DSCI 310: Historical Horse Population in Canada

Tiffany Timbers & Jordan Bourak

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 the Government of Canada's Open Data website (**govCanada2017a?**; **govCanada2017b?**).

Methods

The Python programming language (**vanRossum2009?**) and the following Python packages were used to perform the analysis: pandas (**mckinney2010?**), altair (**vanderPlas2018?**), click (**clickTeam2020?**), as well as Quarto (**allaire2022?**). *Note: this report is adapted from Timbers (2020).*

Results

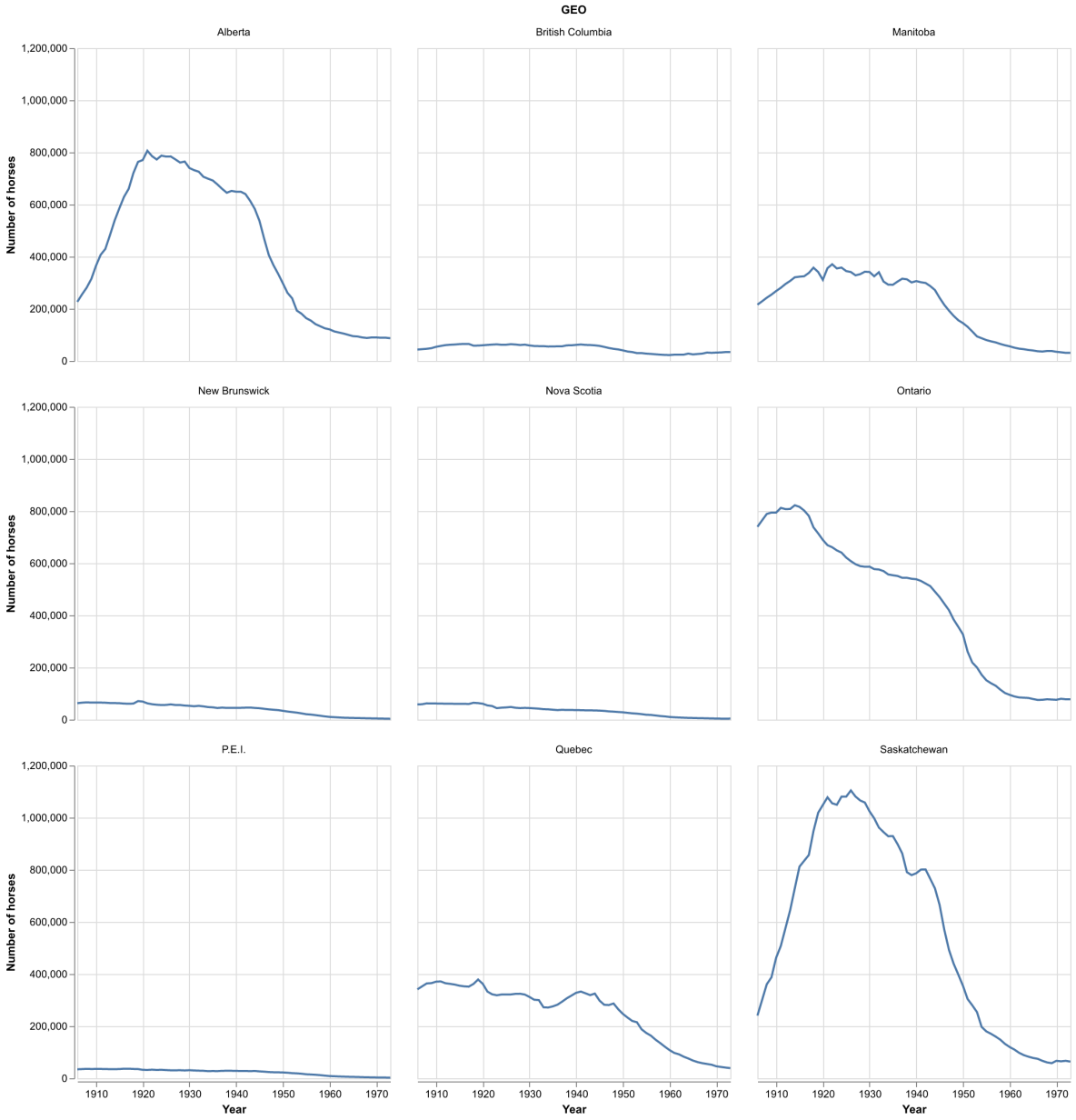

Figure @ref(fig:horsePops) shows the horse populations for all provinces in Canada from 1906 to 1972. We can see from this figure that Ontario, Saskatchewan, and Alberta have had the highest horse populations in Canada. All provinces have experienced a decline in horse populations since 1940, likely due to the rebound of the Canadian automotive industry after the Great Depression and the Second World War. An interesting follow-up visualisation would

be car sales per year for each Province over the time period visualised 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 @ref(tbl:horsePopStdDev).

Province	Std
Saskatchewan	377266
Ontario	266435
Alberta	266063
Manitoba	122404
Quebec	111411
New Brunswick	22019.5
Nova Scotia	19879.3
British Columbia	14945.7
P.E.I.	11355.7

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, in Table @ref(tbl:horsePopStdDev) we consider the sample standard deviation of the number of horses during the same time span as Figure @ref(fig:horsePops).

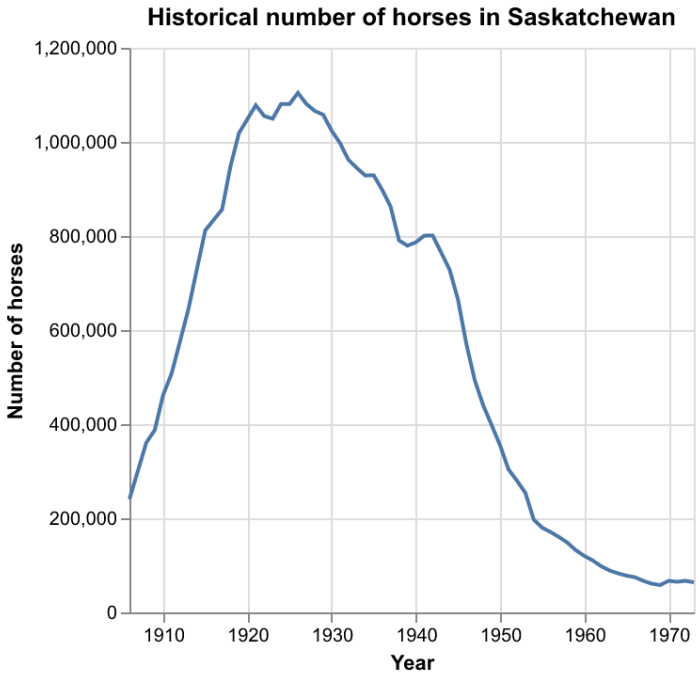

Figure @ref(fig:largestSD) zooms in and looks at the province of `r largest_sd`, which had the largest spread of values in terms of standard deviation.

References