Spring Wheat Analysis in Manitoba and Saskatchewan

Automation & Digital Agriculture
Data Analysis Stream Final Project

Tao Zhe (Evan) Wu
October 2023

Scoping

Nearly a billion people will go hungry tonight, yet this year the U.S. will turn nearly 5 billion bushels of corn into ethanol. That’s enough food to feed 412 million people for an entire year.

How about Canada? Are Canadian farmers lowering wheat production for canola?

Data Collection

- Saskatchewan Yield Data Source (https://dashboard.saskatchewan.ca/agriculture/rm-yields/rm-yields-data#rm-yields-tab)

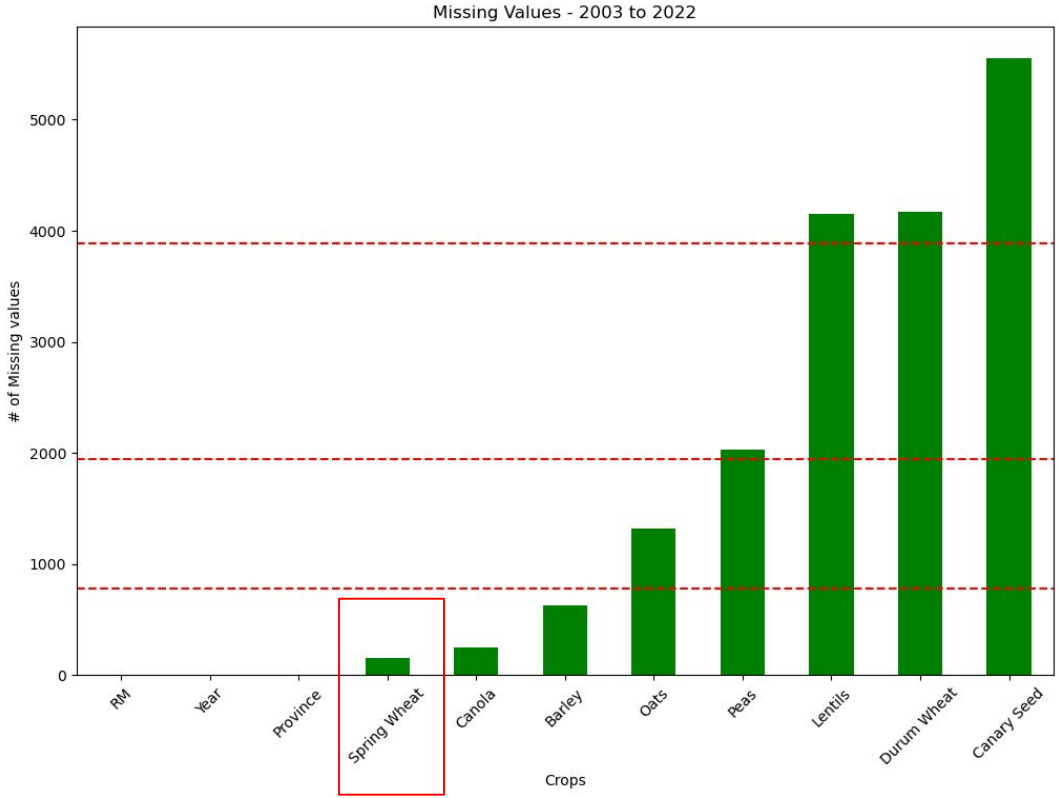
- Manitoba Yield Data Source (https://www.gov.mb.ca/agriculture/markets-and-statistics/crop-statistics/index.html)

- Manitoba and Saskatchewan Shapefiles are provided by Instructor (Ruhid)

Data Transformation

- Yield unit conversion
- String formatting for RM names and column (crop) names to join the datasets
- Combine the joined dataset with geo database for visualization

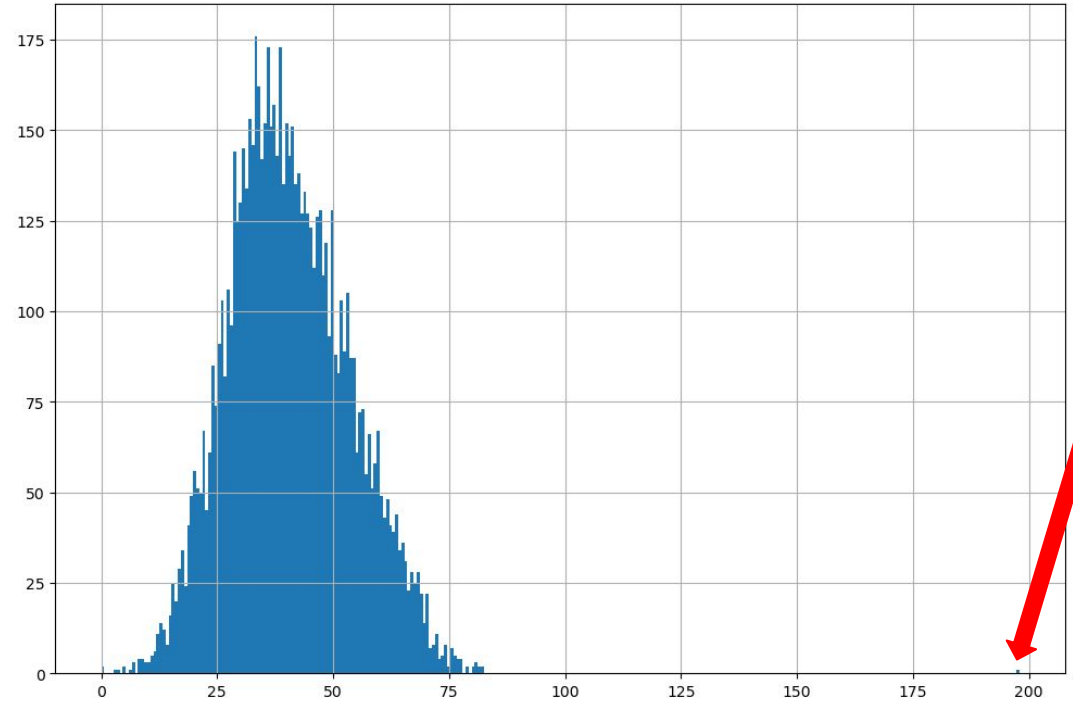
Exploratory Data Analysis (EDA)

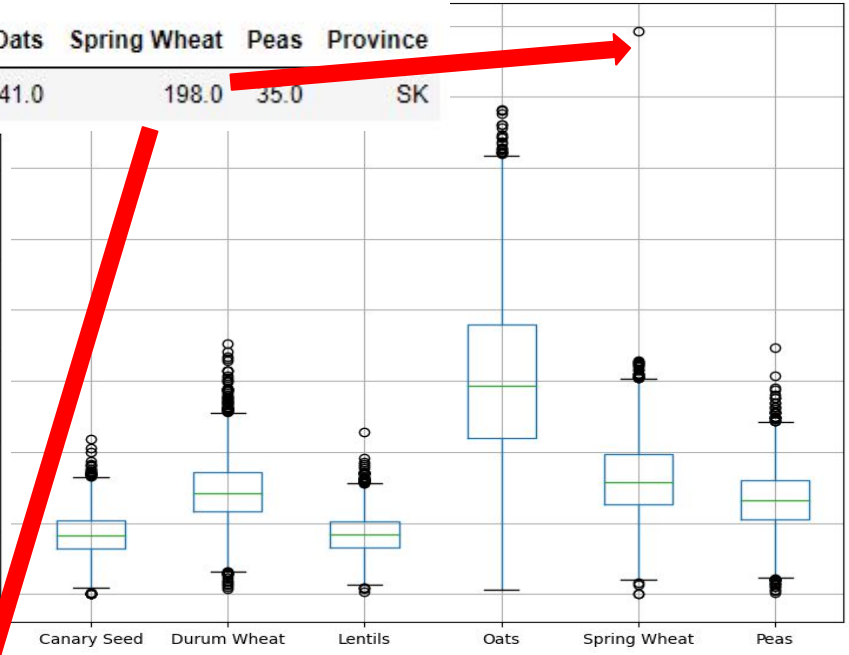


- Spring wheat has the lowest missing values - it's popular in Manitoba and Saskatchewan!

Exploratory Data Analysis (EDA): Distribution

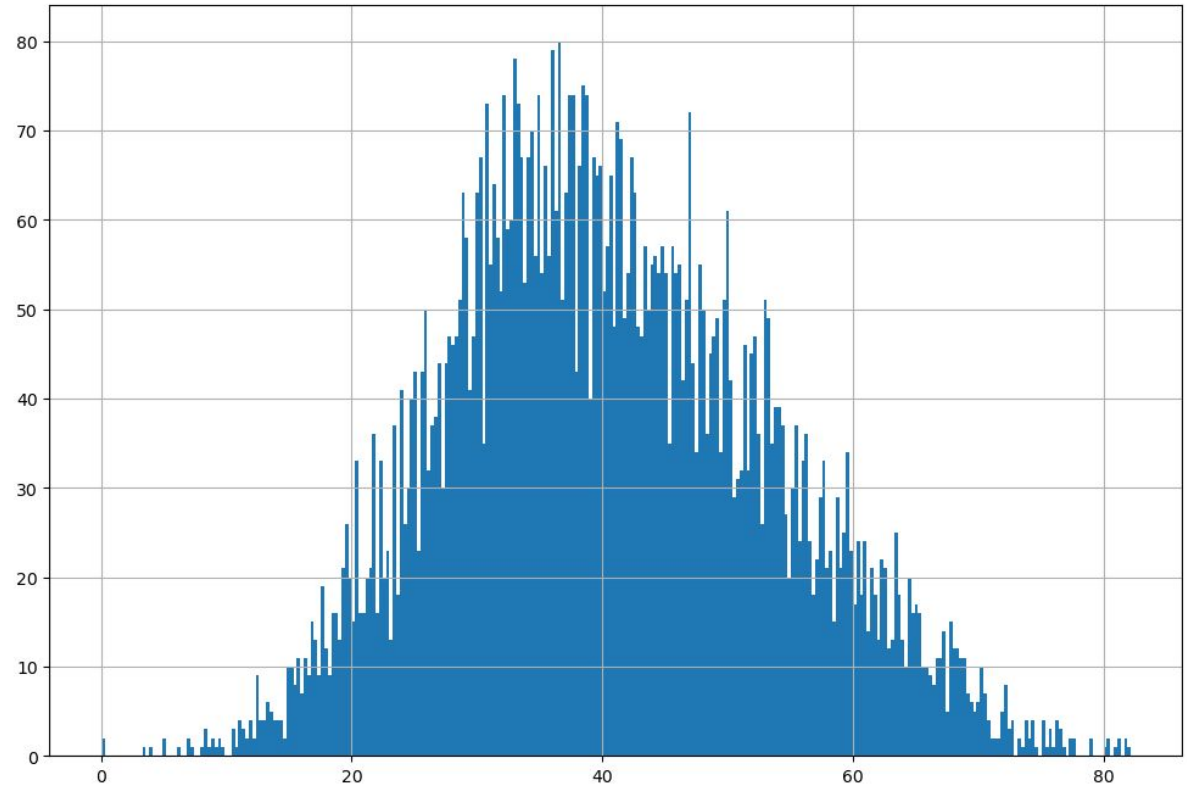
	RM	Year	Canola	Barley	Canary Seed	Durum Wheat	Lentils	Oats	Spring Wheat	Peas	Province
16215	314	2018	33.0	43.5	NaN	NaN	29.033333	41.0	198.0	35.0	SK





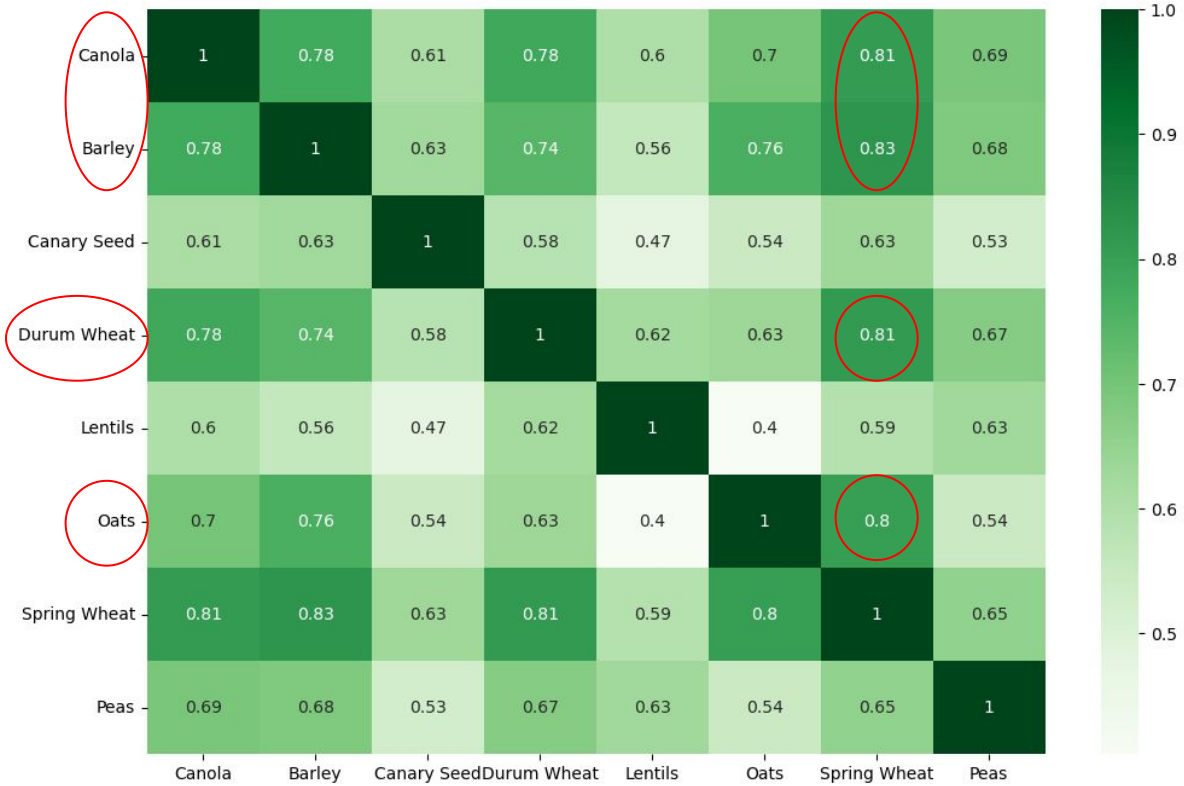
- The average wheat yield of 2017 and 2018 in RM 314 is used to replace the outlier.

Exploratory Data Analysis (EDA)



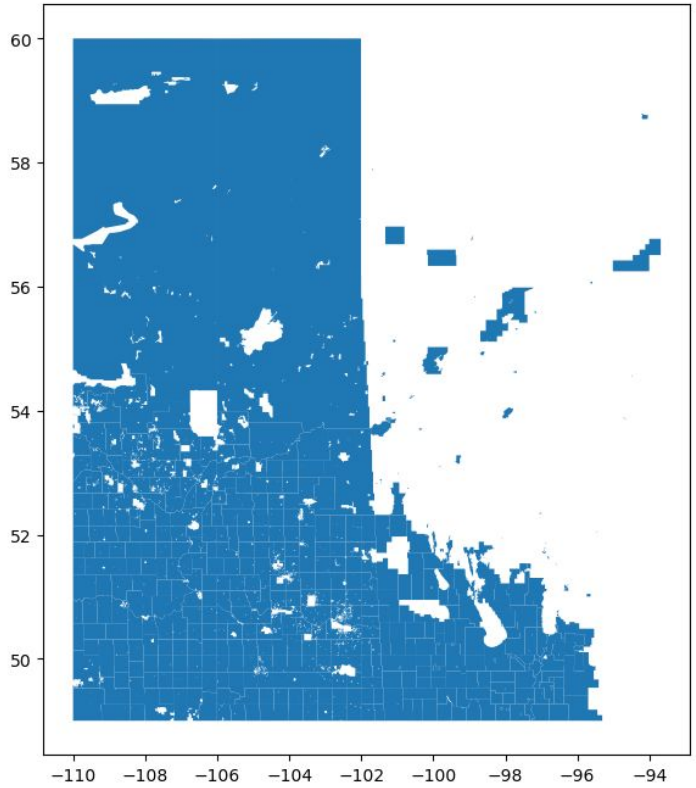
- The wheat yields are now normally distributed

Exploratory Data Analysis (EDA)

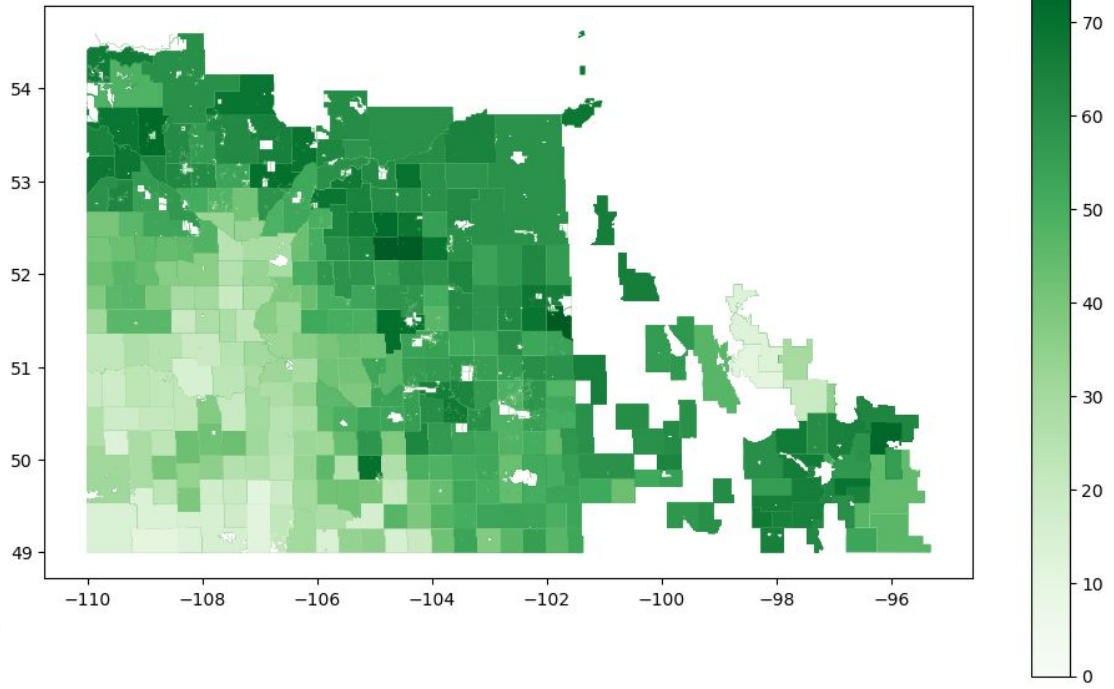


- Canola, Barley, Durum Wheat, and Oats have strong correlation with Spring Wheat.

Exploratory Data Analysis (EDA): Visualizations

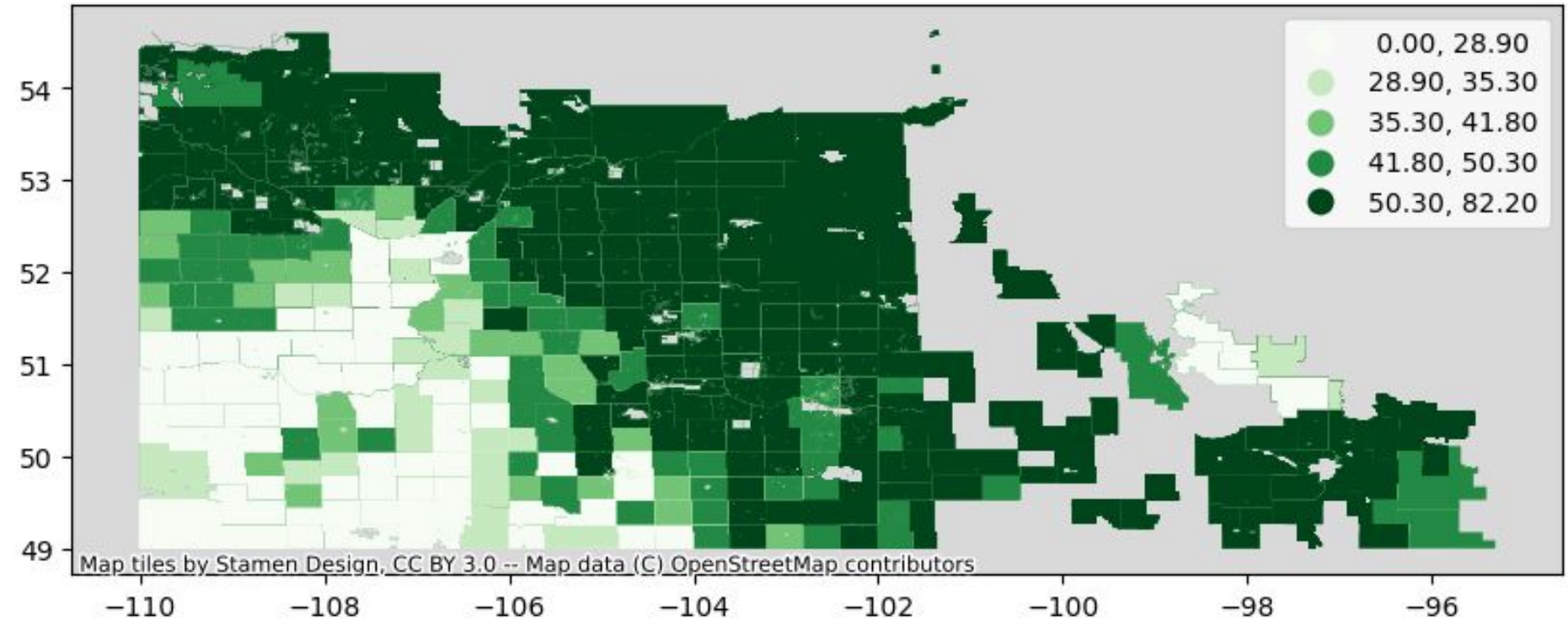


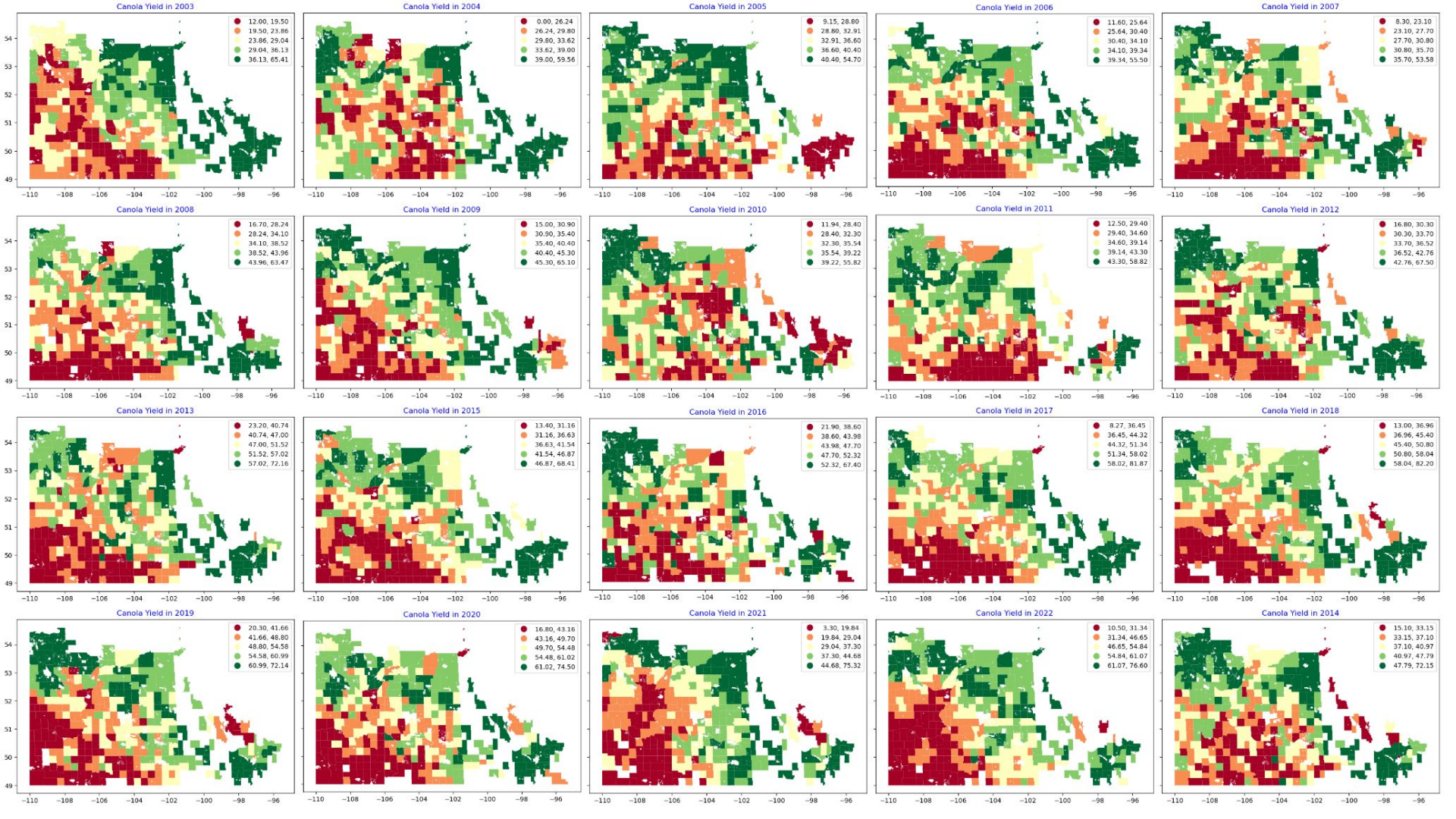
- Spring yields are reported in 349 out of all 481 RMs



Exploratory Data Analysis (EDA)

- The wheat yields are generally lower in southwest, likely due to Canola.

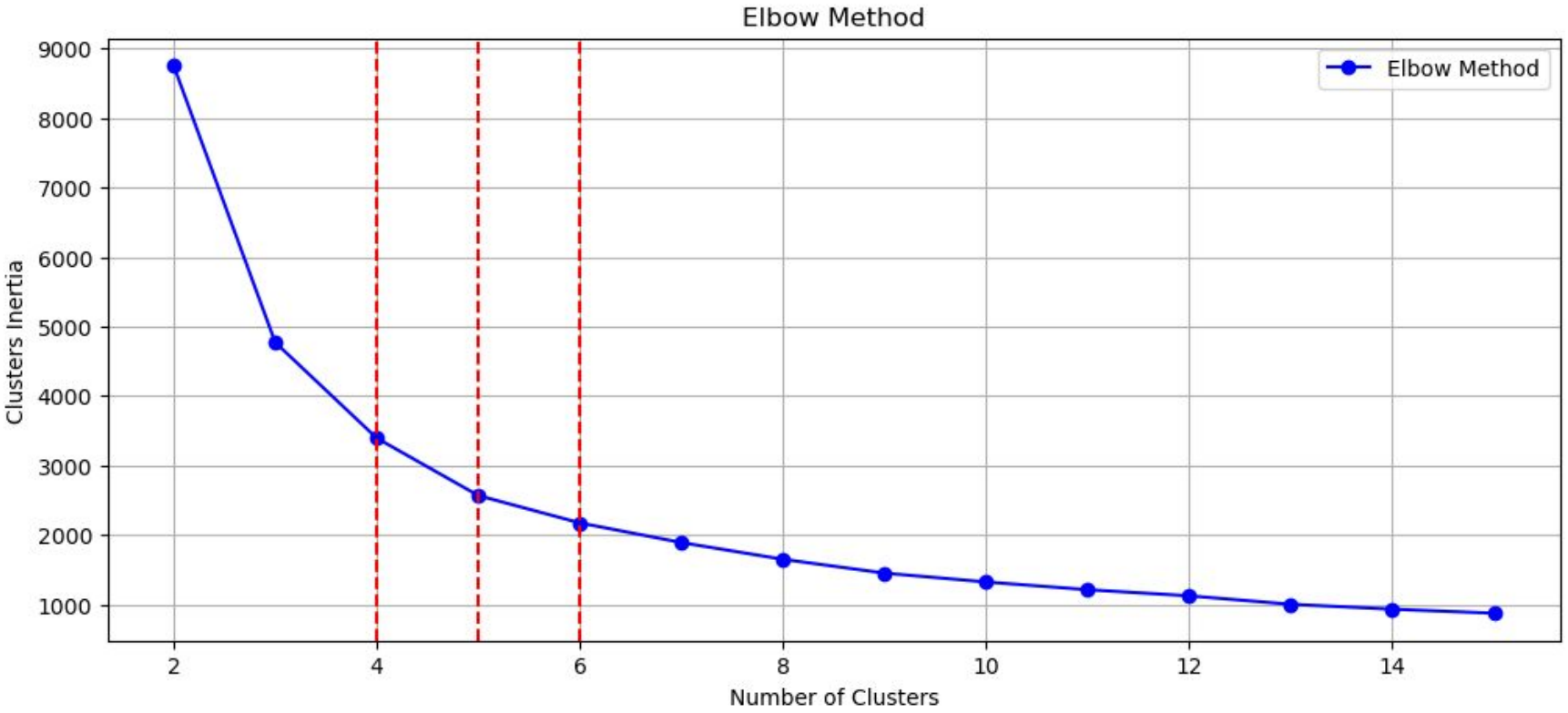



Canola Yield in 2003
12.00, 19.50
19.50, 23.86
23.86, 29.04
29.04, 36.13
36.13, 65.41
Canola Yield in 2004
0.00, 26.24
26.24, 29.80
29.80, 33.62
33.62, 39.00
39.00, 59.56
Canola Yield in 2005
9.15, 28.80
28.80, 32.91
32.91, 36.60
36.60, 40.40
40.40, 54.70
Canola Yield in 2006
11.60, 25.64
25.64, 30.40
30.40, 34.10
34.10, 39.34
39.34, 55.50
Canola Yield in 2007
8.30, 23.10
23.10, 27.70
27.70, 30.80
30.80, 35.70
35.70, 53.58
Canola Yield in 2008
16.70, 28.24
28.24, 34.10
34.10, 38.52
38.52, 43.96
43.96, 63.47
Canola Yield in 2009
15.00, 30.90
30.90, 35.40
35.40, 40.40
40.40, 45.30
45.30, 65.10
Canola Yield in 2010
11.94, 28.40
28.40, 32.30
32.30, 35.54
35.54, 39.22
39.22, 55.82
Canola Yield in 2011
12.50, 29.40
29.40, 34.60
34.60, 39.14
39.14, 43.30
43.30, 58.82
Canola Yield in 2012
16.80, 30.30
30.30, 33.70
33.70, 36.52
36.52, 42.76
42.76, 67.50
Canola Yield in 2013
23.20, 40.74
40.74, 47.00
47.00, 51.52
51.52, 57.02
57.02, 72.16
Canola Yield in 2015
13.40, 31.16
31.16, 36.63
36.63, 41.54
41.54, 46.87
46.87, 68.41
Canola Yield in 2016
21.90, 38.60
38.60, 43.98
43.98, 47.70
47.70, 52.32
52.32, 67.40
Canola Yield in 2017
8.27, 36.45
36.45, 44.32
44.32, 51.34
51.34, 58.02
58.02, 81.87
Canola Yield in 2018
13.00, 36.96
36.96, 45.40
45.40, 50.80
50.80, 58.04
58.04, 82.20
Canola Yield in 2019
20.30, 41.66
41.66, 48.80
48.80, 54.58
54.58, 60.99
60.99, 72.14
Canola Yield in 2020
16.80, 43.16
43.16, 49.70
49.70, 54.48
54.48, 61.02
61.02, 74.50
Canola Yield in 2021
3.30, 19.84
19.84, 29.04
29.04, 37.30
37.30, 44.68
44.68, 75.32
Canola Yield in 2022
10.50, 31.34
31.34, 46.65
46.65, 54.84
54.84, 61.07
61.07, 76.60
Canola Yield in 2014
15.10, 33.15
33.15, 37.10
37.10, 40.97
40.97, 47.79
47.79, 72.15
54
53
52
51
50
49
−110
−108
−106
−104
−102
−100
−98
−96

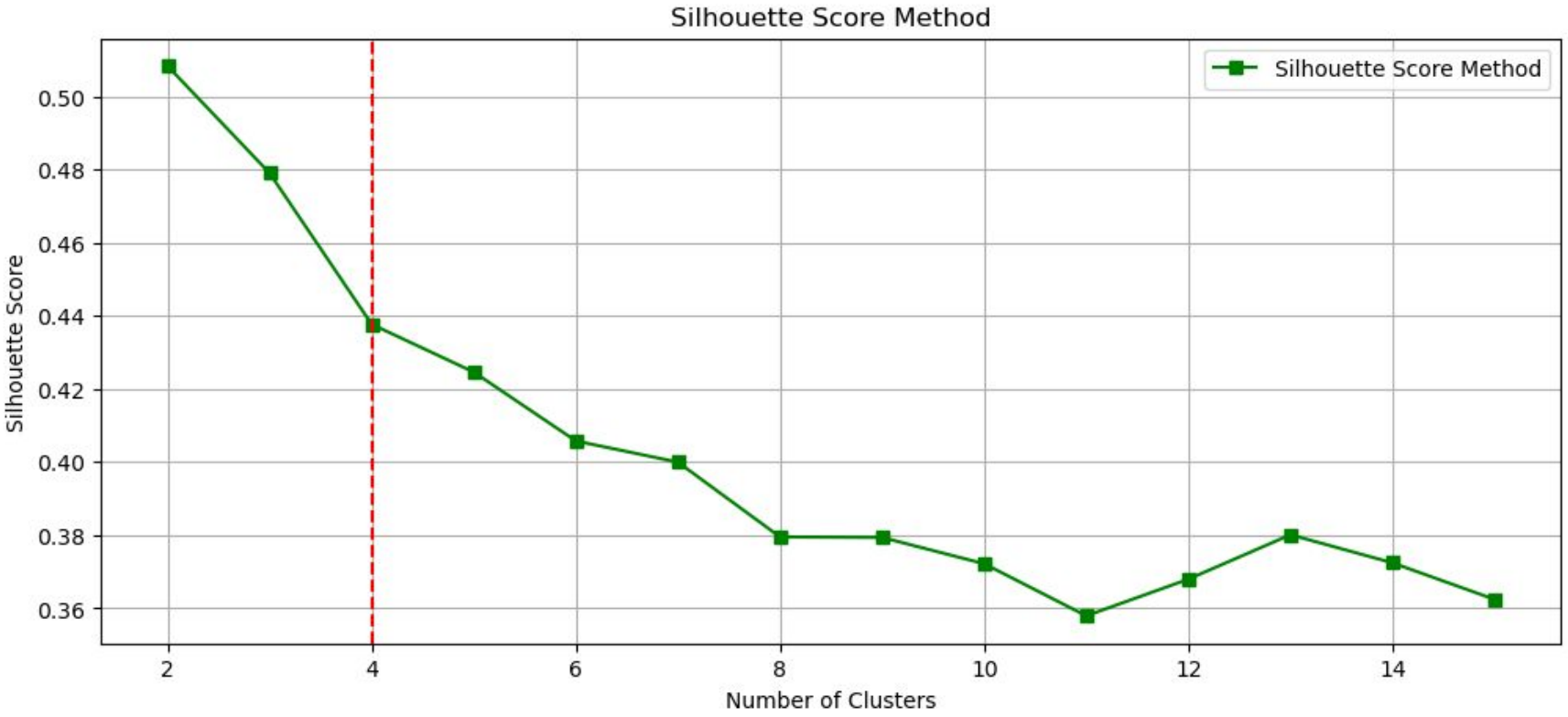
Methodology

- K-Means clustering method is used to classify the RMs
- For this project, only Spring Wheat data are analyzed
- The mean and standard deviation of wheat yields over 2003-2022 in each RM are used for clustering
- The RMs are clustered into 2 to 15 groups, and Elbow Method and Silhouette Scoring are used to determine the proper number of clusters

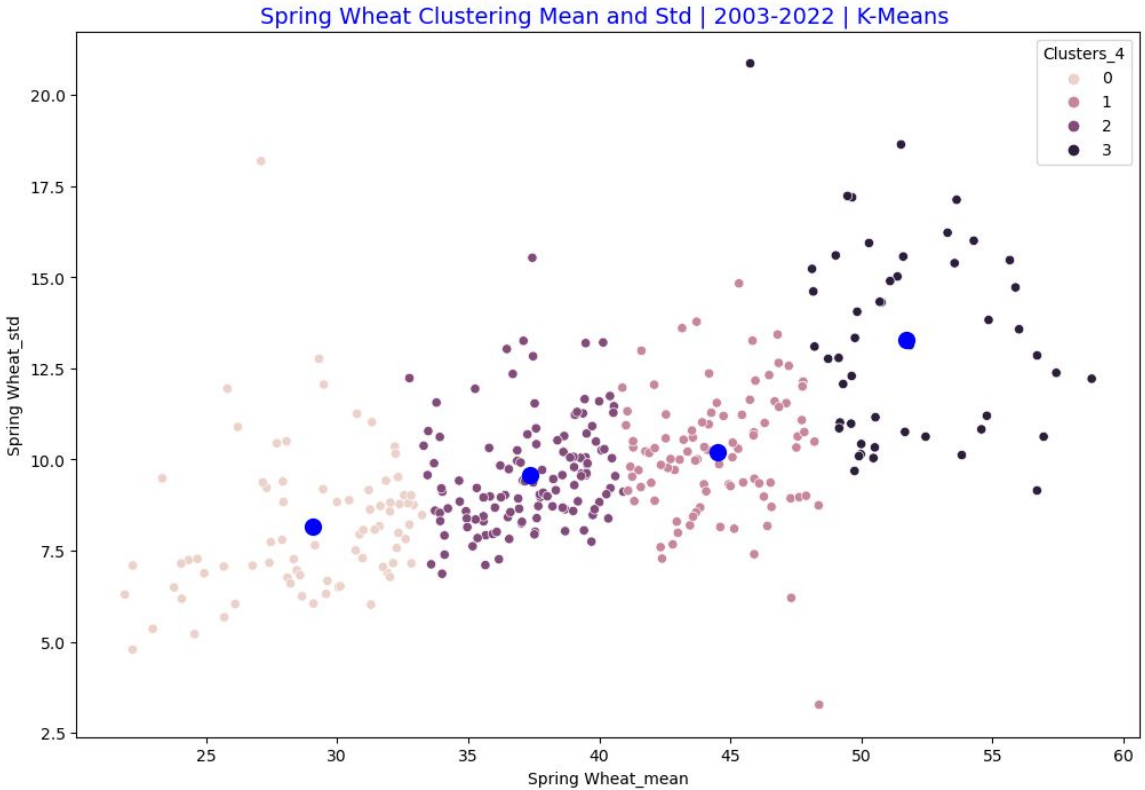
Elbow Method



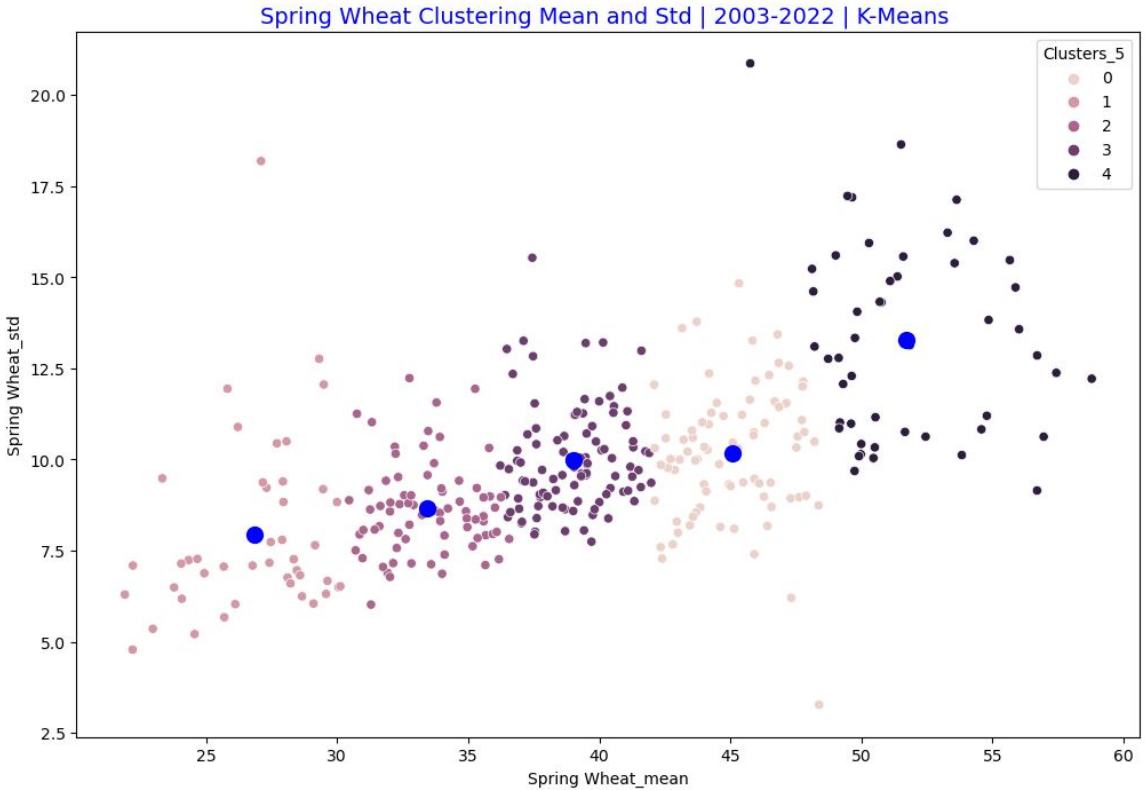
Silhouette Score Method



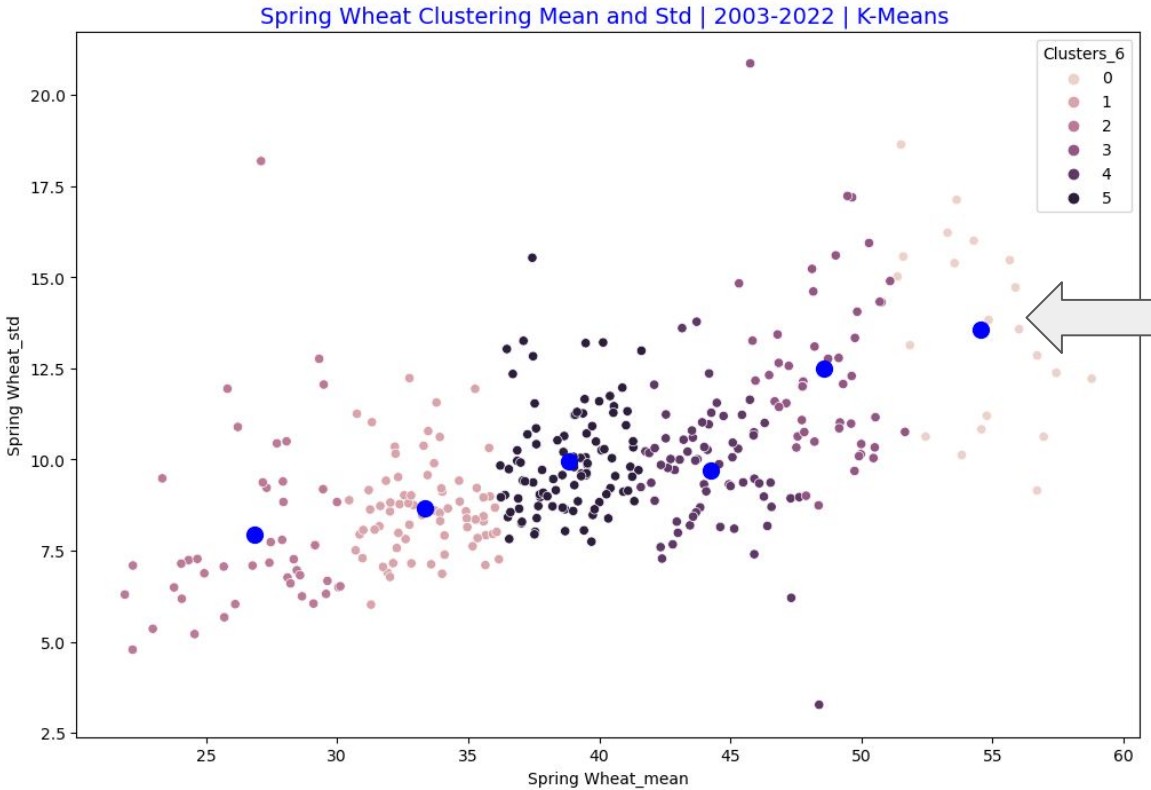
N = 4


Spring Wheat Clustering Mean and Std | 2003-2022 | K-Means
Clusters_4
0
1
2
3
Spring Wheat_std
Spring Wheat_mean
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
25
30
35
40
45
50
55
60

N = 5


Spring Wheat Clustering Mean and Std | 2003-2022 | K-Means
Clusters_5
0
1
2
3
4
Spring Wheat_std
Spring Wheat_mean
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
25
30
35
40
45
50
55
60

N = 6



A bad center

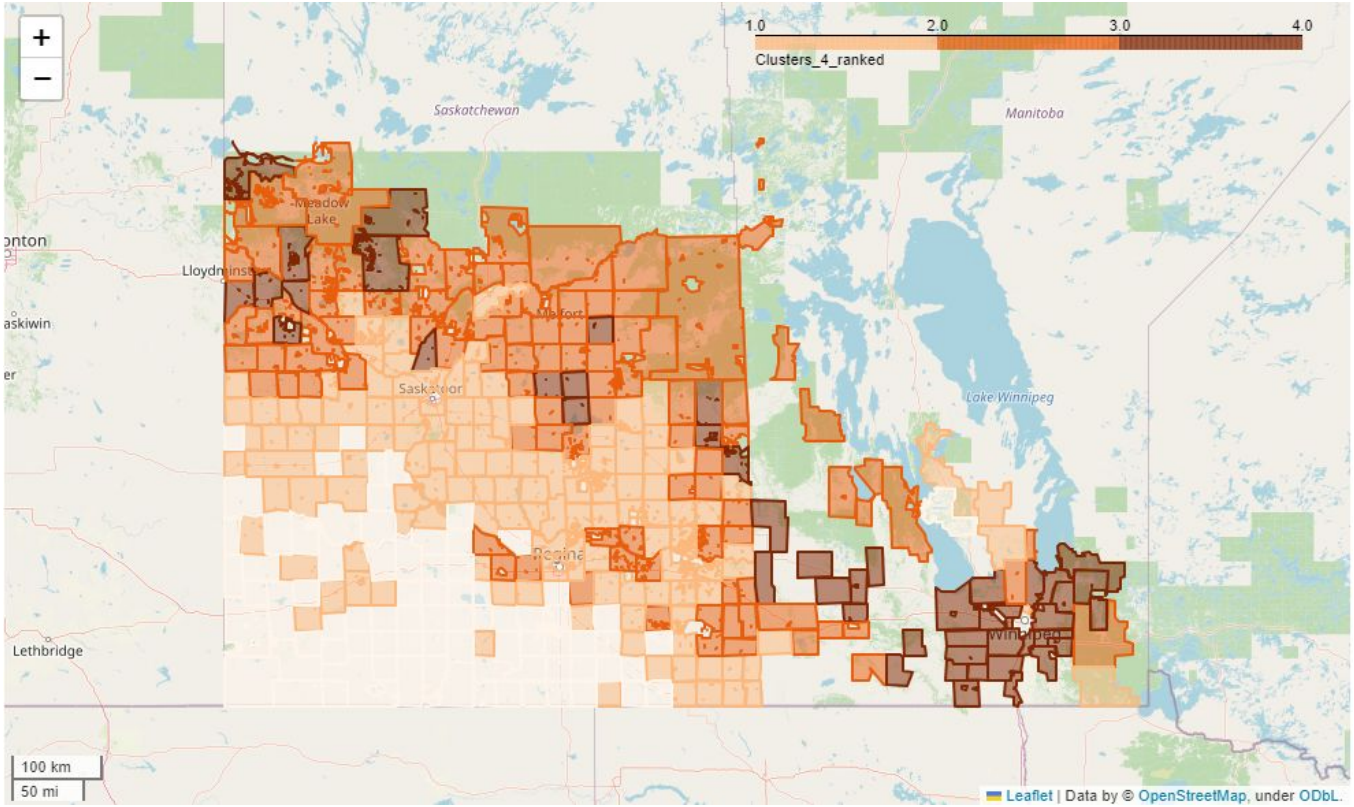
Ranking based on Mean or Standard Deviation

Clusters_4	Spring Wheat_mean	Spring Wheat_std
0	29.076890	8.166517
2	37.354170	9.586894
1	44.498793	10.187763
3	51.716221	13.282003

Clusters_5	Spring Wheat_mean	Spring Wheat_std
1	26.854778	7.954952
2	33.423295	8.653647
3	39.004220	9.989843
0	45.068284	10.172036
4	51.716221	13.282003

Clusters_6	Spring Wheat_mean	Spring Wheat_std
2	26.854778	7.954952
1	33.341542	8.660789
5	38.835044	9.966704
4	44.231024	9.698660
3	48.582343	12.481640
0	54.558429	13.557375

RMs Ranked in Spring Wheat Yielding with N=4



Results & Conclusions

- Canola, barley, durum wheat and oats are found to be highly correlated with spring wheat yielding.

- Yearly spring wheat yields in 351 Manitoba and Saskatchewan RMs from 2003 - 2022 are mapped and visualized.
  - Spring wheat is in general very popular in both provinces, but **southwest Saskatchewan** grows **less and less spring wheat since 2016**

- RMs are clustered using K-Means method based on the mean and deviation of Spring Wheat yields in 2003-2022.
  - N = 4 is the elbow point with highest Silhouette Score.
  - The rankings of RMs are visualized on map.

- Successfully observed the change of spring wheat yielding visually over years, and ranked each RM based on the yield potential and stability.

Future Works

- Analyze other crops to explain the drop in spring wheat yielding in southwest Saskatchewan. Temperature, precipitation, and soil type might be considered as factors.

- Canola, barley, durum wheat and oats are highly correlated with spring wheat production. They might be in a crop rotation cycle. The effectiveness of crop rotation in increasing spring wheat production can be analyzed.