

# Examen:OPI Analytics

Christian Jiménez Beltrán

24 de julio de 2017

## Sección 1: Bebés

Para realizar la estimación de los bebés en la delegación Álvaro Obregón se implementó una regresión logística. Para implementar el modelo se revisó la base de datos del Censo de Población y Vivienda del 2010 en específico el cuestionario ampliado de personas, donde se tomaron las variables de sexo ('sexo'), edad ('edad'), escolaridad acumulada ('escoacum'), número de hijos ('numhij') y el año ('fecnaca') y mes del último hijo nacido ('fecnacm').

La lógica que se siguió para elegir esas variables es: Edad, es bien sabido que las mujeres de edad entre 15 a 30 años están en el proceso donde es muy probable que se embaracen, además, aquí en México de las estadísticas del INEGI se encontró que en el 2014 21 años representa la edad de fecundidad promedio a nivel nacional; Escolaridad acumulada, me pareció que es una variable que está muy estrechamente relacionada con el número de hijos que tiene una mujer, esto se verifica por la ENADID; Número de hijos, es una variable que afecta la probabilidad de tener un hijo más.

La variable dependiente se formó por el año y mes del último hijo nacido, donde se hizo una variable dummy con valor de uno si la mujer tuvo un hijo en el año 2010 entre los meses de enero a junio.

La regresión logística usó las variables antes mencionadas con la variable dependiente, se buscó modelar la probabilidad de que una mujer tenga un bebé en los primeros 6 meses del año 2010. Como no se cuenta con información actual de las variables independientes se pretendió hacer una estimación de la situación actual de las mujeres del 2010, con estos datos ajustados se realizó la predicción donde se encontró que actualmente no hay bebés de 0 a 6 meses en la delegación Álvaro Obregón.

## Sección 2: Ecobici

### 1. ¿En qué horarios hay mayor afluencia y en que estaciones?

Del análisis realizado se encontró que de lunes a viernes existe una mayor afluencia en los horarios de 8 y 9 de la mañana y 6 a 7 de la tarde en las estaciones 18, 266 y . Mientras que, los sábados son a las 2 y 6 de la tarde en las estaciones 27 y 267 respectivamente, y domingos a las 9 de la mañana y 7 de la noche en las estaciones 27 y 266 respectivamente. La siguiente tabla resume lo antes mencionado.

Cuadro 1: *Resultados*

	<i>Crowded time</i>	<i>Station</i>	<i>Second Crowded time</i>	<i>Station</i>
<b>Monday</b>	18	266	8	17
<b>Tuesday</b>	8	18	18	267
<b>Wednesday</b>	8	18	19	267
<b>Thursday</b>	8	27	19	266
<b>Friday</b>	8	18	18	266
<b>Saturday</b>	14	27	18	267
<b>Sunday</b>	19	266	9	27

A continuación se presenta un histograma de la afluencia con respecto a la hora del día.

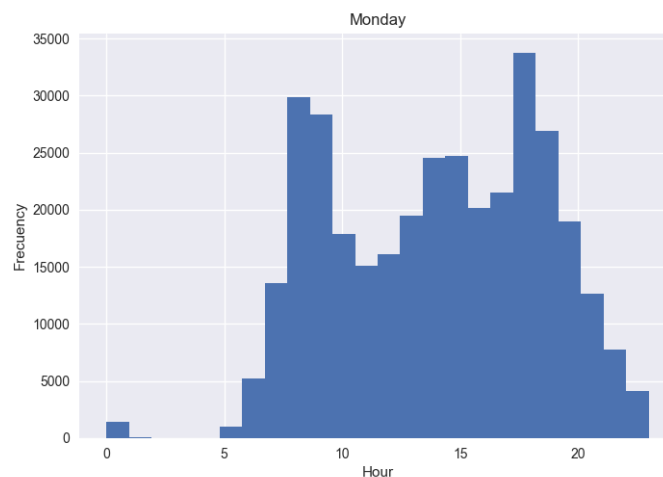


Figura 1: *Histograma de afluencia respecto a la hora del día*

El comportamiento de las horas con respecto a las estaciones de servicio se debe a la ubicación donde están, cerca de paseo de la reforma e insurgentes, donde pueden representar las zonas de trabajo de la gente y donde viven. Se puede decir que a las 8 de la mañana la gente que usa las eco bicis son gente que sale de paseo de la reforma dirigiéndose hacia sus respectivos trabajos y a las 7 de la tarde es gente que regresa a su casa.

### 2. Lo que se hizo fue obtener la serie de tiempo de cada estación y se sacó el cambio con respecto la semana pasada.

3. La siguiente figura muestra una sección de la matriz origen-destino , donde se destacan la ciclo base 1 con arribo a la estación 18 y viceversa.

Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	32	33	34
1	460	251	129	373	297	199	326	27	445	168	251	211	100	226	228	122	255	1.1e+03	143	70	197	80	236	150	244	59	545	171	178	128	91	61	275
2	183	253	45	70	27	160	66	10	367	19	84	101	108	69	126	148	69	101	55	32	83	27	54	218	80	12	299	76	41	47	320	105	148
3	120	91	193	71	65	79	184	31	60	156	34	22	162	106	224	120	234	227	113	55	104	47	92	48	90	113	61	64	172	24	27	16	35
4	493	73	35	207	40	115	75	37	45	56	120	58	31	52	46	52	33	110	32	12	130	4	51	38	41	19	203	63	50	53	256	68	37
5	299	44	22	70	172	50	103	13	78	34	105	77	66	78	77	37	32	91	57	9	57	23	35	64	25	11	272	106	42	23	150	34	94
6	260	103	16	212	104	357	118	24	135	35	65	143	183	98	188	45	74	155	53	18	58	19	79	49	43	20	208	110	94	47	98	41	135
7	410	109	21	102	71	204	961	29	255	39	181	106	86	62	67	82	41	195	122	144	39	36	62	104	80	16	369	200	55	77	141	39	41
8	48	9	20	31	12	22	72	128	19	59	32	6	29	37	103	36	52	35	14	23	34	39	40	29	53	125	78	30	336	18	11	11	87
9	343	146	19	48	28	84	97	16	255	18	125	122	88	51	59	48	17	84	44	22	77	10	16	52	20	21	270	87	39	21	337	80	57
10	225	57	216	89	53	111	96	47	62	282	76	79	130	161	328	175	297	377	429	217	119	41	163	90	105	61	203	79	60	52	32	4	43
11	342	122	44	171	138	107	110	51	101	37	305	47	29	64	202	94	81	94	29	16	31	29	51	104	48	49	69	132	61	17	30	41	87
12	164	98	29	77	92	237	99	11	91	37	43	134	89	57	116	136	125	118	95	16	35	35	82	36	88	17	54	72	93	63	32	10	60
13	179	38	55	32	37	121	70	48	55	52	25	32	149	64	76	190	114	302	103	152	64	50	68	124	213	44	490	58	116	54	35	46	50
14	79	32	38	16	25	61	63	48	34	51	25	26	83	193	32	68	50	188	135	106	45	38	42	89	132	133	349	14	56	74	22	23	28
15	180	91	94	30	45	57	113	93	30	190	106	72	103	59	231	143	97	368	146	319	102	30	74	85	140	99	769	45	134	115	112	69	152
16	140	139	36	49	22	46	183	37	39	98	58	42	211	125	156	220	45	128	108	169	144	16	39	64	53	45	299	75	99	33	247	28	64
17	304	87	53	52	51	109	130	39	37	60	56	64	200	131	176	66	200	164	164	202	89	39	76	22	61	62	416	74	184	44	70	13	182
18	1.47e+03	165	104	131	74	167	310	50	83	146	78	69	440	377	531	189	204	428	139	151	247	96	205	124	69	90	210	261	244	44	150	28	56
19	125	90	47	73	82	102	296	23	73	86	45	72	243	310	380	321	369	271	248	68	239	70	175	127	113	104	82	115	212	65	60	83	113
20	104	140	17	127	100	80	265	27	114	45	31	25	294	250	536	366	520	389	117	251	160	96	185	129	138	57	126	84	123	37	46	13	87
21	233	50	85	118	46	99	46	196	102	100	49	39	28	23	86	141	116	274	256	200	303	212	240	344	422	141	898	61	196	100	66	101	62
22	93	46	51	9	21	49	33	102	15	88	53	30	34	59	25	35	37	140	106	119	207	182	51	100	165	123	669	36	202	98	18	33	48

Figura 2: Sección de la matriz origen destino

Las siguiente figura muestra la ubicación de la estación 1 y 18, se puede observar que son estaciones cercanas. Se puede pensar que la estación 1 está en una zona de casa habitación mientras que la 18 está cerca de una avenida grande, lo que lleva a concluir que es gente que se quiere acercar a la avenida.

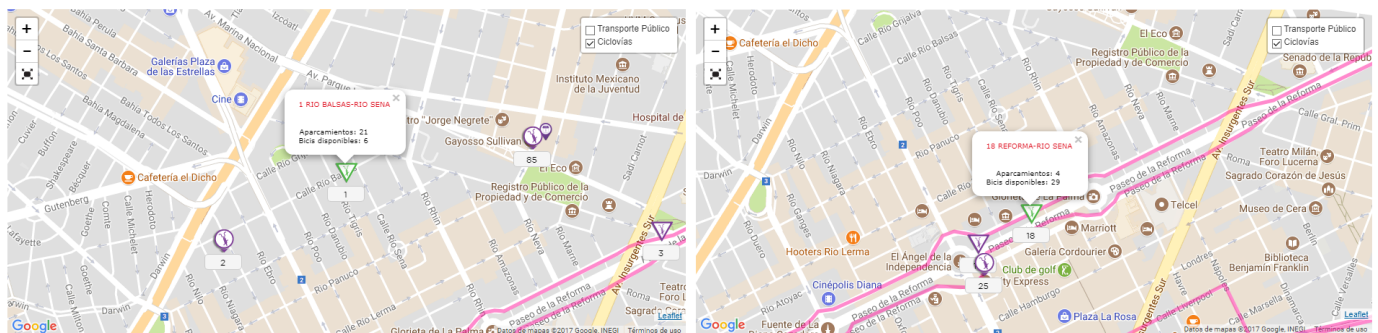


Figura 3: Muestra la relación entre la las estaciones 1 y 18

Un comportamiento similar se puede observar en las estaciones 21 y 27. Ambas estaciones se encuentran a corta distancia y sobre la misma ciclovía.

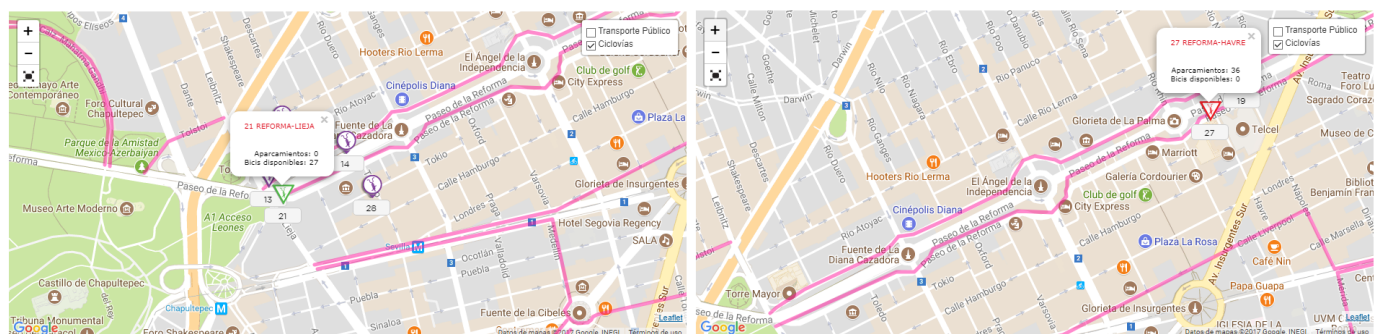


Figura 4: Muestra la relación entre la las estaciones 11 y 27

4. Para obtener los clusters se utilizó la técnica de K mean clustering, la figura que se muestra adelante muestra los diferentes clusters que se encontraron, además, del coeficiente de silhouette. Como se observa el cluster que mejor se acopla es el de  $k=2$ , lo que nos indica que se pueden agrupar las estaciones como las estaciones más recurrentes tanto de salida como de llegada y las que no son tan recurrentes.

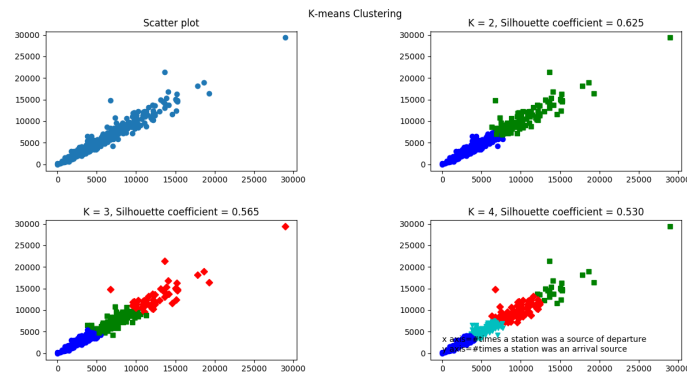


Figura 5: *Clusters de las estaciones de ecobici*

Se utilizo este método porque se quería ver el comportamiento de las estaciones como fuentes de salida y llegada, se pensó desde un principio que deberían de existir 2 clases de estaciones de acuerdo a su recurrencia y este método me pareció para identificar esa homogeneidad entre clusters.

## Referencias

- <http://www.beta.inegi.org.mx/proyectos/ccpv/2010/>
- [http://www.sedesol.gob.mx/work/models/SEDESOL/Sedesol/sppe/upri/dgap1/fais/Calculo/2013/FAIS/Calculo\\_FAIS\\_2013.pdf](http://www.sedesol.gob.mx/work/models/SEDESOL/Sedesol/sppe/upri/dgap1/fais/Calculo/2013/FAIS/Calculo_FAIS_2013.pdf)
- <http://cuentame.inegi.org.mx/poblacion/escolaridad.aspx?tema=P>
- <http://www.beta.inegi.org.mx/temas/natalidad/>
- <http://www.beta.inegi.org.mx/proyectos/registros/vitales/natalidad/>
- <http://www.beta.inegi.org.mx/proyectos/enchogares/especiales/enadid/2014/>
- [http://www.inegi.org.mx/saladeprensa/aproposito/2017/mujer2017\\_Nal.pdf](http://www.inegi.org.mx/saladeprensa/aproposito/2017/mujer2017_Nal.pdf)