DSCI 310: Historical Horse Population in Canada

Tiffany Timbers & Jordan Bourak

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.5.1     v tibble    3.2.1
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.0.4
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

Aim

This project explores the historical population of horses in Canada between 1906 and 1972 for each province.

Data

Horse population data were sourced from Government of Canada (2017a) and Government of Canada (2017b).

Methods

The R programming language (R Core Team, 2019) and the following R packages were used to perform the analysis: **knitr** (Xie (2014)), **tidyverse** (Wickham (2017)), and **Quarto** (Allaire et al. (2022)).
Note: this report is adapted from Timbers (2020).

Results

Horse Population Trends

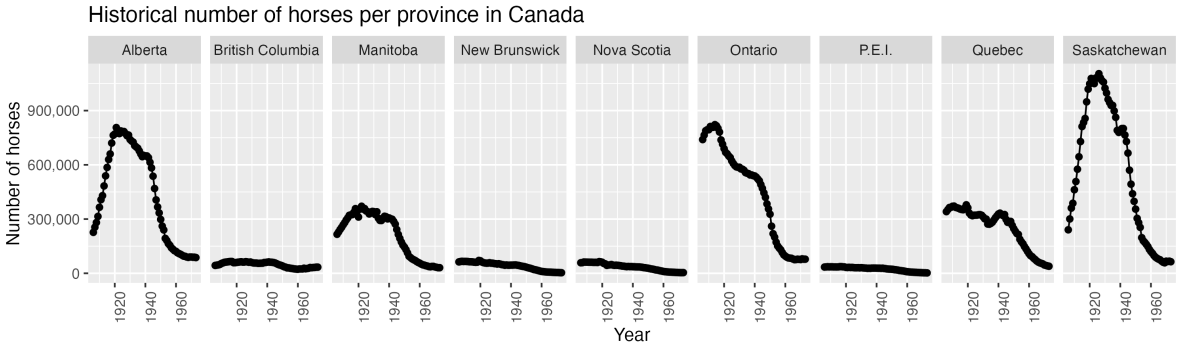

Figure 1: Horse populations for all provinces in Canada from 1906 - 1972.

Figure Figure 1 shows the horse population trends in Canada.

We can see from Figure 1 that Ontario, Saskatchewan, and Alberta have had the highest horse populations in Canada. All provinces have had a decline in horse populations since 1940. This is likely due to the rebound of the Canadian automotive industry after the Great Depression and the Second World War. An interesting follow-up visualization would be car sales per year for each province over the time period visualized above to further support this hypothesis.

Suppose we were interested in looking more closely at the province with the highest spread (in terms of standard deviation) of horse populations. We present the standard deviations in Table 1.

```
Rows: 9 Columns: 2
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (1): Province
dbl (1): Std

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.
```

Table 1: Standard deviation of historical (1906-1972) horse populations for each Canadian province.

Province	Std
Saskatchewan	377265.58
Ontario	266435.32
Alberta	266063.19
Manitoba	122403.87
Quebec	111411.10
New Brunswick	22019.49
Nova Scotia	19879.25
British Columbia	14945.66
P.E.I.	11355.75

Table Table 1 shows the standard deviation of historical (1906-1972) horse populations for each Canadian province.

Note that we define standard deviation (of a sample) as:

$$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \overline{x})^2}{N-1}}$$

Additionally, note that in Table 1, we consider the sample standard deviation of the number of horses during the same time span as Figure 1.

Figure Figure 2 zooms into the province with the highest standard deviation in horse population.

In Figure 2, we zoom in and look at the province of Saskatchewan, which had the largest spread of values in terms of standard deviation.

References

Allaire, J. J., Charles Teague, Carlos Scheidegger, Yihui Xie, and Christophe Dervieux. 2022. "Quarto." https://doi.org/10.5281/zenodo.5960048.

Government of Canada. 2017a. "Horses, Number on Farms at June 1 and at December 1." Open Government - Open Data. https://open.canada.ca/data/en/dataset/a3ecf553-8ec4-4551-a0fe-8df1472c6cf7.

———. 2017b. "Horses, Number on Farms at June 1, Farm Value Per Head and Total Farm Value." Open Government - Open Data. https://open.canada.ca/data/en/dataset/e175ef9c-98f0-49b3-8131-ca0e3895a0cb.

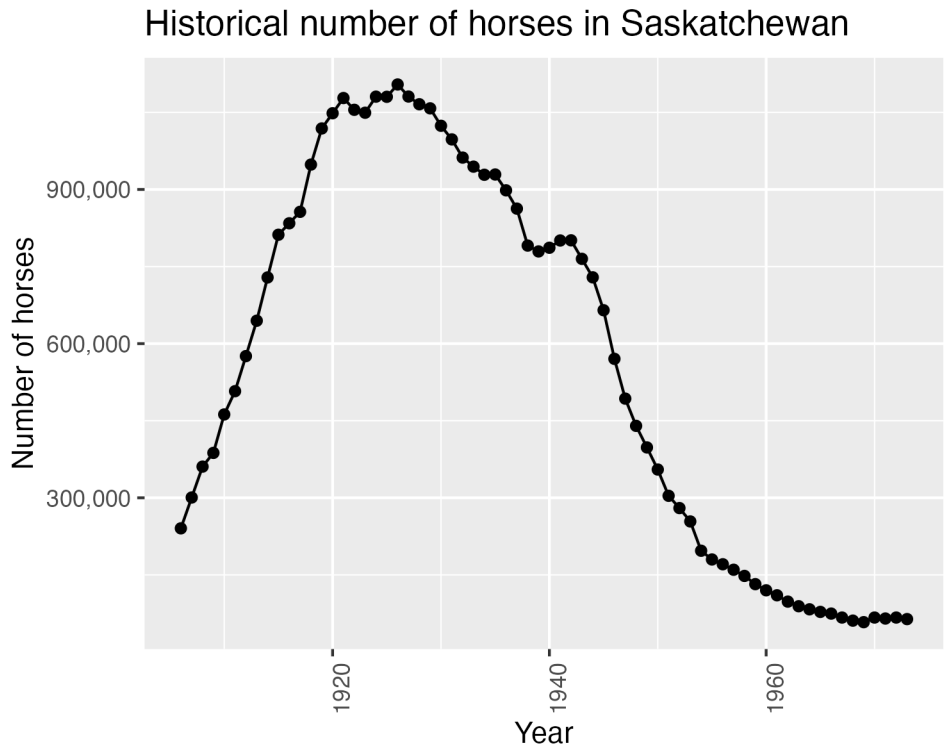

Figure 2: Horse populations for the province with the largest standard deviation.

Timbers, Tiffany. 2020. *Historical Horse Population in Canada.* https://github.com/ttimbers/equine_numbers_value_canada_parameters.

Wickham, Hadley. 2017. *Tidyverse: Easily Install and Load the 'Tidyverse'.* https://CRAN.R-project.org/package=tidyverse.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.