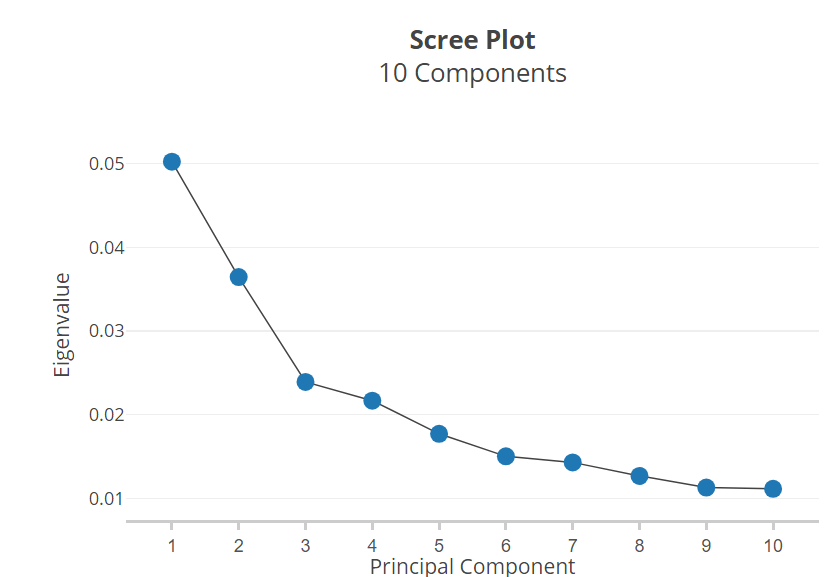
**CAN Treaties – Methods Overview 2** (Douglas Treaties aggregated)

**Using String Kernels:** When we use string kernels to measure the similarity between two texts, we look at common sequences of characters. Using the term “**majesti**” as an example, with a specified length of 5 characters:

1. **Split** “majesti” into all possible 5-character sequences (substrings):
   1. “**majes**”
   2. “**ajest**”
   3. “**jesti**”
2. **Compare Substrings**: We compare these 5-character sequences with the 5-character sequences from another text.
   1. e.g., if the other treaty text also has “majesty”, it would have the same substrings:
      1. “**majes**”
      2. “**ajest**”
      3. “**jesty**” (which shares “jesti” with “majesti”)
3. **Count Common Substrings**: We count how many of these 5-character sequences are common between the two texts. The more common sequences they have, the more similar the texts are considered to be.

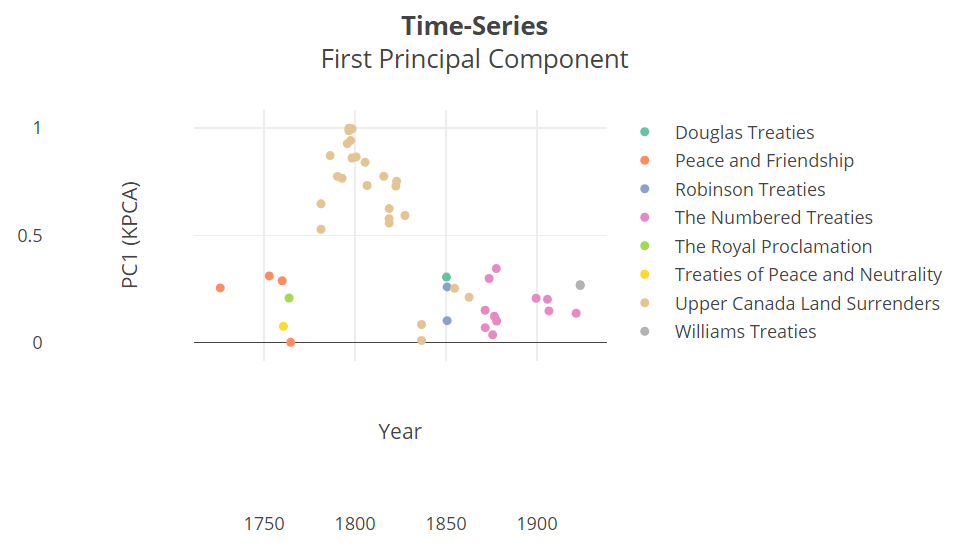
Ultimately, this is how we get the graph(s) of the “thing” we want to look at – the overall theme or commonality across all the treaties – because we’ve computed the **Kernel Principal Component Analysis (KPCA)**. For Spirling this was harshness.

**How many “things” are there? – we can argue for 1 or 2:**

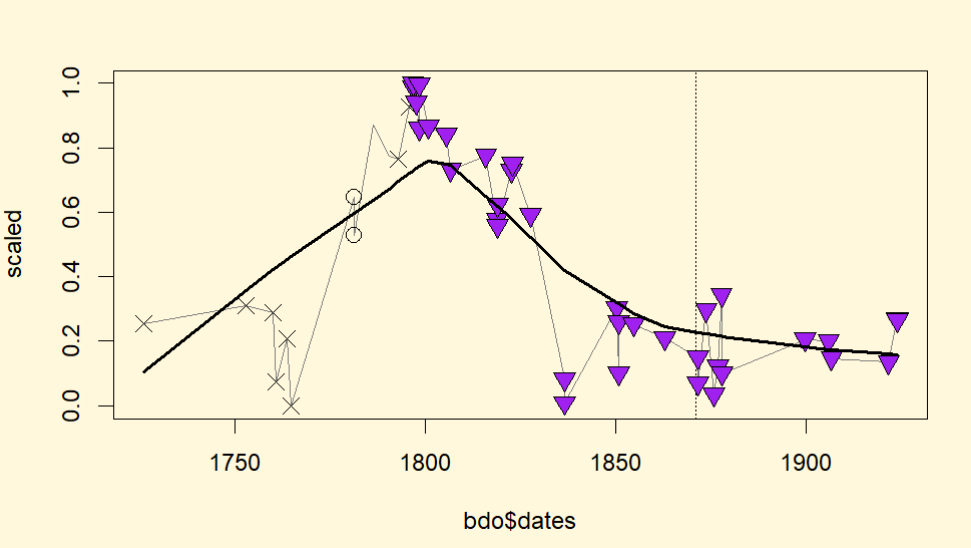
We are *definitively* looking for a **1st** and a **2nd component:**

When you see the elbow, you backtrack at the largest bend (if it’s a large enough bend comparatively to all bends). The biggest bend indicates which component captures the most variance in the data.

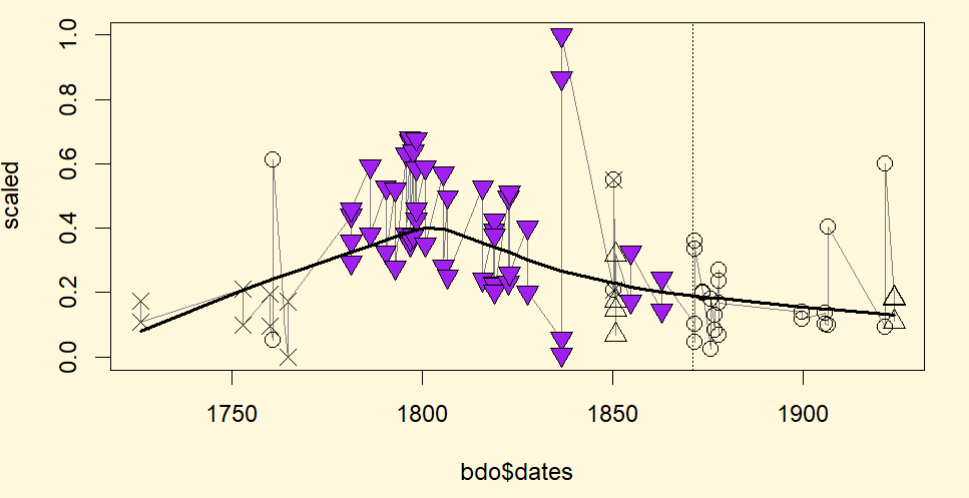
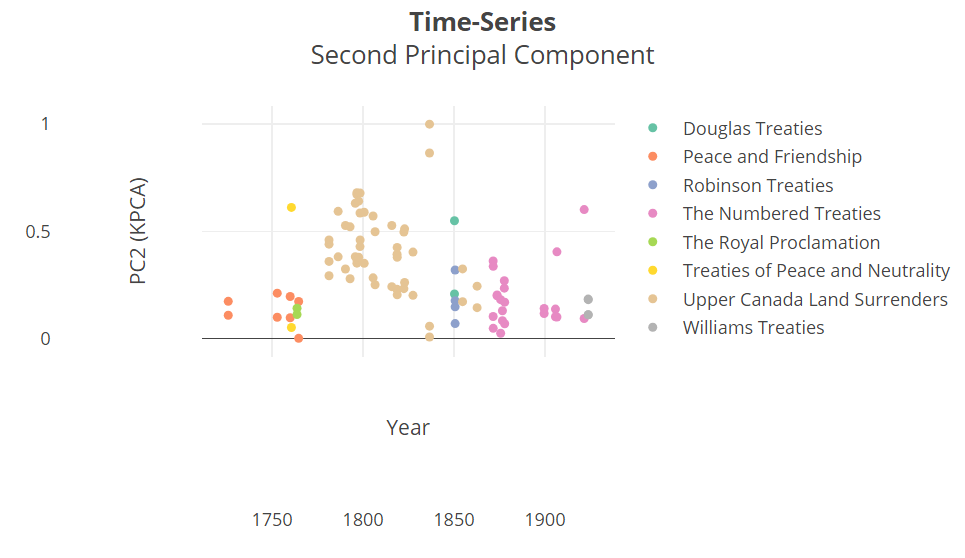
* **Eigenvalues**: indicates the amount of variance in the data that is explained by its corresponding principal component.
  + Higher eigenvalues mean that the principal component captures more variance from the data.

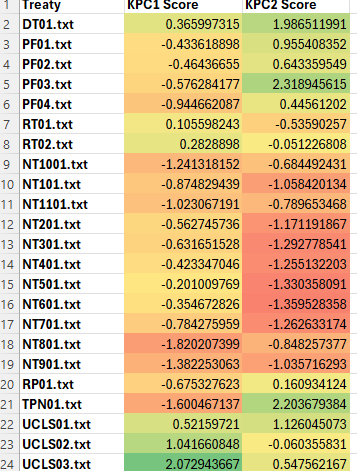
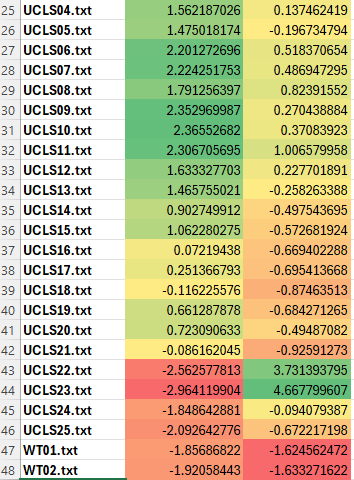
**Component 1 Graph – “the main thing”:**

* Here the early Upper Canada Land Surrenders seem to be the distinguishers; you’re either a UCLS (high on KPC1) or you’re not).

****This is a graph mirroring Spirling’s visual – with a trendline for illustrative purposes.

* We’ll be working on adjusting these graphs in a way that can clearly distinguish each individual treaty in a more visually appealing manner.

**Component 2 Graph – the “secondary thing”:**

  
Here’s the “**loadings**” of each treaty on each of the “things”.

* e.g., UCLS09 is very high on **KPC1**, which UCLS22 is the low on.
* e.g., UCLS23 is very high on **KPC2**, which WT02 is lowest on.
* Components might reflect the varying approaches to treaty-making, where some focus more on legal precision and renunciation of claims, while others emphasize the Crown’s paternalistic role and efforts to “civilize” the Indigenous people (see below).

**Component 1 – Formality and Legal Precision vs. Paternalism and Advisory Tone (?)**

**Highest (UCLS09.txt):**

* Highly formal and legalistic, detailing a land transaction between the Chippewa Nation and the British Crown.
  + Precise geographical descriptions, a thorough outline of the transaction’s terms, and a strong emphasis on the renunciation of any future claims by the Chippewa Nation.

**Themes:** Formal land cession, detailed legal language, renunciation of future claims, strong focus on geographic boundaries and legal titles.

**Lowest (UCLS22.txt):**

* Less formal and more advisory in tone, with the speaker (likely a representative of the Crown) addressing the Indigenous people as “My Children” and discussing the necessity of new arrangements due to encroachment by settlers.
  + Emphasizes the paternalistic role of the Crown, the protection of Indigenous people, and the possibility of relocating to areas where they can be “civilized.”

**Themes:** Paternalism, advisory tone, less formal legal language, focus on the relationship and protection.

**Component 1:** might represent a spectrum from formal, legalistic treaties involving precise land cession (high values) to more paternalistic, advisory communications focused on protection and relocation (low values).

**Component 1 – High vs. Low Comparison**

**UCLS09.txt**

To all whom these presents may come, – GREETING:

WHEREAS we the principal Chiefs, Warriors, and People of the Cheppewa Nation of Indians being desirous for a certain consideration hereinafter mentioned of selling and disposing of a certain parcel or tract of land situated and lying on the north side of the River Thames or River La Tranche and known in the Indian name by Escunnisepe unto His Britannic Majesty King George the Third our great Father.

NOW KNOW YE, that we the said principal Chiefs, Warriors and People of the Chippewa Nation for and in consideration of the sum of twelve hundred pounds Quebec currency value in goods estimated according to the Montreal price now delivered to us, the receipt whereof we hereby acknowledge, have give, granted, sold, disposed of and confirmed, and by these presents do give, grant, sell, dispose of and confirm forever unto Alexander McKee, Esquire, Deputy Superintendent General and Deputy Inspector General of Indians and of their affairs on behalf of His said Britannic Majesty King George the Third His heirs and successors, all that parcel or tract of land situated and lying on the north side of the River Thames as aforesaid, beginning at a certain station on the north bank of the said river about nineteen miles above the Deleware Village following the windings of the said river and about twelve miles distant from the said village in a direct northerly course, being about two miles above a lime stone rock and spring on the said river which station will be more perfectly found by a line run from the main or lower fork at London six miles on a course south, sixty-eight degrees thirty miles; thence north sixty-eight degrees thirty minutes east twelve miles’ thence south twenty-one degrees thirty minutes east till it intersects a right line running from the upper forks of the said river at Oxford to the main or lower forks of the said river at London; thence along the said line to the said upper forks on a course north sixty-eight degrees thirty minutes east; thence down the said River Thames following the several winding and courses with the stream to the place of beginning. To have and to hold the said parcel or tract of land together with all the woods and waters thereon situate lying, and being unto the said Alexander McKee, Esquire, for and on behalf of His said Britannic Majesty King George the Third His heirs and successors for ever free and clean of and from all claims, rights, privileges, or emoluments which we the said Chiefs, Warriors and People of the said Chippawa Nation might have before the execution of these presents, and free and clear of any pretended claim which our children or descendants may hereafter make the same, hereby renouncing and forever absolving ourselves, our children, descendents and posterity of all title to the said parcel or tract of land, the soil, wood, and waters thereof, in favour of the said Alexander McKee, Esquire, for and on behalf of His said Britannic Majesty, His heirs and successors for ever.

IN WITNESS WHEREOF, we have for ourselves and the rest of our Nation hereunto set out respective marks and seals this seventh day of September, in the thirty sixth year of the reign of King George the Third, having first heard this instrument openly read and rehearsed in our own language and fully approved by ourselves and our Nation. And in the year of our Lord one thousand seven hundred and ninety six.

Present at the execution and delivery of this instrument, and witnesses thereto:

**UCLS22.txt**

MY CHILDREN:

Seventy snow seasons have now passes away since we met in Council at the crooked place (Niagara), at which time and place your Great Father, the King, and the Indians of North America tied their hands together by the wampum of friendship.

Since that period of various circumstances have occurred to separate from your Great Father many of his red children, and as an unavoidable increase of white population, as well as the progress of cultivation, have had the natural effect of impoverishing your hunting grounds it has become necessary that new arrangements should be entered into for the purpose of protecting you from the encroachment of the whites.

In all parts of the word farmers seek for uncultivated land as eagerly as you, my red children, hunt in your forest for game. If you would cultivate your land it would then be considered your own property, in the same way as your dogs are considered among yourselves to belong to those who have reared them; but uncultivated land is like wild animals, and your Great Father, who has hitherto protected you, has now great difficulty in securing it for you from the whites, who are hunting to cultivate it.

Under these circumstances, I have been obliged to consider what is best to be done for the red children of the forest, and I now tell you my thoughts.

It appears that these islands on which we are now assembled in Council are, as well as all those on the north shore of Lake Huron, alike claimed by the English, the Ottawas, and the Chippewas.

I consider that from the facilities and form their being surrounded by innumerable fishing island, they might be made a most desirable place of residence for many Indians who wish to be civilized, as well as to be totally separated from the whites; and I now tell you that your Great Father will withdraw his claim to these islands and allow them to be applied for that purpose.

Are you, therefore, the Ottawas and Chippewas, willing to relinquish your respective claims to these islands and make them the property (under your Great Father’s control) of all Indians whom he shall allow to reside on them; if so, affix your marks to this my proposal.

MANTIWANING, 9th August, 1836.

**Component 2 – Promises of Protection and Relocation vs. Legal and Financial Agreements (?)**

**Highest vs. Lowest Treaty**

**Highest (UCLS23.txt):**

* Highly formal, describing the surrender of land by the Sauking tribe to the Crown.
  + Includes a proposal for the Sauking to relocate and be provided with houses and resources to become civilized. It emphasizes the Crown’s control over the lands and the promises made to protect the Indigenous people from encroachments.

**Themes:** Land surrender, formal agreements, promises of protection, relocation and civilization efforts.

**Lowest (WT02.txt):**

* Formal and complex, dealing with the rights, titles, and interests of the Mississauga Tribe in Ontario.
  + Extensive legal descriptions of the land, references to previous treaties, and agreements on compensation.
  + Heavily focused on legal terms and conditions, including specific rights, boundaries, and financial compensations.

**Themes:** Formal legal structure, complex legal terminology, detailed land descriptions, financial compensation.

**Component 2:** might represent a spectrum from treaties focusing on promises of protection and relocation under the Crown’s control (high values) to very formal, complex legal agreements involving detailed descriptions of land rights and financial compensations (low values).

**Component 2 – High vs. Low Comparison**

**UCLS23.txt**

To the Sauking:

MY CHILDREN,

You have heard the proposal I have just made to the Chippewas and Ottawas, by which it has been agreed between them and your Great Father that these islands (Manatoulin), on which we are now assembled, should be made, in Council, the property (under your Great Father’s control) of all Indians whom he shall allow to reside on them.

I now propose to you that you should surrender to your Great Father the Sauking Territory you at present occupy, and that you should repair either to this island or to that part of your territory which lies on the north of Owen Sound, upon which proper houses shall be built for you, and proper assistance given to enable you to become civilized and to cultivate land, which you Great Father engages for ever to protect for you from the encroachments of the whites.

Are you there, the Saukin Indians, willing to accede to this arrangement; if so, affix your marks to this my proposal.

MANITOWANING, 9th August, 1836.

**WT02.txt**

ARTICLES OF A TREATY made and concluded on the fifteenth day of November in the year of Our Lord One thousand nine hundred and twenty-three, between His Most Gracious Majesty, George the Fifth, of the United Kingdom of Great Britain and Ireland, King, Defender of the Faith, Emperor of India, by His Commissioners, Angus Seymour Williams, of the City of Ottawa, in the Province of Ontario, Esquire, Barrister-at-law, and Departmental Solicitor of the Department of Indian Affairs: Robert Victor Sinclair, of the said City of Ottawa, Esquire, one of His Majesty’s Counsel, learned in the law, and Uriah McFadden, of the City of Sault Sainte Marie, in the said Province, Esquire, one of His Majesty’s Counsel learned in the law, the said Angus Seymour Williams, Chairman of the said Commission, representing the Dominion of Canada, and the said Robert Victor Sinclair and Uriah McFadden, representing the Province of Ontario, of the One Part, and the members of the Mississauga Tribe, inhabiting, as members of bands thereof, reserves at Rice Lake, Mud Lake, Scugog Lake and Alderville, all in the Province of Ontario, by their chiefs and headmen, of the Other Part.

WHEREAS, the Mississauga Tribe above described, having claimed to be entitled to certain interests in the lands in the Province of Ontario, hereinafter described, such interests being the Indian title of the said tribe to fishing, hunting and trapping rights over the said lands, of which said rights, His Majesty, through His said Commissioners, is desirous of obtaining a surrender, and for such purpose has appointed the said Commissioners, with power on behalf of His said Majesty, to enquire into the validity of the claims of the said tribe, and, in the event of the said Commissioners determining in favour of the validity thereof, to negotiate a treaty with the said tribe for the surrender of the said rights upon the payment of such compensation therefor as may seem to the said Commissioners to be just and proper:=

AND WHEREAS the said Commissioners, having duly made the said enquiry, have determined in favour of the validity of the said rights.

AND WHEREAS the Indians belonging to the said tribe, having been duly convened in Council, at the respective places named hereunder, and having been requested by the said Commissioners to name certain chiefs and headmen to be authorized on their behalf to conduct negotiations with the said Commissioners for a surrender of the said rights and to sign a treaty in respect thereof and to become responsible to His Majesty for the faithful performance by the said tribe and by the respective bands thereof inhabiting the said reserves, of such obligations as shall be assumed by them under such treaty, the said Indians have therefore appointed for the purposes aforesaid the several chiefs and headmen who have subscribed to this treaty:

AND WHEREAS the said Commissioners, acting under the powers in them reposed as aforesaid, have negotiated the present treaty with the said tribe:

NOW THEREFORE THIS TREATY WITNESSETH that the said tribe and the Indians composing the same, occupying as members of bands the said reserves, by their chiefs and headmen, duly authorized thereunto as aforesaid, do hereby cede, release, surrender and yield up to the Government of the Dominion of Canada for His Majesty the King and His Successors forever, all their right, title, interest, claim, demand and privileges whatsoever, in, to, upon, or in respect of the lands and premises described as follows, that is to say:

FIRSTLY: All that parcel of land situate in the Province of Ontario and described as commencing on the northeasterly shore of Georgian Bay at that mouth of the French River which forms the boundary between the District of Parry Sound and the District of Sudbury; thence southerly and easterly along the shores of Georgian Bay to that point on Matchedash Bay where the land included in the surrender of the eighteenth day of November, 1815, of record in Book of Surrenders, Volume One, is reached, and including all the islands in the Georgian Bay waters in which the Indians making this treaty have any interest; thence along the easterly limit of the said lands purchased in 1815 to the Narrows between Lake Couchiching and Lake Simcoe; thence due east across the said Narrows; thence southerly and easterly following the east side of the Narrows and the north shore of Simcoe to the foot of McPhee Bay off the northerly part of Lake Simcoe; thence by a straight line easterly to a point thirty-three miles north of the northwest corner of the Township of Rawdon measured along the division line between the Counties of Hastings and Peterborough, which point is the most western northwest corner of the parcel surrendered on the twenty-eighth day of November, 1822 (noted in Volume One of the Book of Surrenders as number twenty-seven and one-quarter, 27¼); thence following the north and west boundaries of the last mentioned parcel to the Ottawa River; thence westerly along the interprovincial boundary to the mouth of the Mattawa River; thence westerly along the interprovincial boundary to the mouth of the Mattawa River, Talon Chute and Talon Lake, Turtle Lake, and Trout Lake to the westerly point of Trout Lake; thence to the shore of Lake Nipissing at North Bay; thence by the north shore of Lake Nipissing to the French River; thence by those waters along the division line between the Districts of Parry Sound and Sudbury to the place of commencement: Excepting thereout and therefrom those lands which have already been set aside as Indian reserves. The parcel hereby surrendered contains seventeen thousand, six hundred square miles, more or less.

SECONDLY: All that parcel of land situate in the Province of Ontario and described as parts of the Counties of Northumberland, Durham, Ontario and York, commencing at the point where the easterly limit of that portion of the lands said to have been ceded in 1787, which was confirmed on the first of August, 1805, of record as number thirteen in Volume One, of the Book of Surrenders, intersects the northerly shore of Lake Ontario; thence northerly along the said easterly and northerly limits of the confirmed tract to the Holland River; thence northerly along the Holland River and along the westerly shore of Lake Simcoe and Kempenfelt Bay to the Narrows between Lake Couchiching and Lake Simcoe; thence southeasterly along the shores of Lake Simcoe to the Talbot River; thence easterly along the Talbot River to the boundary between the Counties of Victoria and Ontario; thence southerly along that boundary to the northwest angle of the Township of Darlington; thence along the northerly boundary line of the Townships of Darlington, Clarke, Hope and Hamilton to Rice Lake; thence along the southern shore of said lake to River Trent along the River Trent to the Bay of Quinte; thence westerly and southerly along the shore of the Bay of Quinte to the road leading to Carrying Place and Weller’s Bay; thence westerly along the northern shore of Lake Ontario to the place of beginning: Excepting thereout and therefrom those lands which have already been set aside as Indian Reserves. The land hereby conveyed contains two thousand, five hundred square miles more or less.

AND ALSO all the right, title, interest, claim, demand and privileges whatsoever of the said Indians, in, to, upon or in respect of all other lands situate in the Province of Ontario to which they ever had, now have, or now claim to have any right, title, interest, claim, demand or privileges, except such reserves as have heretofore been set apart for them by His Majesty the King.

TO HAVE AND TO HOLD the same to His Majesty the King and His Successors forever:

AND THIS TREATY FURTHER WITNESSETH that in consideration of the aforesaid surrender, His Majesty, through His said Commissioners, hereby agrees, upon the execution of a treaty similar to this treaty by the Chippewa Tribe inhabiting as members of bands, reserves at Christian Island, Georgina Island and Rama, in the Province of Ontario, to pay to each member of the said Mississauga Tribe, being also a member of one of the said bands, the sum of twenty-five dollars, to be paid through the Indian agents for the respective bands, with a reasonable time after the execution of the said treaties, and a further sum of ---233,425.00 dollars--- to be administered for the said tribe by His Majesty’s Department of Indian Affairs under and pursuant to the provisions of the Indian Act, Revised Statutes of Canada, 1906, Chapter Forty-three and its amendments; making together the sum of 250,000.00 dollars.

AND THE UNDERSIGNED chiefs and headmen, on their own behalf and on behalf of all the Indians whom they represent, do hereby solemnly covenant, promise and agree to strictly observe this treaty in all respects and that they will not, nor will any of them, nor will any of the Indians whom they represent, molest or interfere with the person or property of anyone who now inhabits or shall hereafter inhabit any portion of the lands covered by this treaty, or interfere with, trouble, or molest any person passing or travelling through the said lands or any part thereof, and that they will assist the officers of His Majesty in bringing to justice and punishment any Indian, party to this treaty, who may hereafter offend against the stipulations hereof or infringe the laws in force in the lands covered hereby:

AND IT IS FURTHER UNDERSTOOD that this treaty is subject to an agreement dated the day of April, A.D. 1923, made between the Dominion of Canada and the Province of Ontario, a copy of which is hereto attached.

IN WITNESS WHEREOF, His Majesty’s said Commissioners and the said chiefs and headmen have hereunto set their hands and seals at the places and times hereinafter set forth, in the year herein first above written.

SIGNED AND SEALED at Alderville on the nineteenth day of November, A.D. 1923, by His Majesty’s Commissioners and the undersigned chiefs and headmen in the presence of the undersigned witnesses, after first having been interpreted and explained.

SIGNED AND SEALED at Mud Lake on the fifteenth day of November, A.D. 1923, by His Majesty’s Commissioners and the undersigned chiefs and headmen in the presence of the undersigned witnesses, after first having been interpreted and explained.

SIGNED AND SEALED at Rice Lake on the sixteenth day of November, A.D. 1923, by His Majesty’s Commissioners and the undersigned chiefs and headmen in the presence of the undersigned witnesses, after first having been interpreted and explained.

SIGNED AND SEALED at Scugog Lake on the twenty-first day of November, A.D. 1923, by His Majesty’s Commissioners and the undersigned chiefs and headmen in the presence of the undersigned witnesses, after first having been interpreted and explained.

MEMORANDUM OF AGREEMENT made this of April, 1923,

BETWEEN:

The Government of the Dominion of Canada, acting herein by the Honourable Charles Stewart, Superintendent General of Indian Affairs, of the first part,

AND

The Government of the Province of Ontario, acting herein by the Honourable Beniah Bowman, Minister of Lands and Forests for the said Province, of the other part

WHEREAS certain Indians of the Chippewa and Mississauga tribes claim that the said tribes were and are entitled to a certain interest in lands in the Province of Ontario to which the Indian title has never been extinguished by surrender or otherwise, the said lands being described as parts of the counties of Renfrew, Hastings, Haliburton, Muskoka, Parry Sound and Nipissing, and being bounded on the south and east by the lands included in the surrenders of the Indian title made on the 18th of November, 1815, the 5th of November, 1818, and November, 1822; on the north by the Ottawa and Mattawa Rivers and Lake Nipissing, and on the west by the lands included in the surrender of the Indian title made in 1850, known as the Robinson-Huron surrender, and by the Georgian Bay, the area in question including about 10,719 square miles.

AND WHEREAS a departmental enquiry made by the Department of Indian Affairs indicates that the said claim has such probable validity as to justify and require further investigation, and if found valid to be satisfied on such just and fair terms as may be settled by a treaty of surrender.

NOW THEREFORE THIS AGREEMENT made in pursuance of certain statutes of Canada and of the Province of Ontario, both entitled “an Act for the settlement of certain questions between the Governments of Canada and Ontario respecting Indian Lands,” the Statute of Canada having been passed in the 54th and 55th years of the reign of Her Majesty Queen Victoria and chaptered 5, and the statute of Ontario in the 54th year of Her Majesty’s said reign and chaptered 3.

WITNESSETH THAT the Governments of Canada and of the Province of Ontario have agreed as follows: ---

The Government of Canada will, pursuant to Part I of the Enquiries Act, R.S.C., 1906, c. 104, and amendments, appoint three persons as commissioners to enquire into the validity of the claim of the Chippewa and Mississauga Indians aforesaid, and will empower the said commissioners, in the event of their determining in favour of the validity of the said claim, to negotiate a treaty with the said Indians for the surrender of the said lands upon payment of such compensation as may be fixed by such treaty.

Of the three commissioners so named, one shall be selected by the Government of Canada, who shall be Chairman of the Commission, and the remaining two shall be selected by the Minister of Lands and Forests for the Province of Ontario and notified to the Superintendent General of Indian Affairs.

The question of the validity of said claim may be determined by any two of the said Commissioners and it shall be necessary that at least two of them of whom the chairman shall be one shall concur in any treaty which may be negotiated.

The expenses of the said commission, including the remuneration and expenses of the commissioners and any expenses incurred for securing the attendance of witnesses or otherwise, shall be payable by the Government of Canada, but the rates of remuneration of each of the commissioners selected by the Minister of Lands and Forests for the Province of Ontario shall be agreed upon between him and the Superintendent General of Indian Affairs before the Constitution of the Commission.

In the event of the commissioners negotiating a treaty with the Indians the compensation to be paid to such Indians shall be payable to the Dominion of Canada by the Province of Ontario from time to time in accordance with the terms of the treaty of surrender, and shall be applied by the Dominion of Canada in accordance with the said terms.

In the event of provision being made by such treaty of surrender for the setting apart of reserves for the Indians, the Dominion of Canada will bear the expense to be incurred in the location and survey thereof, and the Province of Ontario will concur in the setting apart of such reserves.

All such reserves shall be administered by the Dominion of Canada for the benefit of the band or bands of Indians to which each may be allotted; portions thereof may, upon their surrender for the purpose by the said band or bands, be sold, leased or otherwise disposed of by letters patent under the Great Seal of Canada, and the proceeds of such sale, lease or other disposition applied for the benefit of such band or bands, provided, however, that in the event of the band or bands to which any such reserve has been allotted becoming extinct, or if for any other reason such reserve or such portion thereof as remains undisposed of is declared by the Superintendent General of Indian Affairs to be no longer required for the benefit of the said band or bands, the same shall thereafter be administered by and for the benefit of the Province of Ontario, and any balance of the proceeds of the sale or other disposition of any portion thereof then remaining under the control of the Dominion of Canada shall, so far as the same is not still required to be applied for the benefit of the said band or bands of Indians, be paid to the Province of Ontario, together with accrued unexpended simple interest thereon.

IN WITNESS WHEREOF these presents have been signed by the parties thereto.

**How do we figure out what the “things” are? – we combine 2 methods:**

1. **Vector-space analysis: interpret string kernels directly**
   * We’ve done this already to determine the **KPCA**, but we could analyze the string kernels directly. However, academics add on a 2nd method to interpret string-kernels because it provides more nuance. Thus, we add in…
2. **Term Document Matrix conversion (stemming, etc.)**
   * Take the words and break them up into root components.
   * Eliminate common words (said, the, etc.)
   * Remove sparse terms (in our case, if the words don’t appear in 90% of documents they aren’t considered).
   * Remove punctuation.

The argument is that using string kernels gives a more accurate *overall* representation of what the things are – because when you use **string kernels** the algorithm *preserves and considers the order that word appear in*. However, it’s easier to add on a **TDM-IDF** process to assess each word’s importance in a vacuum.

* Once we’ve stemmed, removed common words, etc. we feed them to two algorithms – **Random Forest** (Spirling) and **xgboost** (new age extremely powerful competition-winning predictive algorithm that uses Random Forest and **Gradient Descent**).

**Importance Algorithms:**

1. **Random Forest:** you can think of the entire sample of the treaties as the “forest”. The algorithm plants 500-2000 “trees” at random points (word stems) in each treaty. Each tree makes a predictive “bucketing” decision (simplified below) extending throughout the entire corpus before the algorithm aggregates its calculations.
   * e.g., Planting a “**punish**” tree: In Numbered Treaty 10, “punish” might occur frequently, indicating it is high in the “thing.” Conversely, the Douglas Treaties might have few occurrences of “punish,” indicating they are low in the same “thing.”
     + The algorithm then asks: “If we split the treaties based on the occurrence of ‘punish,’ do they fall into the correct buckets?”

Numbered Treaty 10 **SHOULD** go into the bucket of “treaties high in thing,” and the Douglas Treaties would go into the bucket of “low in thing.”

If **punish** is **important**, we should see proper separation = punish is important:

If **punish** is **not important**, we might see inaccurate separation; we already know Douglas Treaties are low in thing, but “punish” is putting them in the wrong bucket = punish not important:

A scrolled paper with red trim

Description automatically generatedA person with arms spread out

Description automatically generated

**Douglas Treaties:  
 “thing” very low**

**Numbered Treaty 10:  
 “thing” very high**



[This Photo](https://www.pngall.com/scroll-png/) by Unknown Author is licensed under [CC BY-NC](https://creativecommons.org/licenses/by-nc/3.0/)

**Numbered Treaty 10**



[This Photo](https://www.pngall.com/scroll-png/) by Unknown Author is licensed under [CC BY-NC](https://creativecommons.org/licenses/by-nc/3.0/)



[This Photo](https://www.pngall.com/scroll-png/) by Unknown Author is licensed under [CC BY-NC](https://creativecommons.org/licenses/by-nc/3.0/)

**Douglas Treaties**

**Low in “thing” bucket**

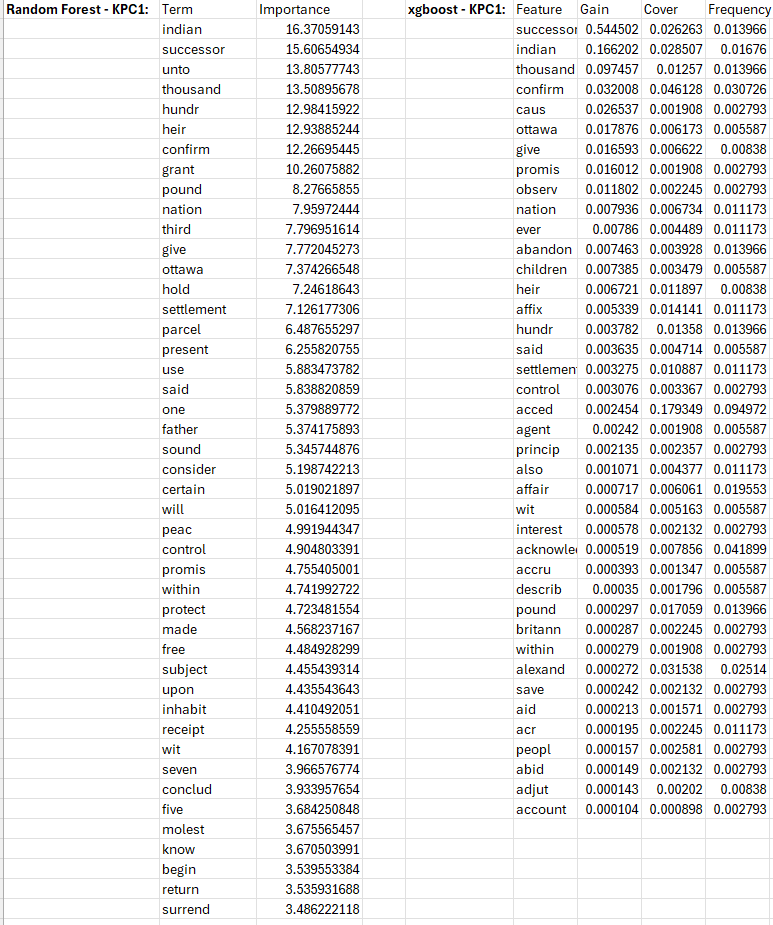
**High in “thing” bucket**

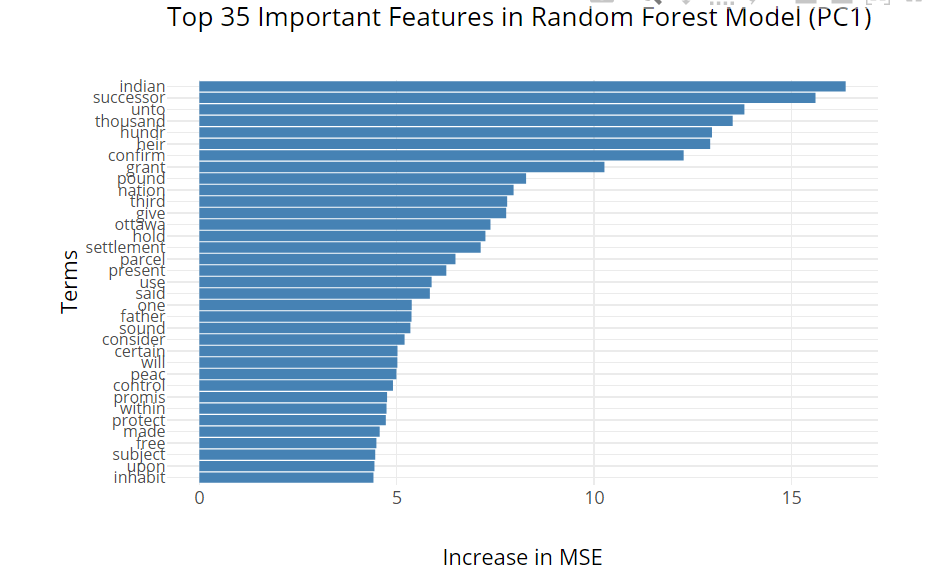
**High in “thing” bucket**

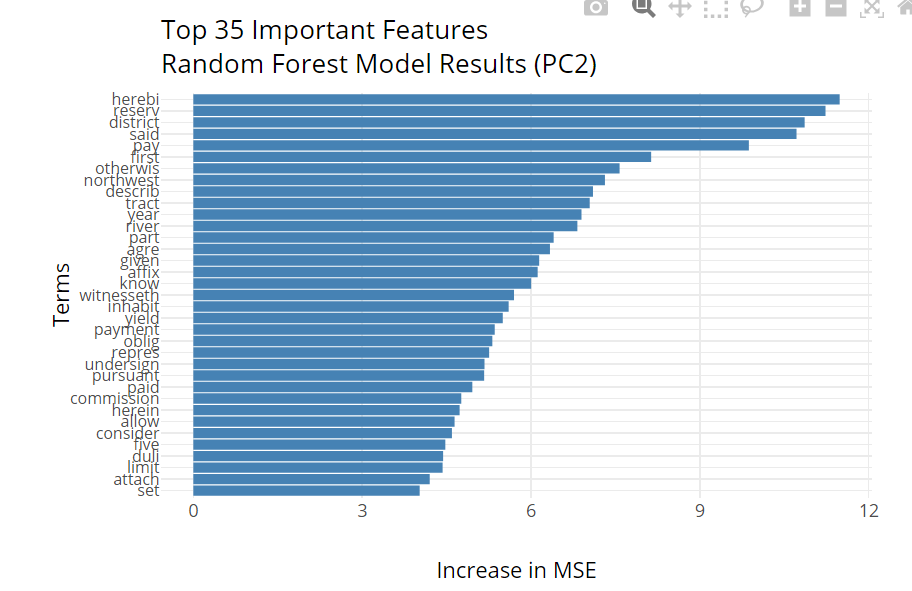
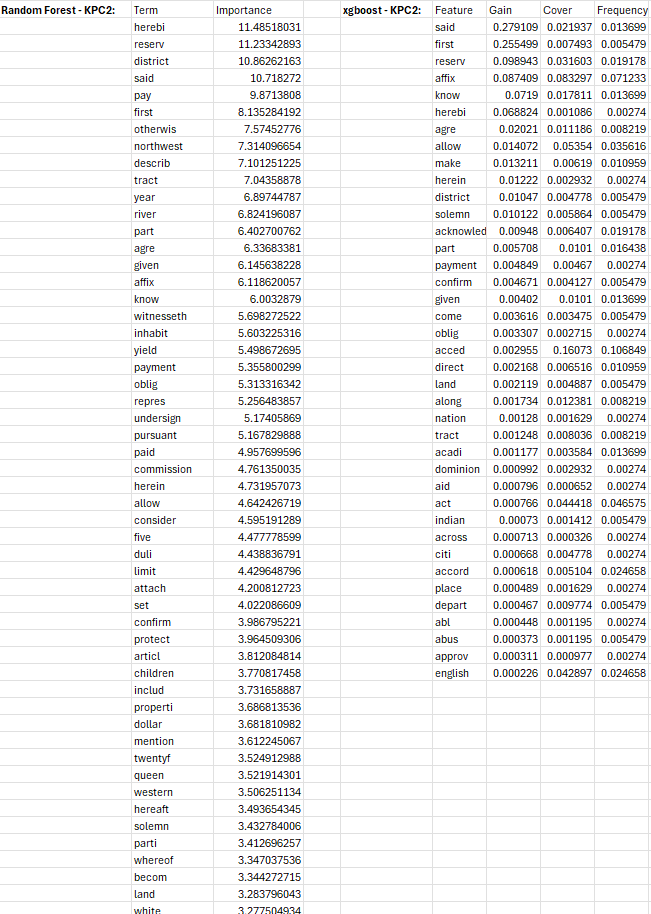
**Low in “thing” bucket**

1. **XGBoost:** Extreme Gradient Boosting builds a model in a stage-wise fashion, where each new tree corrects errors made by previous trees.
   * XGBoost starts with a simple initial model, perhaps predicting the average value of “thing” across all treaties.
   * It calculates the difference (residual) between the actual values of “thing” (e.g., occurrences of important terms) and the predictions made by the initial model for each treaty.
   * A new decision tree is created to predict these residuals.

* e.g., Planting a “**punish**” tree:
  + If “punish” appears often in a treaty, the tree might predict a higher residual (indicating the initial model underestimated the “thing” for this treaty).
  + If “punish” is rare, the tree might predict a lower residual (indicating the initial model overestimated the “thing”).
* The predictions from this new tree are added to the initial model to improve its accuracy.
  + This process of calculating residuals, creating new trees, and updating the model continues iteratively, with each new tree focusing on the remaining errors from the previous iteration.
  + The final model is an ensemble of all the trees, where each tree contributes to refining the predictions.

** Algorithm Results – KPC1 (“main thing”):**

* ****Here we have a graph of the most important features from KPC1 (the words that best separate the treaties into the correct “main thing” buckets), and the corresponding algorithms scores.

**Algorithm Results – KPC2 (“secondary thing”):**

* Here we have a graph of the most important features from KPC2 (the words that best separate the treaties into the correct “secondary thing” buckets), and the corresponding algorithms scores.

**Component (“thing”) Correlations:**

When **stemming** words, we erase word suffixes to obtain the “root” of the word.



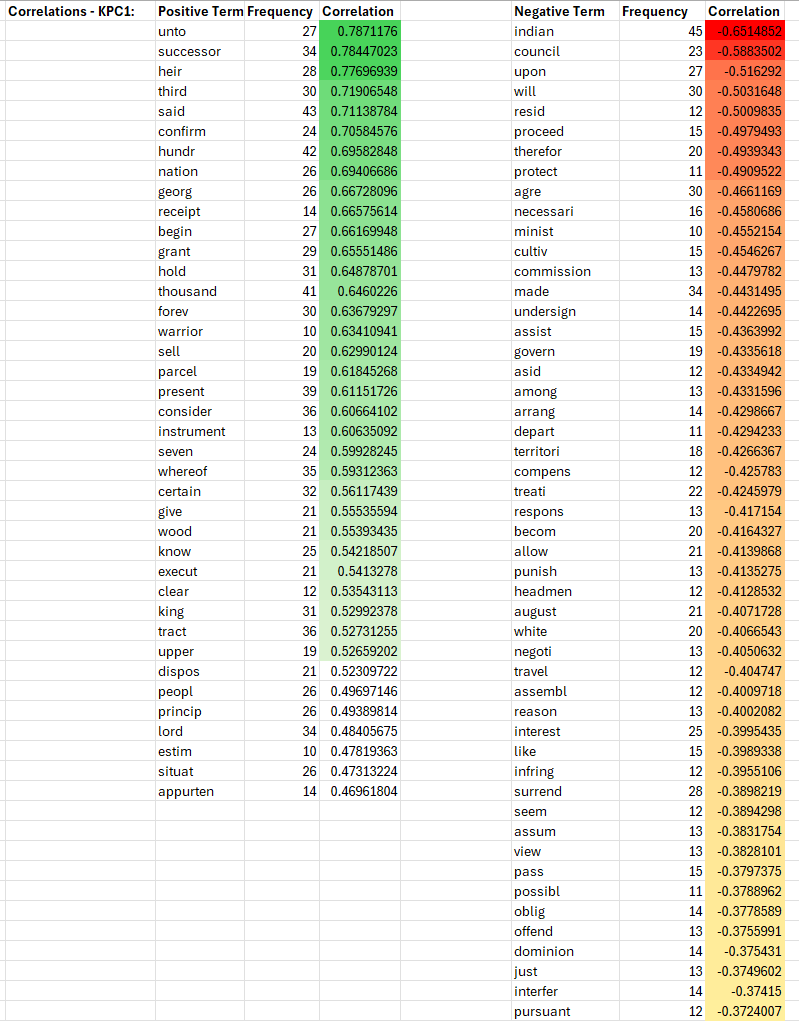
**Word stems with a high positive correlation** appear more frequently/are more prominent in treaties that have high scores on the principal component.

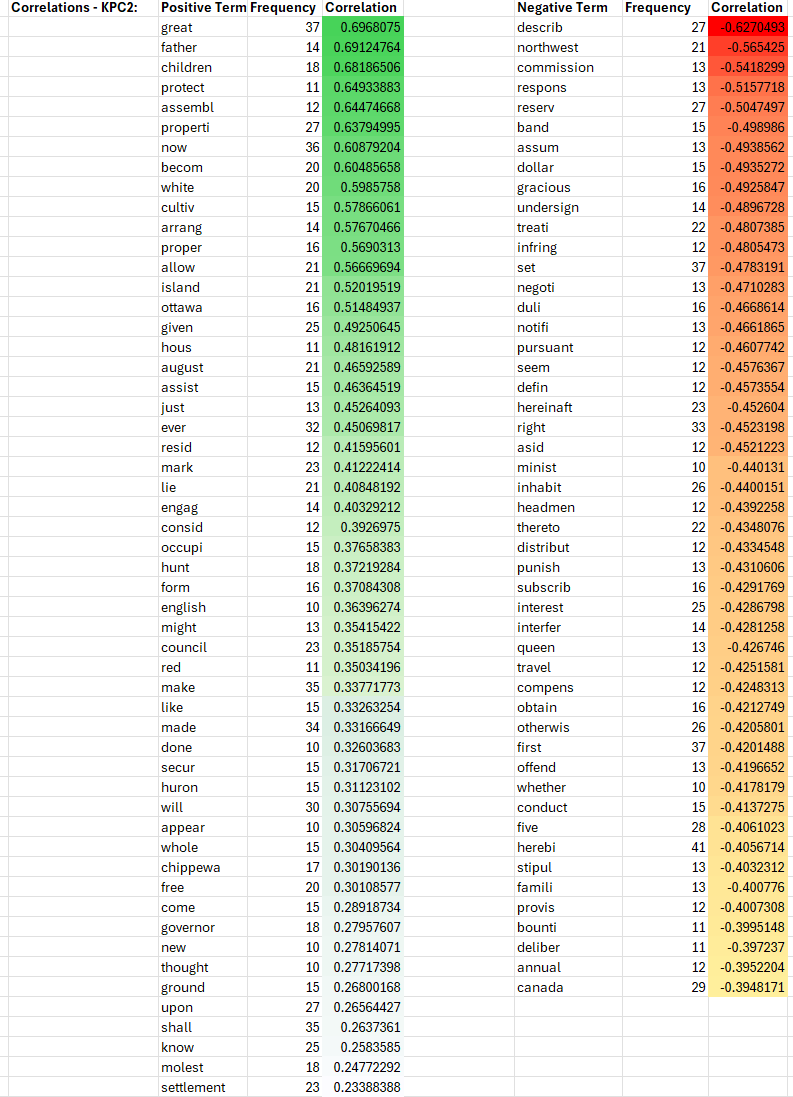
* These stems are characteristic of whatever “thing” the KPC is capturing when it scores high.

**Word stems with a high negative correlation** appear more frequently or are more prominent in treaties that have low scores on the principal component.

* These stems are characteristic of treaties that represent the opposite end of whatever the “thing” we’ve captured is.

**Word Stem correlations can be treated as Kernel Principal Component Loadings** – how much each word affects the “thing(s)”.

****There are 3677 terms captured in the documents (which we have the correlations for) so I’ll just include the most important ones for each of KPC1 & KPC2.

****