

STATS 780 Assignment 1

Annual Crop Inventory of Saskatchewan & Manitoba

Anna Ly (400254496)

2024-09-24

Dataset Sourcing

Here we are working on the annual crop dataset, specifically looking at Manitoba and Saskatchewan between the years 2009 through 2013. Each colour on the map corresponds to different types of crops and terrains; a specific point is labelled by a crop if at least 85% of the 30m (where m represents meters) by 30m map is dominated by said crop. For example, the rgb codes (red = 255, green = 255, blue = 153) correspond to a location that has at least 85% corn. (There is no public documentation regarding what happens if an area is not dominated by any crop or type of terrain.) This data was collected by the Earth Observation Team of the Science and Technology Branch (STB) at Agriculture and Agri-Food Canada (AAFC) to understand how the cropland was used across Canadian provinces.

The dataset was sourced from the Open Government Portal using the following filter options:

Portal Type: Open Data. **Collection Type:** Open Maps. **Resource Type:** Dataset. **Format:** GeoTIF. **Update Frequency:** Annually

The exact geotif files can be found [here](#) (Canada (n.d.)).

Application for Crop Data

Here we are observing habitat types we have in Canada (such as grassland, wetlands, etc.,) as well as the types of crops that are grown in the provinces of Manitoba and Saskatchewan. This allows us to:

1. See how the agricultural landscape changes throughout the years (i.e., how the habitat types are affected and whether they became transformed as crop fields.)
2. See what types of crops are grown throughout the years and whether there are an increasing or decreasing amount of a particular crop grown.

This data helps provide context for future agricultural decisions, such as what crops we may need to grow but also the amount of natural land being transformed into farmland, potentially destroying ecosystems.

In particular, in this study we're going to be analyzing the amount of plot land used to grow corn and soybeans.

Data Transformation and Preprocessing Steps

Although the Annual Crop Data comes with a plethora of geotif files, ranging from 2009 to 2024, I specifically chose to download between 2009 and 2013 for Saskatchewan and Manitoba due to computational limitations; the geotif files (especially those in Ontario) were massive in size and my machine was unable to process all of them. For similar reasons, I narrowed down to analyze specific crops of interest (such as cereals and legumes).

Here are the pre-processing steps I took to clean and prepare for analysis.

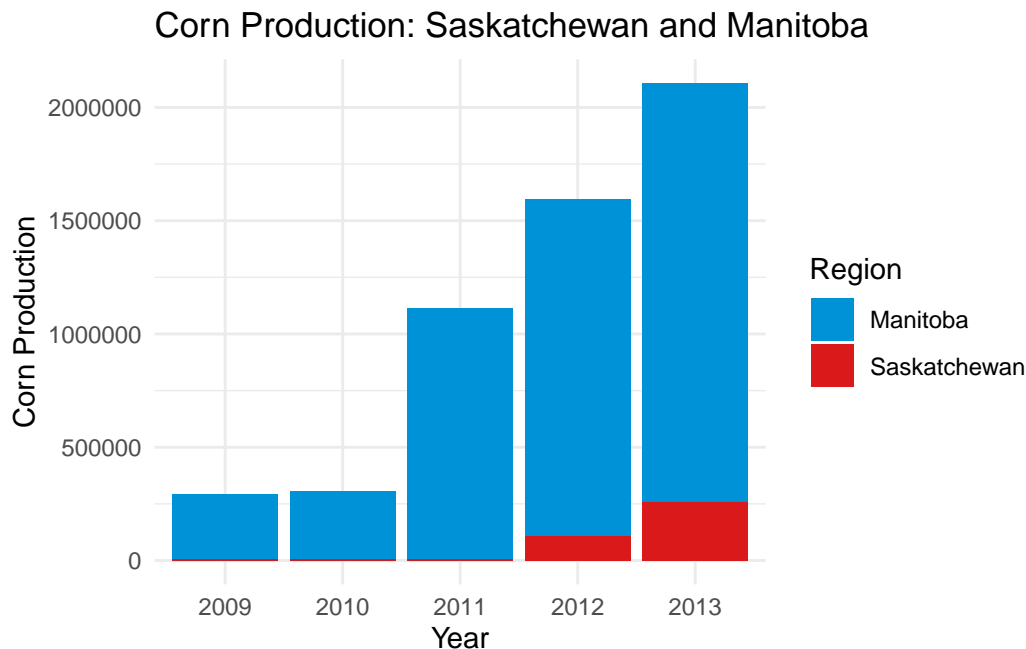
1. Used the R [terra library](#) Hijmans (2024) to create a SpatRaster object from the geotif files.
2. Extracted the values from the SpatRaster object to create a vector which contains the numeric codes (a vector of integers). These numerical codes correspond to a specific geographic region. For example, the number "147" corresponds to "corn". As a result, I did encounter **missing values** as I interpreted them to not be associated to any plot. Furthermore, there were **no outliers** for me to deal with; either the crop was present or it was not.
3. Take the sum of the amount of times a specific numerical code appeared to count the amount of land used for specific crops such as corn or soybeans.

4. Construct a new data frame, which includes the year, the province, and the number of plots used for specific crops. This new data frame can now be used for analysis and constructing graphs.

Single-variable Analysis - The Difference in Amount of Cropland Used for Corn

Research question: how much land is dedicated to **corn** (the variable of interest) in the provinces Manitoba and Saskatchewan throughout the years 2009 to 2013? Are farmers growing more or less corn?

Using the new data frame (the details for its creation is outlined in the previous section), I extracted the counts for the amount of plots containing corn in both Manitoba and Saskatchewan. Here's a stacked bar plot, showing the total amount of corn being planted throughout the years:



As predicted, the amount of land being used to grow corn increases year by year. Hence, farmers are growing more corn and/or there's a larger demand for corn farmers. The exact values (combining Manitoba and Saskatchewan) year to year is 291001, 305179, 1115682, 1594023, 2107593, respectively. More research needs to be done to understand exactly why there are more land used for corn fields. However, a possible hypothesis is that time passing is highly correlated with population growth and the need for more food.

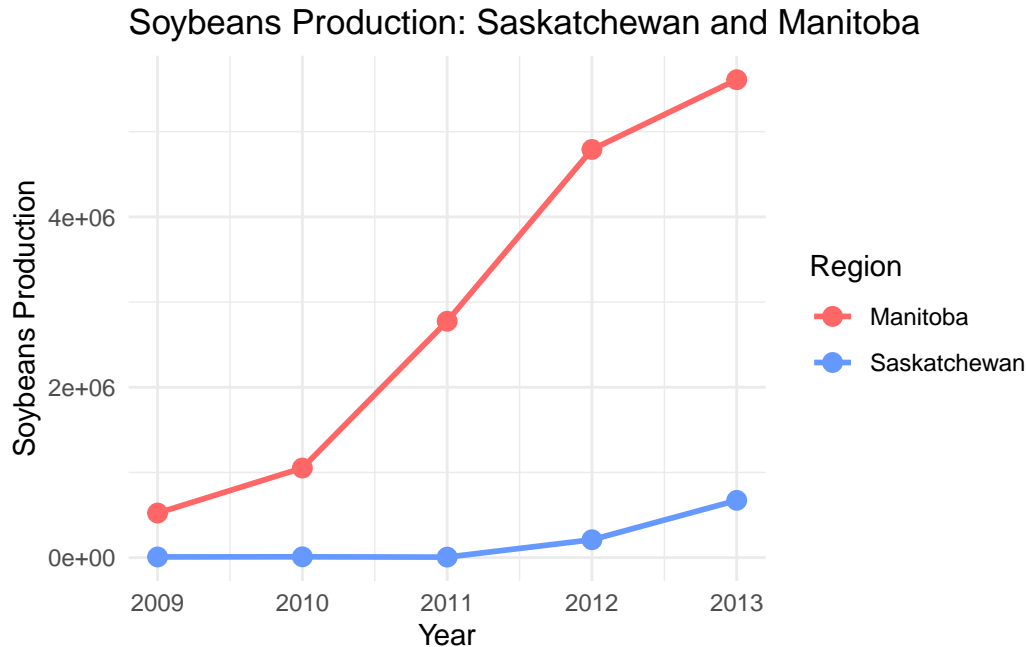
It is interesting to see that most of the corn production is in Manitoba, even though Saskatchewan and Manitoba had comparable populations in 2009. The population in Saskatchewan was 1,023,810 "Saskatchewan Annual Population Report" (n.d.) and the population in Manitoba was 1,214,403 "Manitoba Health and Healthy Living Population Report" (n.d.). Again, more research needs to be done to understand why this occurred, but this caveat is beyond the scope of this report.

Multi-Variable Analysis

Research question: how much land is dedicated to soybeans in the provinces Manitoba and Saskatchewan throughout the years 2009 to 2012? Are farmers growing more or less soybeans?

Variables of focus: amount of crop used for soybeans, province.

Again, using the aforementioned data frame, I extracted the counts for the amount of plots containing soybeans in both Manitoba and Saskatchewan. Here's a line plot, where there's a separate line for Saskatchewan and Manitoba, showing the total amount of corn being planted throughout the years:



Similar to the previous research question, we find that more crops (in this case, soybeans) are being grown in Manitoba in comparison to Saskatchewan. Hence, farmers are growing more soybeans or there's a larger demand for soybeans. Ordering from 2009 to 2013, the amount of 30m by 30m plots for soybeans grown in Manitoba are: 523395, 1052033, 2775207, 4791723, 5610724. Additionally, the amount of soybeans grown for Saskatchewan are: 8258, 9784, 6296, 210324, 672449. Again, more research needs to be done to understand why more crops are being grown in Manitoba compared to Saskatchewan.

Interactive Visualization with Shiny

I have developed an interactive Shiny app Chang et al. (2024) using R to investigate the multi-variable analysis research question. In fact, it explores additional crops that I did not have the space to discuss here, such as oats. The link to the Shiny app: <https://annaly.shinyapps.io/STATS780Homework1/>

I used additional packages, such as shinycssloaders to create beautiful loading screens when creating different graphs Attali and Sali (2024), colourpicker so the user can easily switch the colours of the graphs Attali (2023), and tidyverse to create the beautiful plots Wickham et al. (2019). Occasionally, Github Copilot "GitHub Copilot · Your AI Pair Programmer. GitHub" (2024) was used to assist with writing code for the R Shiny website; in particular, it provided a skeleton to write sections.

Students can select a crop (either grains and legumes), the type of plot they would like to check (either a bar plot or a line plot), whether they would want a comparison or an individual plot, and the colour scheme. In particular, this app helps to show possible visual representations of how much land is used for individual crops throughout the years.

The code for the Shiny app will be included in the supplementary material.

References

- Attali, Dean. 2023. *Colourpicker: A Colour Picker Tool for Shiny and for Selecting Colours in Plots*. <https://CRAN.R-project.org/package=colourpicker>.
- Attali, Dean, and Andras Sali. 2024. *Shinycssloaders: Add Loading Animations to a 'Shiny' Output While It's Recalculating*. <https://CRAN.R-project.org/package=shinycssloaders>.
- Canada, Agriculture {and} Agri-Food. n.d. "Annual Crop Inventory - Open Government Portal." Accessed September 24, 2024. <https://open.canada.ca/data/en/dataset/ba2645d5-4458-414d-b196-6303ac06c1c9>.
- Chang, Winston, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. 2024. *Shiny: Web Application Framework for r*. <https://CRAN.R-project.org/package=shiny>.
- "GitHub Copilot · Your AI Pair Programmer. GitHub." 2024. 2024. <https://github.com/features/copilot/>.
- Hijmans, Robert J. 2024. *Terra: Spatial Data Analysis*. <https://CRAN.R-project.org/package=terra>.
- "Manitoba Health and Healthy Living Population Report." n.d. Accessed September 24, 2024. <https://www.gov.mb.ca/health/population/pr2009.pdf>.
- "Saskatchewan Annual Population Report." n.d. Accessed September 24, 2024. <https://www.saskatchewan.ca/-/media/news-archive/2009/march/26/saskatchewan-grew-by-more-than-15000-people-in-2008/annual-report.pdf>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.