

Práctica 2: Implementación en Shiny

Método de selección de práctica

Cada estudiante debe tener un práctica diferente. Para su distribución debéis enviar un **lista de 2 opciones ordenadas por preferencia** de las aplicaciones que queráis de desarrollar antes del **Lunes 4**

Podéis plantear alternativas si queréis hacer otra cosa. Una buena fuente de inspiración son **otras asignaturas del master**. P.e.: Aplicar un modelo de agrupación o econométrico mediante una aplicación en Shiny.

También podéis hacer modificaciones a las propuestas que os envío. Soy todo oídos.

A continuación se hace una breve descripción escrita de cada opción y se adjunta los storyboards de cada una.

Have Fun! ;)

Birthday paradox

La paradoja del cumpleaños no es realmente una paradoja sino una conclusión matemática contraintuitiva.

La paradoja habla sobre la probabilidad de que dos personas cumplan el mismo día del año en un mismo grupo de personas (normalmente una clase de estadística).

Con sólo 23 personas la probabilidad ya es superior al 50% y con 47 es superior al 90%.

Esta aplicación permite visualizar esta paradoja con una visión de calendario y ejecutando una simulación en la que se van añadiendo sujetos al grupo.

El usuario puede controlar el tamaño de la clase.

Además se puede ejecutar una simulación masiva para calcular mediante una aproximación monte carlo la probabilidad real de coincidencia.

Por último se analiza la probabilidad de otros fenómenos "llamativos" como la probabilidad de doble coincidencia (tres personas cumplen el mismo día), triple (cuatro personas) y cumplir años en días adyacentes.

Para la visualización del calendario se puede usar [rChartsCalMap](https://github.com/ramnathv/rChartsCalmap), [calheatmapR](http://durtal.github.io/calheatmapR/) o [googleVis](https://www.r-bloggers.com/calendar-charts-with-googlevis/).

Más información: [Wikipedia](https://es.wikipedia.org/wiki/Paradoja_del_cumplea%C3%B1os)

Birthday paradox

#1 U.Hi Single simulation

frame #

Population 365

Play There are 3 coincidences

Calendar visualization

Jan Feb March - - - Dec

All

Coincidences:

Name	Birth
Charles	1/13
Fred	5/18
Luca	9/2

Name	Birthday
x	1/13
y	1/13
z	1/13

#2 Multiple simulation

frame #

Single
Multiple

Run [13] simulations per #population

Textbox

Run Button

when its clicked show graph

prob of coinc.

population

#3 See other stats

frame #

Run [1F]

Run

multiple

Probability of double coincidence

Prob of triple

Prob of consecutive birthdays

frame #

frame #

Distribution Painter

Se trata de una herramienta de analítica visual y docencia que mediante interacciones de clicks permite al usuario crear sus propias formas de distribuciones y obtener estadísticos básicos.

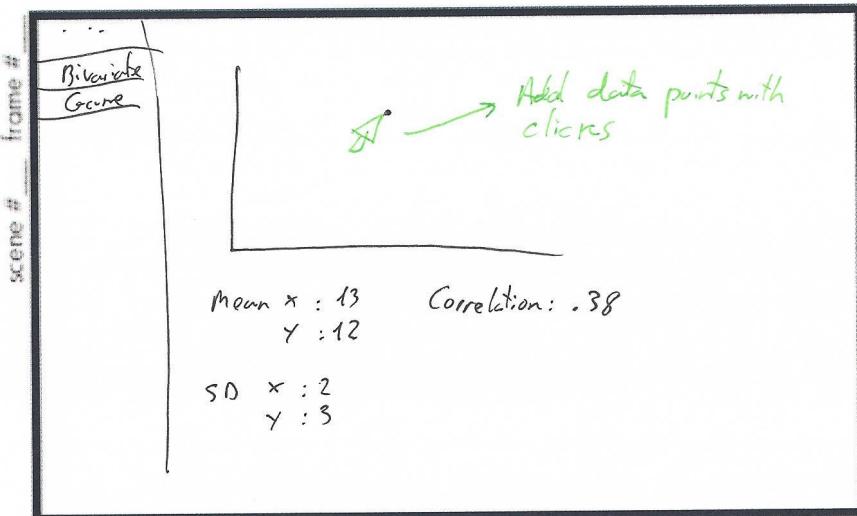
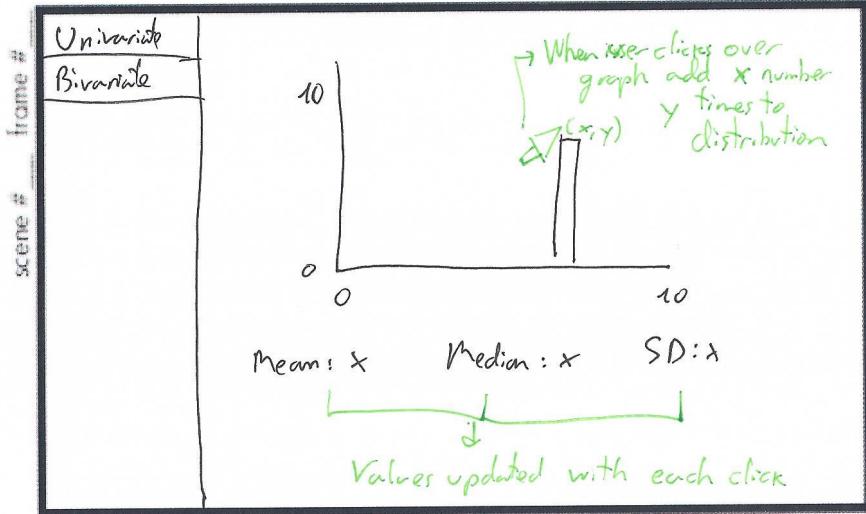
Existe la versión univariada (histograma) y univariada.

También se añade un juego que consiste en crear una distribución que cumpla unos estadísticos descriptivos proporcionados por la aplicación

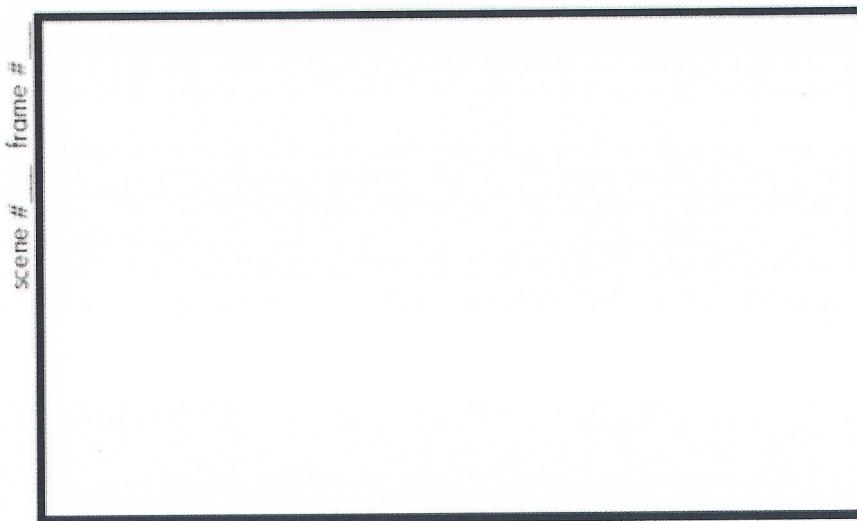
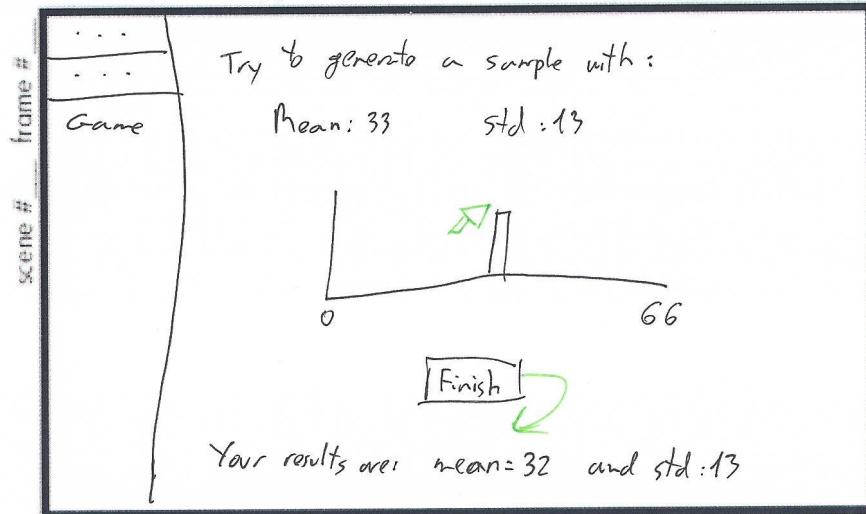
#1 Paint univariate distribution

Distribution painter

#2 Paint bivariate distribution



#3 Game:



Probabilistic Excel

Inspirado en [Caladis \(<http://www.caladis.org/>\)](http://www.caladis.org/). Se trata de un "excel" u hoja de cálculos probabilística. De manera que se le permite al usuario hacer cálculos inciertos sobre distribuciones, p.e:

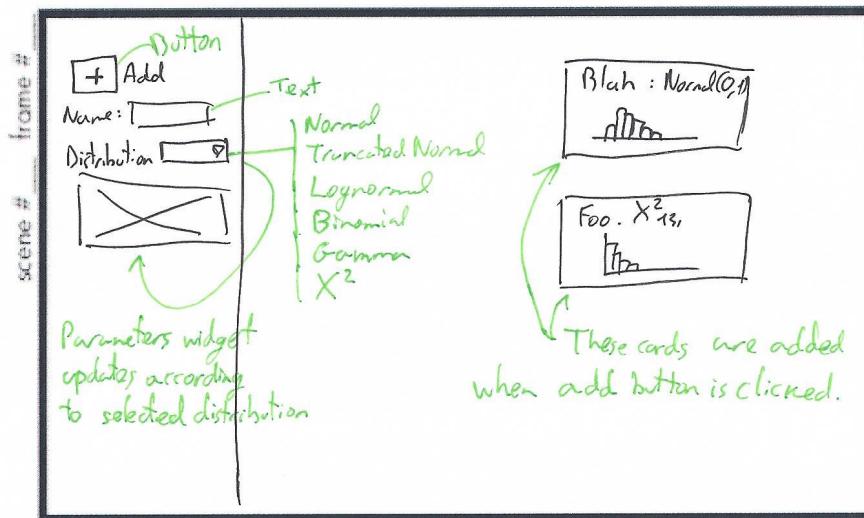
Ventas = N(1000,150) Beneficio por venta = X2(3)

Costes = N(800,13) Beneficios = Ventas * Beneficio por venta - Coste

Se utiliza método monte carlo (simulaciones) para calcular los valores de la distribución.

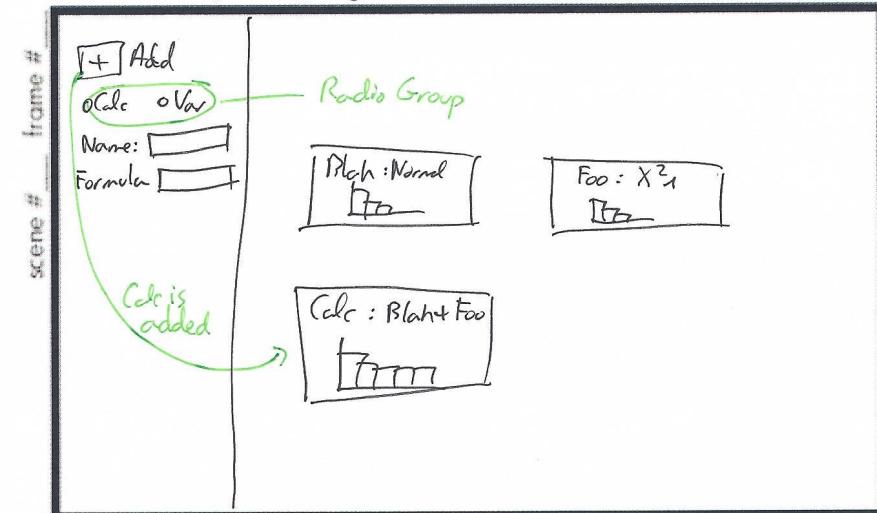
Cada variable y cálculo se acompaña de una tarjeta con su distribución

#1 Create new variable

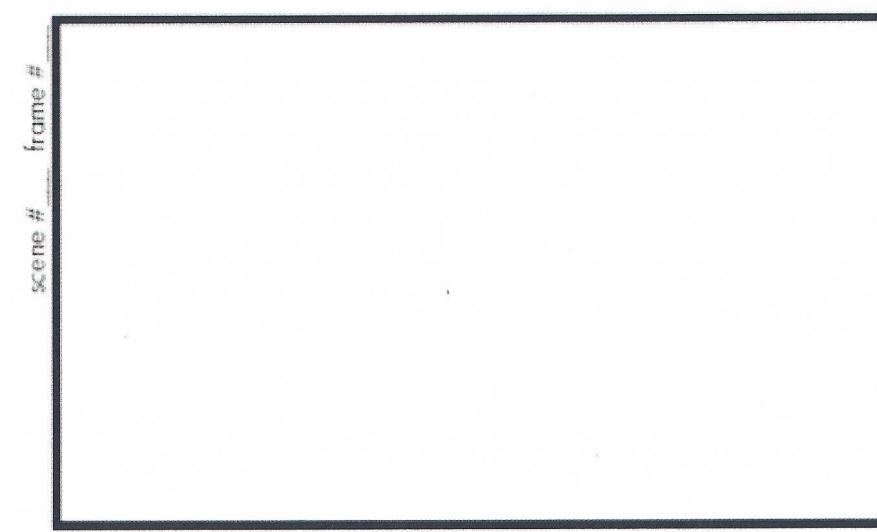
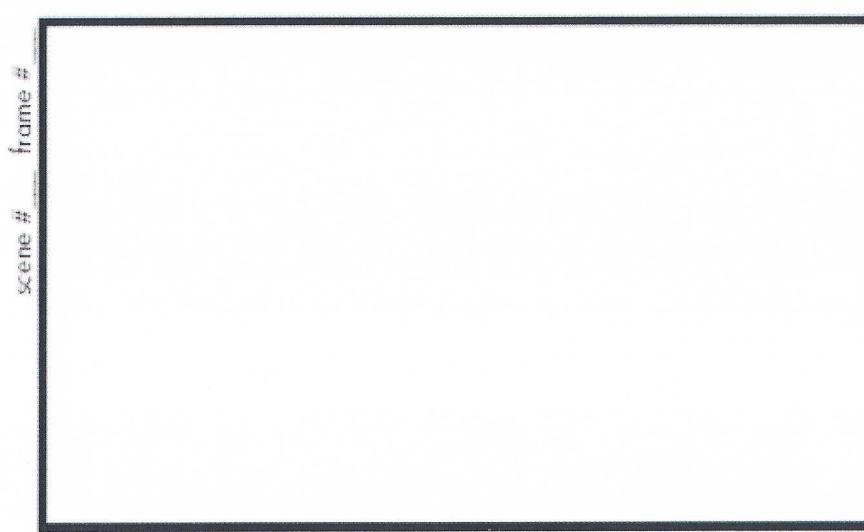


Probabilistic excel

#2 Create new calc



Los cálculos son expresiones de R que se calculan con las variables generadas anteriormente.



Explorador de modelos

Herramienta que permite añadir modelos polinómicos a datos arbitrarios

El usuario puede elegir la semilla y la correlación de una normal multivariada. Además pulsando sobre la gráfica puede añadir nuevos puntos de datos.

Todos los modelos se muestran en la misma gráfica junto a los datos y cualquier cambio en los datos o en los modelos hace que se vuelvan a ajustar.

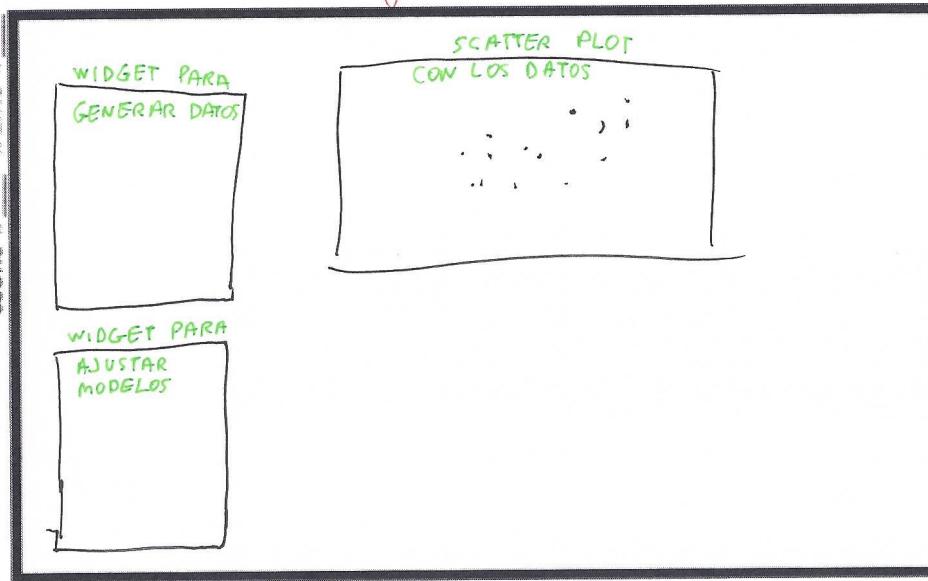
También se añaden unas pestañas con estadísticos de calidad del modelo.

Visión general

Explorador de modelos

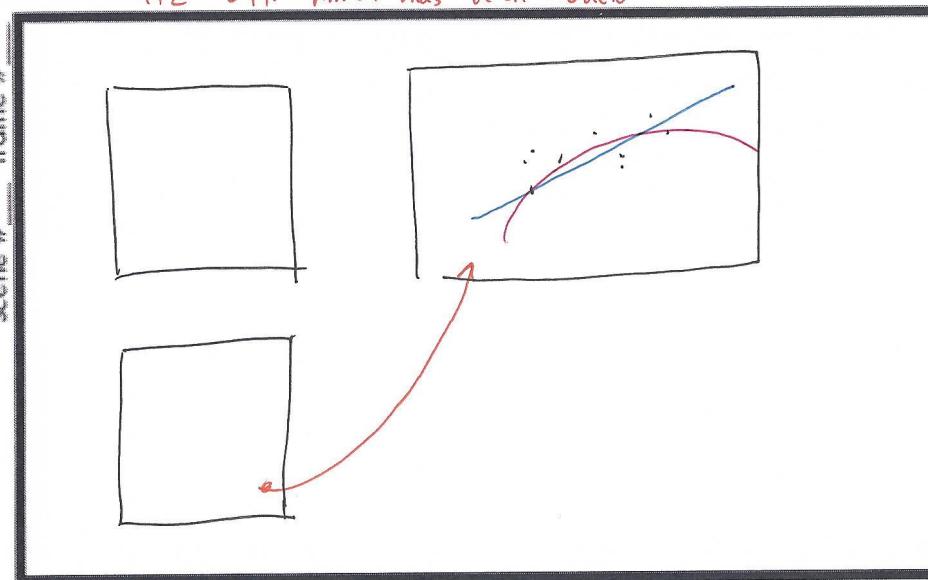
#1 VH: Añadir modelo polinómico

scene # frame #

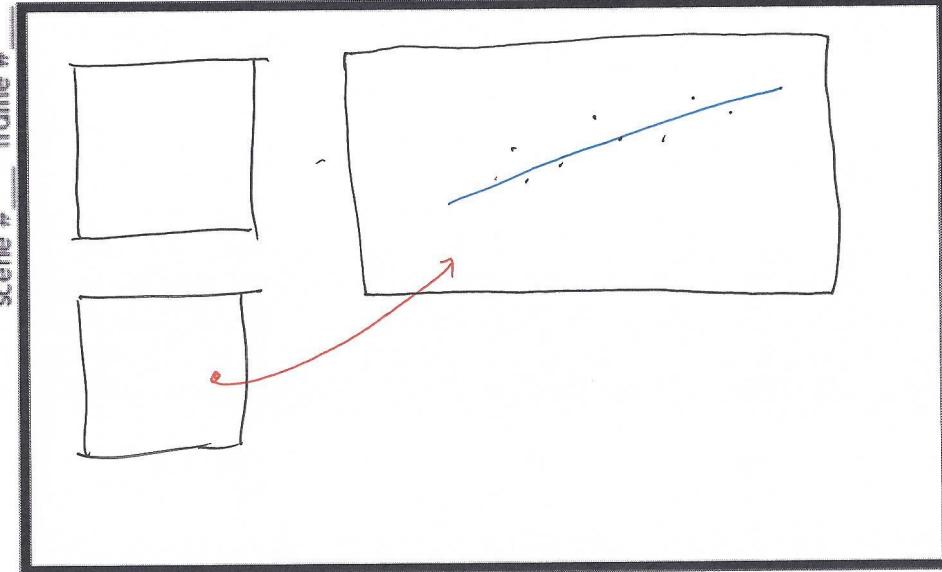


#2 VH. Añadir más de un modelo

scene # frame #

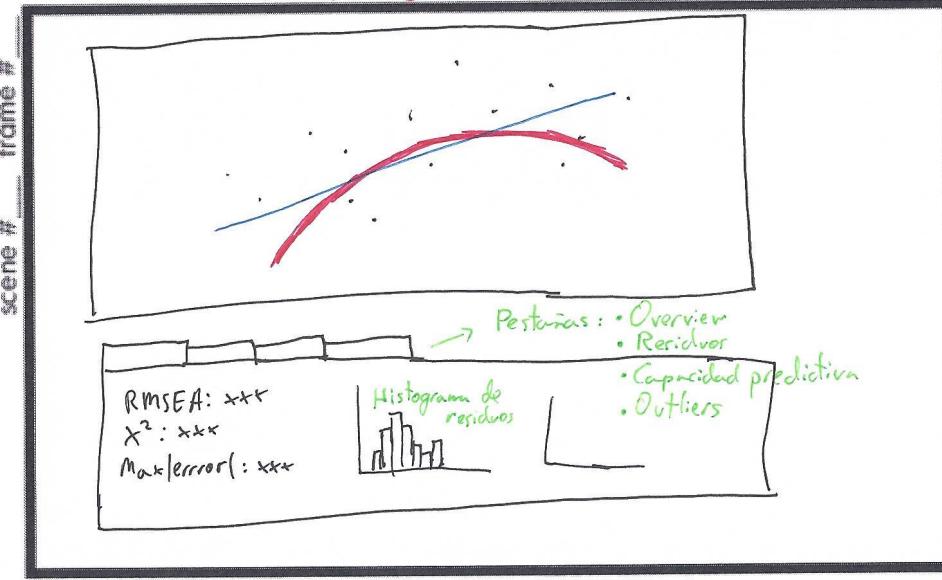


scene # frame #



#3 VH Ver detalle de un modelo

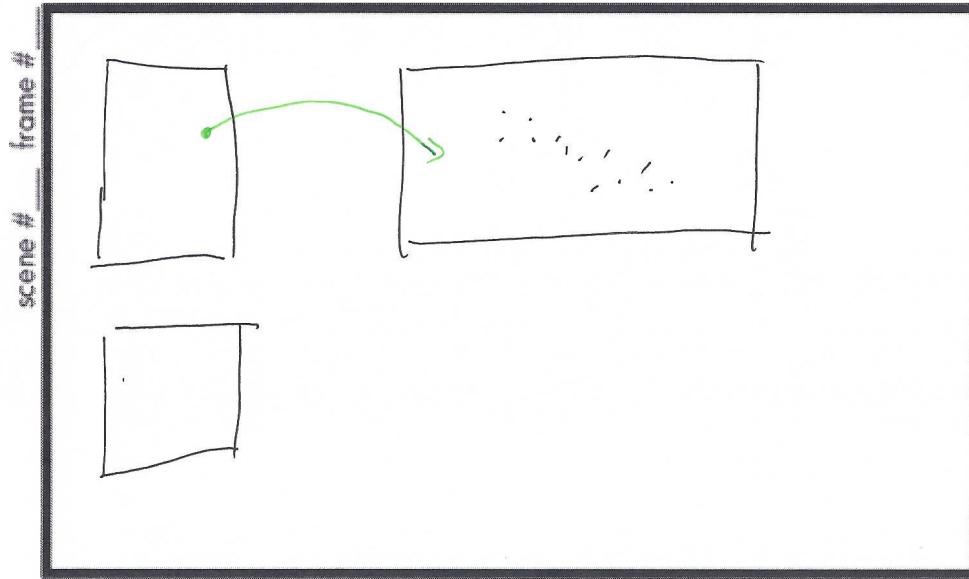
scene # frame #



El modelo que se analiza se destaca en la gráfica superior

Explorador de modelos (II)

#4 VH Generar nuevos datos (Opcional)



Requisitos

- Generar datos

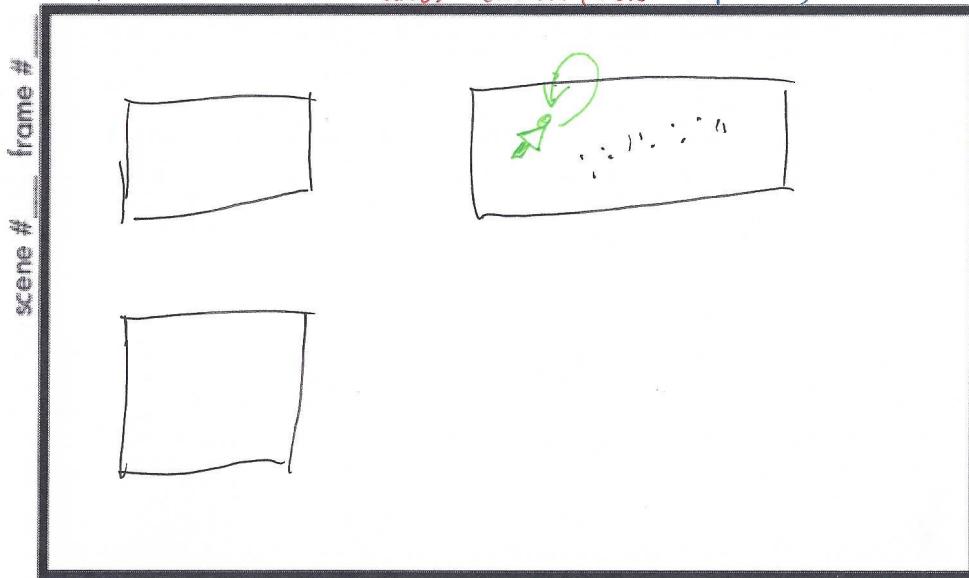
- Puedes especificar la semilla (o no)
- Puedes especificar la correlación (r_{xy})
- Se genera una normal multivariada

- Crear modelos

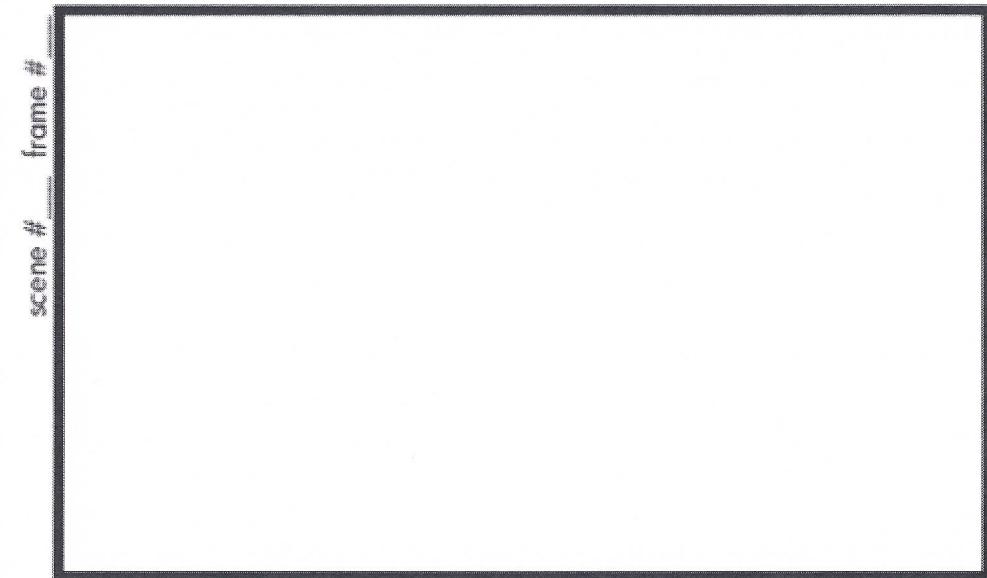
- Polinómicos de orden arbitrario
- Personalizados con código en R (Opcional)

Los modelos se vuelven a ajustar

#5 VH Añadir datos con un click (Opcional)



Cuando el usuario pulse en la gráfica aparece un nuevo dato y los modelos se reentrenan



Flight explorer

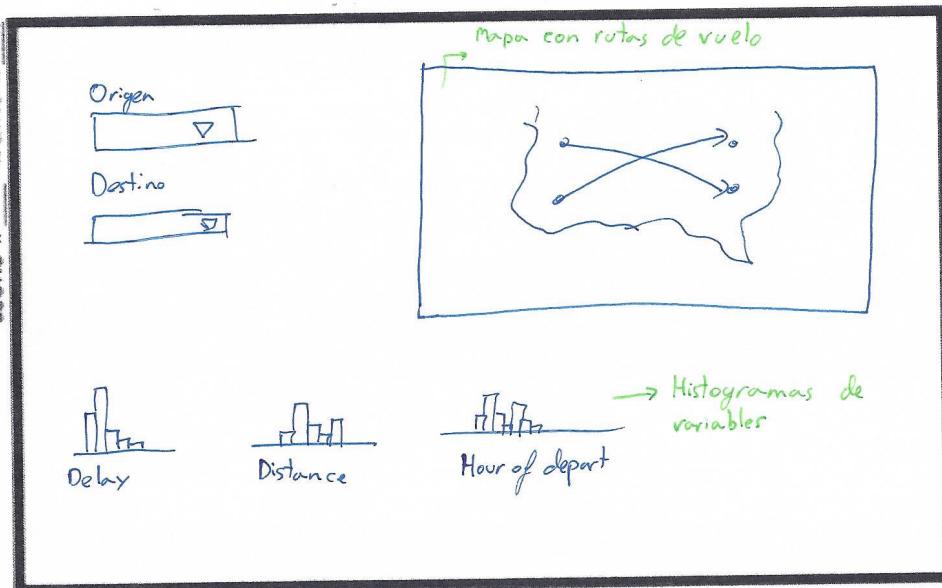
Herramienta que permite explorar el dataset de delays de vuelos con dos filtros por Origen y destino.

Se muestran histogramas de varias variables y en el caso de que los filtros elijan origen y destino las variables constantes (p.e: distancia) se muestran como un número constante en vez de una distribución.

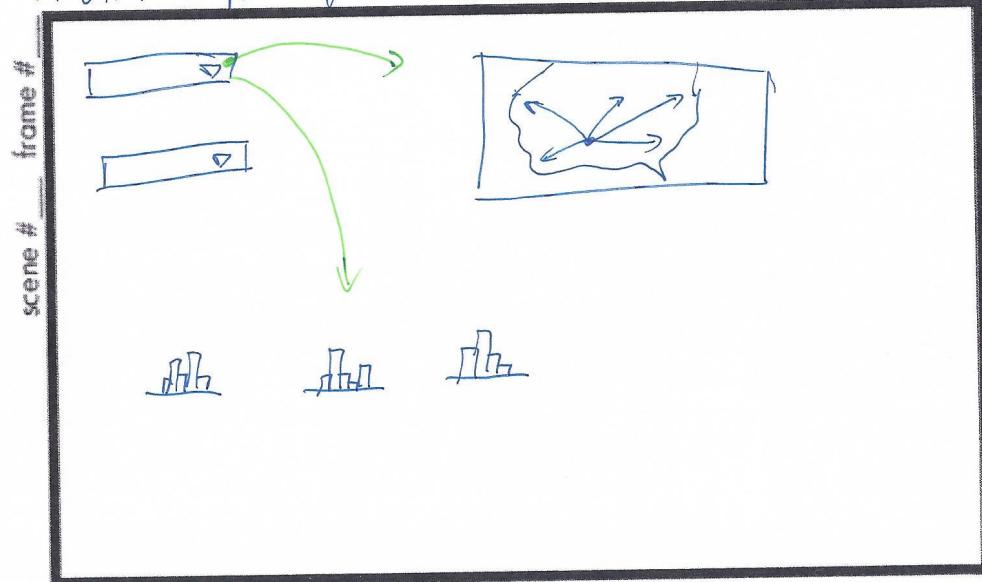
También se puede filtrar haciendo un recuadro en el mapa. Además se puede elegir el mapeo de otra variable alternativa y de una función agregadora para que el tamaño de los puntos de los aeropuertos cambie.

Flight Crossfilter

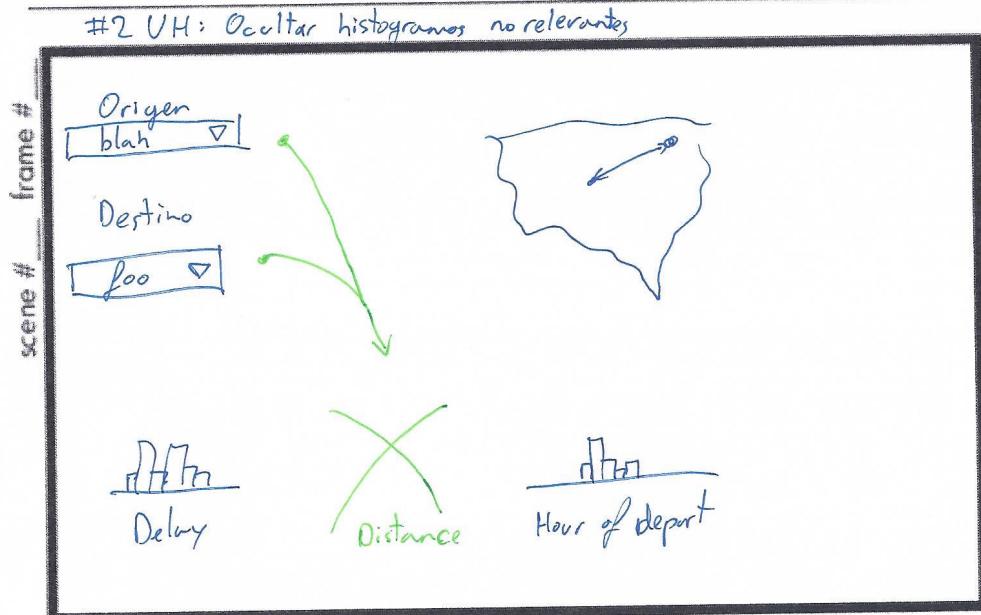
frame # — scene # —



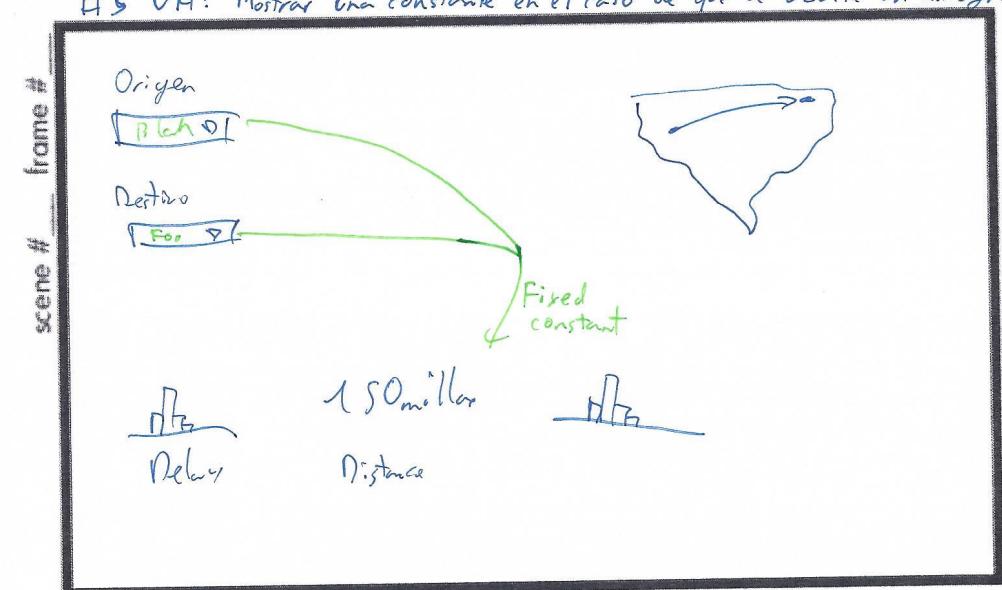
#1 VH Filtrar por origen o destino



Este filtro aplica tanto al mapa



#3 VH: Mostrar una constante en el caso de que se oculte un histograma



Ocultar los histogramas no relevantes cuando se aplican filtros

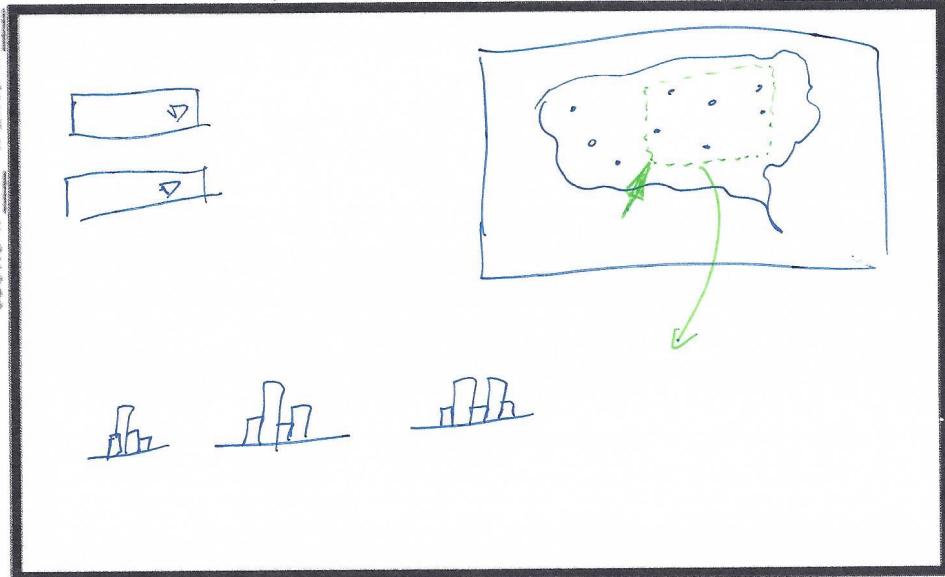
si hay una distribución nula: siendo el origen y el destino no se mueven

no tiene una distancia pre-given por la distancia

Flight crossfilter

#4 VH: Poder filtrar haciendo un recuadro en el mapa (Opcional)

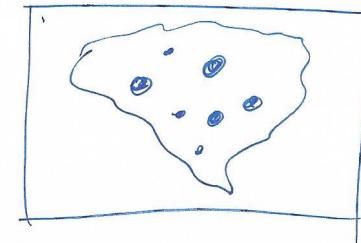
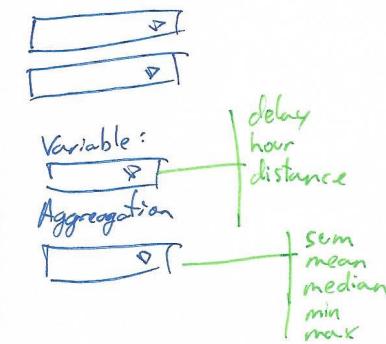
scene # frame #



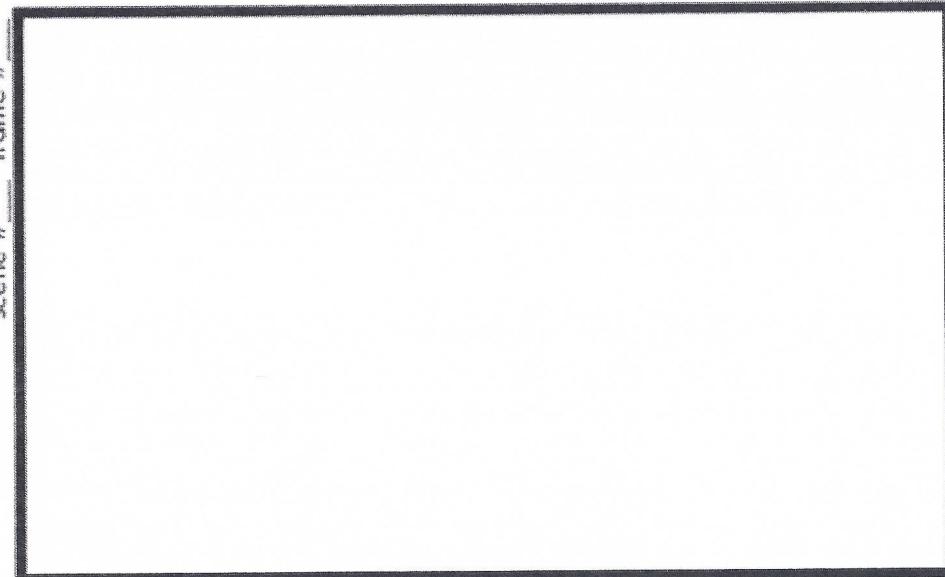
#SVH: Poder elegir una variable y un agregador que se

maapee al tamaño de los círculos:

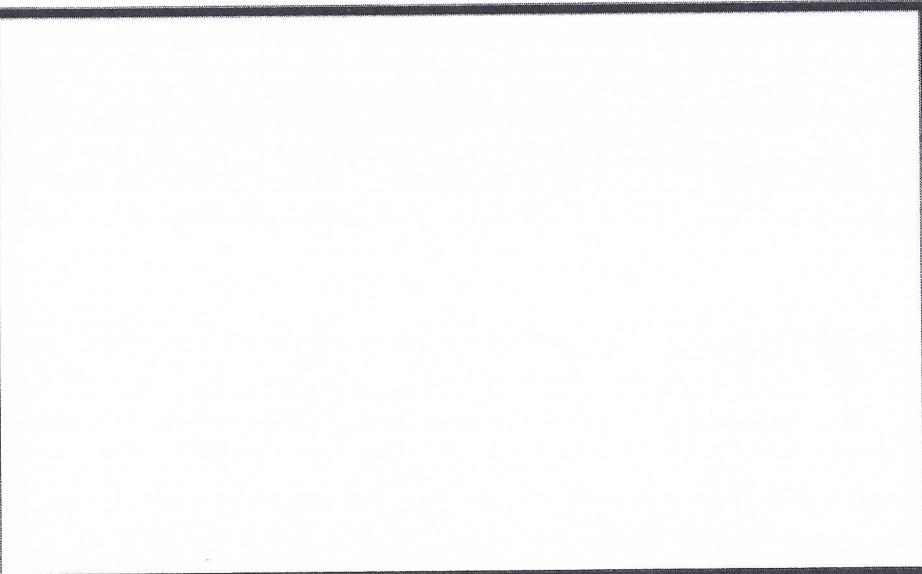
scene # frame #



scene # frame #



scene # frame #



Function explorer

Aplicación interactiva que permite explorar funciones n-dimensionales. Permite cualquier expresión de R arbitraria.

Las variables de la función será cualquier expresión no conocida por R (p.e: x,y, eta o alpha en los ejemplos)

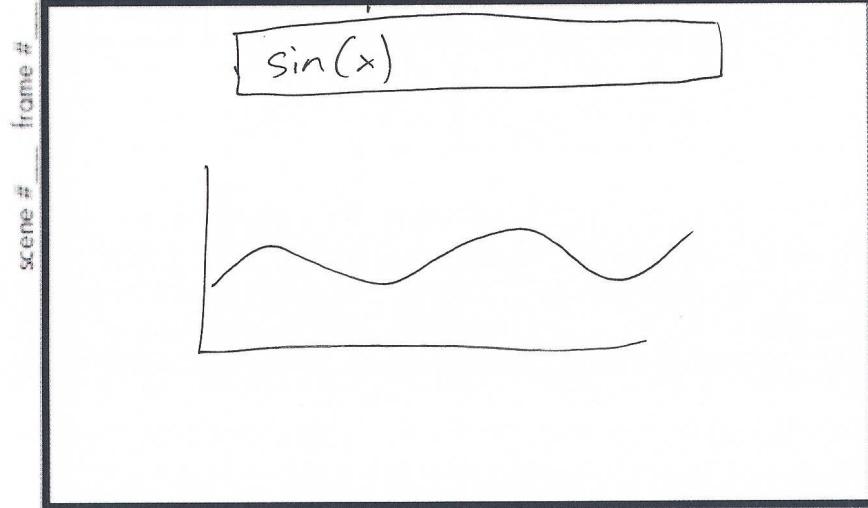
Según sea unidimensional o de dos dimensiones o más se dibujará una gráfica de líneas o una superficie.

Además se añade información sobre la transformada de Fourier de la Función, su dominio y su rango.

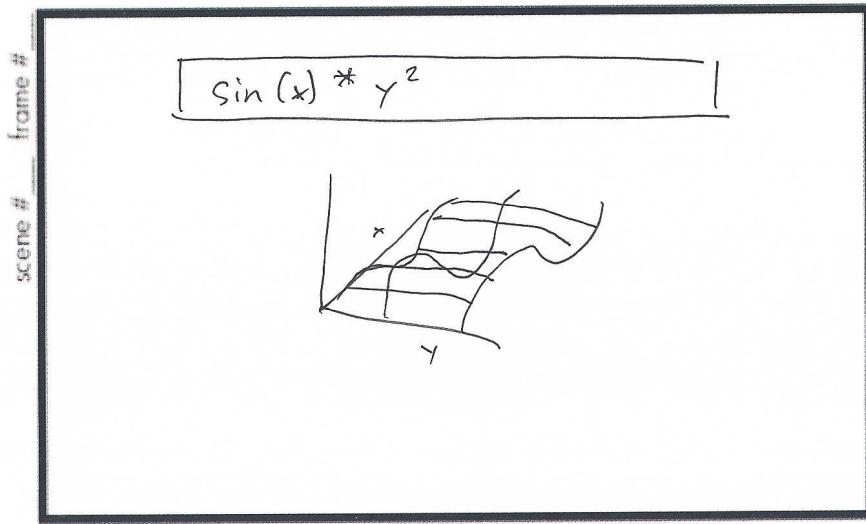
En el caso de que haya más de dos variables, las restantes (cogiendo aquellas con longitud mayor) se permiten modificar mediante sliders que se pueden animar. De manera que el usuario pueda "navegar" a través de la animación viendo el efecto de un parámetro

#1 U.H. Introduce function (univariate)

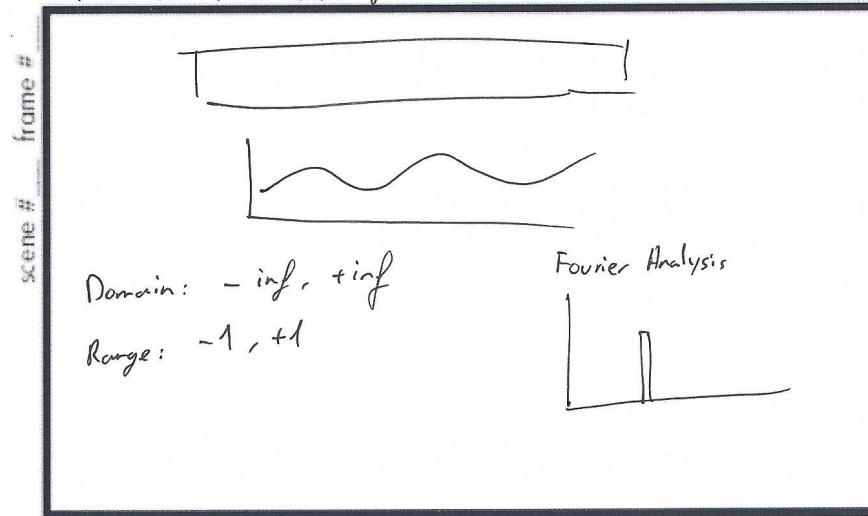
Function explorer



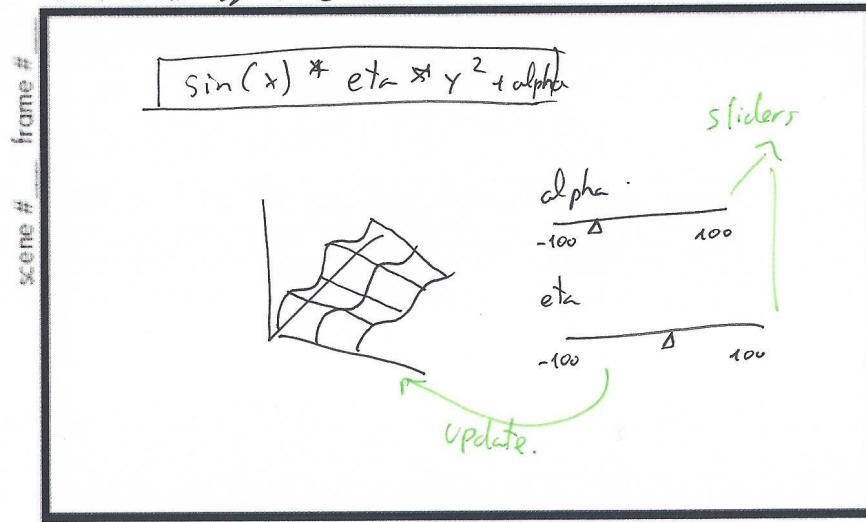
#2 U.H. Introduce bivariate



#3.U.H Show extra info



#4 Show ~~for~~ > 2 variables



GGplot2 themer

Visor de visualizaciones el ggplot2. Permite modificar el código fuente de la visualización y vienen algunos ejemplos por defecto.

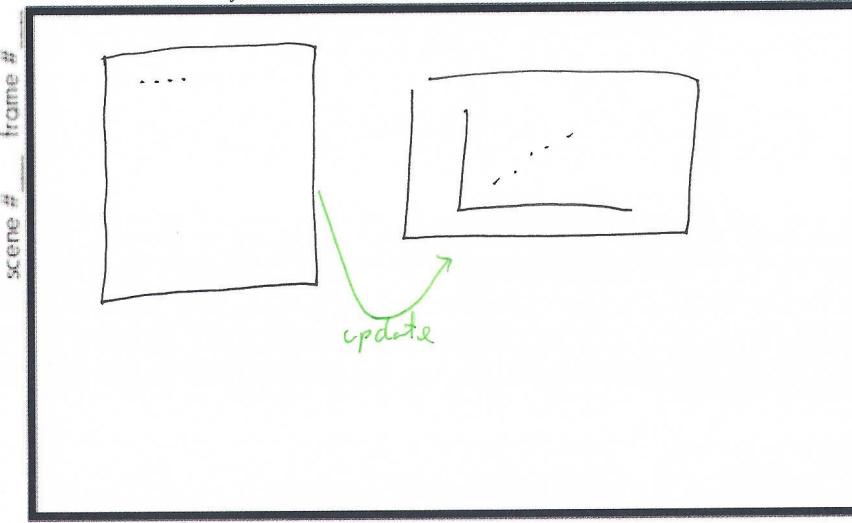
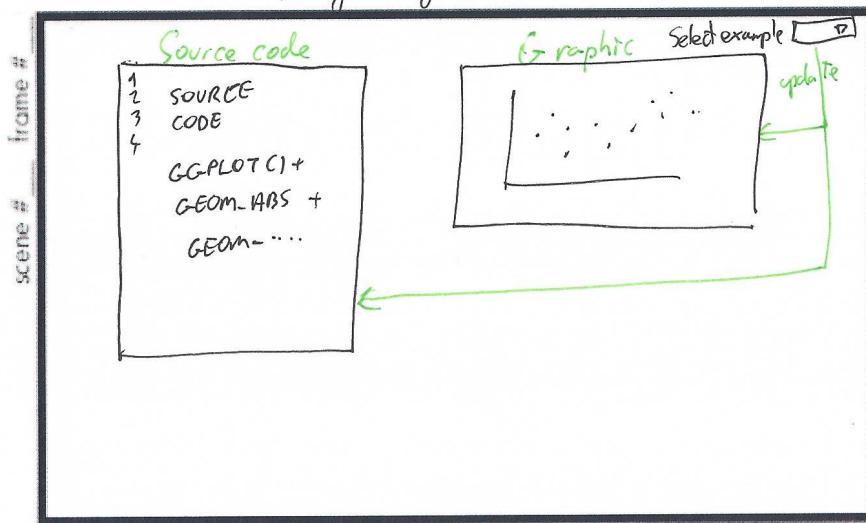
Además tiene una sección de estética en la que se permite modificar parámetros visuales de la gráfica (tamaño de letra, color, etc) y se permite descargar al usuario el código en R que aplica esos estilos a otras gráficas.

#1 U.H. Select predefined graph

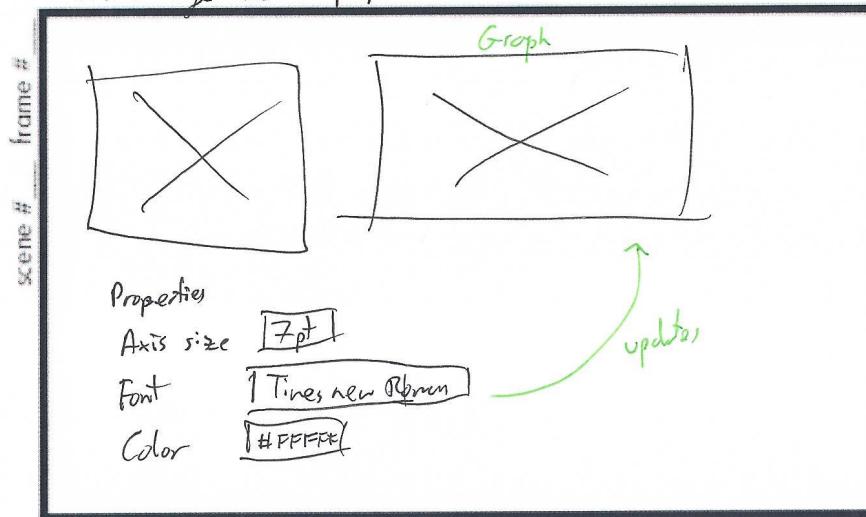
ggplot2

themer

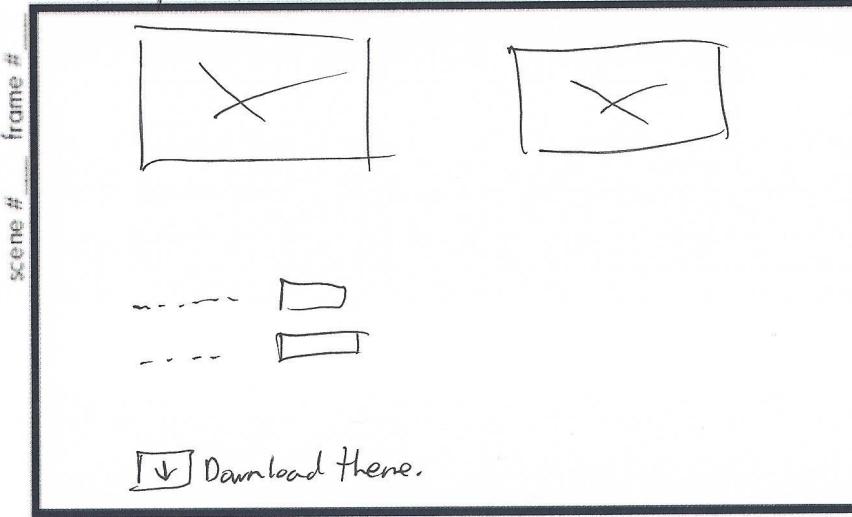
#2 Change source code



#3 Change visual properties



#4 Export theme



Cuando el usuario pulse en download theme se descargará un archivo de R con el código fuente necesario para aplicar el tema visual aplicado

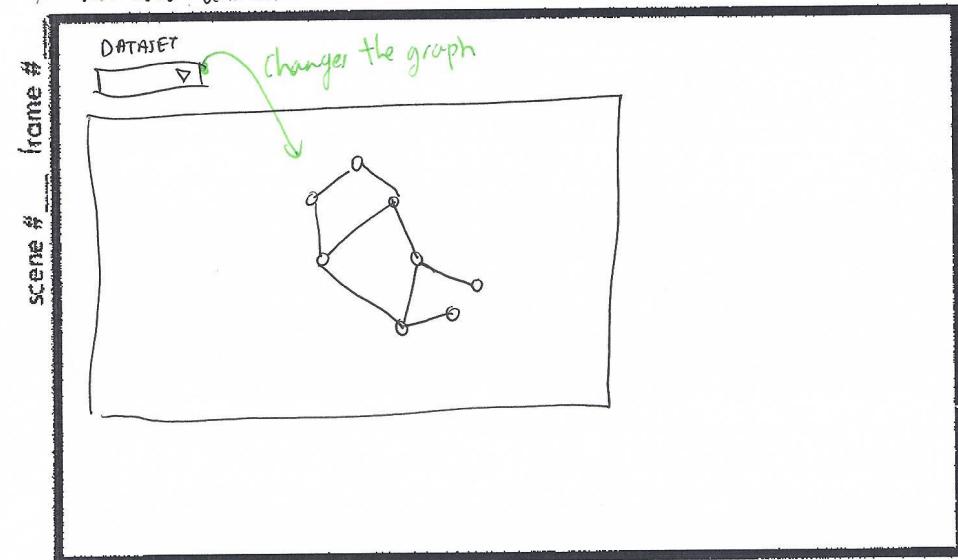
Graph explorer

Herramienta de exploración de grafos. Permite elegir distintos datasets y visualizarlos. Para visualizarlos se puede utilizar [networkD3](#) (http://www.htmlwidgets.org/showcase_networkD3.html) o [visNetwork](#) (http://www.htmlwidgets.org/showcase_visNetwork.html).

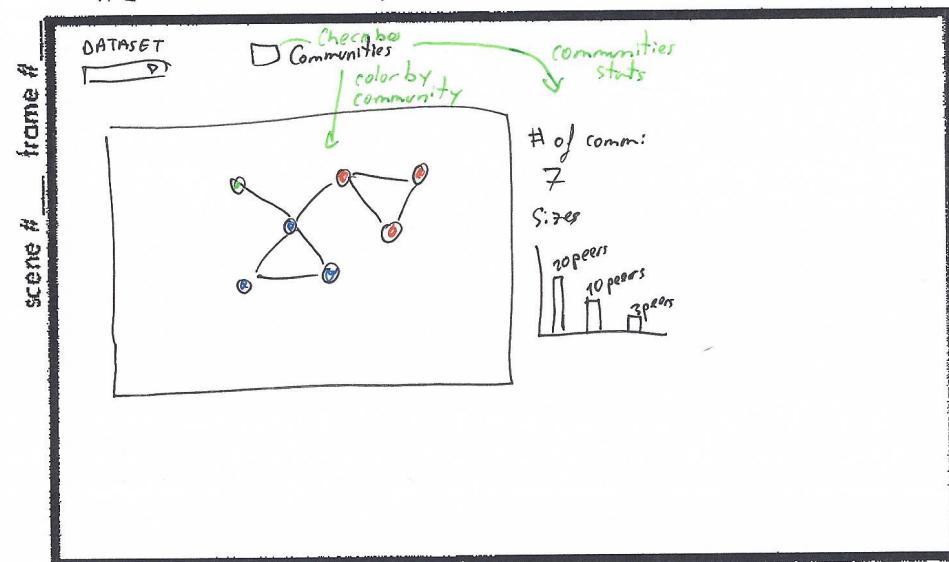
El usuario puede realizar un [análisis de comunidades](#) (https://es.wikipedia.org/wiki/Estructura_de_comunidades), (estudios de centralidad) [<https://en.wikipedia.org/wiki/Centrality>] y posibles rutas de fuga de información entre dos nodos

Graph explorer

#1 Select dataset

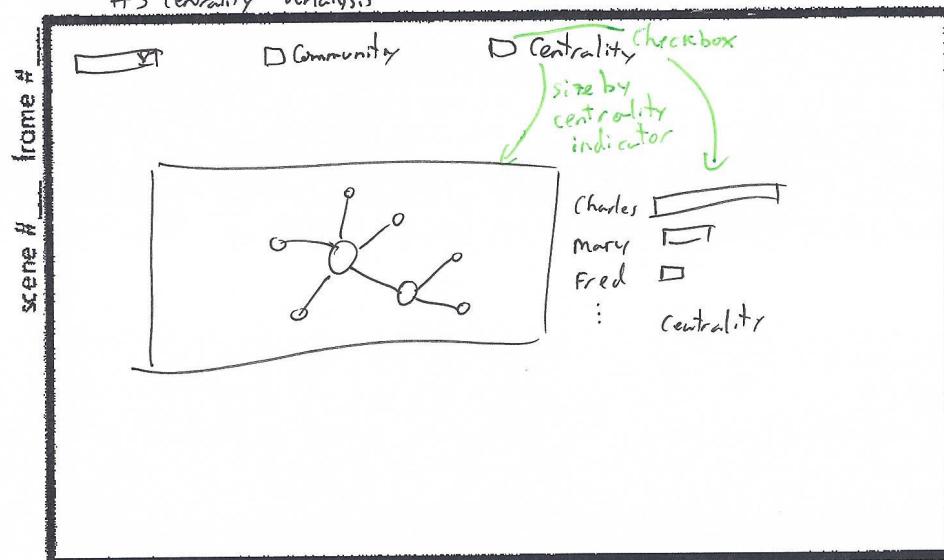


#2 Communities analysis

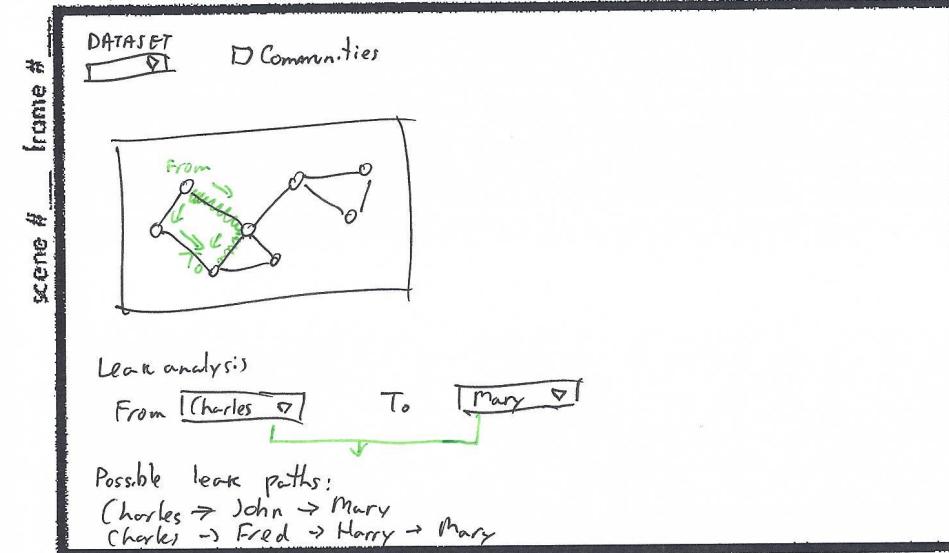


Cuando el usuario marca el checkbox se colorean los nodos por comunidad

#3 Centrality analysis



#4 Leak analysis



Pone el tamaño de los nodos según su centralidad y muestra los 10 nodos más centrales

Si se selecciona un nodo origen y un nodo destino se muestran las 3 rutas más cortas de posible fuga de información

Guess the correlation

Inspirado en (Guess the correlation)[<http://guessthecorrelation.com/>]

Pequeño conjunto de juegos relacionados con correlaciones.

Primer minijuego de adivinar la correlación de un dataset aleatorio.

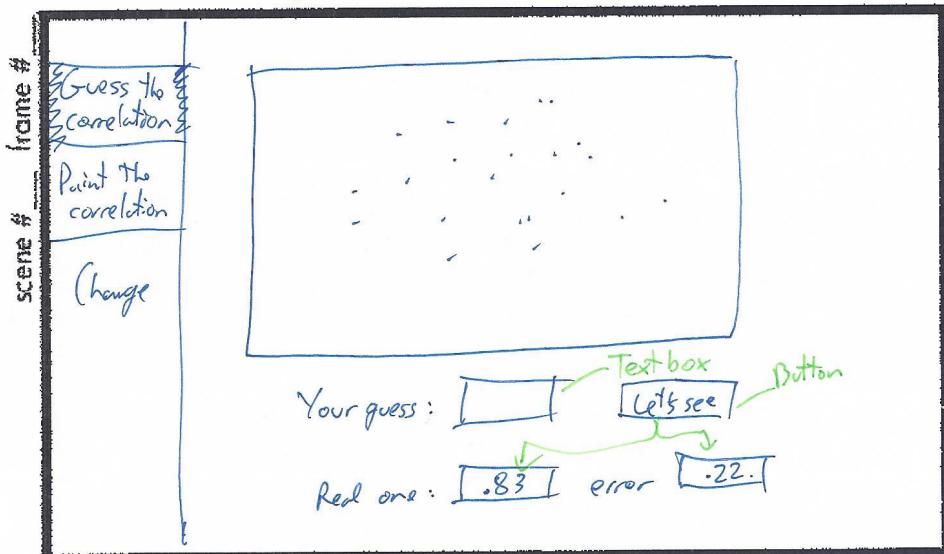
Un segundo de pintar un dataset que cumpla la correlación de pearson solicitada.

Otro juego que consiste en modificar un dataset ya existente para cambiar la correlación

Y otro juego que trata de adivinar las medianas de ambas variables

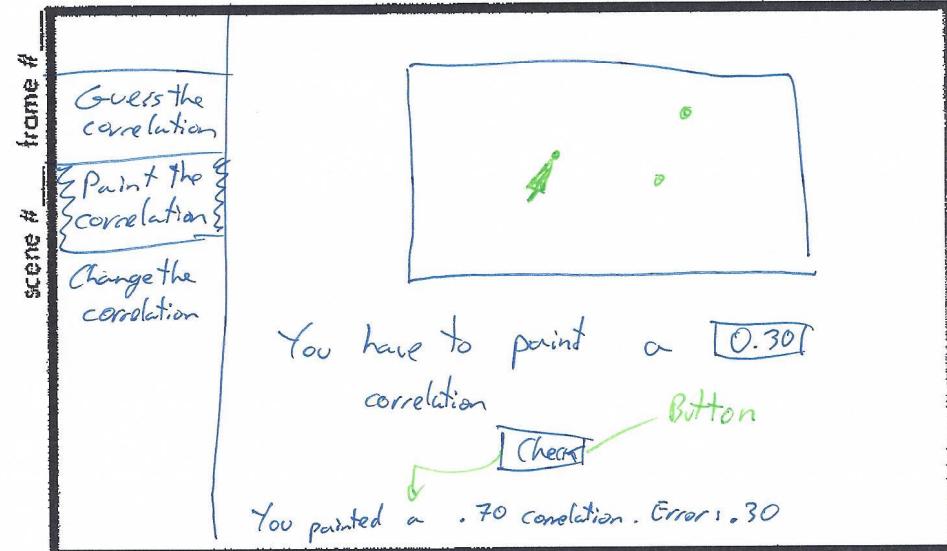
Guess the correlation

#1 U.H. Guess the Correlation



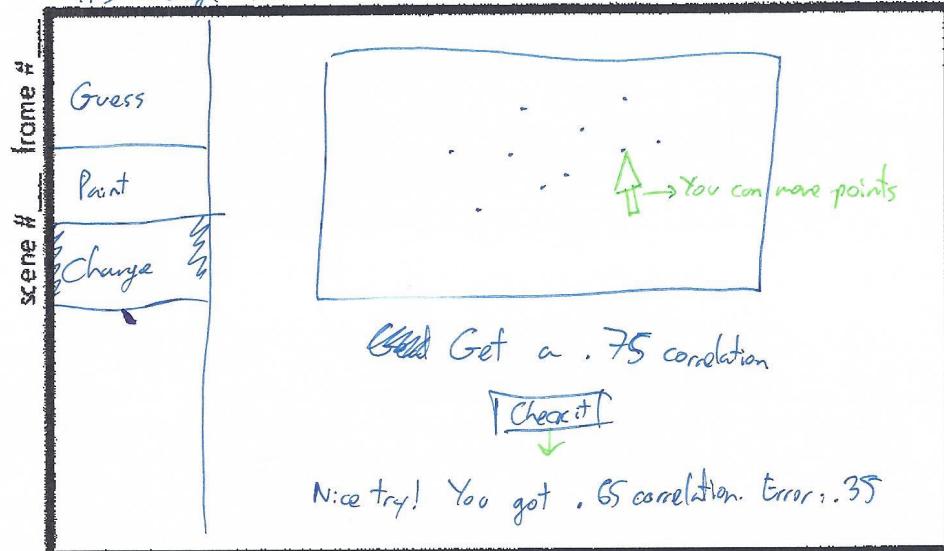
Cuando el usuario pulse "Let's see" se ve la correlación real y el error

#2 U.H. Paint the correlation



Cuando el usuario pulsa sobre la gráfica en blanco se añade un punto de dolor. Si el usuario pulsa "Check" aparece un mensaje con la que ha creado el usuario y el error.

#3 Change the correlation

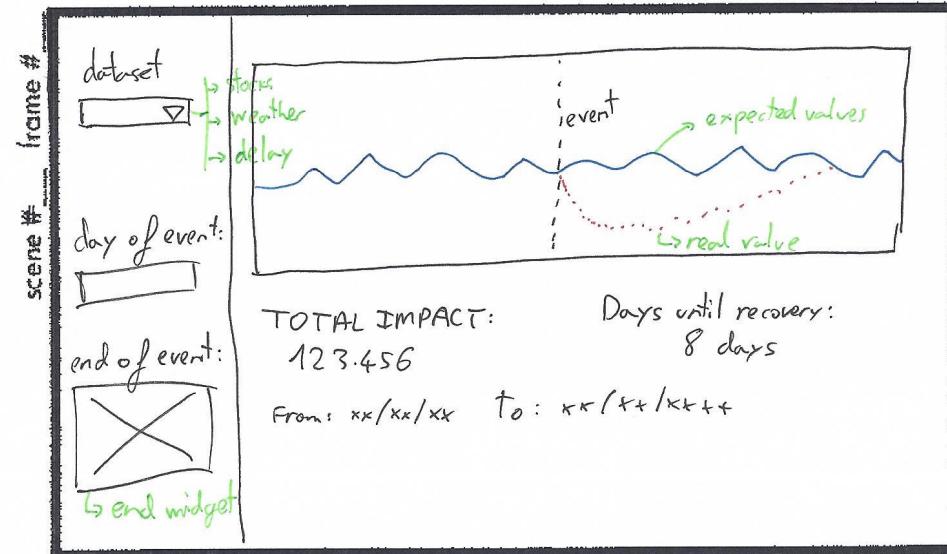


Impact effect model

Herramienta visual para utilizar el modelo estadístico (Causal Impact)[<https://google.github.io/CausalImpact/CausalImpact.html>] con distintos datasets de ejemplo.

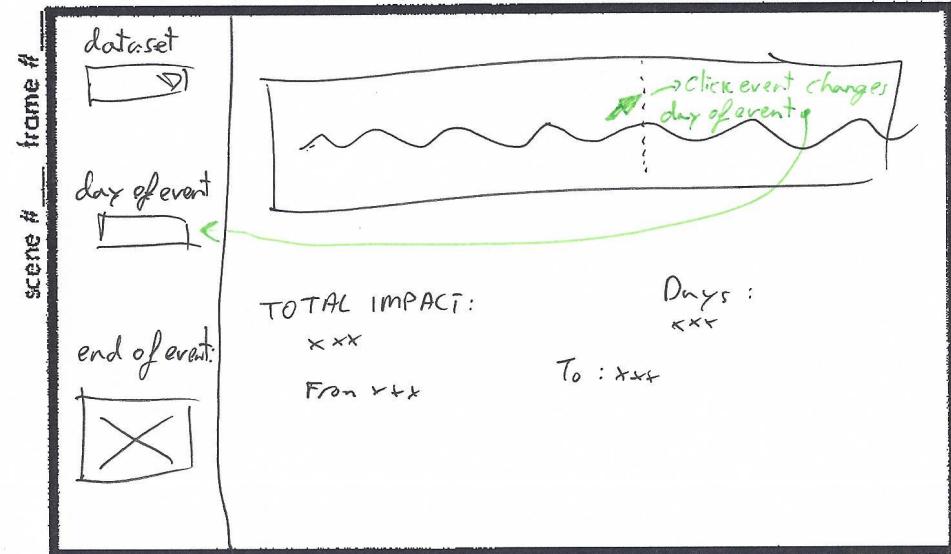
El usuario puede elegir distintos criterios de "parada" del impacto y visualizar distintas métricas (impacto acumulado, puntual y relativo)

#1 U.H: Mostrar estadísticas básicas del evento

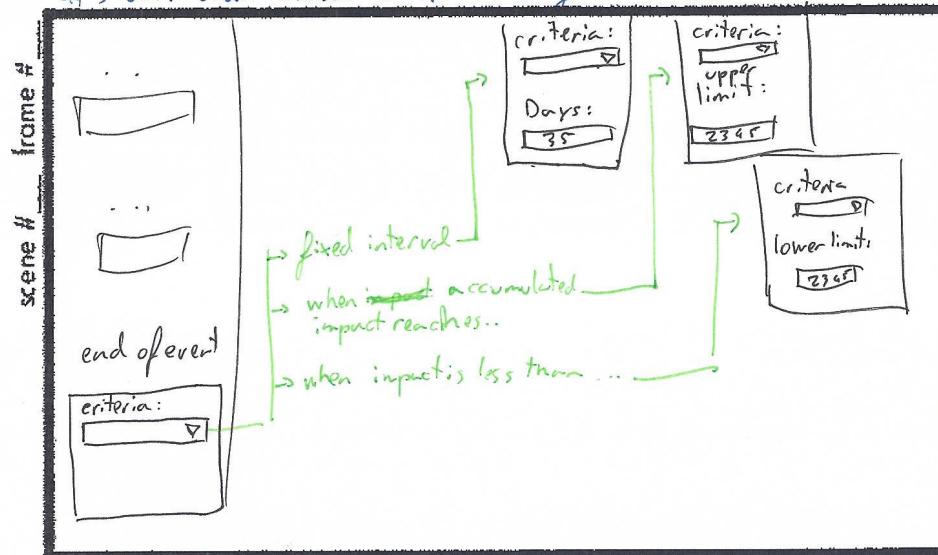


Impact effect model

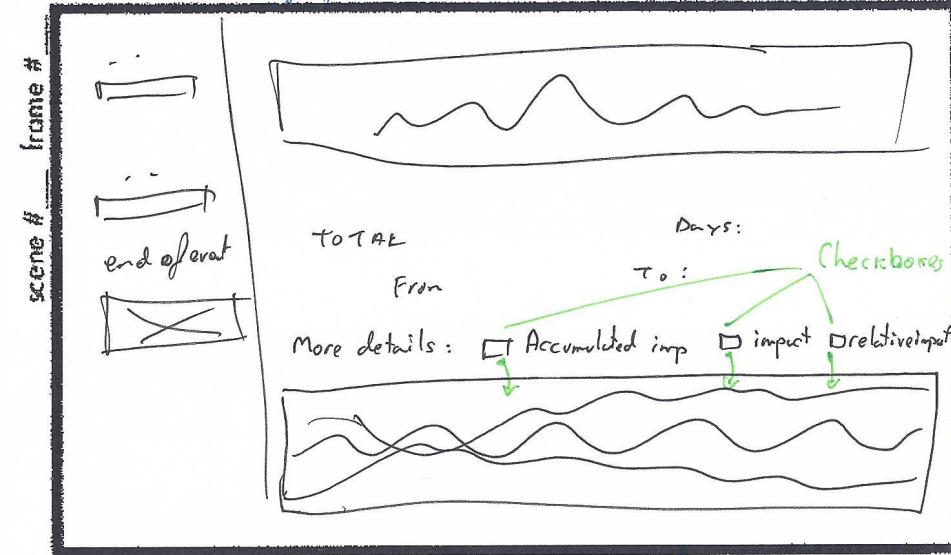
#2. U.H Señalar el evento con un click



#3 U.H: Cambiar el criterio de fin del evento



#4 U.H Ver otras gráficas



El usuario puede marcar que otras gráficas quiere ver pulsando marcar varias simultáneamente.

Influence Analysis

Herramienta de Visual Analytics para diferenciar de una muestra bivariada aquellos datos que son outliers (con alto error) de aquellos que tienen mucha influencia en la regresión.

El usuario puede generar una muestra al azar del tamaño seleccionado.

Hay una visión de outliers en la que se muestra la línea de regresión y de errores fijos. También se codifica en tamaño el error de estimación de cada punto.

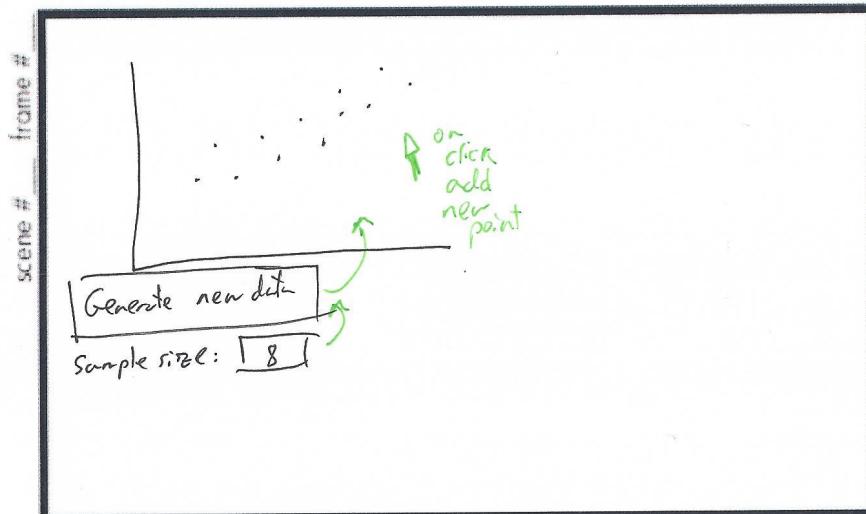
Hay una visión de influencia en la cual el tamaño del círculo se mapea al indicador de influencia y se calculan las regresiones de toda la muestra y la muestra excluyendo este punto.

Ambas regresiones se superponen a los datos y se comparan.

#10.H. Create datasets

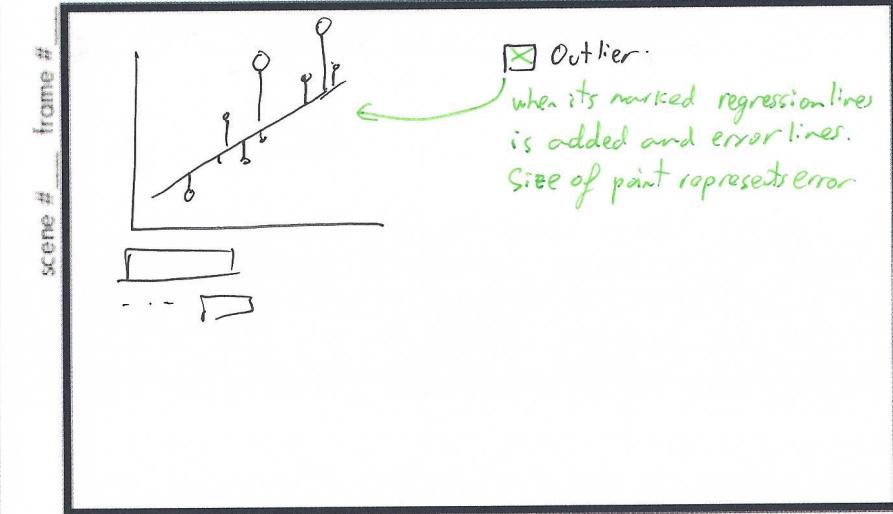
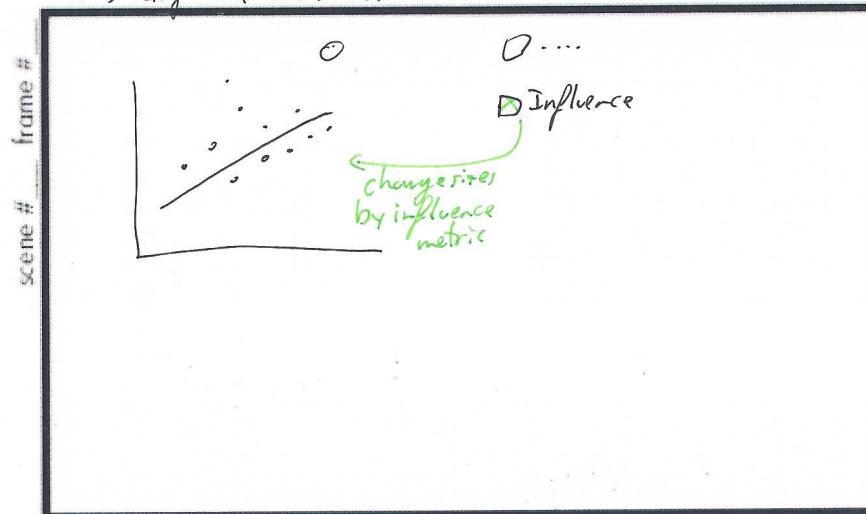
Influence Analysis

2. Outliers detection

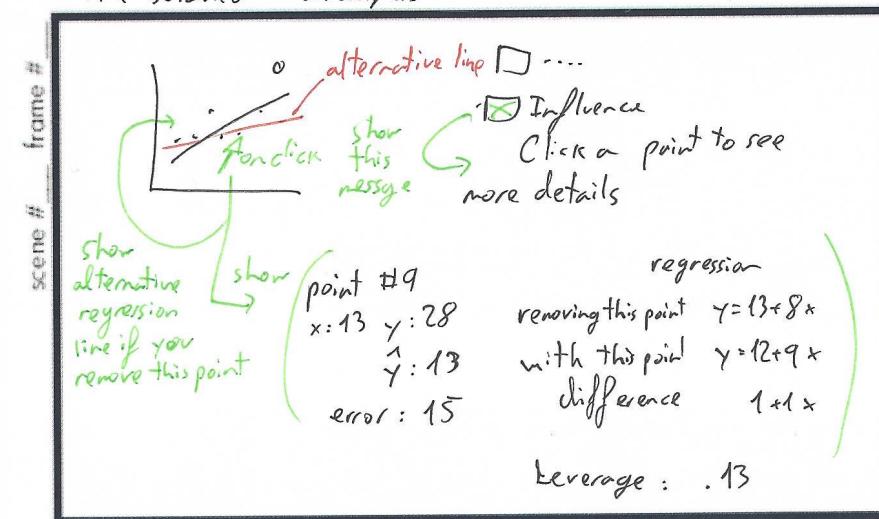


Al abrir la aplica

#3 Influencer detector



#4 Detailed analysis



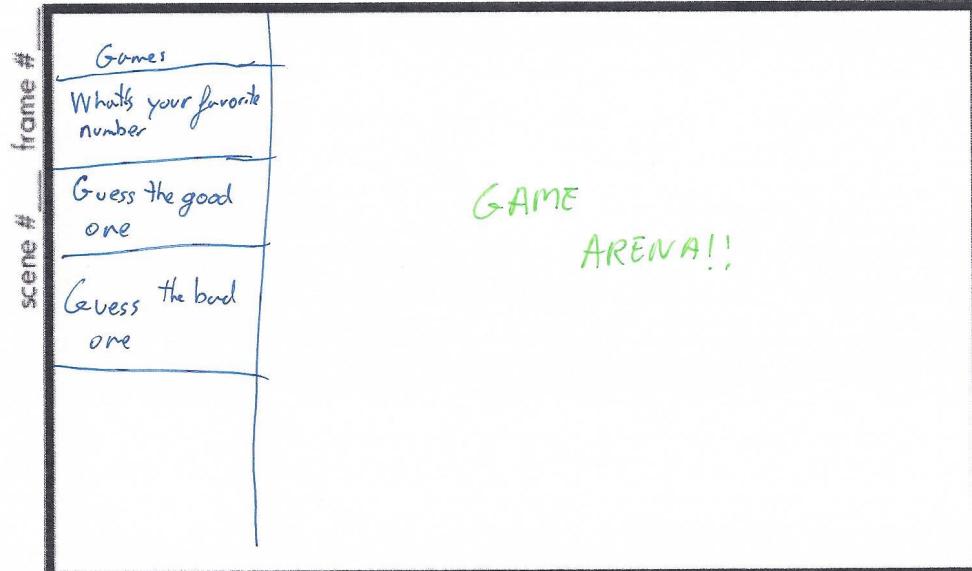
Permitir al usuario seleccionar un punto para ver el detalle de él (predicción, valores, leverage) y la regresión al excluir este punto se visualizará (línea naranja)

Lotto bias

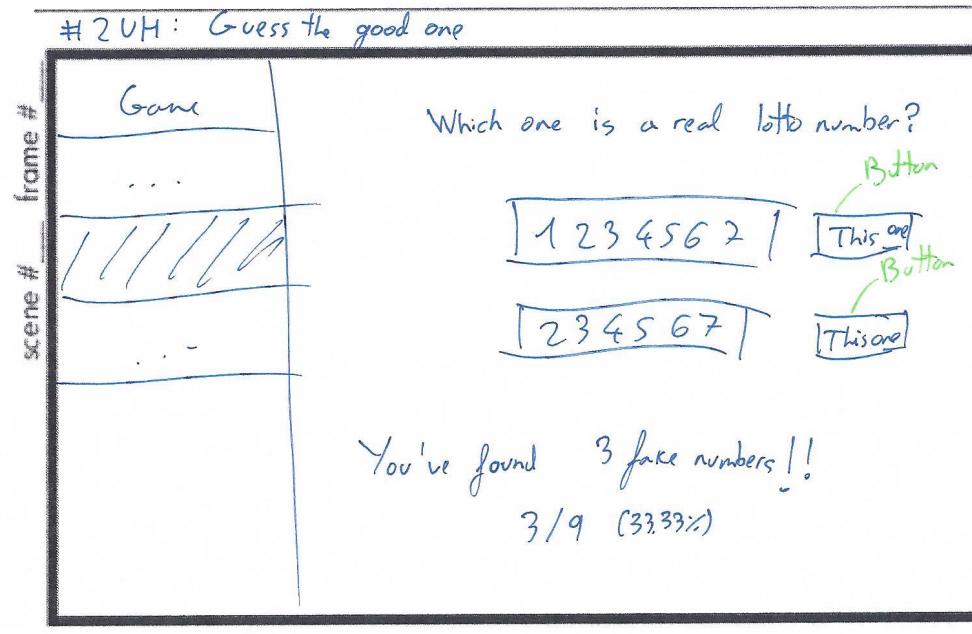
Herramienta para estudiar los sesgos perceptivos de azar mediante un juego de loterías. Para ello se utilizan los datos reales de loterías públicas. Se compone de 3 minijuegos:

1. What's your favourite number?: El usuario selecciona un número y se revisa en el histórico real de loterías cuántas veces ha aparecido ese número mostrando la estadística al usuario
2. Guess the good one: Se muestra un número del dataset real de loterías y otro generado al azar por R. Existe una creencia a que los números que tocan en la lotería cumplen ciertos patrones (repetitividad, recencia negativa, no empiezan por cero, etc.). Como son totalmente aleatorios el sujeto no debería poder distinguirlos entre ambos. Las posiciones del número "real" y el aleatorio se sortean cada vez que el usuario responde.
3. De un dataset "human-trap numbers" (números "bonitos" o aparentemente más probables) se selecciona dos números y se añaden una opción más realmente aleatorias. El usuario debe seleccionar aquella que está diseñada para atraer la atención ("human-trap")

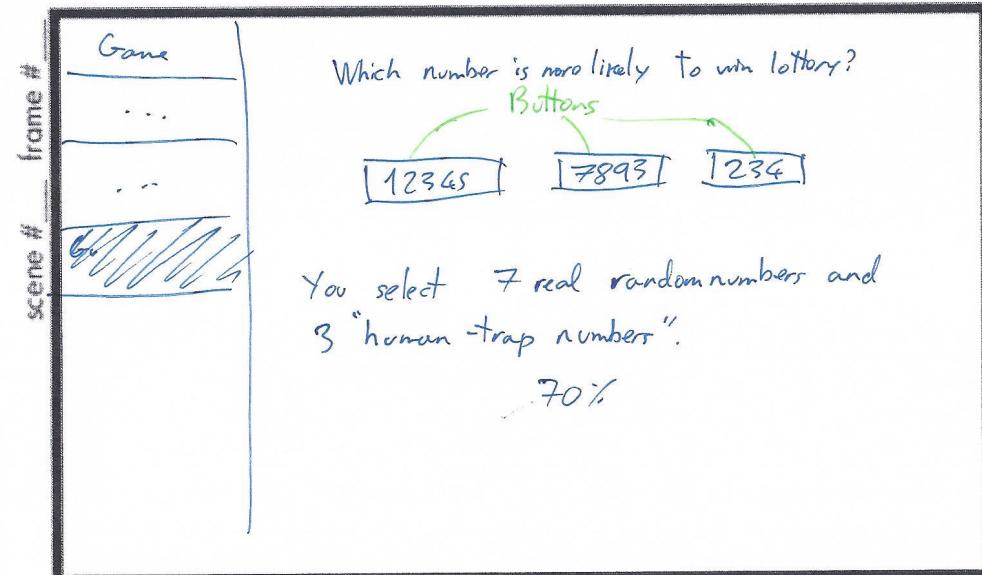
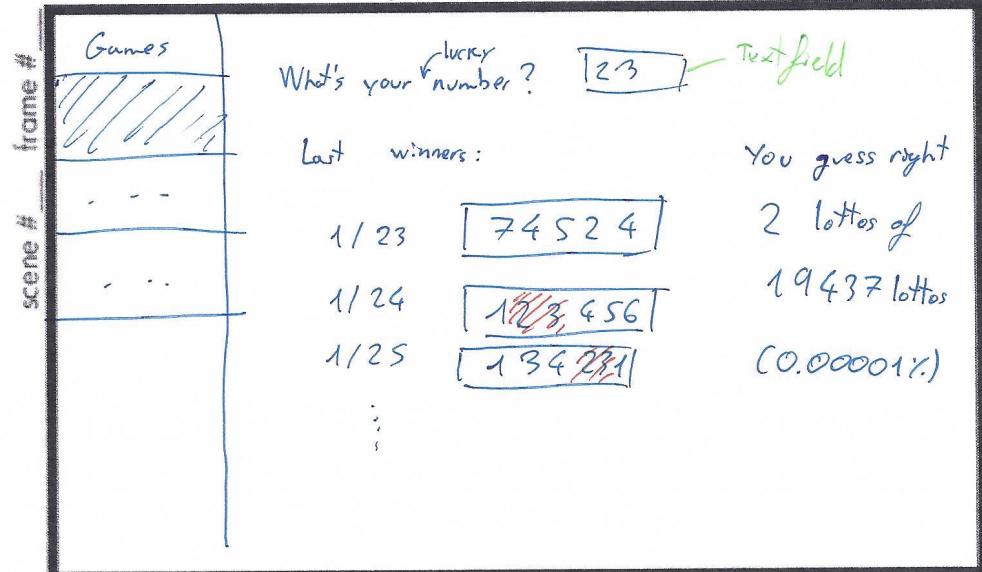
Lotto bias.



El usuario tiene un menú para seleccionar a que juego quiere jugar.



#1 UH: What's your favourite number



Monty hall paradox

La paradoja de Monty Hall es una contraintuición muy popular. Ver funcionamiento del juego en [Problema de Monty Hall \(Wikipedia\)](#) (https://es.wikipedia.org/wiki/Problema_de_Monty_Hall).

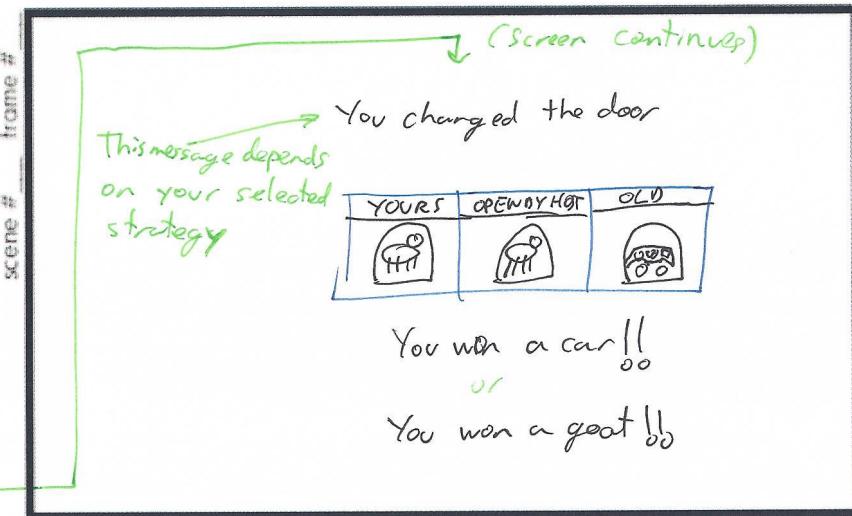
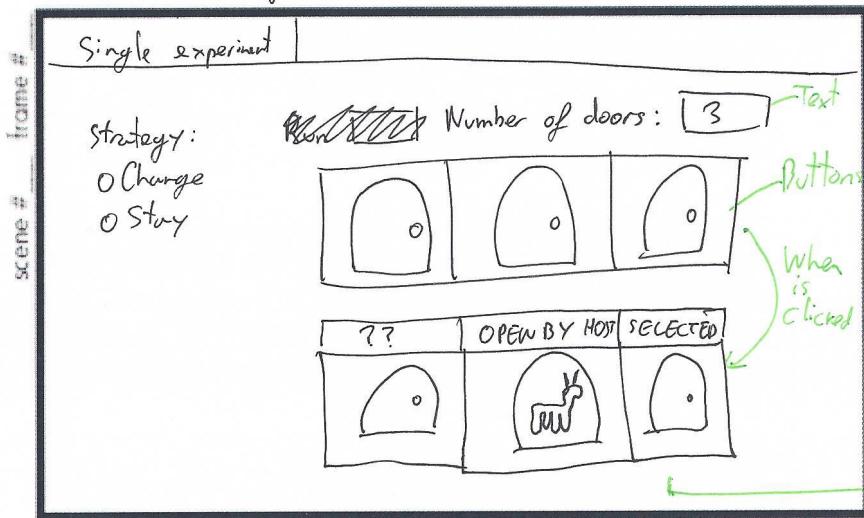
Se trata de reproducir cómo es el juego bajo dos estrategias de juego distintas: mantener la puerta y cambiar la puerta (leer el artículo de la Wikipedia).

El programa permite hacer simulaciones simples para enseñar el funcionamiento al usuario y cambiar el número de puertas (esta es una manera frecuente de enseñar esta paradoja).

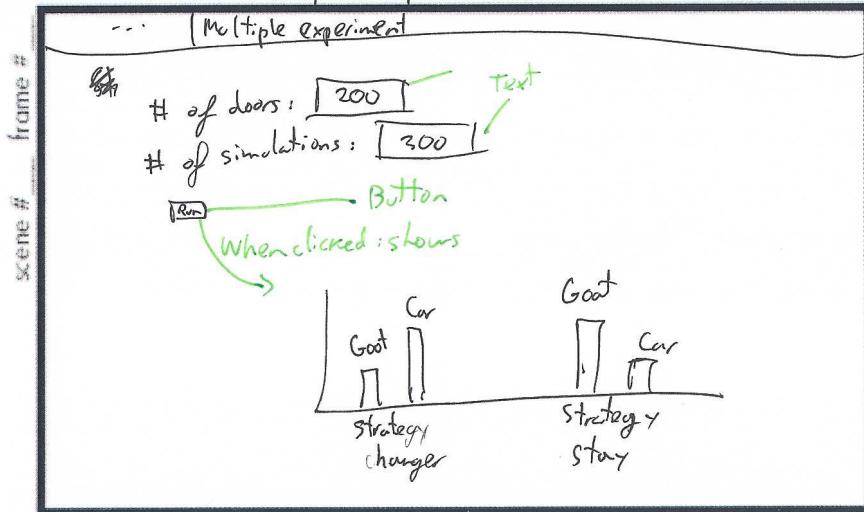
Otro apartado de la aplicación permite ejecutar múltiples simulaciones para comprobar que la estrategia óptima es cambiar de puerta.

Monty Hall Paradox

#1. U.H: Single experiment



#2 U.M Multiple experiment



Nutrients and calories calculator

Una herramienta para calcular las propiedades nutricionales y las calorías de una receta según sus ingredientes.

Para ellos se utilizará un dataset de ingredientes y propiedades nutricionales de ellos.

Se añade un buscador de alimentos altos en un nutriente (elegido por el usuario) y un juego de comparación de nutrientes en el que el usuario tiene que adivinar cuál de las dos propuestas tiene más cantidad de un nutriente.

#10.M: Calc calories

Nutrients and Kcal calculator

scene # frame #

Ingredient	Calories
15g lettuce	17 kcal
20g milk	5 kcal
100g bread	40 kcal

Add new ingredient

7% daily intake

20 minutes of walking

Change

Update stats

Button that opens a new window with a selector to change "reference sport"

#12 Nutrient analysis

scene # frame #

	vit A.	vit C	blah
50g garlic	3g	2g	1g
150g tomatoe	1g	2g	3g

Add 1 ingredient

Percentage of your intake

#3 Find good ingredients

scene # frame #

Find ingredient with high ...

lettuce ...

On click add that ingredient

Lettuce	7 g / 100g	[+]
Tomatoe	4 g / 100g	[+]

#4 Game

Game Calculator

scene # frame #

Which one has more vitamin C

Bottles

Correct: 7
Fail: 2

100g Banana

100g Swordfish

when they'll clicked update marks

Each time they'll click, nutrient and ingredients changes

Perceptual experiment

Utilizando las imágenes de MNIST (-ver Wikipedia-)[https://en.wikipedia.org/wiki/MNIST_database] esta aplicación permite diseñar y ejecutar un diseño experimental factorial.

La tarea a resolver es el reconocimiento del dígito mostrado la imagen original con un poco de ruido. El experimento trata de analizar el efecto del ruido en la percepción de los dígitos.

Para eso hay varios tipos de ruido: - Ruido blanco: en puntos aleatorios de la imagen se coge ruido de una distribución aleatoria - Zeroing: En puntos aleatorios se pone a 0 la casilla (se pone a negro) - Vertical lines: se añaden líneas verticales negras a la imagen al azar - Small squares: en regiones cuadradas al azar de un tamaño determinado se tapa la imagen

La aplicación muestra todas las combinaciones según se hayan activado tipos de ruido. (Es un diseño factorial, por tanto será multiplicar el número de alternativas. p.e: tipos de ruido* porcentajes de ruido*...)

En la sección play el usuario puede realizar el experimento. Se debe permitir que más de un usuario haga el experimento después.

En la sección results se pueden ver los resultados (en forma alargada).

P hacking

El p-hacking es la práctica ilegítima de modificar el diseño de un análisis o los datos con el fin de obtener un resultado estadísticamente significativo.

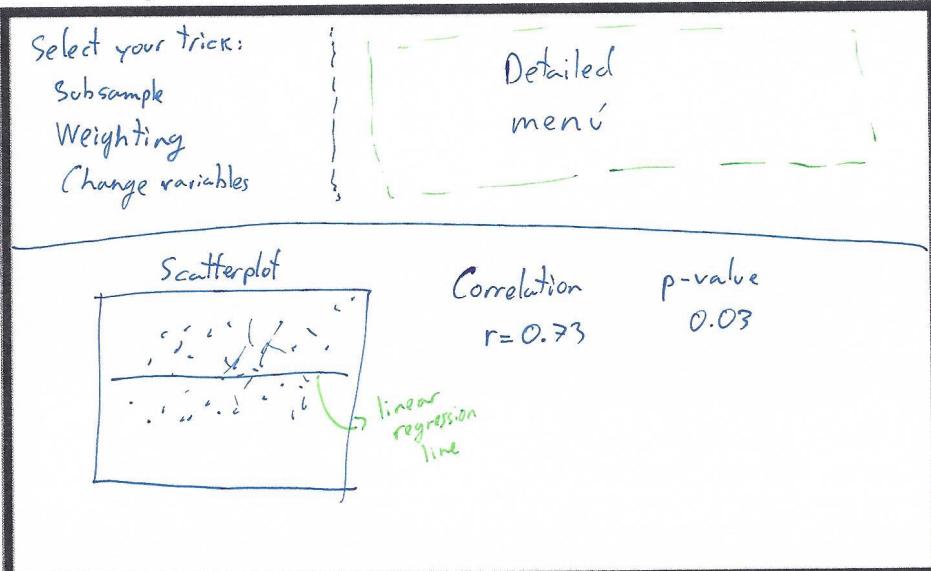
Muchos expertos apuntan a que estas prácticas se ven fomentadas por el sesgo de publicación de las revistas científicas que sólo suelen publicar resultados significativos aunque eso no es siempre relevante.

Esta aplicación permite mostrar distintas estrategias para manipular los valores p de una correlación:

- Subsample: coger una muestra más pequeña hasta que obtengamos un resultado significativo
- Weighting: ponderar los casos por una tercera variable
- Change variables: elegir otras variables del dataset que correlacionen más

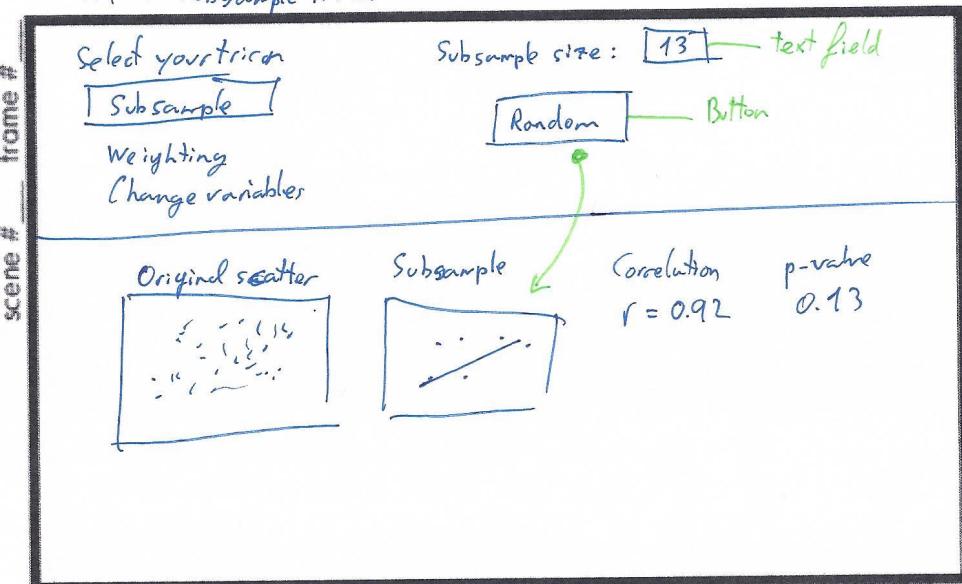
p hacking

frame # scene #



#1 : Subsample trick

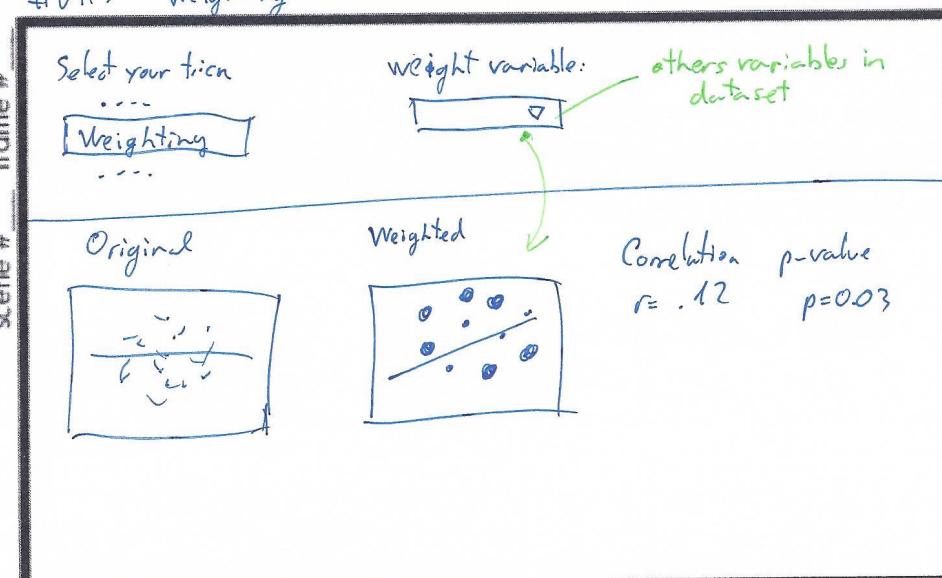
frame # scene #



En esta aplicación se dibuja una población bivariada en un scatterplot, la correlación de pearson y el p-value asociado

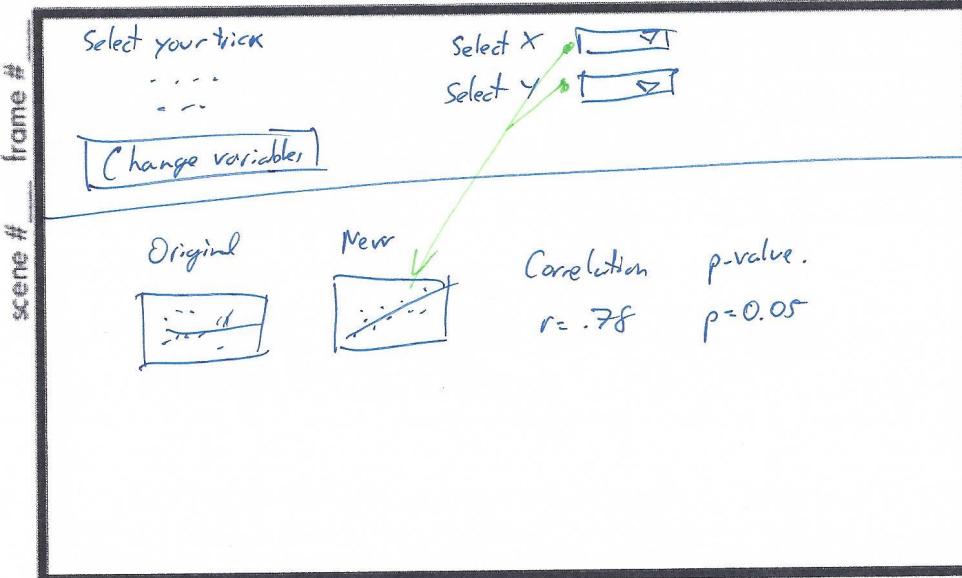
#1/H3: Weighting

frame # scene #



Cuando el usuario elige una variable para pesar los datos se muestra un scatterplot en el que el tamaño de los círculos y la regresión se actualizan con esos pesos

frame # scene #



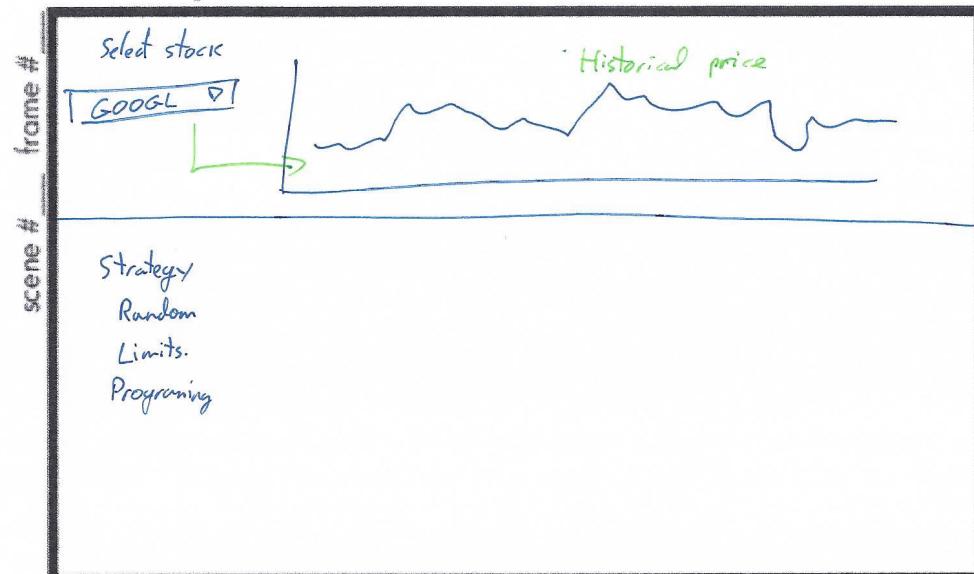
Stock

Aplicación para estudiar el rendimiento contra histórico de valores de stock.

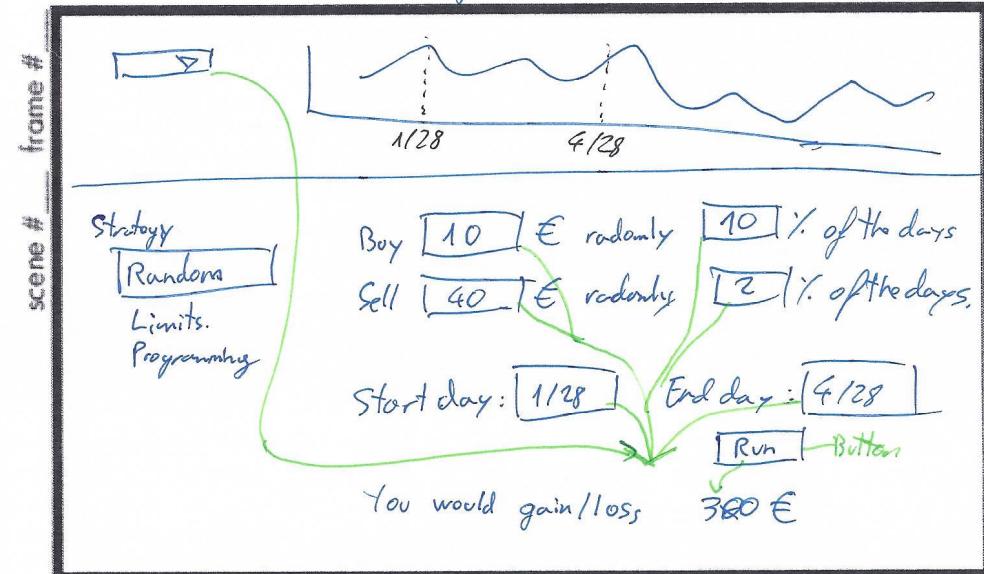
Hay tres estrategias: - Random: compra y venta al azar de una cantidad fija. - Limits: comprar y vender al alcanzar un límite superior e inferior - Programmatic: puedes programar tu propio algoritmo de trading en R y evaluarlo

Stock market simulator

#1 U.H. Random strategy

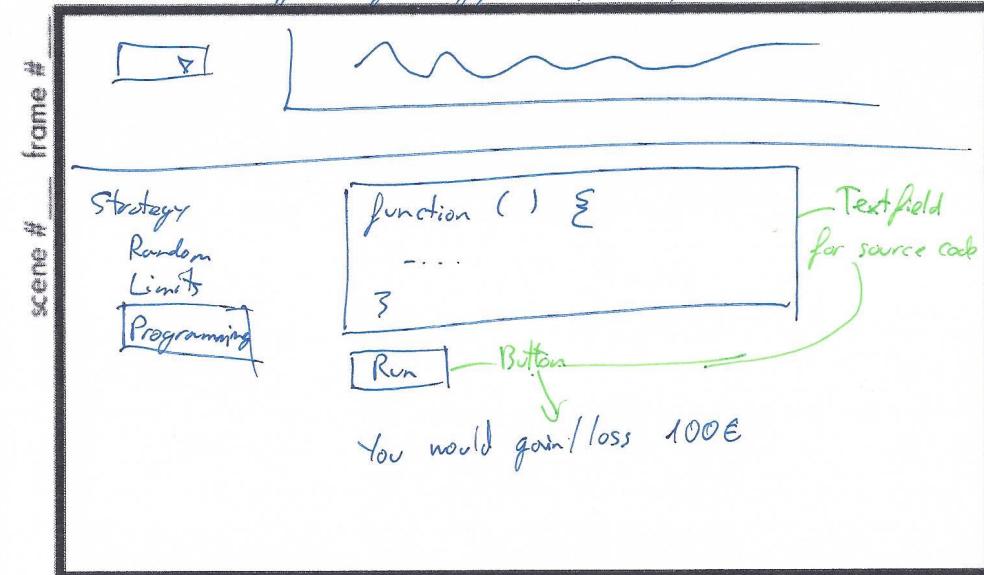
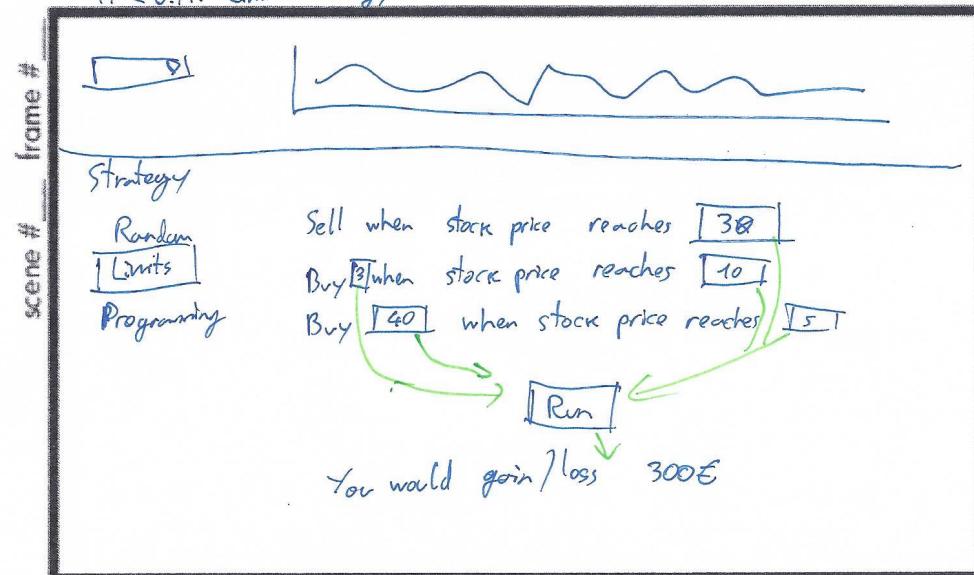


Cuando el usuario cambia en el selector el stock la grafica del historico se actualiza



Cuando el usuario pulse en "Run" ejecutaremos una simulación de cuanto hubieras ganado con esa estrategia. También cuando el usuario cambia de stock

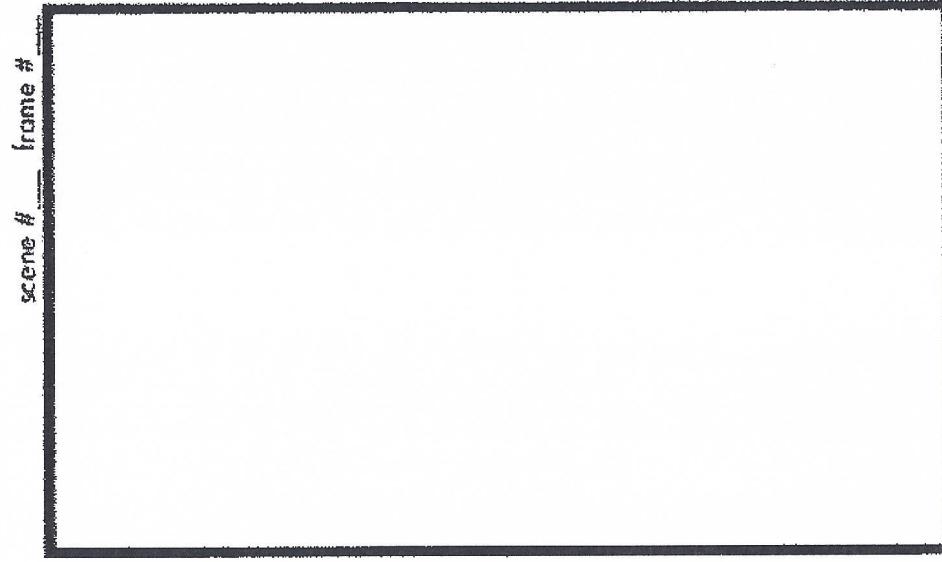
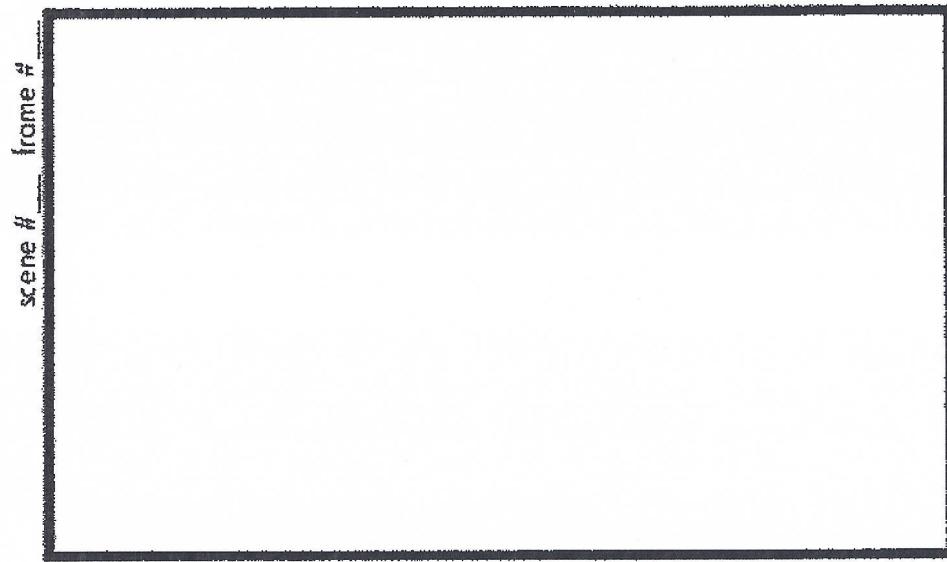
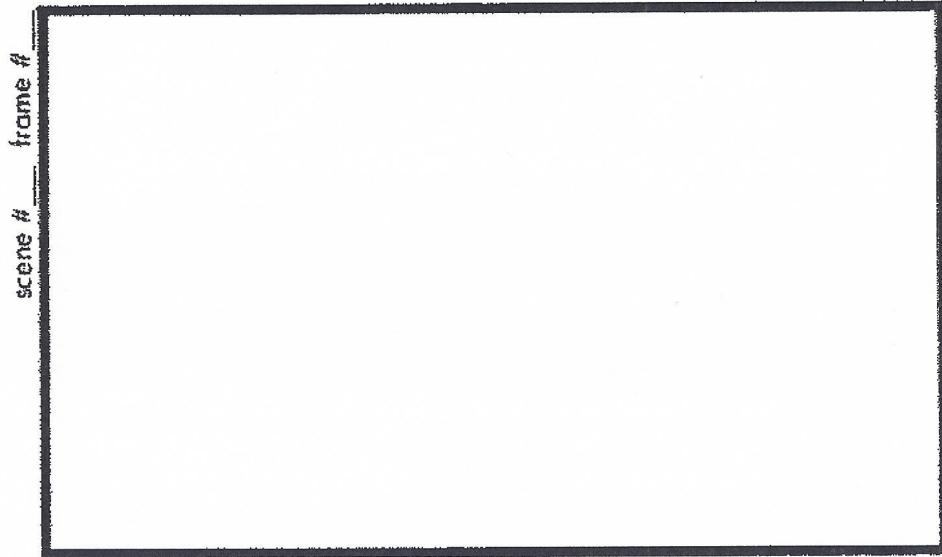
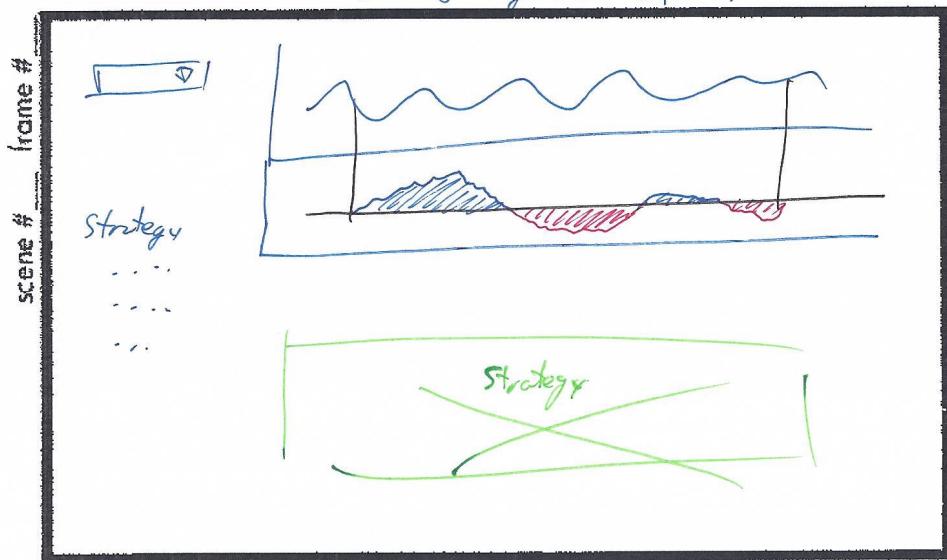
#2 U.H. Limits strategy



El usuario puede programar su propia estrategia en R

#4 U.H: Visualize your gaining historical.(optional)

Stock market simulator (II)



Viz designer

Herramienta que imita a una interfaz amigable para crear visualizaciones.

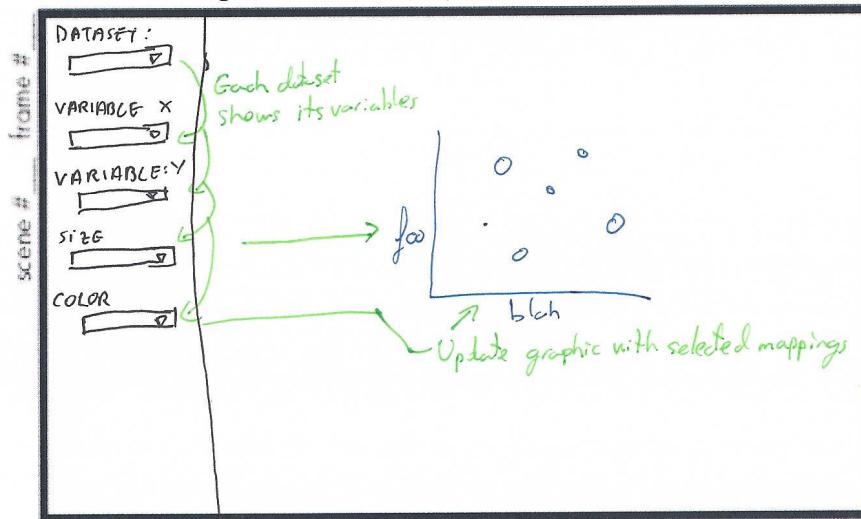
Te permite elegir el dataset y las variables a mapear al eje X, Y, color y tamaño.

También permite animar la visualización mediante un slider y hacer facets de la misma visualización.

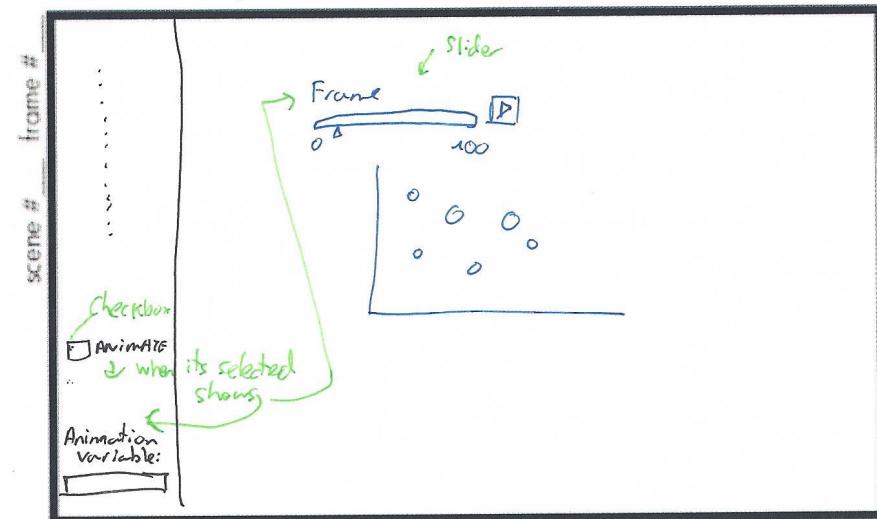
Opcionalmente se puede hacer que la aplicación te permita ver y descargar el código fuente de la aplicación.

Viz designer

#1 U.H. Create new visualization



#2 Create animation



#3 Add facets to graph

