



# Modelamiento Basado en Agentes como técnica contra el Lavado de Dinero (ABM-AML)

Agent-Based Modelling (ABM) en una topología de Red como técnica Anti-Money Laundering (AML)

Milton Palacin Grijalva

1 de septiembre de 2021



# Contenido

- 1 Introducción
- 2 Problema
- 3 Modelado
- 4 Experimentos
- 5 Resultados
- 6 Demostración y Revisión de Código
- 7 Conclusiones

# Introducción

---



# Introducción - ABM

El modelamiento y la simulación para predecir el comportamiento humano en la construcción de teorías sociales, económicas y biológicas han evolucionado del modelamiento matemático antes de los 50's (a partir del comportamiento del sistema), desarrollo de la Dinámica de Sistemas en los 50's (Jay W. Forrester), Micro Simulación en los 50's, Autómatas Celulares en los 60's y Modelos Basados en Agentes (ABM - Agent Bases Modelling) en los 90's.

ABM es una técnica potente para el modelamiento y la simulación de complejos sistemas estocásticos, cubriendo brechas de incompletitud de modelos matemáticos y de la inteligencia artificial. ABM considera como unidad básica de medida a los eventos aleatorios contrario a todo lo que se puede determinar bajo un riguroso análisis: “ Everything should be made as simple as possible, but no simpler”.



# Introducción - AML

Es tipo de fraude que consiste en proceso de ocultamiento de la existencia, la fuente ilegal o la aplicación de ingresos obtenidos provenientes de actividades criminales, y el subsiguiente ocultamiento de la fuente de esos ingresos para hacerla parecer legítimos. También es cualquier tentativa por ocultar o disfrazar la identidad de los fondos obtenidos ilegalmente de manera que aparezcan como originados en fuentes legítimas.

La legitimación de activos no es otra cosa que tomar el producto de una actividad ilícita y darle apariencia de legitimidad. La corrupción y el terrorismo son dos casos que consisten en uso de cierto poder y el terror respectivamente para obtener objetivos económicos, políticos, religioso e ideológicos.

# Problema

---



# Antecedentes

- ▷ Los Informes de Inteligencia Financiera (IIF) de la UIF-Perú desde de enero del 2012 a mayo del 2021 involucran US\$ 14,478 millones, procedentes de minería ilegal (57 %), defraudación tributaria (10 %), tráfico ilícito de drogas (10 %), delitos contra la administración pública (10 %) y entre otros (13 %). En el periodo junio 2020 a mayo 2021 fue de UD\$ 1,566 millones.
- ▷ Perú se encuentra en la posición 104 de 141 países en índice “Anti-lavado de Dinero – AML 2020”, que publica el Instituto de Gobernabilidad de Basilea.’
- ▷ Las actividades de Lavado de Dinero tienen un costo entre el 2 % y 5 % del PBI mundial (entre US\$800 billones a USD\$ 2 trillones aproximadamente).
- ▷ 95 % de los sistemas de alertas contra el Lavado de Dinero resultan ser falsos positivos.
- ▷ Las Naciones Unidas estimó que para 2020-2021, 90 % del dinero lavado permanecerán sin ser detectados.



# Definición del problema

En la lucha contra el Lavado de Dinero no se cuenta con herramientas de simulación para la detección de patrones en el comportamiento de los agentes dentro las redes de Lavado de Dinero. Como resultado no se cuenta con información que permita mejorar políticas y procedimientos para la identificación temprana, acorde a los sectores económicos, de operaciones sospechosas y castigo posterior a los agentes que incurran en delitos. Las herramientas de Business Intelligence así como de IA cubren en su mayoría aspectos determinísticos: datos históricos y predicción a corto plazo, sumando a que todos los datos son confidenciales y protegidas por las leyes, hacen difícil la generación de herramientas que puedan medir y analizar los eventos aleatorios en las redes de Lavado de Dinero.



# Modelado

---

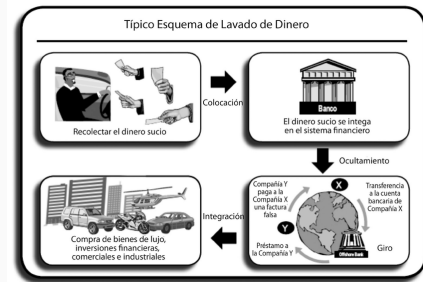


# Descripción AML

El Lavado de Dinero inicia cuando por medio de un ilícito se produce algún activo del cual es necesario ocultar su procedencia. Se puede reconocer las siguientes etapas principales:

- *Colocación*: Mezclado con dinero lícito es llevado a una entidad financiera para adquirir un producto o servicio financiero.
- *Estratificación*: Dentro del sistema financiero comienza a ocultar su origen ilícito a través de transacciones que vulneran el control de anti-lavado.
- *Integración*: El dinero con apariencias legítima se inserta nuevamente a la economía para adquirir bienes y servicios, mantener vidas lujosas o para re-invertir en actos ilícitos.

+





# ABM

- ▷ *Agentes*: existen tres tipos de agentes. Primero; **individuos** comunes. Segundo; **empresas** con fines de lucro, empresas sin fines de lucro, empresas de fondos fiduciario y empresas fantasma. Tercero; **intermediarios** financieros, formales como los bancos o informales como los “Hawala”. Cada tipo de agente se representará como un “Autómata Finito No Determinístico” (AFND).‘
- ▷ *Ambiente*: La topología será una red tipo ‘grafo no dirigido’ donde cada nodo representa a un agente y cada enlace representa la conexión a sus vecinos más cercanos. Estos enlaces son necesarios para las operaciones de dinero que fluye en la red.
- ▷ *Comunicación*: un individuo o empresa solo se comunicará con otros individuos o empresas. Individuos y empresa se comunicarán con los intermediarios entidades financieros simulando poseer una cuenta o billetera.



# Agente Individuo (1) y Empresa (4)

▷ *Estados:*

- ① ( $q_1$ ) Balance de situación/estado inicial
- ② ( $q_2$ ) En espera de alguna operación.
- ③ ( $q_3$ ) Realizando operación.
- ④ ( $q_4$ ) Cuentas congeladas.

- ▷ *Autonomía:* El agente individuo y empresa tienen una predisposición al fraude ("*predisposition-Fraud*"), además es responsable de su libro contable mayor ("*ledger*"). Con base en estos dos principales atributos toma decisiones para realizar operaciones legítimas o ilegítimas.
- ▷ *Atributos:* Sí cuenta esta cancelada ("*isFrozenAccounts*"), libro contable mayor ("*ledger*"), predisposición al fraude ("*predispositionFraud*"), ubigeo de origen ("*place*"), datos que corresponden a AFND.



# Agente Individuo (1) y Empresa (4)

▷ Estados:  $Q = \{q_1, q_2, q_3, q_4\}$

▷ Alfabeto:  $\Sigma = \{a_1, a_2, a_3\}$

- $a_1 \Rightarrow$  Esperar operación.
- $a_2 \Rightarrow$  Realizar operación.
- $a_3 \Rightarrow$  Congelar cuenta.

▷ Palabras aceptadas:

$$\mathcal{L} = \{(a_1^+), (a_1^+ a_2^+ a_1^+), (a_1^+ a_2^+ a_3), (a_1^+ a_2^+ a_1^+ a_2^+ a_3), \dots\}$$

▷ Estado Inicial:  $q_1$

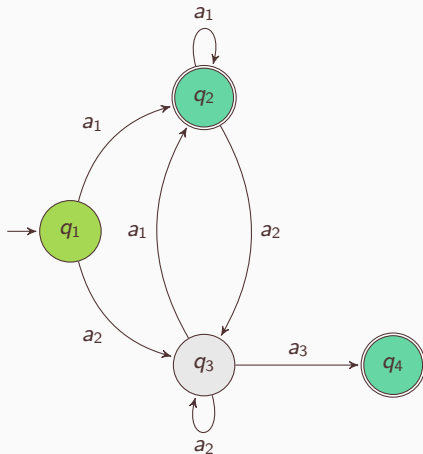
▷ Estados de aceptación:  $\mathcal{F} = \{q_2, q_4\}$

Transiciones:

$\delta :$	$a_1$	$a_2$	$a_3$
$q_1$	$\{q_2\}$	$\emptyset$	$\emptyset$
$q_2$	$\{q_2\}$	$\{q_3\}$	$\emptyset$
$q_3$	$\{q_2\}$	$\{q_3\}$	$\{q_4\}$
$q_4$	$\emptyset$	$\emptyset$	$\emptyset$



# Agente Individuo (1) y Empresa (4)



Dentro de la Red solo se mueve dinero y se crean algunos enlaces. En consecuencia, los estados de los agentes son referenciales y son cubiertos por las operaciones con dinero o transacciones.



# Agente Intermediario (1)

▷ *Estados:*

- ① ( $q_1$ ) Actualización/creación de cuentas.
- ② ( $q_2$ ) En espera de alguna operación.
- ③ ( $q_3$ ) Realizando operación.

▷ *Autonomía:* El intermediario está asociado a tipo de entidad financiera formal o informal.

▷ *Atributos:* Entidad financiera asociada ("*financialEntity*", que puede ser formal o informal), ubigeo de origen ("*place*"), datos AFND.



# Agente Intermediario (1)

▷ Estados:  $Q = \{q_1, q_2, q_3\}$

▷ Alfabeto:  $\Sigma = \{a_1, a_2\}$

- $a_1 \Rightarrow$  Esperar operación.
- $a_2 \Rightarrow$  Realizar operación.

▷ Palabras aceptadas:

$$\mathcal{L} = \{(a_1^+), (a_1^+ a_2^+ a_1^+), (a_2^+ a_1^+), \dots\}$$

▷ Estado Inicial:  $q_1$

▷ Estados de aceptación:  $\mathcal{F} = \{q_2\}$

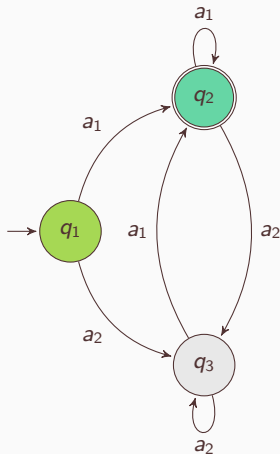
Transiciones:

$\delta :$	$a_1$	$a_2$
$q_1$	$\{q_2\}$	$\emptyset$
$q_2$	$\{q_2\}$	$\{q_3\}$
$q_3$	$\{q_2\}$	$\{q_3\}$





# Agente Intermediario (1)



Similar al tipo de agente anterior, estos estados son referenciales. Un individuo o empresa establece un enlace con un intermediario similar a crear una cuenta bancaria en entidad financiera.



# Ambiente - Creación de la Red

- ▷ Existen dos tipos de enlace: 1) de un intermediario a otros agentes y 2) entre individuos y empresas.
- ▷ Existe un parámetro (p.e. 40 %) para que cada intermediario esté enlazado con los otros intermediarios.
- ▷ Existe un parámetro (p.e.  $[1, 4]$ ) para que cada individuo y empresa se enlace con un cierto número de intermediario.
- ▷ Existe un parámetro, de una distribución exponencial de media 2, para que cada individuo y empresa se enlacen entre ellos (vecinos).
- ▷ Un individuo o empresa que tiene una predisposición mayor (p.e a 60 %) se le crea un enlace adicional con una empresa fantasma o de fondo fiduciario.
- ▷ Un individuo o empresa que tiene una predisposición al fraude mayor (p.e a 90 %) se le crea un enlace adicional con una empresa fantasma o de fondo fiduciario y otra con un agente (individuo o empresa) que tiene una predisposición al fraude mayor (p.e a 60 %).



# Ambiente - Transacciones (Movimiento)

- ▷ Existe un parámetro (p.e. 5 %) para que individuos o empresa inicie una tracción (operación de dinero).
- ▷ Durante el proceso de una simulación el dinero de las transacciones se mantienen contabilizado.
- ▷ Existe un parámetro de monto máximo (p.e. S/ 100,000) para las transacciones que realicen los agentes.
- ▷ Cada agente origen elige, de sus vecinos cercanos (agentes enlazados), al destinatario. Existe un parámetro (p.e. 10 %) para que el agente elija a otro que no sea su vecino, por ende crea un nuevo enlace.



# Ambiente - Transacciones (Movimiento)

- ▷ Agente origen y destinatario deben estar enlazados al mismo intermediario o a diferentes pero que estén conectados, caso contrario el agente origen creará un enlace hacia un intermediario en común.
- ▷ Cada agente origen decide como estructurar la transacción, si ambos tienen baja predisposición o el monto es menor al umbral de “operación sospechosa” se envía todo el dinero en una sola transacción. Caso contrario el agente origen dividirá todo el monto en partes que no superen el umbral y creará tantas transacciones a otros agentes vecinos para hacer llegar todo el dinero al destinatario de manera indirecta. Si el agente origen no tiene el número de vecinos necesario para enviar el monto total, este creará nuevos enlaces con otros agentes (individuos o empresas).



# Ambiente - Adicional

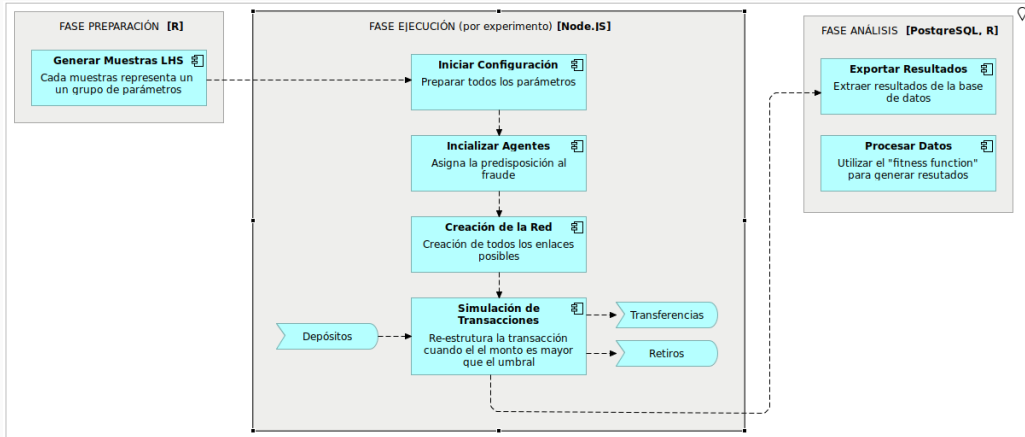
- ▷ Se actualizará de manera constante una lista de observación de operaciones sospechosas (**Watchlist**)
- ▷ Cuando se creen los agentes en la Red, existe un parámetro (p.e. 10 %) para ponerlos en el Watchlist.
- ▷ Existe un parámetro (p.e. [2 – 5]) para congelar todas las cuentas agentes (individuos y empresas) cuando estas pasan mucho tiempo de manera consecutiva en el Watchlist.
- ▷ Para el umbral de operación sospechosa se asumirá S/ 9,000.
- ▷ Se realizará “N” experimentos, por cada una se ejecutará “D” simulaciones de transacciones, una simulación representa un día.
- ▷ Todas las transacciones se llegan a completar en el día, no se deja nada pendiente.

# Experimentos

---



# Procedimiento





# Principales Parámetros

Orden	Descripción	Nombre	Valores LHS	Distribución	Valores Final	Pre-procesado
1	Número de experimentos	<i>totalIterations</i>	$\emptyset$	$\emptyset$	200	No
2	Número de días por simulación	<i>totalTimes</i>	$\emptyset$	$\emptyset$	120	No
3	Población total	Solo para LHS	[200, 2000]	Uniforme	De la muestra	No
4	Población de individuos	<i>numPopIndividual</i>	[0.5, 0.8]	Uniforme	$\emptyset$	Sí
5	Población de intermediarios	<i>numPopIntermediary</i>	[10, 74]	Uniforme	$\emptyset$	Sí
6	Población de empresas	Solo para LHS	[0.1, 0.4]	Uniforme	De la muestra	No
7	Población de empresa con fines de lucro	<i>numPopProfitBusiness</i>	[0.1, 0.8]	Uniforme	$\emptyset$	Sí
8	Población de empresa sin fines de lucro	<i>numPopNoProfitBusiness</i>	[0.1, 0.8]	Uniforme	$\emptyset$	Sí
9	Población de empresa de fondo fiduciario	<i>numPopTrustBusiness</i>	[0.1, 0.8]	Uniforme	$\emptyset$	Sí
10	Población de empresa/negocio fantasma	<i>numPopShellBusiness</i>	[0.1, 0.8]	Uniforme	$\emptyset$	Sí
11	Población de individuos con alta predisposición al fraude	<i>numPopHighPropensityFraud</i>	[0.1, 0.8]	Uniforme	$\emptyset$	Sí
12	Porcentaje de población de individuos y empresas a incluir en el Watchlist	<i>perPopWatchList</i>	[0, 0.3]	Uniforme	$\emptyset$	Sí
13	Máximo de días consecutivos en el Watchlist para congelar la cuenta de un individuo o empresa	<i>maxTimesWatchList</i>	[15, 20]	Uniforme	De la muestra	No
14	Máximo de días consecutivos que no esta en Watchlist para reiniciar el conteo	<i>maxTimesCleanWatchList</i>	[2, 10]	Uniforme	De la muestra	No
15	Predisposición de fraude a asignar a <i>numPopHighPropensityFraud</i>	<i>maxPropensityFraud</i>	[0.6, 0.9]	Uniforme	De la muestra	No
16	Alto valor de la predisposición al fraude para los enlaces extras	<i>maxHighPropensityFraud</i>	$\emptyset$	$\emptyset$	$\min(\maxPropensityFraud + 0.2, 0.9)$	Sí
17	Porcentaje de enlaces entre intermediarios	<i>perLinkedIntermediary</i>	[0.4, 0.8]	Uniforme	De la muestra	No
18	Número de enlaces entre individuos y empresas	<i>numMaxLinkedNoIntermediary</i>	[1, 16]	Uniforme	De la muestra	No
19	Número de enlaces de individuos/empresas con los intermediarios	<i>numMaxLinkedIndBusInter</i>	[2, 6]	Uniforme	De la muestra	No
20	Porcentaje de individuos y empresas que realizan transacciones de tipo depósito	<i>perExecuteDeposit</i>	[0, 0.3]	Uniforme	De la muestra	No
21	Porcentaje de individuos y empresas que realizan transacciones de tipo transferencia	<i>perExecuteTransfer</i>	[0.3, 0.7]	Uniforme	De la muestra	No
22	Porcentaje de individuos y empresas que realizan transacciones de tipo retiro	<i>perExecuteWithdrawal</i>	[0.3, 0.7]	Uniforme	De la muestra	No
23	Umbral para determinar si es una operación sospechosa	<i>amountSuspiciousOperation</i>	$\emptyset$	$\emptyset$	9000	No
24	Valor máximo para ser seleccionado de manera aleatoria en una transacción	<i>rangeAmountTransaction</i>	[500, 110000]	Uniforme	[0, De la muestra]	Sí
25	Porcentaje para seleccionar de manera aleatoria un agente que no es vecino para una transacción	<i>rangeNewLinkTransact</i>	[0, 0.2]	Uniforme	De la muestra	No





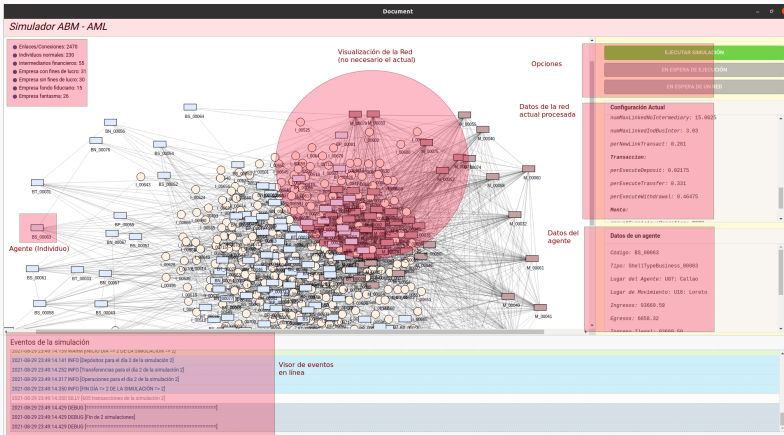
# Experimento

- ▷ Se realizó 200 (N) experimentos.
- ▷ En cada experimento se realiza 120 simulaciones (días) de transacciones.
- ▷ Se realizó tres grupos de experimentos; en promedio se tomó 5.5 horas cada una.
- ▷ Debido a la cantidad de información generada se reducirá el caso de estudio con el siguiente **fitness function**:  $ILD$  = Indicador de lavado de dinero,  $DI$  = Dinero Ilegal,  $DL$  = Dinero legal.

$$\%ILD = \frac{\sum_{i=1}^N DI_i}{\sum_{i=1}^N (DI_i + DL_i)} \quad (1)$$



# Experimento - Programa/Aplicación

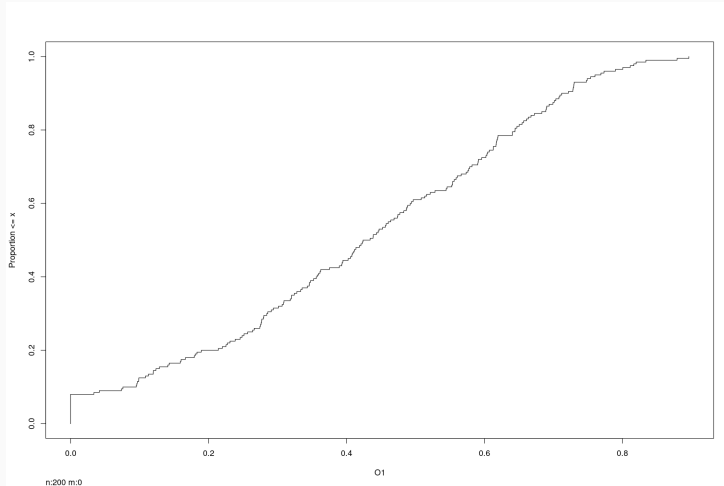


## Resultados

---



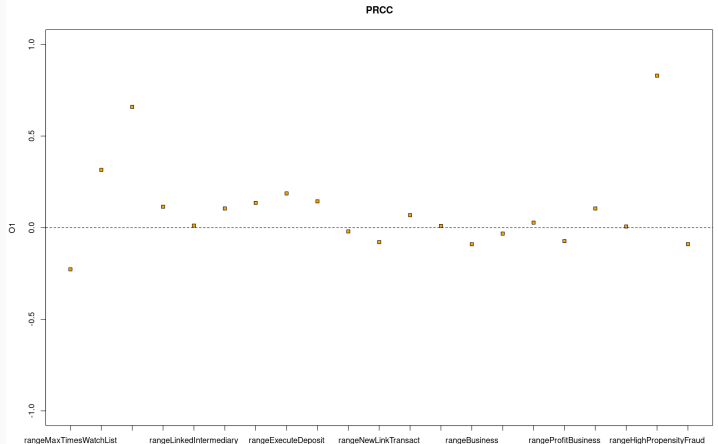
# Distribución acumulado de los resultado





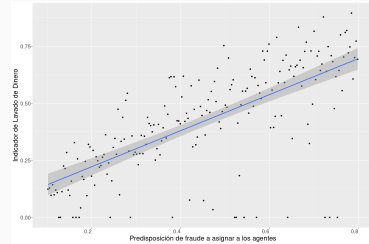
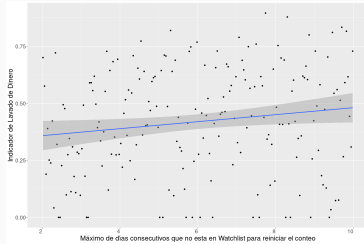
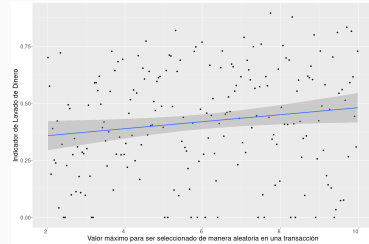
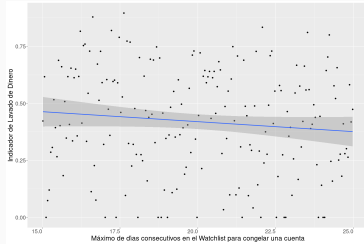
# PRCC

Parámetro	PRCC
rangeMaxTimesWatchList	-0.227546195
rangeMaxTimesCleanWatchList	0.315445893
rangeMaxAmountTransaction	0.658932723
rangePropensityFraud	0.114238044
rangeLinkedIntermediary	0.011139089
rangeMaxLinkedNoIntermediary	0.104609436
rangeMaxLinkedIndBusInter	0.135285659
rangeExecuteDeposit	0.186759631
rangeExecuteTransfer	0.143579939
rangeExecuteWithdrawal	-0.020701911
rangeNewLinkTransact	-0.079030242
rangePopulation	0.068174119
rangeIndividual	0.009458012
rangeBusiness	-0.090407205
rangeIntermediary	-0.032441369
rangeNoProfitBusiness	0.027471308
rangeProfitBusiness	-0.073413330
rangeTrustBusiness	0.104533005
rangeShellBusiness	0.006335857
rangeHighPropensityFraud	0.829617586
rangeWatchList	-0.090226869





# Análisis de Sensibilidad

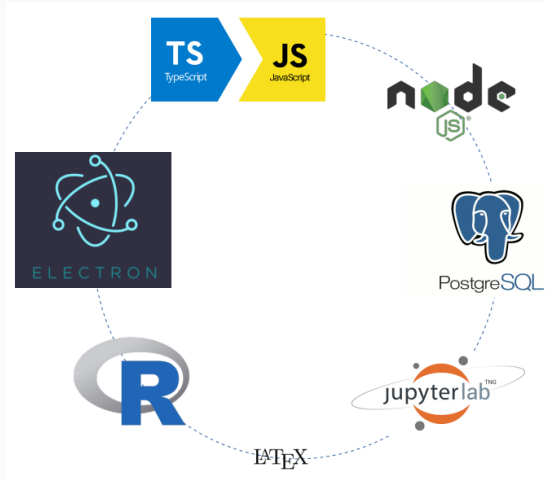


## **Demostración y Revisión de Código**

---



# Software Utilizado





## Conclusiones

---



# Conclusiones

- ▶ AML es un problema complejo, en el desarrollo de este trabajo solo se consideró las principales variables. Sin embargo faltó incluir información de sectores económicos, predisposición al fraude por el lugar donde vive agente, entre mucho otros.
- ▶ En el presente trabajo se muestra que modelar AML bajo el enfoque de ABM, puede adelantar el análisis de sensibilidad de variables que son difíciles de medir con la data histórica o aún no son considerados bajo estudio. Esto sucede dado que la información generada ha sido sometida a la aleatoriedad de la incertidumbre sistémica que presenta AML.
- ▶ Contar con una herramienta, como el desarrollado en este proyecto, es de uso necesario cuando se quiere explorar más allá de los datos con los que se cuenta. Este tipo de enfoque sirven para generar datos sintéticos que ayuden a mejorar el rendimiento de los modelos de Machine Learning/IA.

**¡Gracias!**