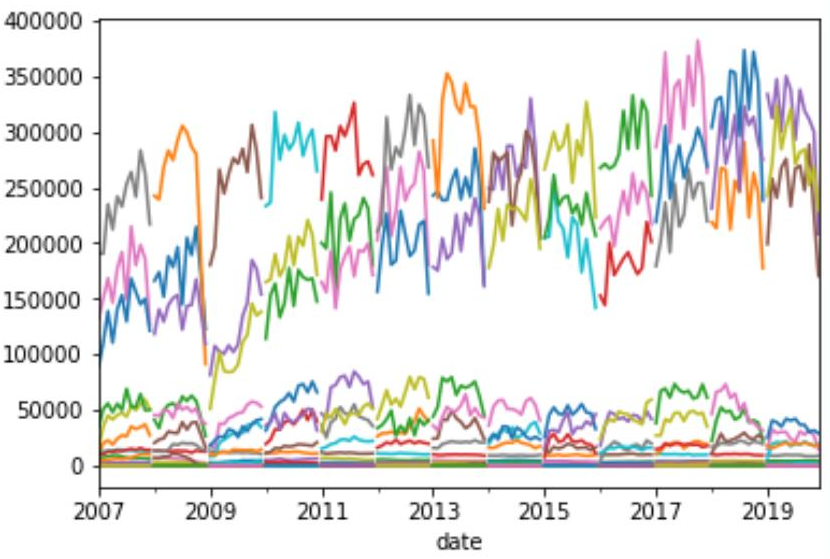
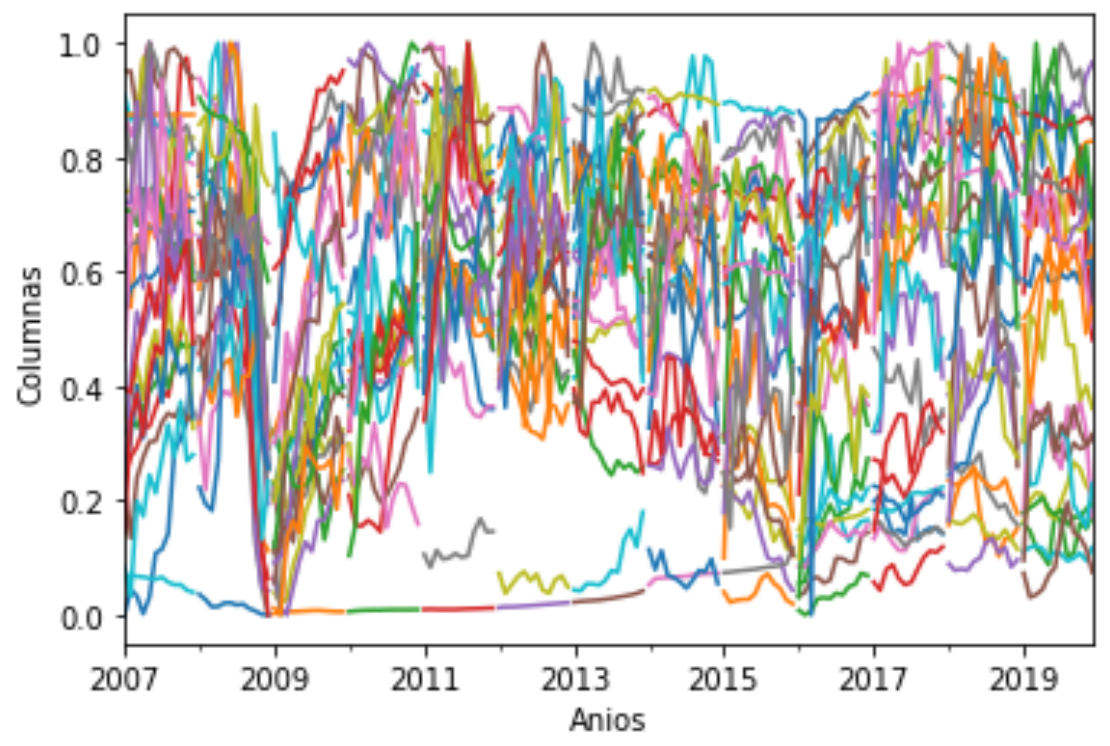
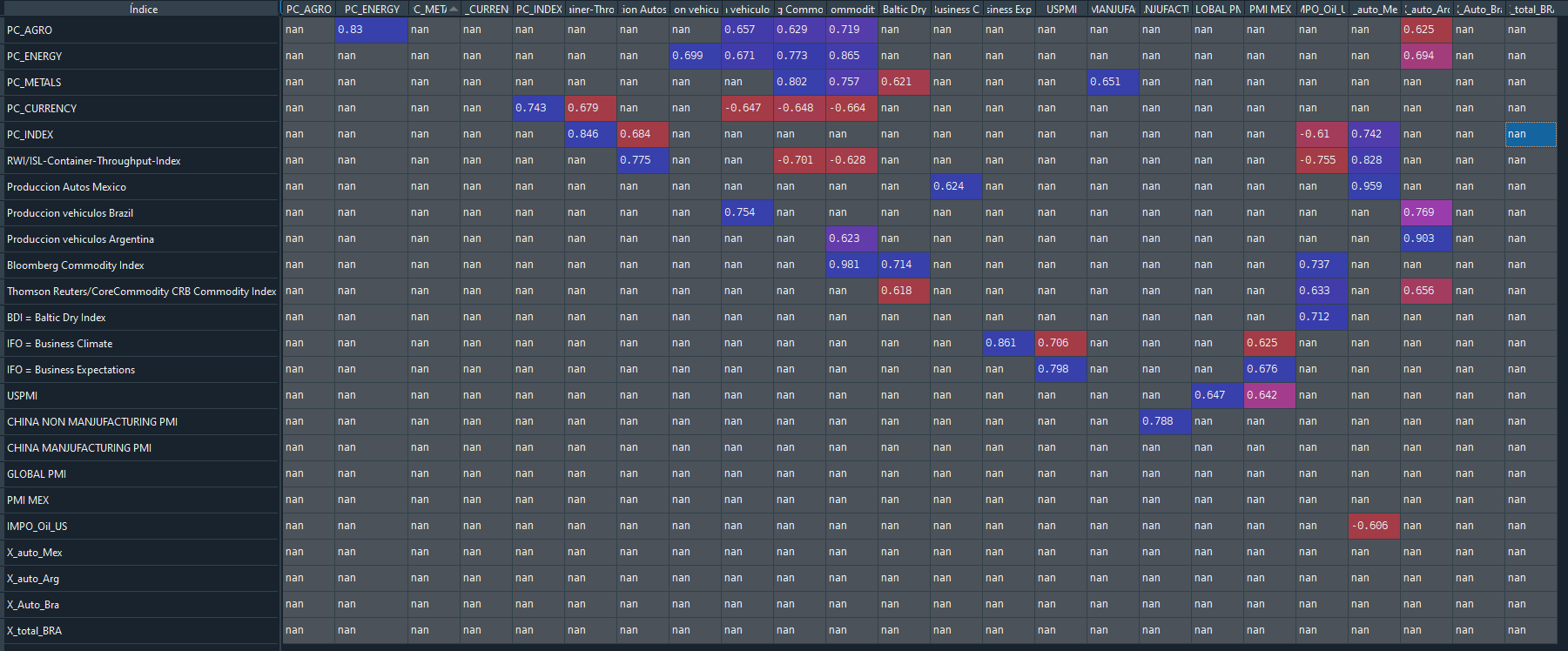
1. Definir un problema de negocio: En economía es muy valioso conocer el nowcast sin tener que esperar uno o dos meses hasta que todos los datos se publiquen fehacientemente.
2. Definir el dataset original: Contamos con 24 variables de diferentes indicadores económicos, a nivel mensual, que se obtienen a partir de diversos métodos: scraping, descarga de fuentes una vez publicadas, etc. Una vez que se descargan los datos (a nivel diario o mensual según corresponda) se corre un proceso de PCA programado en STATA y se unifica para cada una de estas variables un único valor representativo a nivel mensual.
3. Definir el kpi a maximizar: Nowcast para la variable que queremos predecir y analizaremos si el modelo es viable para predecir tendencias para los n meses siguientes (forecast).
4. Definir la variable a predecir: Para determinar la variable a predecir primero analizaremos correlación entre las variables, para lo que evaluaremos la correlación directa y la correlación entre diferentes transformaciones de las variables, para elegir una que tenga buen nivel de correlación y nos permita entender lo que va sucediendo.
5. Crear modelos
6. Evaluarlo
7. Explicarlo

Analisis de correlacion:

Sin normalizar:

Normalizado:





A esta matriz ya le liquidamos el triángulo inferior y la diagonal y borramos todos los valores cuyo módulo es menor a 0.6 indicando entonces una correlación baja. Hay que analizarla por fila: Vemos que las variables que más se correlacionan son:

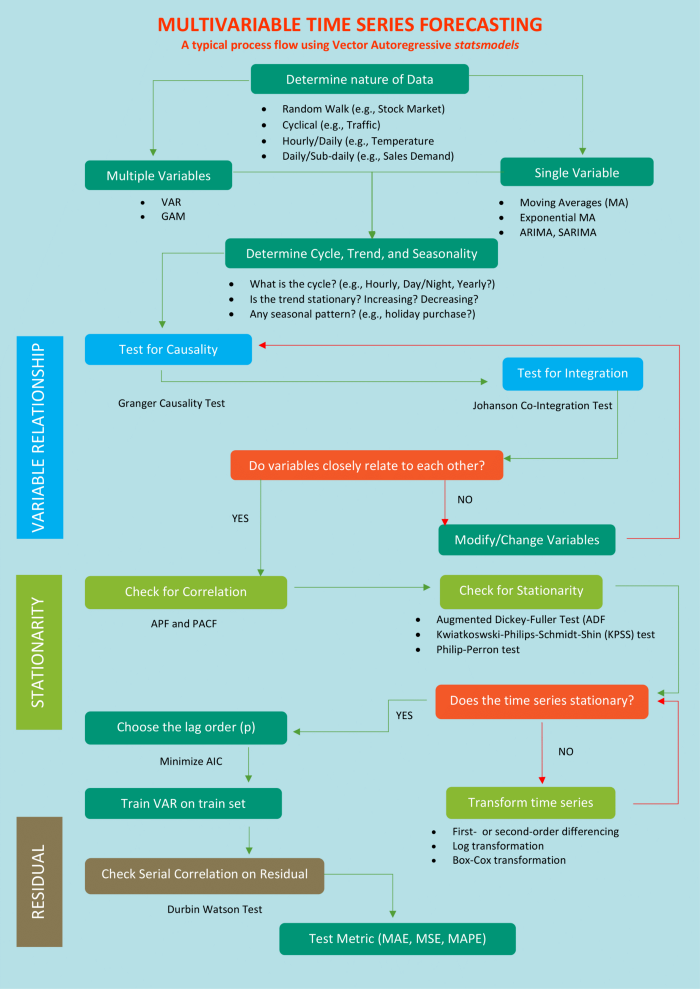
PC\_AGRO (5 incidencias)

PC\_ENERGY (5 incidencias)

PC\_CURRENCY (5 incidencias)

RWI (5 incidencias)

Pipeline a seguir:



Al análisis final agregar U de Theil (sugerencia de Marco)

Extraido de:

https://medium.com/@taufik.azri/forecasting-through-economic-uncertainty-multivariable-time-series-analysis-with-var-and-prophet-e6b801962acb

Comenzar por Nowcast para LAM.

Roadmap:

10 regiones:

- Mercosur

- Brazil

- Caribe

- AL10 -> No lo usamos

- RAP -> Resto de Alianza Pacifico sin mexico

- RM -> Resto de Mercosur sin brazil

- México

- Centroamérica

- SUDAMÉRICA

- LAM -> Latinoamerica (siempre se corre esta)

5 categorias de productos (variables):

- agricola

- energy

- metales

- currency

- index -> Acciones, tasas de int., futuro. (variables financieras)

Con las variables de PCA > DATA\_IN:

CMDTY\_AGRICOLA: Tiene 10 sheets > Hay que unificarlas en agricola

CMDTY\_ENERGY: Idem

CMDTY\_METALS: Idem

FINANCIAL\_CURRCY: Idem

INDEX: Idem

cada uno de estos archivos tiene varias hojas con diferentes

indicadores dentro de cada commodity/bien

En base a esos valores se construyen las variables de

Leading > TRAINING\_CODE > data\_in que son las que usamos

para entrenar los modelos que hagamos

Prefijos:

- px: Son campos estáticos con el ultimo precio negociado

para un valor

- CBOT: Chicago Board Of Trade

Aparentemente se scrapean los precios internacionales de las

commodities según Chicago Board of trade, día a día y luego

se hace algún cálculo para tenerlo mes a mes en Leading > TRAINING\_CODE > data\_in

idea: Comparar el rendimiento de 4 o 5 modelos.

Arima (bueno para val. no muy volatiles), redes neuronales, etc.

(tener cuidado con chi cuadrado)

Opciones:

- U de theil

- https://facebook.github.io/prophet/

Meterle a series de tiempo

! Testear logit

Analizar las transformaciones que se hacen: log, logit, cuadrado, etc

Predecir el mes tomando hasta el mes anterior y luego uno menos y asi sucesivamente

para determinar cuantos meses puede predecir el modelo y cuan bien

Guia a seguir: https://medium.com/@taufik.azri/forecasting-through-economic-uncertainty-multivariable-time-series-analysis-with-var-and-prophet-e6b801962acb

analisis inicial:

transformar las variables y ver cual transformacion tiene mayor correlacion o si es mejor sin transformar

analisis final:

test metrics:

mae, mse, mape, u de theil