



Trabajo Fin de Máster  
TECI 2013-2014

*“Experimental design to measure the  
impact of online advertising on the sales  
of a car manufacturer’s dealer network”*

**AUTOR: Juan José Carín Mansilla**

**TUTORA: Macarena Estévez (CONENTO)**

**PONENTE: Francisco Ballesteros Olmo (DMAT-UPM)**

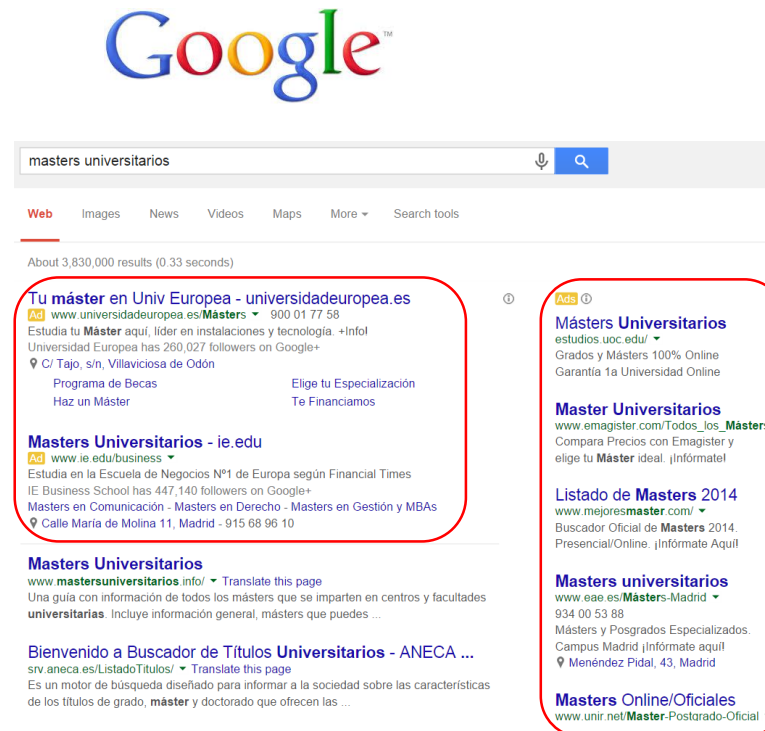
**DIRECTORA: María del Mar Fenoy Muñoz (EIO-UCM)**

# 1 – Introducción

Problema, Objetivo y Diseño  
Experimental

# Problema

- Medir la **eficacia** de la **publicidad**.
  - En concreto, de la publicidad **online**.

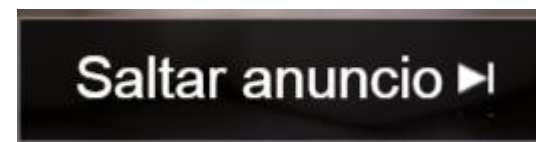


# Problema

- Medir la **eficacia** de la **publicidad**.
  - En concreto, de la publicidad *online*.



- Y más en concreto de:



- **Reto** para los **anunciantes** (y las empresas asociadas del sector publicitario):
  - ¿Aumentan sus clientes por la exposición a una acción publicitaria?
  - ¿En qué medida? ¿Cómo medirlo?

# Objetivo

- Medir el impacto de la publicidad *online* en las ventas [*offline*] de una empresa (*retailer*).
  - **Impacto** = Aumento en las ventas.
  - Modalidad de publicidad: **YouTube**.
  - *Retailer*: red de concesionarios de un **fabricante de coches**.
  - Ámbito de estudio: España.

?



# Experimento

- Segmentar una Muestra de Ciudades en 2 Grupos de Test y Control, cuyo comportamiento (principalmente en cuanto a la Variable Objetivo) sea lo más similar posible.
- **Tratamiento o Estímulo** del Grupo de Test: publicidad en YouTube del modelo de coche bajo estudio (*MUS*).
- Tamaño de la muestra: el máximo posible dadas las condiciones impuestas.
  - *GMAs*: Google puede dirigir su publicidad a determinadas áreas geográficas.

# Fases del Experimento

## 1. Construcción de la Muestra

- A partir de Concesionarios y *GMA*s

## 2. Creación de los Grupos de Test y Control

## 3. Propuesta de Inversión Publicitaria en YouTube

- A partir de Modelo Econométrico

## 4. Periodo de Prueba

## 5. Análisis de resultados

- Comparativa del Grupo de Test con:
  - Resultados en el Grupo de Control
  - Predicciones en ausencia del Estímulo

# Variables y Unidades de Análisis

- Dinámicas (Series Temporales):**

Semana 18 (dom.30 abr.) de 2012 a Semana 22 (lun.1 jun.) de 2014.

Para el Modelo bajo estudio y para el total de Modelos:

- Fabricante (por Concesionario):**

Ventas, Pedidos, Matriculaciones... y **Tráfico**

Variable  
Objetivo

	A Coruña	Albacete	Algeciras	Alicante	Avila	Barcelona	Teruel	Tortosa	Valencia	Valladolid	Vitoria	Zaragoza	
w1	0.008	0.004	0.002	0.010	0.002	0.071	0.001	0.000	0.030	0.006	0.004	0.013	<b>0.437</b>
w2	0.001	0.000	0.000	0.001	0.000	0.009	0.000	0.000	0.003	0.001	0.000	0.001	<b>0.043</b>
w3	0.001	0.000	0.000	0.001	0.000	0.007	0.000	0.000	0.001	0.000	0.000	0.001	<b>0.029</b>
w4	0.004	0.002	0.001	0.006	0.002	0.041	0.001	0.000	0.015	0.005	0.001	0.006	<b>0.265</b>
w5	0.013	0.004	0.002	0.018	0.002	0.108	0.002	0.000	0.047	0.011	0.005	0.021	<b>0.664</b>
w6	0.031	0.011	0.006	0.039	0.004	0.235	0.002	0.001	0.114	0.025	0.013	0.050	<b>1.512</b>
w104	0.003	0.001	0.001	0.005	0.000	0.021	0.001	0.000	0.012	0.002	0.002	0.006	<b>0.159</b>
w105	0.003	0.001	0.000	0.003	0.000	0.018	0.000	0.000	0.009	0.002	0.001	0.003	<b>0.118</b>
w106	0.001	0.001	0.000	0.002	0.000	0.014	0.000	0.000	0.006	0.002	0.001	0.004	<b>0.096</b>
w107	0.014	0.009	0.005	0.023	0.003	0.262	0.002	0.001	0.083	0.016	0.012	0.033	<b>1.145</b>
w108	0.037	0.018	0.008	0.037	0.007	0.523	0.004	0.002	0.168	0.038	0.025	0.077	<b>2.464</b>
w109	0.025	0.013	0.005	0.027	0.005	0.327	0.004	0.001	0.114	0.021	0.018	0.053	<b>1.550</b>
	<b>1.798</b>	<b>0.783</b>	<b>0.365</b>	<b>2.079</b>	<b>0.297</b>	<b>17.910</b>	<b>0.194</b>	<b>0.065</b>	<b>7.055</b>	<b>1.674</b>	<b>0.832</b>	<b>3.020</b>	<b>100.000</b>



# Variables y Unidades de Análisis

- **Dinámicas (Series Temporales):**

Semana 18 (dom.30 abr.) de 2012 a Semana 22 (lun.1 jun.) de 2014.

Para el Modelo bajo estudio y para el total de Modelos:

- **Fabricante (por Concesionario):**

Ventas, Pedidos, Matriculaciones... y **Tráfico**

Variable  
Objetivo

- **Agencia de Medios (por Ciudad):**

**Visitas** a la Web (por origen: SEM, SEO, **Display**, Redes Sociales...)

Variable  
independiente  
principal

# Variables y Unidades de Análisis - 2

- **Variables Estáticas** (por Ciudad):

A 1 de enero de 2013.

- INE:

Población

- Anuario Económico de La Caixa:

Líneas de Teléfono y de Banda Ancha, Coches, Indicadores Económicos (Cuota de Mercado, Índices de Actividad Económica, Comercial—mayorista y minorista—, Industrial...), etc.

- **Unidades de Análisis:**

- Google (principal): **GMA**s

- Fabricante (secundarias): Concesionarios y sus Sedes (*POS*s)

# 2 – Limpieza de los Datos y Selección de la Muestra

# Limpieza de los Datos y Selección de la Muestra

- Situación de partida ( $+32 \cdot 10^6$  datos en +250 ficheros):
  - 178 concesionarios
- Primer filtro:
  - Concesionarios de los que no hay suficientes datos ( $<104$  semanas)
  - ... o no son significativos (Tráfico nulo  $>26$  semanas)
  - 178  $\rightarrow$  119 Concesionarios (con 163 *POSs* en 140 Ciudades)



# Limpieza de los Datos y Selección de la Muestra

- Segundo filtro → Muestra final:
  - Concesionarios con *POSs* en más de 1 Ciudad
  - *POSs* no ubicados en 1 *GMA*
  - 69 Concesionarios (con 76 *POSs* en **57 *GMAs***)
    - Su Tráfico representa el 56% de la muestra inicial (y el 61% de la anterior).

A Coruña, Albacete, Algeciras, Alicante, Avila, Barcelona, Bejar, Burgos, Caceres, Camas, Cartagena, Cordoba, Coslada, Eibar, El Puerto de Santa Maria, Fuenlabrada, Gandia, Getafe, Gijon, Granada, Guadalajara, Huelva, Jerez de la Frontera, Leganes, Leioa, Lleida, Lugo, Madrid, Mahon, Malaga, Manacor,



Marbella, Mataro, Mollet del Valles, Murcia, Orihuela, Ourense, Palencia, Palma de Mallorca, Pamplona, Parla, Pontevedra, Requena, Reus, San Sebastian, Sant Cugat del Valles, Santander, Santiago de Compostela, Segovia, Seville, Tarragona, Teruel, Tortosa, Valencia, Valladolid, Vitoria, and Zaragoza.

# 3 – Análisis Exploratorio

# Análisis Descriptivo y Ajustes de Distribución

	Standard Error	Median	Kurtosis	Skewness	Range	Minimum
MUS-Visits-All	0.132	-0.070	0.217	0.441	4.527	-1.851
Slope MUS-Visits-All	0.132	0.124	50.448	-6.872	8.404	-7.204
MUS-Visits-AFF	0.132	-0.245	0.536	1.063	3.917	-0.907
Slope MUS-VISITS-AFF	0.132	0.229	43.168	-6.356	7.217	-6.929
MUS-VISITS-DIS	0.132	0.030	0.282	0.425	4.711	-1.847
Slope MUS-VISITS-DIS	0.132	-0.182	37.948	5.796	7.718	-0.995
MUS-VISITS-DRT	0.132	-0.218	9.087	2.567	5.889	-1.252
Slope MUS-VISITS-DRT	0.132	0.144	56.733	-7.524	7.650	-7.409
MUS-VISITS-EML	0.132	-0.080	0.561	0.752	4.260	-1.678
Slope MUS-VISITS-EML	0.132	0.215	26.606	-4.783	7.139	-6.145
MUS-VISITS-REF	0.132	-0.164	-0.778	0.289	3.895	-1.818
Slope MUS-VISITS-REF	0.132	-0.194	41.953	6.146	7.671	-0.777
MUS-VISITS-SEA	0.132	-0.099	-0.535	0.243	4.113	-1.882
Slope MUS-VISITS-SEA	0.132	0.224	19.767	-3.672	8.099	-5.695
MUS-VISITS-SEO	0.132	-0.190	-0.285	0.349	4.182	-1.818
Slope MUS-VISITS-SEO	0.132	0.237	34.608	-5.397	8.022	-6.586
MUS-VISITS-SOC	0.132	-0.047	0.098	0.680	4.208	-1.624
Slope MUS-VISITS-SOC	0.132	-0.258	41.757	6.141	7.287	-0.399
MUS-SEO	0.132	-0.117	0.198	0.521	4.735	-1.761
Slope MUS-SEO	0.132	0.240	39.202	-5.980	7.106	-6.771
MUS-SEM	0.132	-0.248	-0.041	0.513	4.493	-1.789
Slope MUS-SEM	0.132	0.251	33.545	-5.496	7.297	-6.497
Manufacturer-VISITS-ALL	0.132	-0.041	0.845	0.580	4.745	-1.844
Slope Manufacturer-VISITS-ALL	0.132	0.086	51.126	-6.913	8.686	-7.226
Manufacturer-VISITS-AFF	0.132	-0.194	-0.298	0.379	4.169	3.918
Slope Manufacturer-VISITS-AFF	0.132	0.236	33.785	-5.583	6.838	-6.486
Manufacturer-VISITS-DIS	0.132	0.053	0.459	0.477	4.710	-1.810
Slope Manufacturer-VISITS-DIS	0.132	-0.277	28.415	5.045	6.631	-0.392
Manufacturer-VISITS-DRT	0.132	-0.239	6.328	2.134	5.632	-1.286
Slope Manufacturer-VISITS-DRT	0.132	0.138	56.895	-7.540	7.632	-7.414
Manufacturer-VISITS-EML	0.132	-0.103	33.700	5.190	7.572	-1.017
Slope Manufacturer-VISITS-EML	0.132	0.063	14.662	1.322	8.449	-3.323
Manufacturer-VISITS-REF	0.132	-0.112	-0.378	0.252	4.338	-1.835
Slope Manufacturer-VISITS-REF	0.132	-0.316	19.240	4.162	6.234	-0.782
Manufacturer-VISITS-SEA	0.132	-0.028	-0.519	0.215	4.127	-1.869
Slope Manufacturer-VISITS-SEA	0.132	0.230	30.923	-5.045	7.930	-6.406
Manufacturer-VISITS-SEO	0.132	-0.063	0.414	0.519	4.703	-1.805
Slope Manufacturer-VISITS-SEO	0.132	0.157	33.303	-5.043	8.329	-6.519
Manufacturer-VISITS-SOC	0.132	-0.215	0.186	0.839	4.269	-1.682
Slope Manufacturer-VISITS-SOC	0.132	-0.234	43.604	6.331	7.363	-0.406
MUS-TRAFFIC	0.132	-0.216	17.857	3.882	6.355	-0.827
Slope MUS-TRAFFIC	0.132	-0.240	7.194	2.332	5.568	-1.209

	Population	Extension (km2)	Unemployment rate	Market share
Population	1.000			
Extension (km2)	0.152	1.000		
Unemployment rate	-0.090	0.254	1.000	
Market share	1.000	0.146	-0.090	1.000
Land lines	0.997	0.131	-0.112	0.998
Change in #land lines since 2007 (%)	-0.223	-0.022	0.086	-0.221
Broadband lines	0.998	0.138	-0.102	0.999
Cars	0.994	0.157	-0.069	0.994
Change in # cars since 2007 (%)	-0.314	0.294	0.193	-0.319
MUS-Visits-All	0.986	0.103	-0.121	0.988
Slope MUS-Visits-All	-0.791	-0.146	0.041	-0.789
MUS-Visits-AFF	0.970	0.105	-0.140	0.972
Slope MUS-VISITS-AFF	-0.963	-0.108	0.134	-0.964
MUS-VISITS-DIS	0.985	0.098	-0.124	0.987
Slope MUS-VISITS-DIS	0.428	-0.085	-0.217	0.437

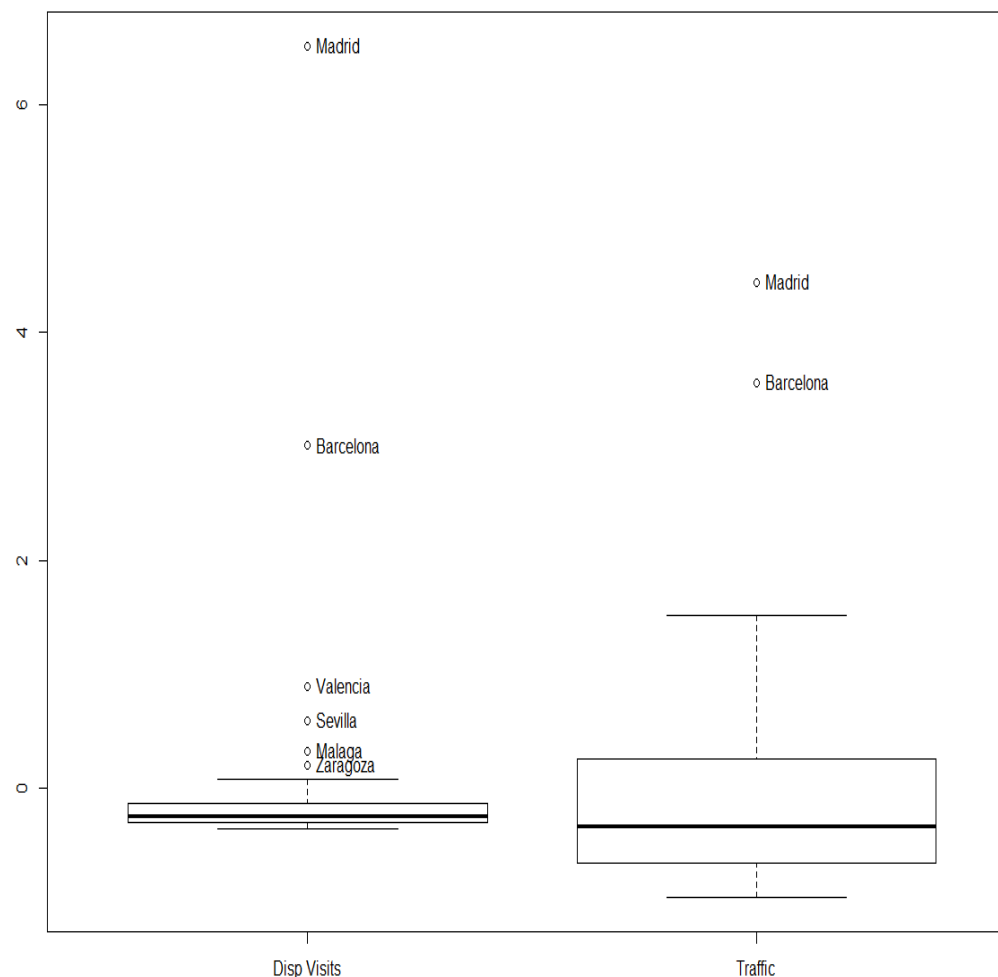
```
shapiro.test(MUS$Traffic)$p.value
```

```
## [1] 1.375e-08
```

```
shapiro.test(MUS$Disp_Visits)$p.value
```

```
## [1] 9.849e-15
```

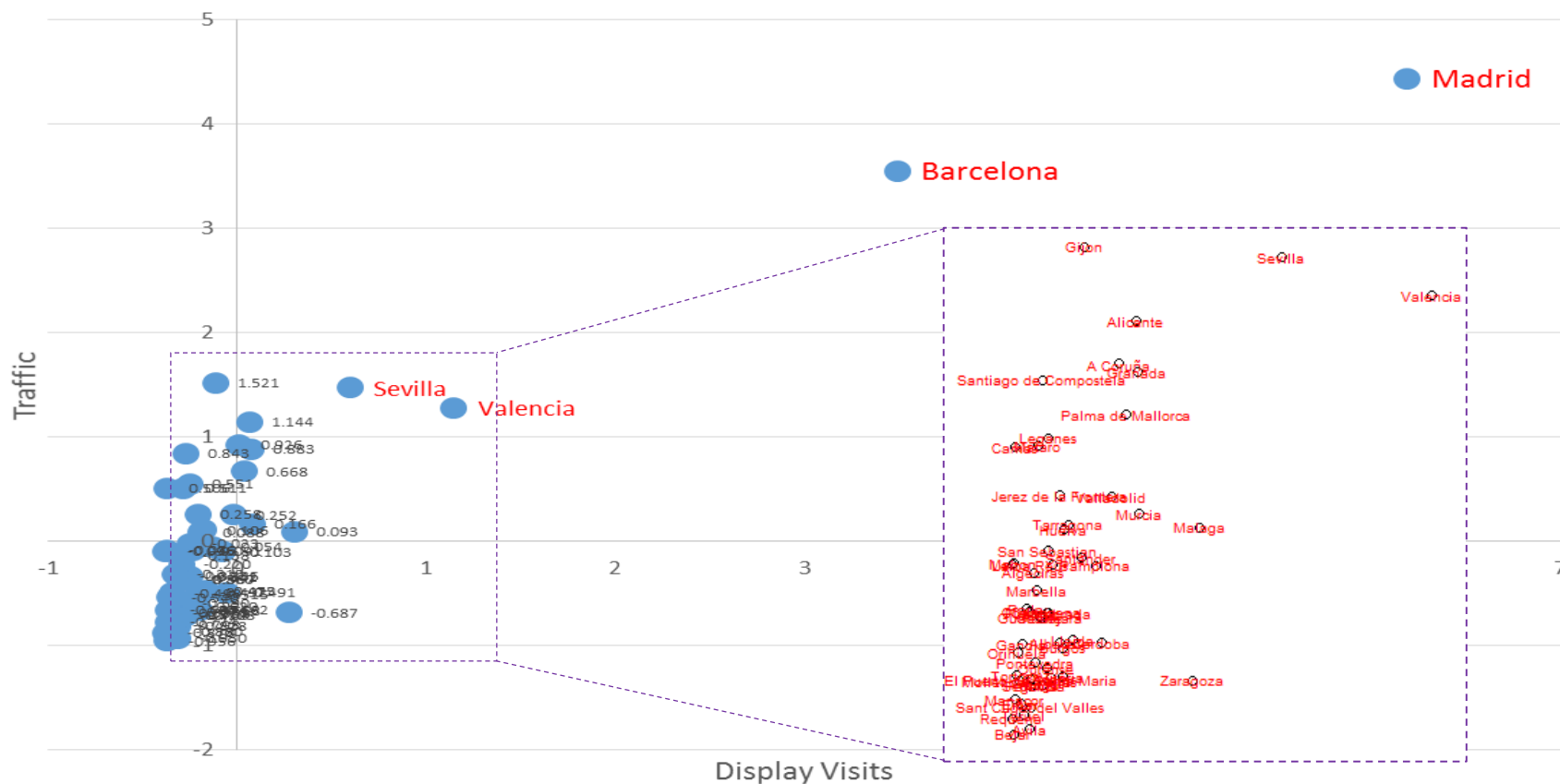
# Análisis Descriptivo y Ajustes de Distribución





# Análisis Descriptivo y Ajustes de Distribución

Display Visits vs Traffic (Correlation **0.84**)





# 4 – Análisis de *Clusters*

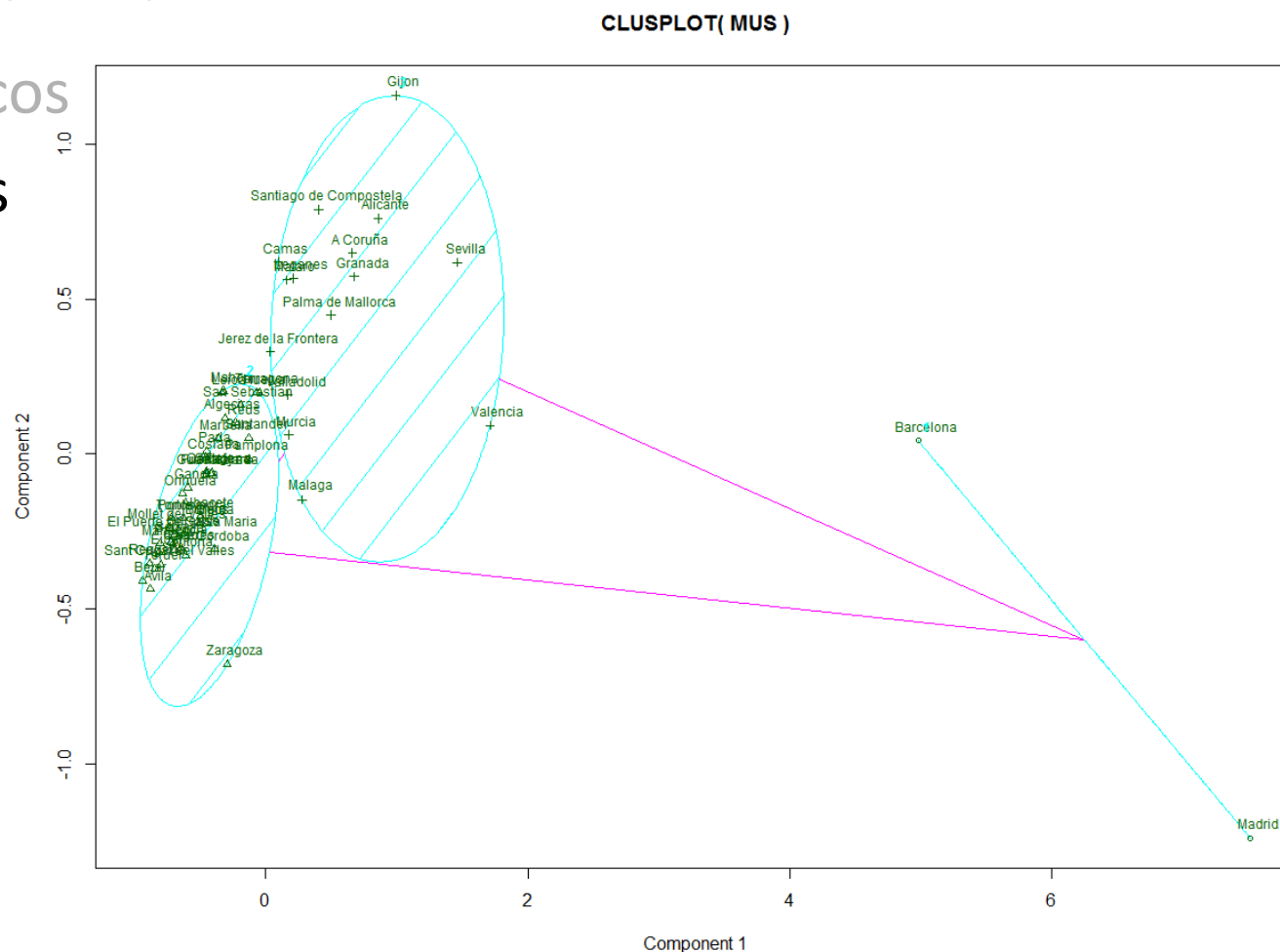
Etapla intermedia, no definitiva



# Clusters según Tráfico y Visitas

## Display

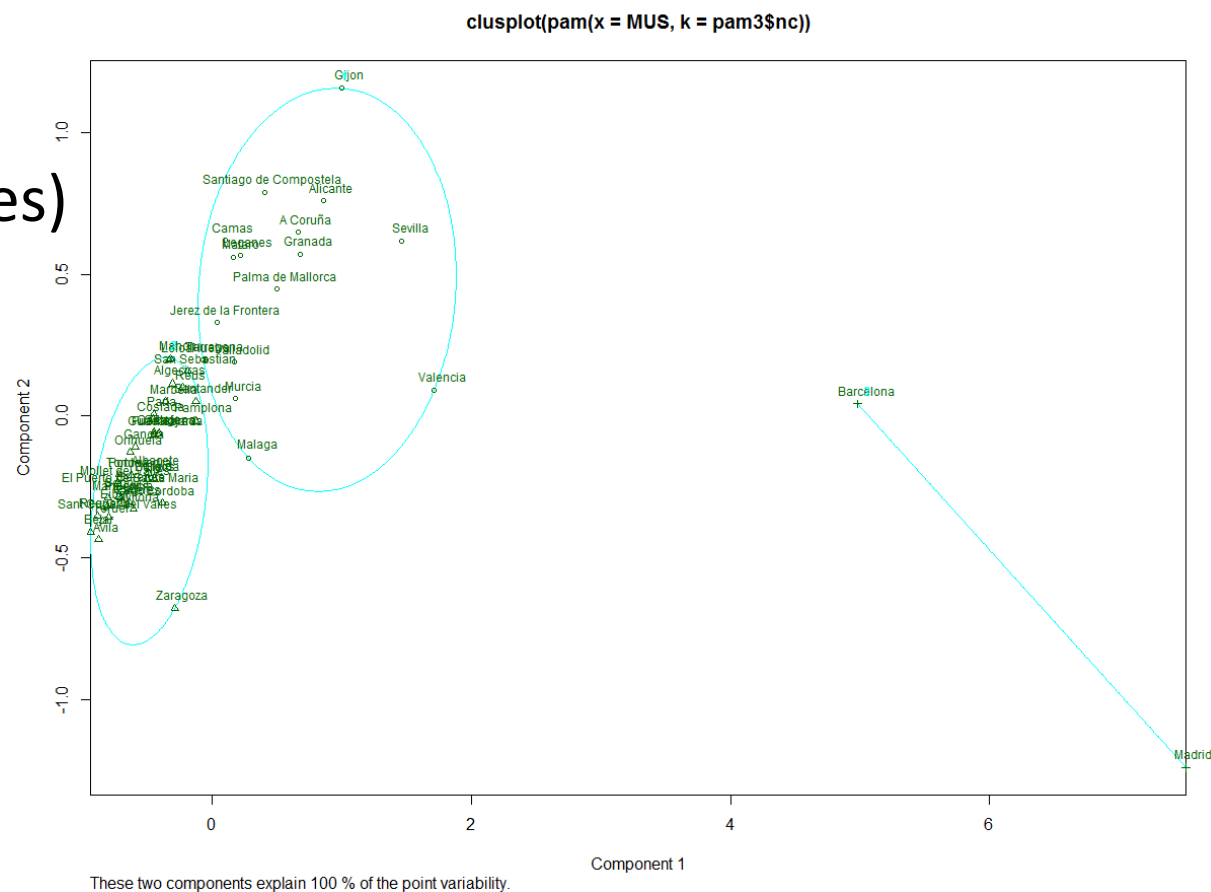
- Jerárquicos
- k-medias



# Clusters según Tráfico y Visitas

## Display

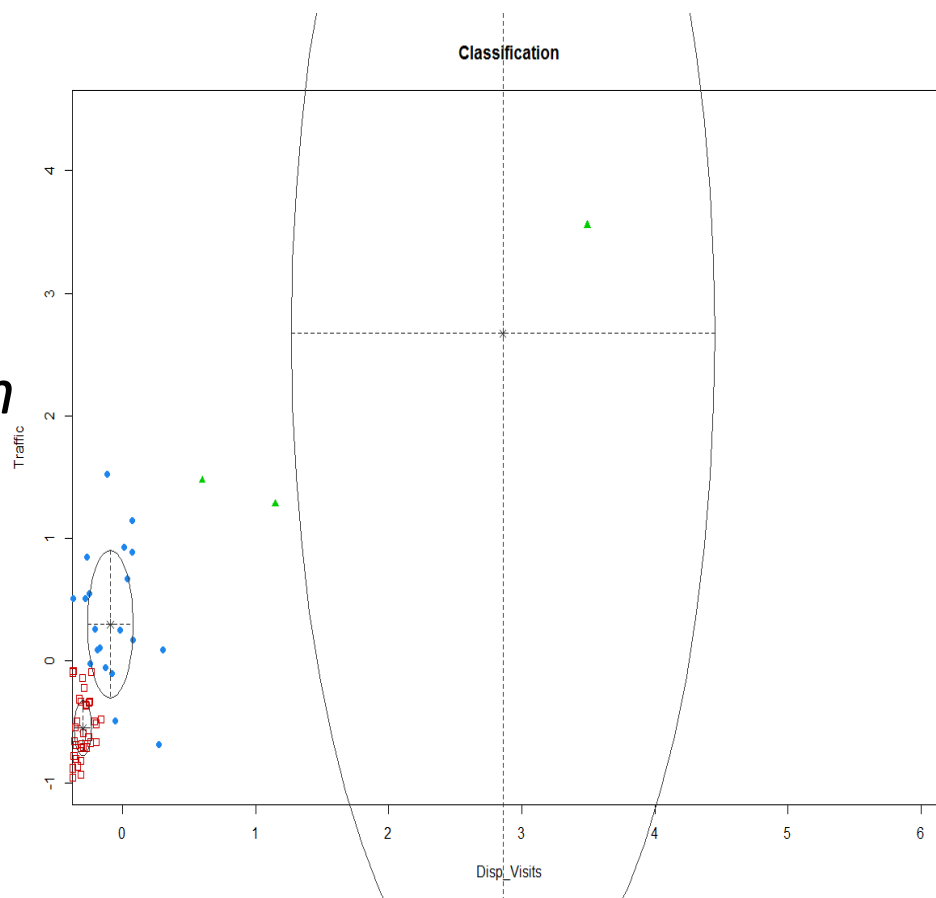
- Jerárquicos
- k-medias
- PAM (k-medioides)



# Clusters según Tráfico y Visitas

## Display

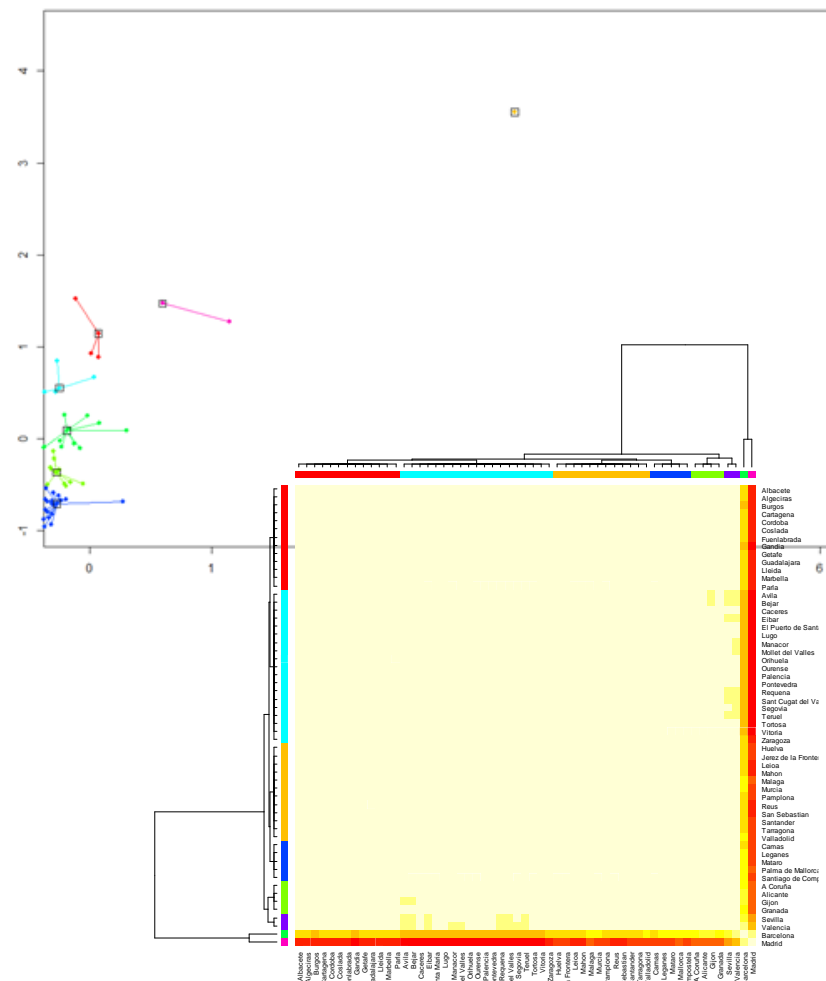
- Jerárquicos
- k-medias
- PAM (k-medioides)
- *Expectation Maximization*



# Clusters según Tráfico y Visitas

## Display

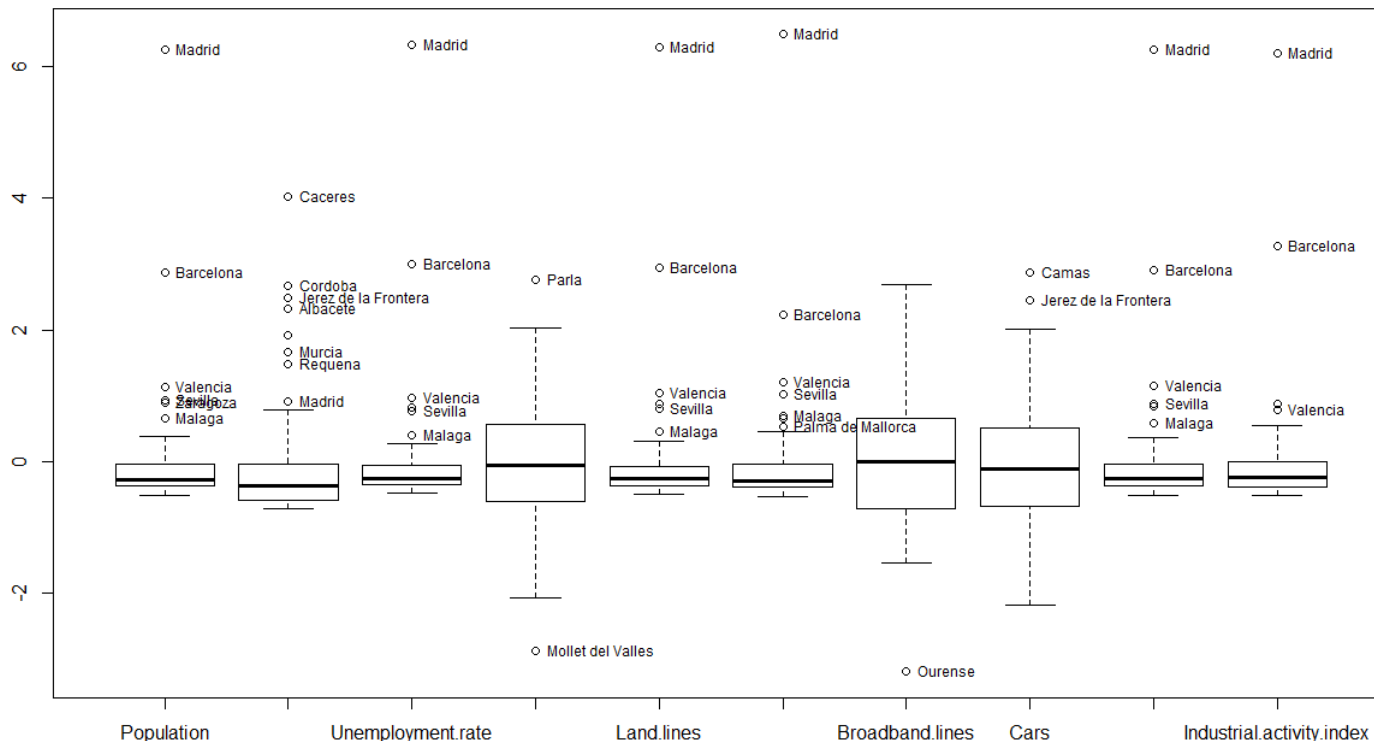
- Jerárquicos
- k-medias
- PAM (k-medioides)
- *Expectation Maximization*
- *Affinity Propagation*





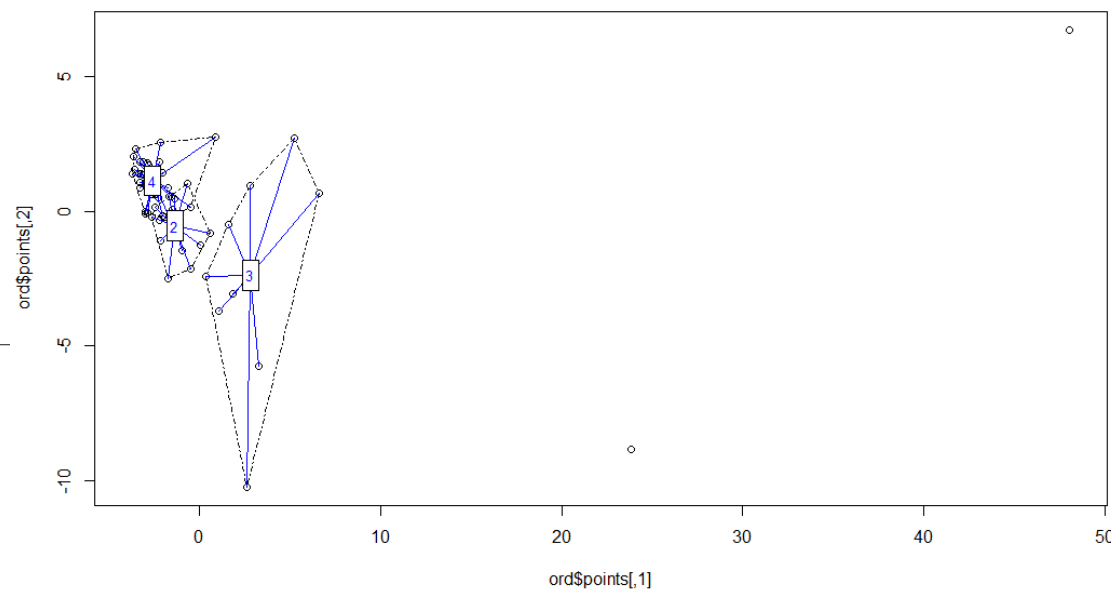
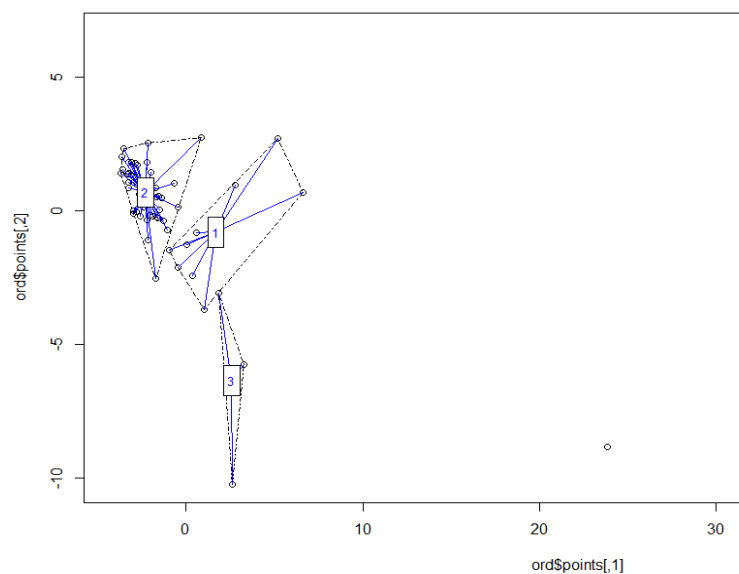
# Clusters considerando todas las Variables

- ACP sobre grupos de variables relacionadas:
  - Demografía, Economía, Internet, Web y Concesionario.



# Clusters considerando todas las Variables

- Resultados similares a los obtenidos considerando sólo Tráfico y Visitas Display.



# ¿Por qué no?

- Los *outliers* forman “su propio *cluster*”.
- Alta variación intra-*cluster*.
  - Ciudades diferentes en un mismo *cluster*.
  - Si se asignan aleatoriamente a cada Grupo, éstos serían disimilares.
- Muchas ciudades sin asignar.
- Se pierde mucha información sobre la evolución temporal de cada ciudad.

# 5 – Distancia entre Series Temporales y Vecino más Próximo

Etapas final

# ¿Por qué?

- Métrica de la similaridad entre dos series temporales (ciudades), que considere todas sus características.
- Se buscan parejas de “ciudades gemelas” (mínima distancia o disimilaridad), en lugar de conglomerados (de mayor tamaño).
- Igual que antes, se asigna un miembro de cada pareja a cada Grupo (Test y Control).
- Al agregar la Variable Objetivo en cada Grupo, se van compensando las posibles diferencias entre los miembros de cada pareja.

# Medidas de distancias entre series temporales

- Posibles métricas:
  - DTW (*Dynamic Time Warping*)
  - Correlación temporal + DTW
  - **Wavelets (respuesta en frecuencia): DWT**
- Se analiza una única variable (Tráfico), obteniéndose un espacio de dimensión 57.

row.names	A.Coru	Albacete	Algeciras	Alicante	Avila	Barcelona	Bejar	Burgos
A.Coru	0.00000000	0.211993163	0.16504905	0.08225513	0.27542747	0.3860368	0.27873733	0.21575057
Albacete	0.21199316	0.00000000	0.05837873	0.25695999	0.06908688	0.5925827	0.07002201	0.03813601
Algeciras	0.16504905	0.058378726	0.00000000	0.22077782	0.12473128	0.5436209	0.12494490	0.08147452
Alicante	0.08225513	0.256959991	0.22077782	0.00000000	0.31163663	0.3665533	0.31670883	0.25053901
Avila	0.27542747	0.069086879	0.12473128	0.31163663	0.00000000	0.6555205	0.01048438	0.06325555
Barcelona	0.38603679	0.592582681	0.54362090	0.36655332	0.65552051	0.0000000	0.65949745	0.59404920
Bejar	0.27873733	0.070022013	0.12494490	0.31670883	0.01048438	0.6594975	0.00000000	0.06934295
Burgos	0.21575057	0.038136007	0.08147452	0.25053901	0.06325555	0.5940492	0.06934295	0.00000000



# Detección de Vecino más Próximo (y Parejas)

- Se busca el vecino más próximo de cada *GMA*.
- Se forman parejas en que esta relación sea mutua.



##	Ciudad	Vecina
## 1	A.Coruña	Granada
## 2	Albacete	Gandia
## 3	Avila	Bejar
## 4	Barcelona	Madrid
## 5	Camas	Mataro
## 6	Eibar	Sant.Cugat.del.Valles
## 7	Fuenlabrada	Parla
## 8	Huelva	Malaga
## 9	Lugo	Mollet.del.Valles
## 10	Mahon	Santander
## 11	Murcia	Valladolid
## 12	Ourense	Pontevedra
## 13	Palencia	Tortosa
## 14	Sevilla	Valencia



# Asignación a los Grupos de Test y Control

- Un miembro de la pareja se asigna al Grupo de Test, y el otro al de Control.
- Objetivo: minimizar la distancia entre las series de Tráfico agregado de ambos Grupos.

- Para  $n$  parejas, se prueban los  $\sum_{k=0}^{n-1} \binom{n-1}{k} = 2^{n-1}$  posibles Grupos de Test.

- Para  $\{a_1, b_1\}, \{a_2, b_2\}, \{a_3, b_3\}$ :

$\{a_1, a_2, a_3\}$  y  $\{b_1, b_2, b_3\}$

$\{a_1, a_2, b_3\}$  y  $\{b_1, b_2, a_3\}$

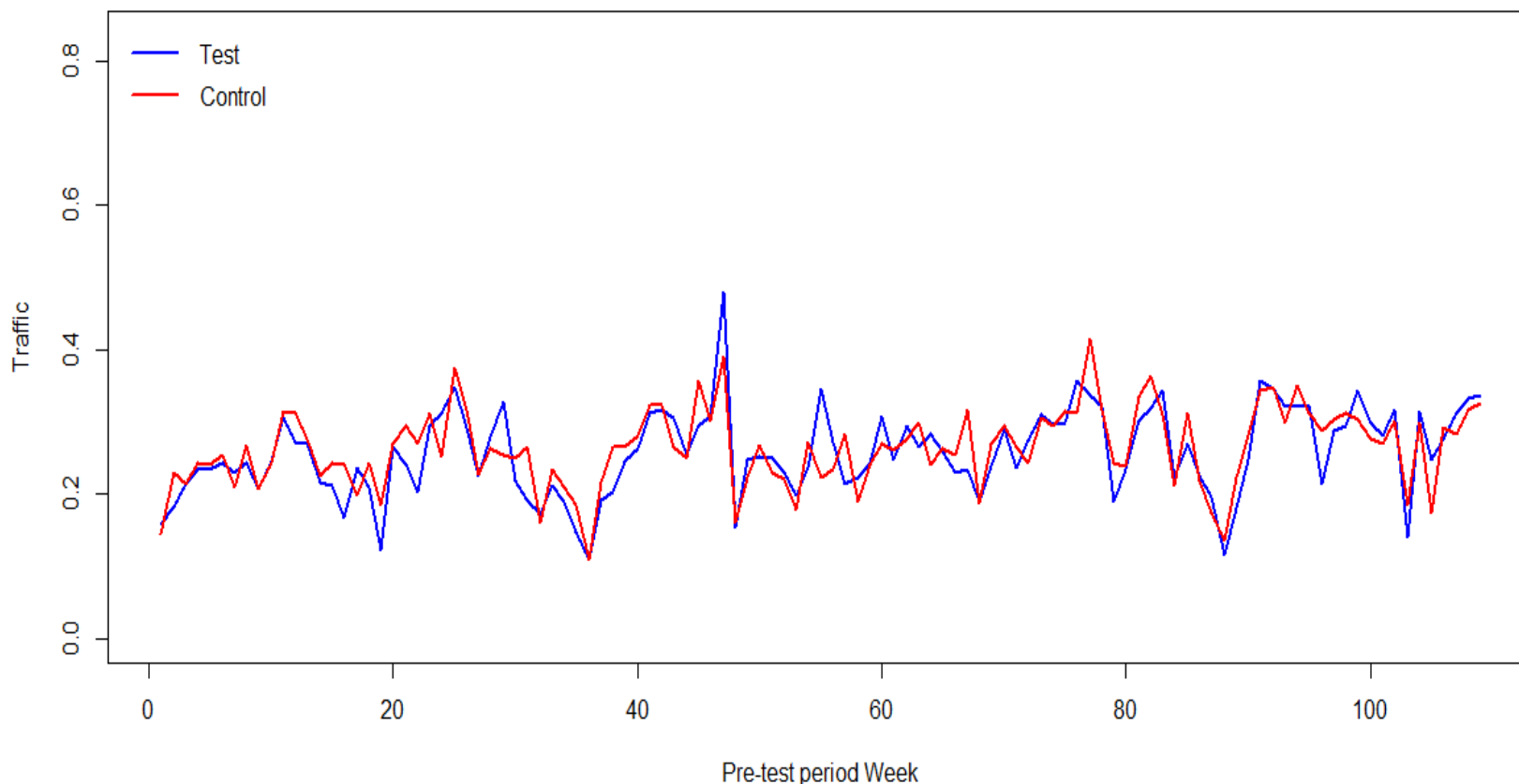
$\{a_1, b_2, a_3\}$  y  $\{b_1, a_2, b_3\}$

$\{a_1, b_2, b_3\}$  y  $\{b_1, a_2, a_3\}$

TEST			CONTROL			OUT		
AVERAGE WEEKLY			AVERAGE WEEKLY			AVERAGE WEEKLY		
CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE
A Coruña	18.30	0.0481	Granada	17.91	0.0536	Algeciras	8.87	-0.0331
Albacete	5.72	-0.0173	Gandia	5.67	-0.0013	Alicante	20.23	0.1979
Avila	1.85	0.0171	Bejar	1.62	0.0019	Burgos	5.52	0.0496
Barcelona	41.58	0.1547	Madrid	49.42	0.1240	Caceres	4.16	0.0192
Camas	14.57	-0.0001	Mataro	14.62	0.0067	Cartagena	7.06	-0.0223
Eibar	3.01	0.0008	Sant Cugat del Valles	2.84	-0.0056	Cordoba	5.73	0.0935
Mahon	9.33	0.0562	Santander	9.61	0.0452	Coslada	7.15	0.0060
Malaga	10.91	0.0086	Huelva	10.87	0.0078	El Puerto de Santa M	4.08	-0.0097
Mollet del Valles	4.01	0.0319	Lugo	3.81	0.0418	Getafe	6.89	0.0064
Parla	7.31	0.0192	Fuenlabrada	7.12	0.0119	Gijon	23.57	0.0086
Pontevedra	4.86	0.0147	Ourense	4.59	0.0149	Guadalajara	6.84	0.0299
Sevilla	23.13	0.0304	Valencia	21.39	-0.0156	Jerez de la Frontera	12.38	-0.0198
Tortosa	4.27	-0.0354	Palencia	3.85	-0.0280	Leganes	14.97	0.2598
Valladolid	12.32	0.0620	Murcia	11.56	0.0380	Leioa	9.23	0.0435
						Lleida	5.88	-0.0094
						Manacor	3.22	0.0346
						Marbella	8.13	0.0240
						Orihuela	5.30	-0.0310
						Palma de Mallorca	16.01	0.0454
						Pamplona	9.17	0.0661
						Requena	2.31	0.0000
						Reus	9.28	0.0112
						San Sebastian	9.88	0.0030
						Santiago de Compost	17.56	-0.0130
						Segovia	3.78	0.0515
						Tarragona	11.03	0.0285
						Teruel	2.46	0.0160
						Vitoria	4.22	0.0667
						Zaragoza	4.00	0.0279
Sum: 161.14 (28%) Mean: 0.0279			Sum: 164.86 (29%) Mean: 0.0211			Sum: 248.90 (43%) Mean: 0.0328		

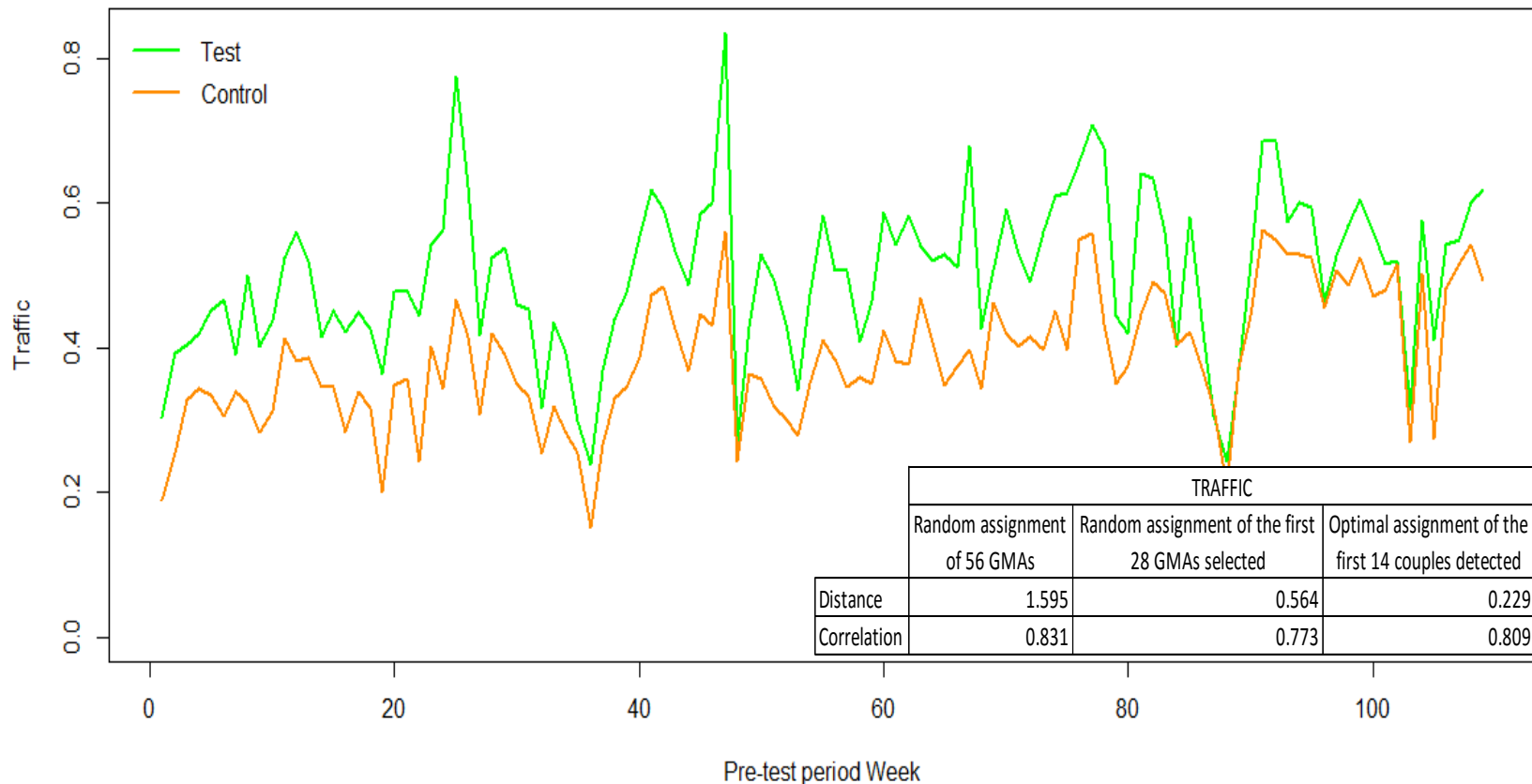
# Asignación a los Grupos de Test y Control

**Traffic of Test & Control groups**  
**Optimal assignment of the 28 best matching GMAs**



# Asignación a los Grupos de Test y Control

Traffic of Test & Control groups  
Random assignment of 56 of the 57 GMAs





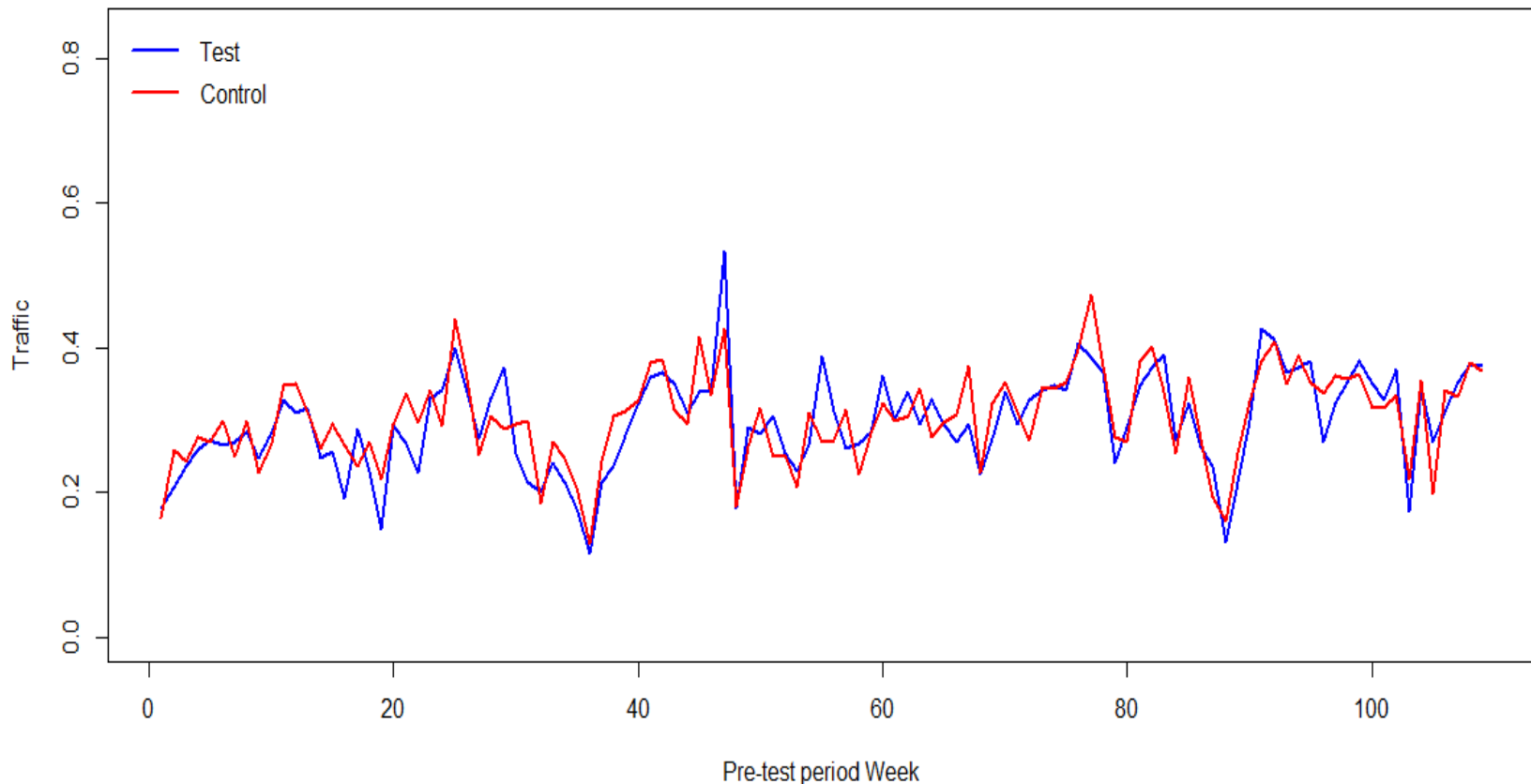
# Obtención de más parejas: Reducción de dimensionalidad

- Objetivo: aumentar el tamaño de los Grupos, manteniendo la similaridad entre ellos.
  - Se utiliza MDS (GoF = .969) para hallar la matriz de distancias euclídeas sobre los pares de coordenadas.
  - Se detectan los Vecinos más Próximos y Parejas.
  - Se añaden las nuevas Parejas “compatibles”.
  - Se reasignan los miembros de cada Pareja a cada Grupo.

TEST			CONTROL			OUT		
AVERAGE WEEKLY TRAFFIC			AVERAGE WEEKLY TRAFFIC			AVERAGE WEEKLY TRAFFIC		
CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE
A Coruña	18.30	0.0481	Granada	17.91	0.0536	Algeciras	8.87	-0.0331
Albacete	5.72	-0.0173	Gandia	5.67	-0.0013	Alicante	20.23	0.1979
Avila	1.85	0.0171	Bejar	1.62	0.0019	Burgos	5.52	0.0496
Barcelona	41.58	0.1547	Madrid	49.42	0.1240	Cartagena	7.06	-0.0223
Camas	14.57	-0.0001	Mataro	14.62	0.0067	Cordoba	5.73	0.0935
Eibar	3.01	0.0008	Sant Cugat del Valles	2.84	-0.0056	Coslada	7.15	0.0060
Mahon	9.33	0.0562	Santander	9.61	0.0452	El Puerto de Santa M	4.08	-0.0097
Malaga	10.91	0.0086	Huelva	10.87	0.0078	Getafe	6.89	0.0064
Mollet del Valles	4.01	0.0319	Lugo	3.81	0.0418	Gijon	23.57	0.0086
Pamplona	9.17	0.0661	Leioa	9.23	0.0435	Guadalajara	6.84	0.0299
Parla	7.31	0.0192	Fuenlabrada	7.12	0.0119	Jerez de la Frontera	12.38	-0.0198
Pontevedra	4.86	0.0147	Ourense	4.59	0.0149	Leganes	14.97	0.2598
Requena	2.31	0.0000	Teruel	2.46	0.0160	Lleida	5.88	-0.0094
San Sebastian	9.88	0.0030	Reus	9.28	0.0112	Manacor	3.22	0.0346
Sevilla	23.13	0.0304	Valencia	21.39	-0.0156	Marbella	8.13	0.0240
Tortosa	4.27	-0.0354	Palencia	3.85	-0.0280	Orihuela	5.30	-0.0310
Valladolid	12.32	0.0620	Murcia	11.56	0.0380	Palma de Mallorca	16.01	0.0454
Zaragoza	4.00	0.0279	Caceres	4.16	0.0192	Santiago de Compost	17.56	-0.0130
						Segovia	3.78	0.0515
						Tarragona	11.03	0.0285
						Vitoria	4.22	0.0667
Sum: 186.51 (32%) Mean: 0.0271			Sum: 189.99 (33%) Mean: 0.0214			Sum: 198.40 (35%) Mean: 0.0364		

# Obtención de más parejas: Reducción de dimensionalidad

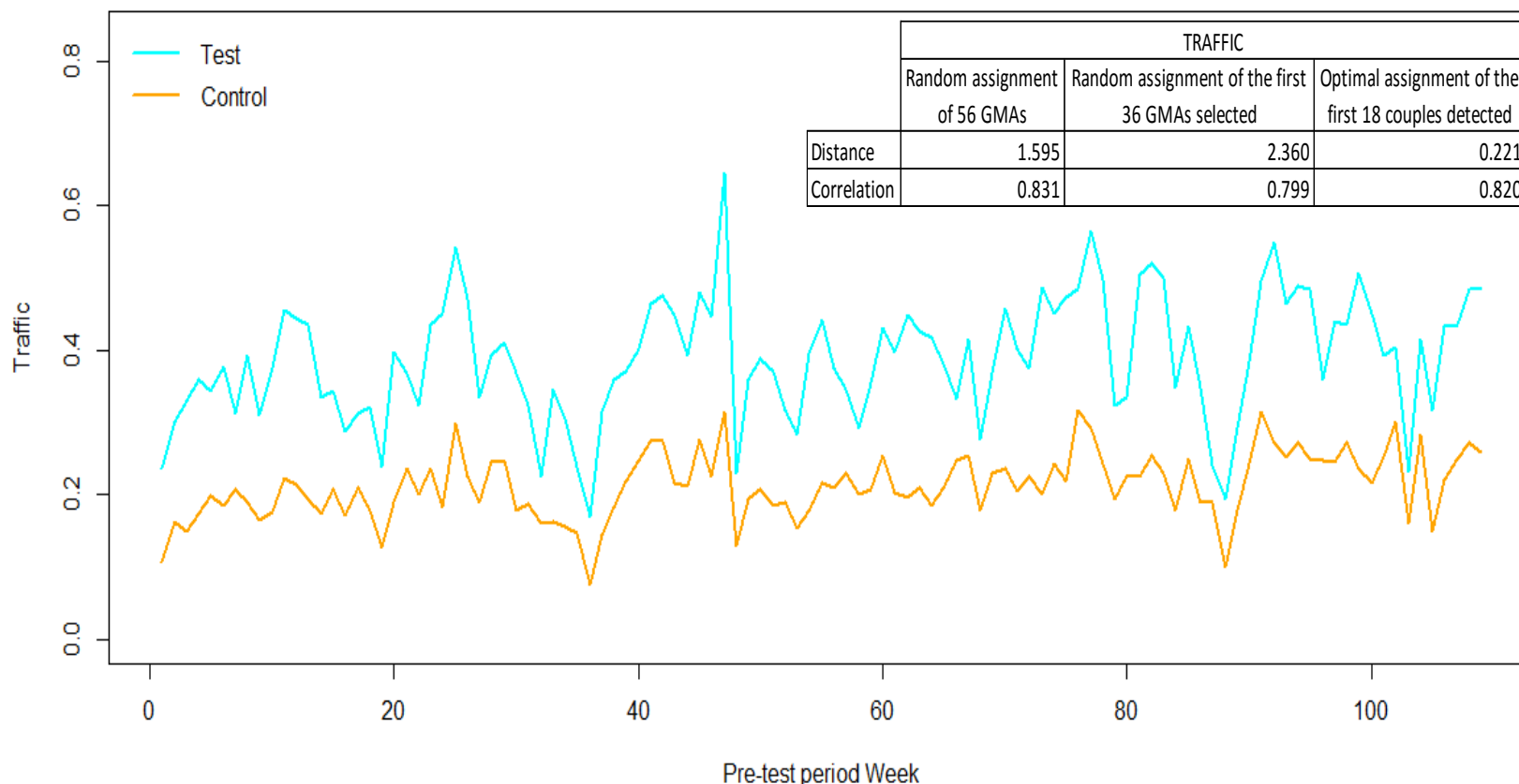
**Traffic of Test & Control groups**  
**Optimal assignment of the 36 best matching GMAs**





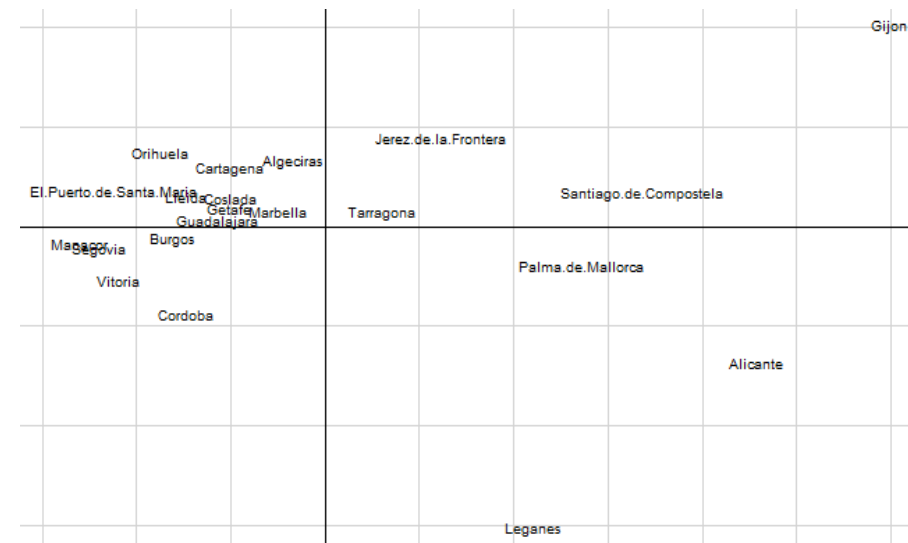
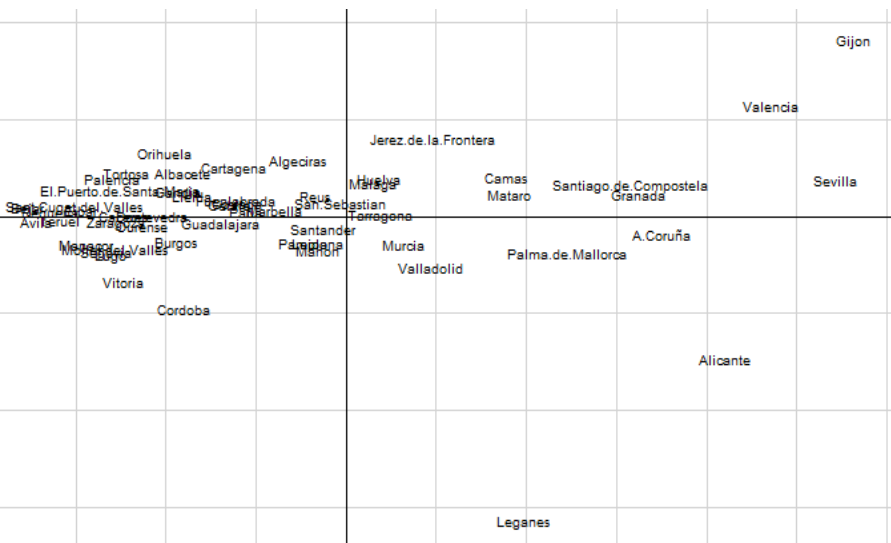
# Obtención de más parejas: Reducción de dimensionalidad

**Traffic of Test & Control groups**  
**Random assignment of the 36 best matching GMAs**



# Obtención de más parejas: Uso de sólo ciudades restantes

- Se descartan las (36) ciudades ya emparejadas y asignadas.

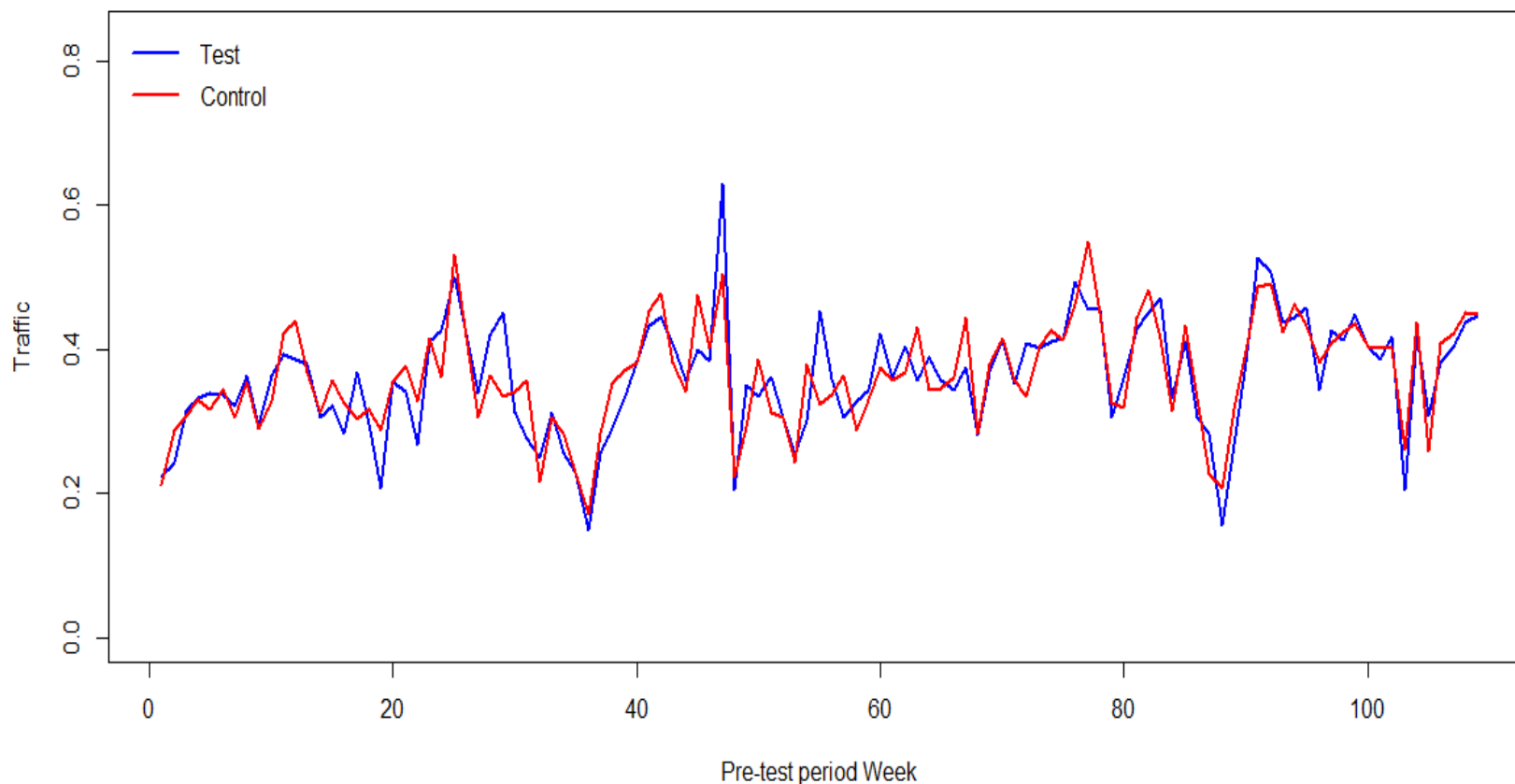


- Se realiza la Detección de Vecinos más Próximos y Parejas, como al principio.
- Se añaden las nuevas Parejas.
- Se reasignan los miembros de cada Pareja a cada Grupo.

TEST			CONTROL			OUT		
AVERAGE WEEKLY			AVERAGE WEEKLY			AVERAGE WEEKLY		
CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE
A Coruña	18.30	0.0481	Granada	17.91	0.0536	Algeciras	8.87	-0.0331
Albacete	5.72	-0.0173	Gandia	5.67	-0.0013	Alicante	20.23	0.1979
Avila	1.85	0.0171	Bejar	1.62	0.0019	Burgos	5.52	0.0496
Barcelona	41.58	0.1547	Madrid	49.42	0.1240	Cordoba	5.73	0.0935
Camas	14.57	-0.0001	Mataro	14.62	0.0067	Coslada	7.15	0.0060
Eibar	3.01	0.0008	Sant Cugat del Valles	2.84	-0.0056	El Puerto de Santa M	4.08	-0.0097
Getafe	6.89	0.0064	Cartagena	7.06	-0.0223	Gijon	23.57	0.0086
Jerez de la Frontera	12.38	-0.0198	Tarragona	11.03	0.0285	Guadalajara	6.84	0.0299
Mahon	9.33	0.0562	Santander	9.61	0.0452	Leganes	14.97	0.2598
Malaga	10.91	0.0086	Huelva	10.87	0.0078	Lleida	5.88	-0.0094
Manacor	3.22	0.0346	Segovia	3.78	0.0515	Marbella	8.13	0.0240
Mollet del Valles	4.01	0.0319	Lugo	3.81	0.0418	Orihuela	5.30	-0.0310
Pamplona	9.17	0.0661	Leioa	9.23	0.0435	Vitoria	4.22	0.0667
Parla	7.31	0.0192	Fuenlabrada	7.12	0.0119			
Pontevedra	4.86	0.0147	Ourense	4.59	0.0149			
Requena	2.31	0.0000	Teruel	2.46	0.0160			
San Sebastian	9.88	0.0030	Reus	9.28	0.0112			
Santiago de Compost	17.56	-0.0130	Palma de Mallorca	16.01	0.0454			
Sevilla	23.13	0.0304	Valencia	21.39	-0.0156			
Tortosa	4.27	-0.0354	Palencia	3.85	-0.0280			
Valladolid	12.32	0.0620	Murcia	11.56	0.0380			
Zaragoza	4.00	0.0279	Caceres	4.16	0.0192			
Sum: 226.56 (39%) Mean: 0.0225			Sum: 227.87 (40%) Mean: 0.0222			Sum: 120.48 (21%) Mean: 0.0502		

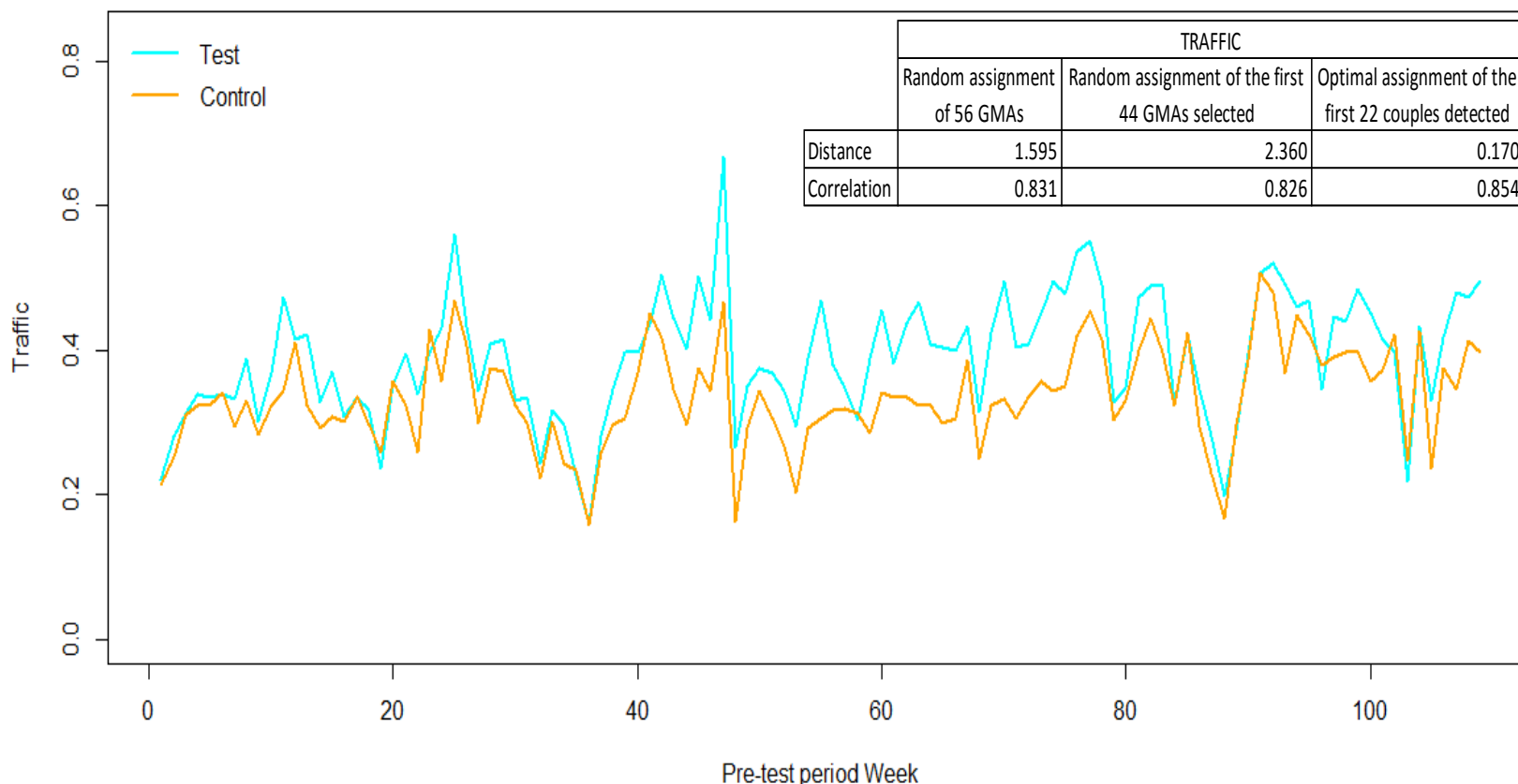
# Obtención de más parejas: Uso de sólo ciudades restantes

**Traffic of Test & Control groups**  
**Optimal assignment of the 44 best matching GMAs**



# Obtención de más parejas: Uso de sólo ciudades restantes

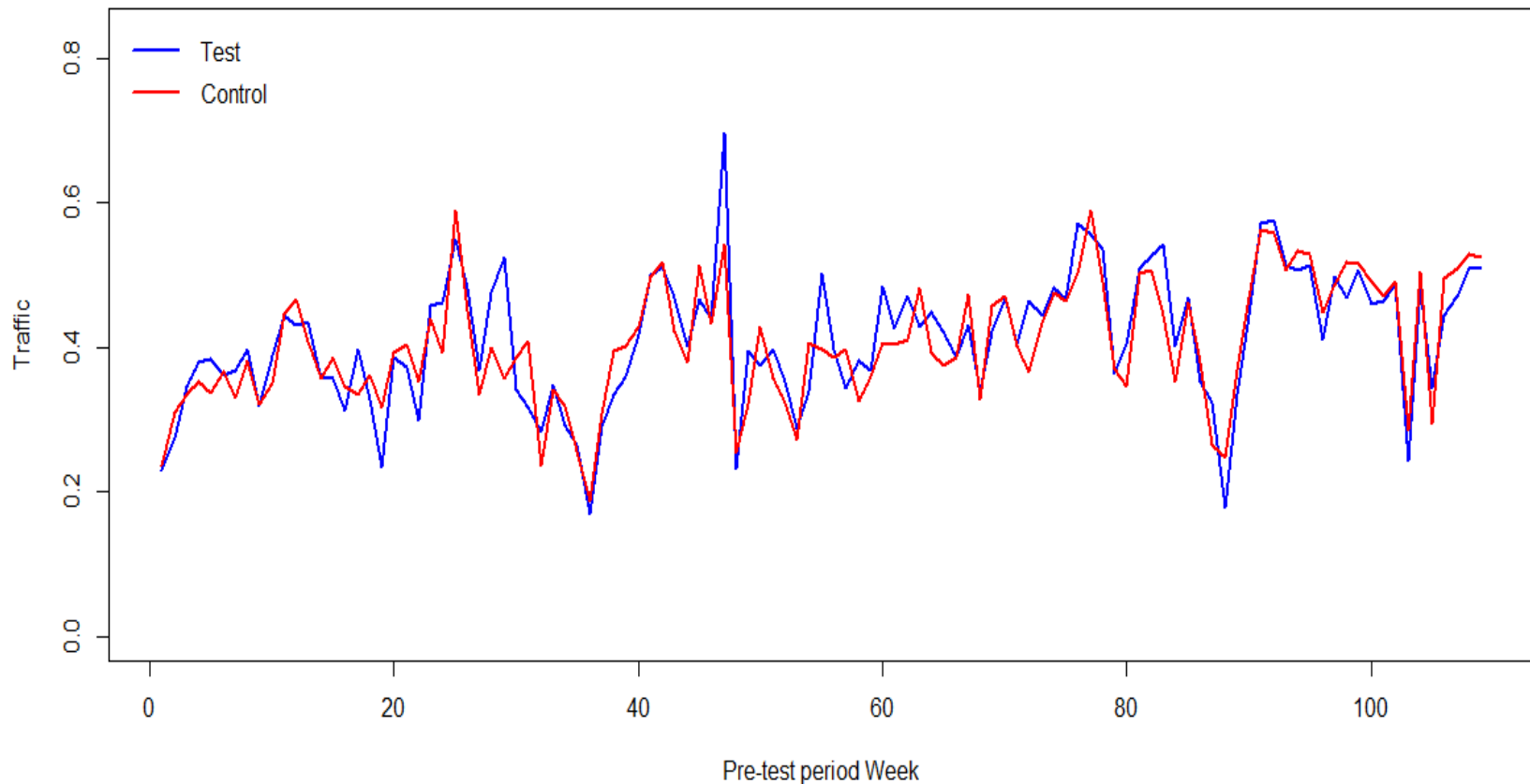
**Traffic of Test & Control groups**  
**Random assignment of the 44 best matching GMAs**



1 - INTRODUCCIÓN			2 - LIMPIEZA DE LOS DATOS Y SELECCIÓN DE LA MUESTRA			3 - ANÁLISIS EXPLORATORIO		
4 - ANÁLISIS DE CLUSTERS			5 - DISTANCIA SERIES TEMPORALES Y VECINO + PRÓXIMO			6 - CONCLUSIONES		
TEST			CONTROL			OUT		
AVERAGE WEEKLY			AVERAGE WEEKLY			AVERAGE WEEKLY		
CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE
Alicante	20.23	0.1979	Leganes	14.97	0.2598	Algeciras	8.87	-0.0331
A Coruña	18.30	0.0481	Granada	17.91	0.0536	Burgos	5.52	0.0496
Albacete	5.72	-0.0173	Gandia	5.67	-0.0013	Cordoba	5.73	0.0935
Avila	1.85	0.0171	Bejar	1.62	0.0019	Gijon	23.57	0.0086
Barcelona	41.58	0.1547	Madrid	49.42	0.1240	Lleida	5.88	-0.0094
Camas	14.57	-0.0001	Mataro	14.62	0.0067	Marbella	8.13	0.0240
Coslada	7.15	0.0060	Guadalajara	6.84	0.0299	Vitoria	4.22	0.0667
Eibar	3.01	0.0008	Sant Cugat del Valles	2.84	-0.0056			
El Puerto de Santa M	4.08	-0.0097	Orihuela	5.30	-0.0310			
Getafe	6.89	0.0064	Cartagena	7.06	-0.0223			
Jerez de la Frontera	12.38	-0.0198	Tarragona	11.03	0.0285			
Mahon	9.33	0.0562	Santander	9.61	0.0452			
Malaga	10.91	0.0086	Huelva	10.87	0.0078			
Manacor	3.22	0.0346	Segovia	3.78	0.0515			
Mollet del Valles	4.01	0.0319	Lugo	3.81	0.0418			
Pamplona	9.17	0.0661	Leioa	9.23	0.0435			
Parla	7.31	0.0192	Fuenlabrada	7.12	0.0119			
Pontevedra	4.86	0.0147	Ourense	4.59	0.0149			
Requena	2.31	0.0000	Teruel	2.46	0.0160			
San Sebastian	9.88	0.0030	Reus	9.28	0.0112			
Santiago de Compost	17.56	-0.0130	Palma de Mallorca	16.01	0.0454			
Sevilla	23.13	0.0304	Valencia	21.39	-0.0156			
Tortosa	4.27	-0.0354	Palencia	3.85	-0.0280			
Valladolid	12.32	0.0620	Murcia	11.56	0.0380			
Zaragoza	4.00	0.0279	Caceres	4.16	0.0192			
Sum: 258.02 (45%)    Mean: 0.0276			Sum: 254.97 (44%)    Mean: 0.0299			Sum: 61.91 (11%)    Mean: 0.0285		

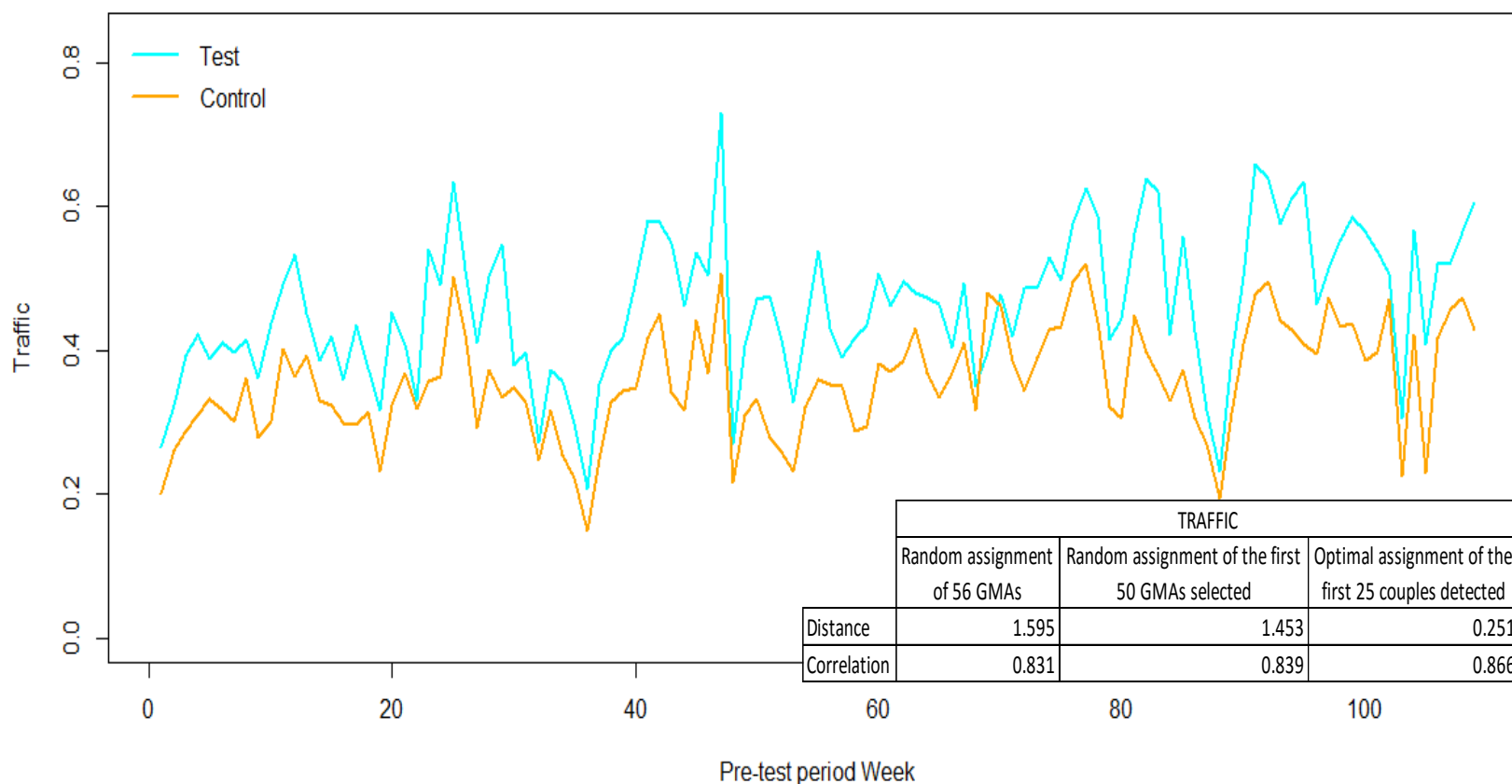
# Sucesivas iteraciones (1)

**Traffic of Test & Control groups**  
**Optimal assignment of the 50 best matching GMAs**



# Sucesivas iteraciones (1)

**Traffic of Test & Control groups**  
**Random assignment of the 50 best matching GMAs**

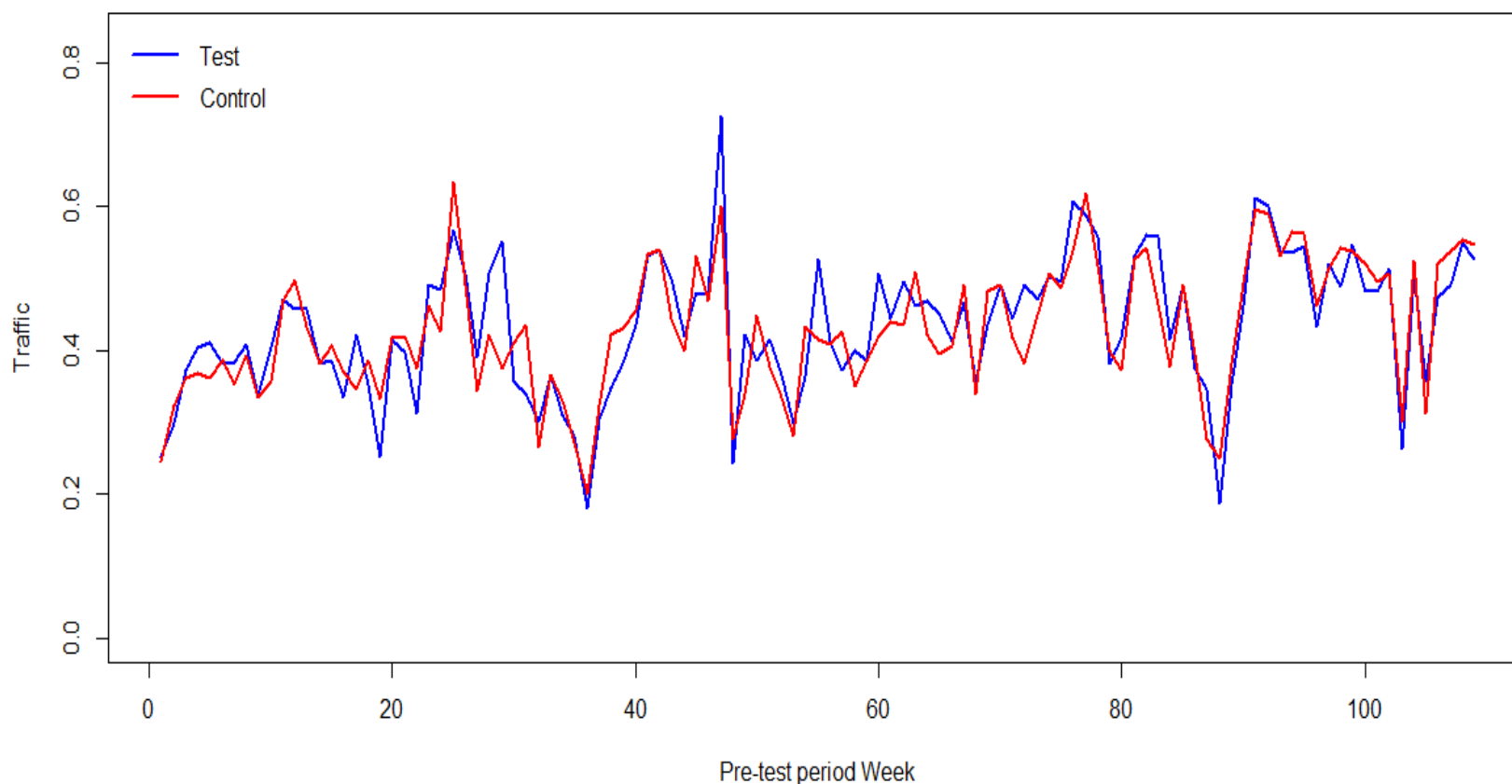




TEST			CONTROL			OUT		
AVERAGE WEEKLY TRAFFIC			AVERAGE WEEKLY TRAFFIC			AVERAGE WEEKLY TRAFFIC		
CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE
Alicante	20.23	0.1979	Leganes	14.97	0.2598	Cordoba	5.73	0.0935
A Coruña	18.30	0.0481	Granada	17.91	0.0536	Gijon	23.57	0.0086
Albacete	5.72	-0.0173	Gandia	5.67	-0.0013	Vitoria	4.22	0.0667
Avila	1.85	0.0171	Bejar	1.62	0.0019			
Barcelona	41.58	0.1547	Madrid	49.42	0.1240			
Camas	14.57	-0.0001	Mataro	14.62	0.0067			
Coslada	7.15	0.0060	Guadalajara	6.84	0.0299			
Eibar	3.01	0.0008	Sant Cugat del Valles	2.84	-0.0056			
El Puerto de Santa M	4.08	-0.0097	Orihuela	5.30	-0.0310			
Getafe	6.89	0.0064	Cartagena	7.06	-0.0223			
Jerez de la Frontera	12.38	-0.0198	Tarragona	11.03	0.0285			
Lleida	5.88	-0.0094	Burgos	5.52	0.0496			
Mahon	9.33	0.0562	Santander	9.61	0.0452			
Malaga	10.91	0.0086	Huelva	10.87	0.0078			
Manacor	3.22	0.0346	Segovia	3.78	0.0515			
Marbella	8.13	0.0240	Algeciras	8.87	-0.0331			
Mollet del Valles	4.01	0.0319	Lugo	3.81	0.0418			
Pamplona	9.17	0.0661	Leioa	9.23	0.0435			
Parla	7.31	0.0192	Fuenlabrada	7.12	0.0119			
Pontevedra	4.86	0.0147	Ourense	4.59	0.0149			
Requena	2.31	0.0000	Teruel	2.46	0.0160			
San Sebastian	9.88	0.0030	Reus	9.28	0.0112			
Santiago de Compost	17.56	-0.0130	Palma de Mallorca	16.01	0.0454			
Sevilla	23.13	0.0304	Valencia	21.39	-0.0156			
Tortosa	4.27	-0.0354	Palencia	3.85	-0.0280			
Valladolid	12.32	0.0620	Murcia	11.56	0.0380			
Zaragoza	4.00	0.0279	Caceres	4.16	0.0192			
Sum: 272.03 (47%) Mean: 0.0261			Sum: 269.36 (47%) Mean: 0.0283			Sum: 33.52 (6%) Mean: 0.0562		

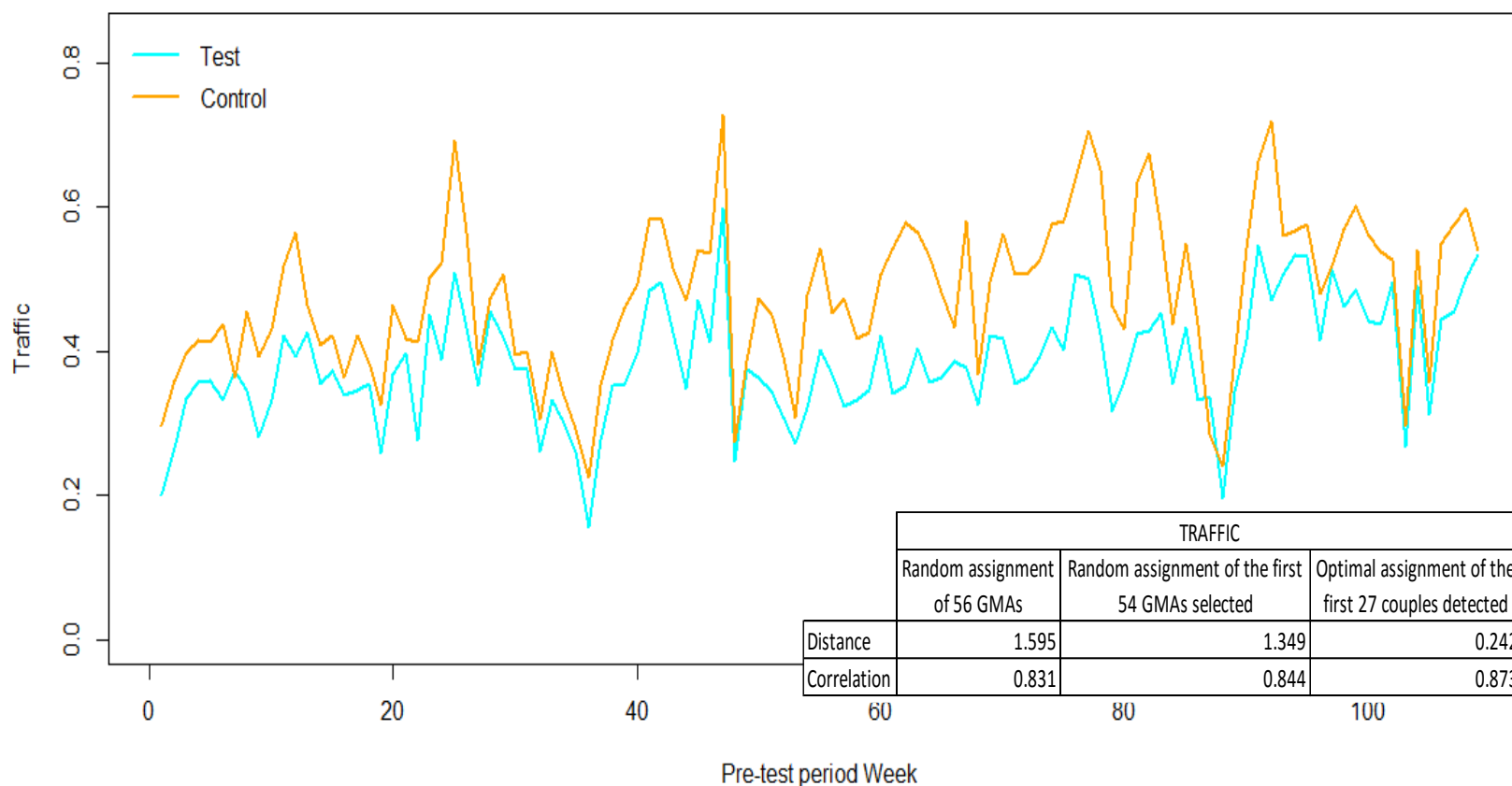
# Sucesivas iteraciones (2)

**Traffic of Test & Control groups**  
**Optimal assignment of the 54 best matching GMAs**



# Sucesivas iteraciones (2)

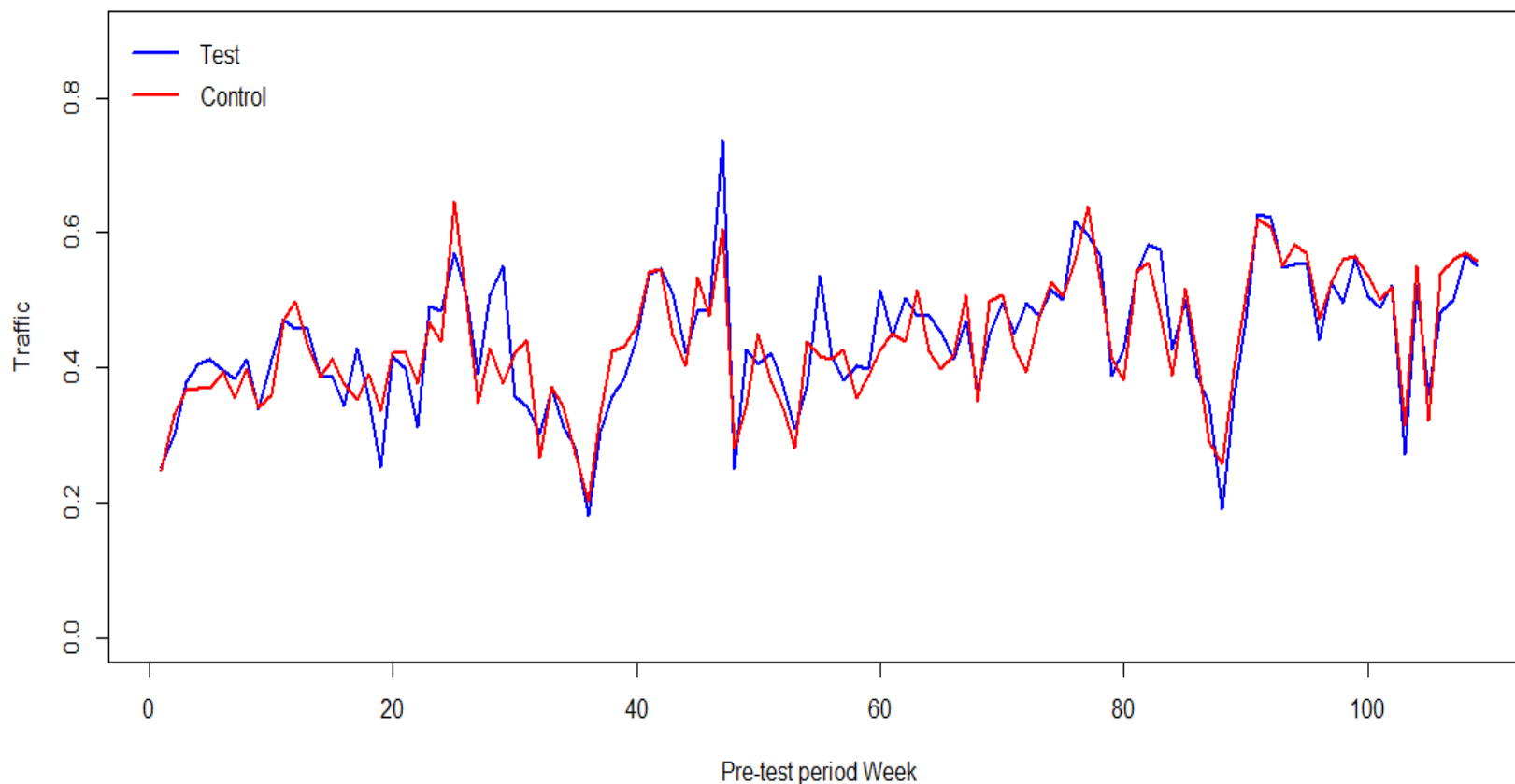
**Traffic of Test & Control groups**  
**Random assignment of the 54 best matching GMAs**



TEST			CONTROL			OUT		
AVERAGE WEEKLY			AVERAGE WEEKLY			AVERAGE WEEKLY		
CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE	CITY	TRAFFIC	TRAFFIC SLOPE
Alicante	20.23	0.1979	Leganes	14.97	0.2598	Gijon	23.57	0.0086
A Coruña	18.30	0.0481	Granada	17.91	0.0536			
Albacete	5.72	-0.0173	Gandia	5.67	-0.0013			
Avila	1.85	0.0171	Bejar	1.62	0.0019			
Barcelona	41.58	0.1547	Madrid	49.42	0.1240			
Camas	14.57	-0.0001	Mataro	14.62	0.0067			
Coslada	7.15	0.0060	Guadalajara	6.84	0.0299			
Eibar	3.01	0.0008	Sant Cugat del Valles	2.84	-0.0056			
El Puerto de Santa M	4.08	-0.0097	Orihuela	5.30	-0.0310			
Getafe	6.89	0.0064	Cartagena	7.06	-0.0223			
Jerez de la Frontera	12.38	-0.0198	Tarragona	11.03	0.0285			
Lleida	5.88	-0.0094	Burgos	5.52	0.0496			
Mahon	9.33	0.0562	Santander	9.61	0.0452			
Malaga	10.91	0.0086	Huelva	10.87	0.0078			
Manacor	3.22	0.0346	Segovia	3.78	0.0515			
Marbella	8.13	0.0240	Algeciras	8.87	-0.0331			
Mollet del Valles	4.01	0.0319	Lugo	3.81	0.0418			
Pamplona	9.17	0.0661	Leioa	9.23	0.0435			
Parla	7.31	0.0192	Fuenlabrada	7.12	0.0119			
Pontevedra	4.86	0.0147	Ourense	4.59	0.0149			
Requena	2.31	0.0000	Teruel	2.46	0.0160			
San Sebastian	9.88	0.0030	Reus	9.28	0.0112			
Santiago de Compost	17.56	-0.0130	Palma de Mallorca	16.01	0.0454			
Sevilla	23.13	0.0304	Valencia	21.39	-0.0156			
Tortosa	4.27	-0.0354	Palencia	3.85	-0.0280			
Valladolid	12.32	0.0620	Murcia	11.56	0.0380			
Vitoria	4.22	0.0667	Cordoba	5.73	0.0935			
Zaragoza	4.00	0.0279	Caceres	4.16	0.0192			
Sum: 276.25 (48%)    Mean: 0.0212			Sum: 275.09 (48%)    Mean: 0.0221			Sum: 23.57 (4%)    Mean: 0.0086		

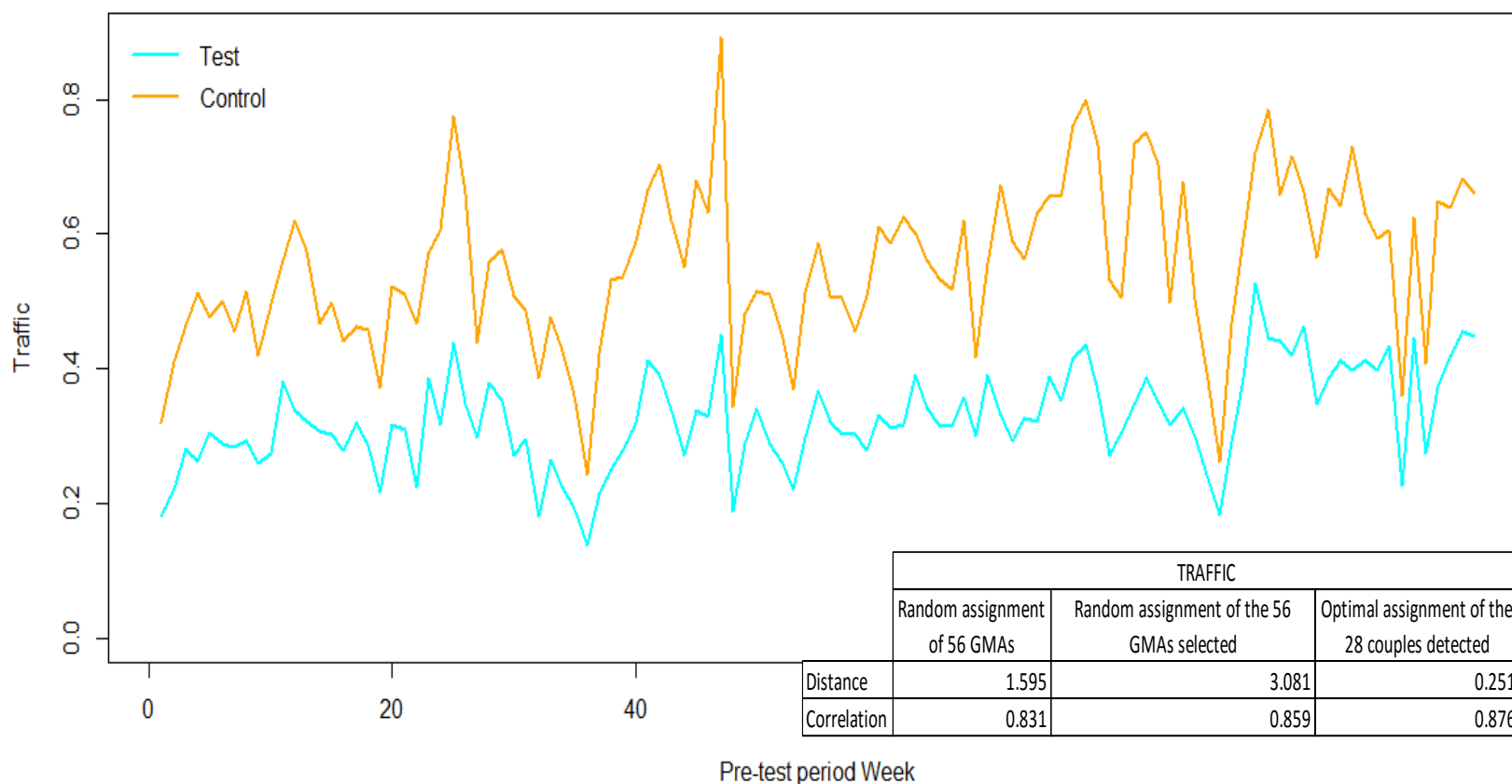
# Sucesivas iteraciones (3)

**Traffic of Test & Control groups**  
**Optimal assignment of the 56 best matching GMAs**



# Sucesivas iteraciones (3)

**Traffic of Test & Control groups**  
**Random assignment of the 56 best matching GMAs**



# 6 – Conclusiones

Etapa final

# Conclusiones

- A partir de cierto punto, añadir más parejas aumenta la distancia entre los Grupos de Test y Control.
  - En cada nuevo paso las parejas que se añaden son menos parecidas.

TRAFFIC							
Optimal assignment							Random assignment
	28 GMAs	36 GMAs	44 GMAs	50 GMAs	54 GMAs	56 GMAs	56 GMAs
Distance	0.229	0.221	0.170	0.251	0.242	0.251	1.595
Correlation	0.809	0.820	0.854	0.866	0.873	0.876	0.831

- Solución de compromiso (Tamaño vs. Similaridad de los Grupos)
  - Lanzar la campaña en el Grupo de Test más grande posible (28 ciudades).
  - Análisis doble, considerando también el subgrupo de 22.



# Conclusiones

- **Medida de la efectividad de la publicidad en YouTube**  
= Medida del aumento de la distancia (con respecto al Tráfico) entre los Grupos de Test y Control, durante el periodo de la campaña
  - Asumiendo que no hay ninguna otra diferencia de estímulos entre ambos Grupos.

Muchas gracias.  
¿Preguntas?