

---

# PROGRAMACIÓN CIENTÍFICA

*Trabajo Final*

Autor:

Alanis Fernandez, Eder Ismael

Del Toro Peña, Arnoldo

19 de octubre de 2021

# Introducción

Como lo mencionaba el psicólogo estadounidense Abraham Maslow, uno de los pilares de nuestras necesidades mas basicas es el alimento. Por tal motivo una de las prioridades de un gobierno prospero o que tiene como objetivo seguir en el poder es tener la capacidad de satisfacer las necesidades de sus ciudadanos entre las que van incluidas las alimentarias y los productos derivados de los animales como por ejemplo pieles, cintos de piel, lana de oveja es decir la producción pecuaria. El gobierno de México a tenor de su transparencia con los ciudadanos a puesto en disposición del público en general los datos historicos concretamente relacionados con la producción pecuaria en este pais para los años 1980 hasta 2020.

El proposito de este trabajo consite en hacer diferentes analisis estadisticos empleando varias herramientas tecnologicas entre las que mencionamos a Orange y OverLeaf y como datos usaremos las bases de datos de producción pecuaria como mencionamos arriba, cabe señalar que estas bases de datos vienen en 2 apartados, 1980 a 2005 y de 2006 a 2020. Cabe aclarar que el uso de Orange en si es para el analisis y transformación de los datos, el Overleaf sera nuestro editor de textos y el github sera para poder hacer diversos cambios contando con la seguridad de poder regresarnos si cometemos uno o varios errores.

# Objetivo

1. Ordenar las entidades federativas según el crecimiento promedio para cada año durante el periodo que se estudia para la especie “Bovina”.
2. Ordenar las entidades federativas según el valor promedio de su producción pecuaria para cada año el periodo que se estudia para la especie “Bovina”.
3. Grafique la relación entre los valores para el crecimiento promedio y el valor promedio establecidos en las tareas anteriores para cada una las entidades: Veracruz, Sonora, Tamaulipas, Sinaloa, Oaxaca. Que observa?
4. Determine coeficientes de correlación de Spearman para cada entidad federativa seleccionada en la actividad anterior. Que puede decir?
5. Estime el crecimiento de la producción bovina de las entidades federativas Veracruz y Tamaulipas para el año 2020 (empleando los datos hasta 2019). Cuan precisa es la estimación al comparar con los datos reales de 2020?

# Hipótesis

No existe una correlación entre ninguno de los valores de la producción agrícola.

# Evaluación de Riesgos

Estimar cuanto vamos a producir de manera errónea puede tener graves consecuencias tanto si estimamos de más, como de menos.

Si nuestras estimaciones son bajas, y la producción del próximo año está basada en cuanto vamos a necesitar, dado que por error produciremos menos, y la demanda del producto es muy superior habrá incremento de los precios, escasez de productos y descontento por parte de la población hacia el gobierno y los productores locales y extranjeros.

En dado caso de que produzcamos más de lo que se vaya a necesitar, es decir hay más oferta que demanda, cada productor al tratar de vender sus productos por sobre la competencia tendrá que bajar sus precios para ser más atractivo es decir como diferenciador de la competencia, esta práctica al ser replicada por los demás traerá pérdidas significativas para los productores.

Notese que esto tiene mayor peso (este tipo de prácticas) en productos perecederos por que tiene un tiempo muy limitado en el que pueden ser comercializados factores externos como la refrigeración y/o congelación aumentan su costo de producción y dado que se está manejando un precio bajo para venderse, no ayuda en nada tener costos más altos en los productos.

## Método

Como ya se mencionó antes, utilizaremos el software **Orange** (en su versión gratuita) para estimar el año 2020 (datos ya existentes), además compararemos algunos estados, tanto específicamente en años como en precios.

## Materiales utilizados

Los materiales utilizados se mencionarán a continuación:

1. El complemento para Google Chrome Batch link Downloader.
2. Software Libre Calc.
3. Orange (Versión Libre).
4. Git.
5. Overleaf.

## Pasos

Los pasos que seguimos se enumerarán enseguida, estos van en orden a los objetivos.

1. **Creación de un repositorio:** Primero lo que hicimos fue crear un repositorio en la página de Git-Hub, para suerte ya contabamos con una, después creamos un proyecto nuevo junto con una rama, cada uno siguió una rama durante el resto del proyecto.
2. **Creación del Documento en Overleaf:** Arnold por su parte creó una cuenta en Overleaf, yo por mi parte ya tenía una, después procedimos a crear un documento.
3. **Descarga de Documentos:** En este paso descargamos los documentos con la ayuda del ya mencionado Batch link Downloader y lo guardamos en una carpeta llamada data.
4. **Orange:** En este paso fue en el que estuvimos más tiempo entretenidos, primero creamos un archivo nuevo y con la ayuda de multifile cargamos todos los archivos que habíamos descargado, una vez cargados procedimos a resolver cada uno de los objetivos, que veremos enseguida:

Objetivo 1. Para el objetivo uno que es: **Ordenar las entidades federativas según el crecimiento promedio para cada año durante el periodo que se estudia para la especie “Bovina”**, tuvimos un problema que era que al parecer algunos de los archivos tenían diferentes datos, lo cual ya se nos había recomendado dividir en dos grupos, pero después nos encontramos con otro problema con que algunos de los datos de Asacrificados estaban divididos, por lo cual Orange lo tomaba como datos perdidos, por lo cual seleccionando filas y columnas terminamos por concatenar ambas columnas en una sola; sin embargo nos surgió otro problema el cual era que algunos de los nombres no estaban bien escritos (debido a los acentos) entonces utilizando la búsqueda de Orange encontramos los años en los cuales estaban mal escritos y los corregimos, después de estos procedimientos por fin teníamos un data table sin datos perdidos, por lo cual pasamos a descargar el add-ons Spectroscopy, para obtener los promedios de nuestra tabla.

Antes de obtener los promedios filtramos por especie.

Se adjuntan las imágenes de lo mencionado anteriormente, en la figura 1 podemos observar el proceso en **Orange** de la concatenación y filtrado para obtener todos los datos correspondientes mientras que en la figura

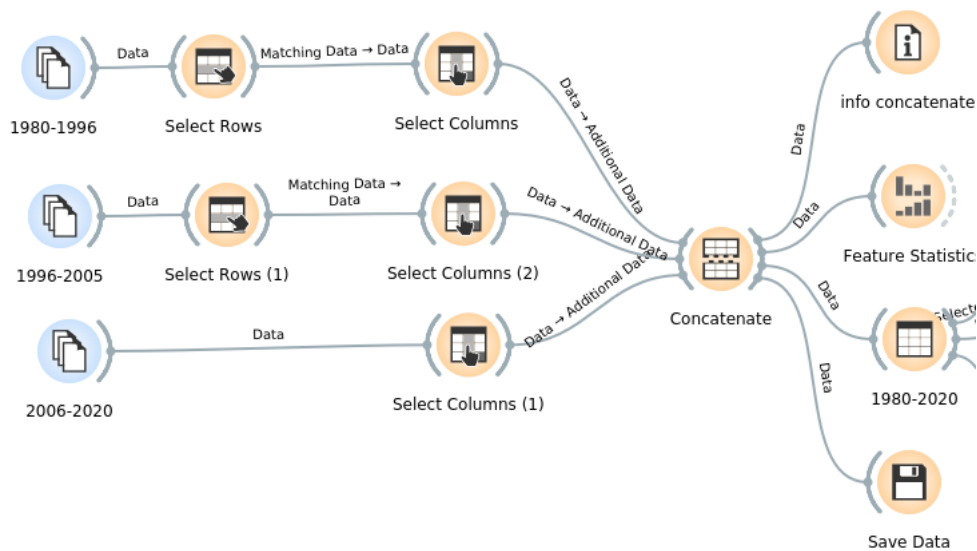


Figura 1: Concatenación de archivos

Estado											
Info	Anio	Cveestado	Cveespecie	Nomespecie	Precio	Nomestado	Cveproducto	Nomproducto	Volumen	Valor	Asacrificado
34 instances (no missing data) 11 features No target variable. No meta attributes	34 2018.50	9	4	Bovino	37.5381	Ciudad de M...	3.33	Ganado en pie	685.03290	8829.43281	121.76190
	33 2017	9	4	Bovino	37.8657	Ciudad de M...	3.33	Ganado en pie	756.21790	9388.82014	126.80952
	9 2006.43	9	4	Bovino	19.2542	Distrito Federal	3.33	Ganado en pie	5167.52543	10751.36412	260.44118
	16 2011.47	17	4	Bovino	22.6772	Morelos	3.33	Ganado en pie	1043.80942	8379.76003	398.81725
	20 2015.30	20	4	Bovino	27.2492	Oaxaca	3.33	Ganado en pie	762.08845	8414.25556	411.56947
	21 2012.33	21	4	Bovino	22.3871	Puebla	3.08	Ganado en pie	2016.94740	12101.63403	415.59012
	29 2011.92	29	4	Bovino	24.5524	Tlaxcala	3.33	Ganado en pie	1824.15659	11773.46078	437.06898
	17 2012.21	15	4	Bovino	26.2783	México	3.33	Ganado en pie	4431.55842	25643.70289	804.97288
	13 2012.08	13	4	Bovino	27.0719	Hidalgo	3.28	Ganado en pie	4608.83561	29275.76007	883.83886
	31 2011.98	31	4	Bovino	31.3751	Yucatán	2.31	Cera	1147.09545	16720.15814	1045.40824
	23 2008.93	23	4	Bovino	28.0038	Quintana Roo	2.86	Cera	1322.32256	19948.37586	1115.46750
	12 2012.05	12	4	Bovino	24.9613	Guerrero	3.18	Ganado en pie	2177.50586	24855.85066	1428.94514
	15 2012.20	16	4	Bovino	23.3506	Michoacán	3.33	Ganado en pie	3804.78164	32876.11403	1568.06286
	32 2011.91	32	4	Bovino	24.5967	Zacatecas	3.33	Ganado en pie	4089.29342	40348.94210	2066.69777
	24 2011.92	24	4	Bovino	26.6976	San Luis Potosí	3.32	Ganado en pie	4606.58489	57111.58817	2124.07607
	8 2009.45	6	4	Bovino	24.9777	Colima	3.33	Ganado en pie	5136.55297	52705.86633	2125.06667
	11 2011.67	11	4	Bovino	25.7749	Guanajuato	3.29	Ganado en pie	14287.80679	82355.26434	2272.44878
	5 2012.20	7	4	Bovino	20.3219	Chiapas	3.31	Ganado en pie	4602.50204	39495.18068	2350.98865
	3 2006.93	3	4	Bovino	20.5445	Baja Californi...	3.20	Ganado en pie	6292.55593	57299.20412	2556.14022
	19 2011.62	19	4	Bovino	27.8592	Nuevo León	3.08	Ganado en pie	3018.57354	51466.34904	2578.35559
	30 2012.33	30	4	Bovino	24.3497	Veracruz	3.35	Ganado en pie	5457.75810	57038.29989	2662.92220
	18 2010.86	18	4	Bovino	24.8786	Nayarit	3.31	Ganado en pie	4930.25469	48036.86987	2849.72627
	26 2012.03	26	4	Bovino	31.9500	Sonora	3.18	Ganado en pie	4066.78888	55839.65352	3142.98382
	22 2010.62	22	4	Bovino	25.4741	Querétaro	3.24	Ganado en pie	13684.42955	107914.82333	3293.35616
	28 2011.60	28	4	Bovino	26.0136	Tamaulipas	3.08	Ganado en pie	4435.38115	57980.31384	3495.24709
	7 2011.53	5	4	Bovino	24.7924	Coahuila	3.25	Ganado en pie	23539.75914	131599.14592	3501.95438
	6 2011.94	8	4	Bovino	29.0510	Chihuahua	2.95	Cera	13727.56608	101050.25365	3608.11939
	1 2009.69	1	4	Bovino	23.9913	Aguascalientes	3.33	Ganado en pie	25212.11668	159235.51105	3612.77654
	14 2012.21	14	4	Bovino	26.2094	Jalisco	3.33	Ganado en pie	15229.86704	109869.54443	3781.66609
	4 2010.03	4	4	Bovino	20.0568	Campeche	3.24	Ganado en pie	5417.53658	62220.00306	3807.99002
	10 2011.56	10	4	Bovino	25.5957	Durango	3.30	Ganado en pie	21024.12393	141954.98538	5466.27009
	25 2010.92	25	4	Bovino	22.3668	Sinaloa	3.34	Ganado en pie	11613.05246	162905.80280	8179.99161
	27 2010.59	27	4	Bovino	22.4743	Tabasco	3.33	Ganado en pie	14110.76351	158507.79690	8485.91128

Figura 2: Estados Federativos ordenados por Sacrificados

Anio	Cveestado	Cveespecie	Nomespecie	Precio	Nomestado	Cveproducto	Nomproducto	Volumen	Valor	Asacrificado
16	2011.47	17	4 Bovino	22.6772	Morelos	3.33	Ganado en pie	1043.80942	8379.76003	398.81725
20	2015.30	20	4 Bovino	27.2492	Oaxaca	3.33	Ganado en pie	762.08845	8414.25556	411.56947
34	2018.50	9	4 Bovino	37.5381	Ciudad de M...	3.33	Ganado en pie	685.03290	8829.43281	121.76190
33	2017	9	4 Bovino	37.8657	Ciudad de M...	3.33	Ganado en pie	756.21790	9388.82014	126.80952
9	2006.43	9	4 Bovino	19.2542	Distrito Federal	3.33	Ganado en pie	5167.52543	10751.36412	260.44118
29	2011.92	29	4 Bovino	24.5524	Tlaxcala	3.33	Ganado en pie	1824.15659	11773.46078	437.06898
21	2012.33	21	4 Bovino	22.3871	Puebla	3.08	Ganado en pie	2016.94740	12101.63403	415.59012
31	2011.98	31	4 Bovino	31.3751	Yucatán	2.31	Cera	1147.09545	16720.15814	1045.40824
23	2008.93	23	4 Bovino	28.0038	Quintana Roo	2.86	Cera	1322.32256	19948.37586	1115.46750
12	2012.05	12	4 Bovino	24.9613	Guerrero	3.18	Ganado en pie	2177.50586	24855.85686	1428.94514
17	2012.21	15	4 Bovino	26.2783	México	3.33	Ganado en pie	4431.55842	25643.70289	804.97288
13	2012.08	13	4 Bovino	27.0719	Hidalgo	3.28	Ganado en pie	4608.83561	29275.76007	883.83886
15	2012.20	16	4 Bovino	23.3506	Michoacán	3.33	Ganado en pie	3804.78164	32876.11403	1568.06286
5	2012.20	7	4 Bovino	20.3219	Chiapas	3.31	Ganado en pie	4602.50204	39495.18068	2350.98865
32	2011.91	32	4 Bovino	24.5967	Zacatecas	3.33	Ganado en pie	4089.29342	40348.94210	2066.69777
18	2010.86	18	4 Bovino	24.8786	Nayarit	3.31	Ganado en pie	4930.25469	48036.86987	2849.72627
19	2011.62	19	4 Bovino	27.8592	Nuevo León	3.08	Ganado en pie	3018.57354	51466.34904	2578.35559
8	2009.45	6	4 Bovino	24.9777	Colima	3.33	Ganado en pie	5136.55297	52705.86633	2125.06667
26	2012.03	26	4 Bovino	31.9500	Sonora	3.18	Ganado en pie	4066.78888	55839.65352	3142.98382
30	2012.33	30	4 Bovino	24.3497	Veracruz	3.35	Ganado en pie	5457.75810	57038.29989	2662.92220
24	2011.92	24	4 Bovino	26.6976	San Luis Potosí	3.32	Ganado en pie	4606.58489	57111.58817	2124.07607
3	2006.93	3	4 Bovino	20.5445	Baja Californi...	3.20	Ganado en pie	6292.55593	57299.20412	2556.14022
28	2011.60	28	4 Bovino	26.0136	Tamaulipas	3.08	Ganado en pie	4435.38115	57980.31384	3495.24709
4	2010.03	4	4 Bovino	20.0568	Campeche	3.24	Ganado en pie	5417.53658	62220.00306	3807.99002
11	2011.67	11	4 Bovino	25.7749	Guanajuato	3.29	Ganado en pie	14287.80679	82355.26434	2272.44878
6	2011.94	8	4 Bovino	29.0510	Chihuahua	2.95	Cera	13727.56608	101050.25365	3608.11939
22	2010.62	22	4 Bovino	25.4741	Querétaro	3.24	Ganado en pie	13684.42955	107914.82333	3293.35616
14	2012.21	14	4 Bovino	26.2094	Jalisco	3.33	Ganado en pie	15229.86704	109869.54443	3781.66609
7	2011.53	5	4 Bovino	24.7924	Coahuila	3.25	Ganado en pie	23539.75914	131599.14592	3501.95438
10	2011.56	10	4 Bovino	25.5957	Durango	3.30	Ganado en pie	21024.12393	141954.98538	5466.27009
27	2010.59	27	4 Bovino	22.4743	Tabasco	3.33	Ganado en pie	14110.76351	158507.79690	8485.91128
1	2009.69	1	4 Bovino	23.9913	Aguascalientes	3.33	Ganado en pie	25212.11668	159235.51105	3612.77654
25	2010.92	25	4 Bovino	22.3668	Sinaloa	3.34	Ganado en pie	11613.05246	162905.80280	8179.99161

Figura 3: Estados Federativos ordenados por valor

Objetivo 2. Para el objetivo 2 que es: **Ordenar las entidades federativas según el valor promedio de su producción pecuaria para cada año el periodo que se estudia para la especie “Bovina”,** hicimos el mismo procedimiento que el anterior, obteniendo claro la siguiente imagen en la figura 3.

Objetivo 3. En el objetivo 3:

**Grafique la relación entre los valores para el crecimiento promedio y el valor promedio establecidos en las tareas anteriores para cada una las entidades: Veracruz, Sonora, Tamaulipas, Sinaloa, Oaxaca. Que observa?**, primero tuvimos que filtrar los estados por medio de **Orange** para solo seleccionar los estados que nos pedían.

Una vez que teníamos el filtro procedimos a obtener la siguiente gráfica en la cual se ve que: Sinaloa fue el estado con un promedio más alto tanto en producción como en valor, enseguida se adjuntan las imágenes de lo obtenido.

Se observa que Sinaloa es el mayor en ambas gráficas.

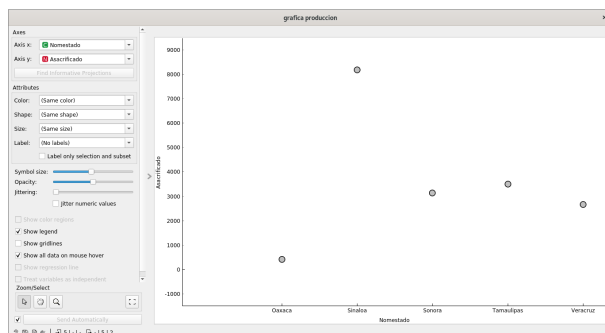


Figura 4: Gráfica Estados ordenados por sacrificio

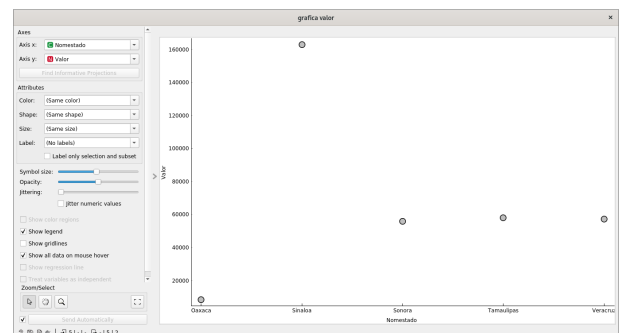


Figura 5: Gráfica Estados ordenados por valor



Objetivo 4. En este objetivo es en el que decidiremos nuestra hipótesis, la cual si recordamos era que no existía ninguna correlación entre los valores de nuestros registros.

Una vez que teníamos nuestra tabla filtrada por los datos pedidos en el objetivo anterior pasamos a obtener su coeficientes de correlación de Spearman, los cuales se muestran en la siguiente imagen

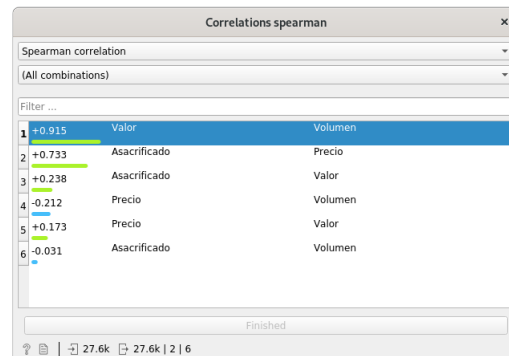


Figura 6: Correlación de Spearman

Como podemos ver existen varias correlaciones, pero las más importantes son las existentes entre Valor-Volumen y Asacrificado-Precio, por lo cual podemos rechazar nuestra hipótesis.

Objetivo 5. Para este objetivo tuvimos que filtrar de nuevo como se ve en la figura 7, ya que solo nos pedían los estados de Veracruz y Tamaulipas y además teníamos que dividir el año 2020; después de esto pasamos a crear un guardado e incorporarlo en el documento, para hacer una predicción como se muestra en la figura 8, de igual manera cargamos los datos reales y los graficamos enseguida se adjuntan ambas imágenes.

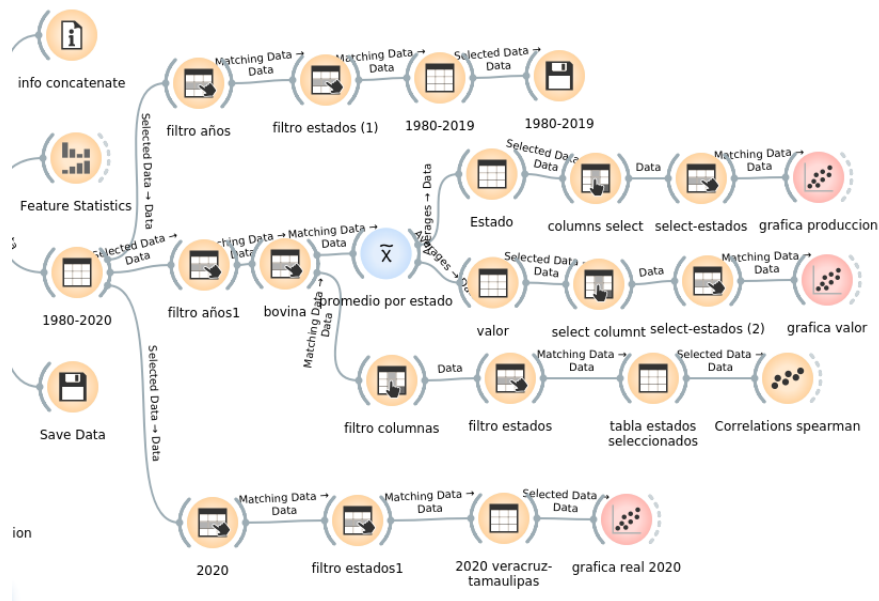


Figura 7: Filtros Nuevos

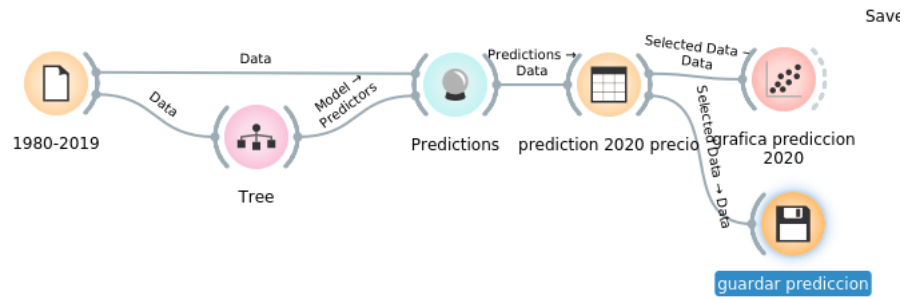


Figura 8: Tercer filtro

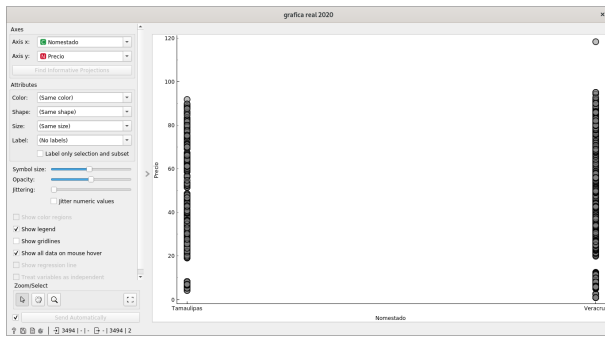


Figura 9: Gráfica real 2020

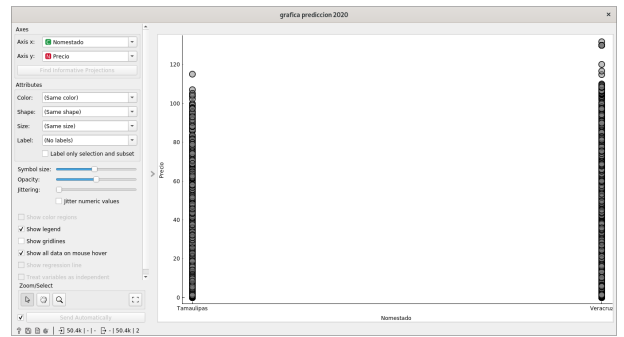


Figura 10: Gráfica predicción 2020

Como podemos ver la gráfica entre ambas es demasiado similar. Entonces de cierto modo podemos decir que nuestra predicción es demasiado cerca a los datos reales.

## Diagrama

El diagrama se modificó demasiadas veces, pero finalmente llegamos a esto:

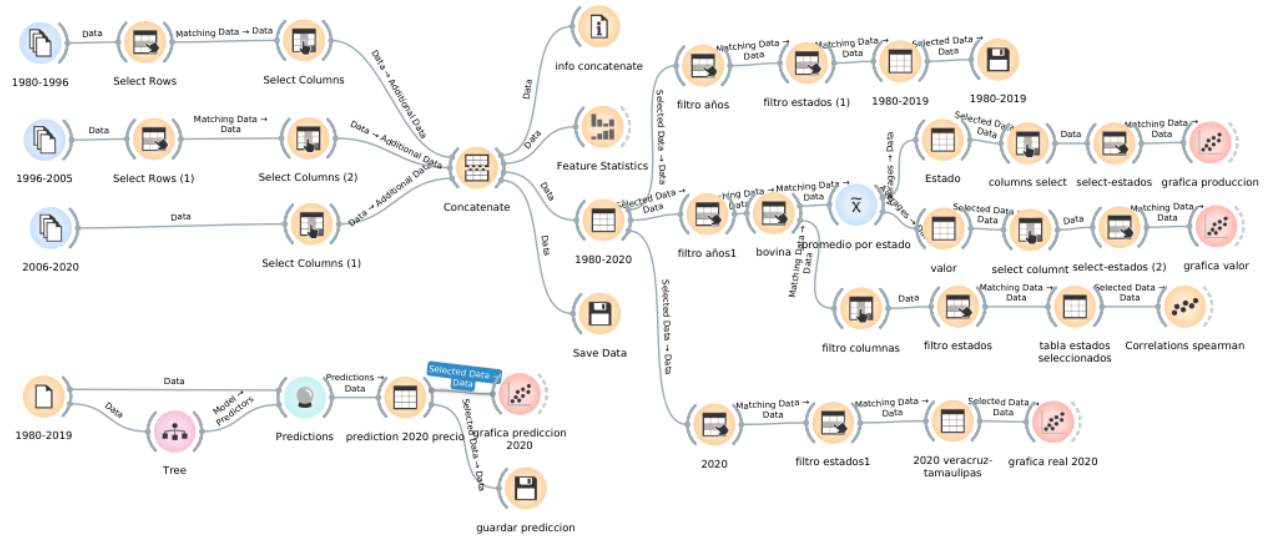


Figura 11: Diagrama final.

Llegamos a este diagrama después de varios intentos, obtamos por crear varias ramificaciones ya que se facilitaba visualizar las gráficas sin necesidad de mover los filtros o los ejes.

## Resultados

Los resultados que obtuvimos son los siguientes:

- I. De las gráficas obtenidas al ordenar las entidades federativas, observamos que Sinaloa tiene un crecimiento demasiado desproporcionado a las demás.
- II. Observando las correlaciones que obtuvimos podemos observar que el volumen y valor están muy relacionados entre sí, así como el número de sacrificados y su precio.
- III. Y por último la predicción que hicimos del año, si bien no es la misma, podemos observar que están en el mismo rango de valores y son casi idénticos en este último.

## Discusión

La discusión sería:

¿Qué podemos hacer para decidir si nuestra predicción es realmente eficaz?

Se nos ocurrió que tal vez en un futuro, podíamos hacer un análisis entre los valores reales y los valores de la predicción, que por falta de tiempo no pudimos realizar. ¿Realmente es suficiente el análisis de correlación para determinar una estrecha dependencia de las variables?

Siguiendo este análisis es casi a simple vista la relación que existe e inclusive podemos decir que es una relación que se esperaba.

## Conclusión

En primer lugar que una buena predicción sobre cuanto es necesario producir para tratar de estar lo mas apegado a la demanda estimada del proximo año es evitar desperdicios de recursos y capital, y que la economia no sea alterada para mal por este tipo de errores.

Por otro lado al comparar la producción REAL de las entidades federativas de Varacruz y Tamaulipas con respecto a lo que nosotros obtuvimos por medio de nuestro analisis estadistico observamos que fueron muy similares por lo que concluimos que hay consistencia de lo producido este año con respecto a los años pasados.