 [70%] (Sampling)

```
Chain 4: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 4: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 4: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 4:
Chain 4: Elapsed Time: 0.282 seconds (Warm-up)
Chain 4:           0.418 seconds (Sampling)
Chain 4:           0.7 seconds (Total)
Chain 4:
```

```
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 [C00-C97]"
    = "malignant neoplasms",
    "cause_of_death_icd_10Diseases of heart [I00-I09,...]"
    = "Diseases of heart",
    "cause_of_death_icd_10Malignant neoplasms of trac..."
    = "Respiratory malignant neoplasms",
    "cause_of_death_icd_10Dementia [F010-F019, F03]"
    = "Dementia",
    "cause_of_death_icd_10COVID-19 [U07.1, U07.2, U10.9]"
    = "COVID-19",
    "cause_of_death_icd_10Major cardiovascular diseas..."
    = "Major cardiovascular diseases",

    "cause_of_death_icd_10Ischaemic heart diseases [I..."
    = "Ischaemic heart diseases",

    "cause_of_death_icd_10Unspecified dementia [F03]"
    = "Unspecified dementia",

    "cause_of_death_icd_10Other forms of chronic isch..."
```



```

    = "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)

```

```

modelsummary(
  list(
    "Poisson" = cause_of_death_poisson,
    "Negative binomial" = cause_of_death_neg_binomial
  ))

```

```

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)

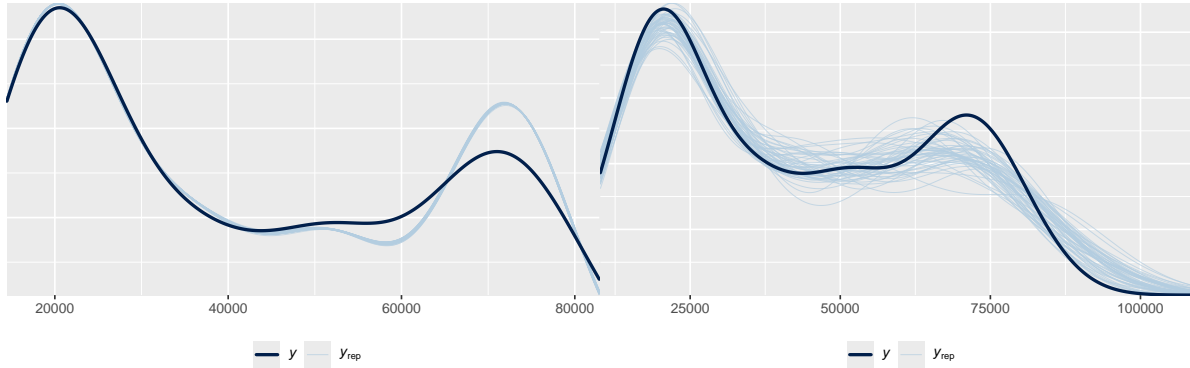
loo_compare(poisson, neg_binomial)

```

	elpd_diff	se_diff
cause_of_death_neg_binomial	0.0	0.0
cause_of_death_poisson	-14104.5	1746.0

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

	Poisson	Negative Binomial
malignant neoplasms	1.479	1.473 (0.086)
Diseases of heart	1.137	1.130 (0.086)
Respiratory malignant neoplasms	0.123	0.117 (0.083)
Dementia	0.345	0.339 (0.097)
Major cardiovascular diseases	1.450	1.443 (0.084)
Ischaemic heart diseases	0.769	0.762 (0.084)
Unspecified dementia	0.258	0.256 (0.099)
Other chronic ischchaemic heart diseases	0.162	0.155 (0.084)
Num.Obs.	153	153
Log.Lik.	-14 947.831	-1464.116
ELPD	-15 573.4	-1468.9
ELPD s.e.	1750.0	7.1
LOOIC	31 146.8	2937.9
LOOIC s.e.	3500.1	14.3
WAIC	31 866.8	2937.6
RMSE	3153.30	3153.43



(a) Poisson model

(b) Negative binomial model

Figure 1: Comparing posterior prediction checks for Poisson and negative binomial models

Table 4: Modeling the most prevalent cause of deaths in Canada, 2001-2020

	Poisson	Negative binomial
(Intercept)	9.723	9.729 (0.079)
cause_of_death_icd_10Dementia [F010-F019, F03]	0.345	0.339 (0.097)
cause_of_death_icd_10Diseases of heart [I00-I09,...	1.137	1.130 (0.086)
cause_of_death_icd_10Ischaemic heart diseases [I...	0.769	0.762 (0.084)
cause_of_death_icd_10Major cardiovascular diseas...	1.450	1.443 (0.084)
cause_of_death_icd_10Malignant neoplasms [C00-C97]	1.479	1.473 (0.086)
cause_of_death_icd_10Malignant neoplasms of trac...	0.123	0.117 (0.083)
cause_of_death_icd_10Other forms of chronic isch...	0.162	0.155 (0.084)
cause_of_death_icd_10Unspecified dementia [F03]	0.258	0.256 (0.099)
Num.Obs.	153	153
Log.Lik.	-14 947.831	-1464.116
ELPD	-15 573.4	-1468.9
ELPD s.e.	1750.0	7.1
LOOIC	31 146.8	2937.9
LOOIC s.e.	3500.1	14.3
WAIC	31 866.8	2937.6
RMSE	3153.30	3153.43