

El seguimiento diario por encuesta en el Distrito Federal

Ricardo de la Peña.

Presidente Ejecutivo de
ISA Investigaciones Sociales Aplicadas

Por su condición de entidad capital, el Distrito Federal ha sido el espacio donde se han desarrollado la mayoría de las empresas de investigación demoscópica. Este hecho propicia que sea aquí donde primero se publican encuestas electorales en fecha próxima a una elección, a mediados de la década de los ochenta del siglo pasado.

La primera oleada de estudios demoscópicos, que atendían a los procesos federales tanto para el Ejecutivo como para el Legislativo en la entidad capital, se desarrolló en el marco de profundos cambios de la realidad política nacional: la emergencia de condiciones propiamente democráticas en la competencia electoral, la diversificación de medios de comunicación y el desarrollo de tecnologías que permitían la realización de estudios por muestreo en tiempos y costos relativamente reducidos.

Debido a las características esencialmente urbanas de la entidad y a la relativa homogeneidad en las distribuciones de preferencias electorales, comparado con lo que ocurre en otras entidades del país, el Distrito Federal no solamente fue un espacio observado regularmente por encuestas, sino que éstas eran generalmente exitosas para anticipar los resultados.

En 1997 se permitió por vez primera a la ciudadanía del Distrito Federal la elección de su gobierno. A partir de ese año, cada ocasión en que es elegido el Jefe de Gobierno de la entidad se ha contado con mediciones por encuesta que han ido dando cuenta del estado de la contienda.

Los resultados electorales de las tres primeras elecciones de Jefe de Gobierno, excluyendo votos nulos, habían sido los siguientes (bajo el rubro del partido mayor se incluyen los resultados conjuntos de los partidos que fueron en alianza):

Año	PAN	PRI	PRD	Otros
1997	15.6	25.6	48.1	10.7
2000	33.9	23.1	38.3	4.7
2006	27.7	21.9	47.1	3.3

Para la elección de 2012 era esperable que se ahondara en el esfuerzo por dar cuenta de las preferencias de los ciudadanos y sus cambios en el tiempo. La experiencia previa permitía suponer que estaban dadas las condiciones para que medios y encuestadoras

acordaran llevar adelante programas de mediciones más frecuentes que las tradicionales series mensuales.

Y eso fue lo que ocurrió: un grupo editorial resolvió llevar adelante un seguimiento diario mediante encuesta: Milenio, contratando a la firma GEA-ISA. Esto, como una estrategia paralela a un ejercicio similar para el seguimiento cotidiano de las preferencias de la ciudadanía a nivel nacional rumbo a la elección presidencial.

Sobre la metodología del seguimiento.

El objetivo de operar una encuesta de seguimiento diario sobre preferencias para las elecciones de 2012 fue una inquietud atendida por la casa encuestadora desde un año antes de los comicios.

Se sabía que era inviable técnica y financieramente realizar dicho ejercicio recurriendo a tecnologías y metodologías convencionales para estudios por encuesta. Además, dada la cobertura de la telefonía domiciliaria en el país, este modo de aproximación estaba descartado si lo que se pretendía era disponer de estimadores relativos al universo de los electores. Así, el único método de aproximación válido era la entrevista personal en vivienda.

Se tendrían que adoptar métodos empleados para encuestas de seguimiento en otras naciones, mediante estimaciones de medias móviles para períodos de tres días, lo que reducía el número de casos a observar a la tercera parte. Además, adoptando sistemas de rotación diaria de las unidades en muestra y no sorteos independientes para cada jornada, se lograría una relativa parsimonia en el cambio de las estimaciones.

Para la definición de la estrategia de operación en una encuesta de seguimiento, había que tomar en cuenta asuntos operativos adicionales. Dado que es imposible lograr un acopio manual de cuestionarios en varios puntos de entrevista por la misma persona, tenía que suprimirse toda intermediación y adoptar mecanismos orientados a la trasmisión directa y en tiempo real de las entrevistas, pero sin menoscabo del control de calidad, que garantizara la efectiva y adecuada aplicación.

Ello sólo era posible arribando al empleo de tecnologías de comunicación en línea mediante teléfonos inteligentes. Pero, para el control de calidad se requería disponer de plataformas de control de ubicaciones geográficas de operadores, por lo que se solicitó y obtuvo del IFE la información para la referencia geográfica de las secciones electorales.

Así, todo caso transmitido podía ser checado de manera automática, para cotejar que las coordenadas de la trasmisión correspondieran con el perímetro de operación elegido; caso contrario, el caso se invalidaba.

Adicionalmente, se disponía del domicilio y descripción de la vivienda, por lo que se podía consultar de inmediato su existencia en cartografía y corroborar que la descripción coincidiera con archivos públicos de imágenes. Asimismo, al acopiar números de teléfonos para el posterior contacto con un entrevistado, se podían efectuar rutinas de verificación mediante llamadas a los supuestos entrevistados.

Luego, si bien se recurrió a procedimientos convencionales de supervisión en campo, mediante itinerarios que permitían la cobertura paulatina y sin previo aviso de los operadores, el control de calidad de los estudios no tuvo que reducirse a estos procedimientos.

Con base en estas decisiones, se hicieron las estimaciones sobre los requerimientos tecnológicos, se hizo un estudio del mercado de oferentes de servicios de transmisión de datos, conociendo y evaluando sus plataformas y mecanismos de comunicación; con base en dicho estudio, se contrató los servicios con un proveedor que aportaba condiciones tecnológicas y económicas idóneas para la operación.

Esto permitió disponer de una primera aproximación a los costos involucrados en el proyecto. Con base en ello, se iniciaron acercamientos con potenciales patrocinadores, para conocer la posibilidad de trabajar un estudio de esta naturaleza orientado a su difusión pública, lo que concluiría en el convenio celebrado con el Grupo Milenio, para que éste difundiera los resultados y se responsabilizara cabalmente del patrocinio.

En febrero de 2012 se llevó a cabo el diseño definitivo y selección de la muestra para el estudio en la entidad capital, incluyendo reemplazos potencialmente necesarios y criterios para enfrentar las dificultades esperables en este tipo de operativos.

Al mismo tiempo, se avanzó en la definición de los contenidos e instrumentos de recuperación de datos, considerando la información mínima deseable y la limitación de tiempos de entrevista por el recurso a nuevas opciones tecnológicas para el acopio y transmisión de datos.

Además, para la puesta en punto del proyecto, se realizaron diversos ejercicios para probar la tecnología, medir tiempos y efectividad de la toma y transmisión de datos, checar la plataforma de recuperación que alimentaría la base de datos y corroborar la corrección de las rutinas de traslado de la información a programas convencionales. Asimismo, se efectuaron mediciones para cotejar el impacto de la adopción de procedimientos de entrevista asistida por computadora contra lo observado con procedimientos tradicionales de papel y lápiz, detectándose que ello no provocaba ningún sesgo significativo en los resultados.

El perfil de los operadores seleccionados y capacitados finalmente para el operativo continuo fue el siguiente: la mitad eran hombres y la otra mitad mujeres; su edad iba

de los veinte a los sesenta años, con una media próxima a los treinta años; dos tercios contaba con estudios superiores y el resto de nivel medio.

En paralelo, se definieron algunos principios básicos de carácter metodológico para la operación del proyecto: tamaño deseable de la muestra y dispersión requerida. Conforme estas estimaciones, se decidió que se requeriría disponer de una muestra de veinticuatro secciones electorales elegidas con probabilidad proporcional al tamaño, definido por el número de electores registrados en el Listado Nominal.

Como segunda etapa, se elegirían ocho viviendas por sección, mediante recorrido aleatorio, con arranque a partir del domicilio de la casilla básica en la elección federal pasada e intervalo constante por cada sección en muestra. Para la selección del informante, se trataría de tomar datos a la persona que, formando parte de la población objetivo, celebrara el próximo cumpleaños de entre las presentes al momento de la entrevista.

En caso de no poderse realizar la entrevista en el domicilio elegido o a la persona seleccionada, se procedería a entrevistar en una vivienda contigua, hasta su realización, llevando control de la tasa de rechazo, asumiendo que ésta correspondería al cociente de los casos de entrevista completa transmitidas automáticamente entre contactos realizados para lograr dichas entrevistas, incluyendo entrevistas completas, incompletas, rechazos y casos en que se hubiera logrado el contacto, pero en los que en la vivienda seleccionada no residieran personas con las características demandadas en la población a observar. Por razones prácticas, se excluirían del cálculo los casos en que no se lograra el contacto o no existiera información que permitiera determinar la elegibilidad de los residentes en una vivienda.

Las entrevistas serían personales (cara a cara), en viviendas particulares, aplicando de manera asistida por computadora un cuestionario estructurado, cuyo contenido estaría exclusivamente vinculado a este ejercicio. La transmisión de datos sería en línea para su compilación; en caso de no poder transmitirse de manera automática los datos por cobertura o cualquier otro motivo, se recuperarían en papel, para su posterior transmisión.

Para fines del reporte público de estimaciones sobre preferencias electorales, se agregarían los 576 casos tomados durante tres días consecutivos en 72 distintas secciones electorales, lo que permitiría cumplir los requerimientos para disponer de datos con un margen de error teórico de ± 4 por ciento al 95 por ciento de confianza.

El procedimiento para el cálculo de proporciones por candidato adoptaría dos modalidades: la estimación a partir del total de observaciones, incluyendo los casos en que el entrevistado no se definiera por un candidato o partido por cualquier razón; y la estimación excluyendo los casos en que el entrevistado no se definiera por alguien, mediante la simple exclusión de los casos no definidos para fines de cálculo.

Los resultados presentados no serían en ningún caso frecuencias simples, sino estimaciones basadas en el empleo de factores de expansión muestral, calculados como el inverso de la probabilidad de selección de cada individuo en muestra, ajustados para hacer acorde la distribución diaria de casos por sexo y edad con la reportada en el Listado Nominal de Electores.

Todo ello respondía a un doble requerimiento, propio de las encuestas de seguimiento cotidiano: buscar la estabilidad básica, pero una sensibilidad a los cambios en la coyuntura. De lo que se trata en estos ejercicios es de mostrar las variaciones en el día al día, en el marco de una distribución básica que puede ser sostenida a lo largo del tiempo.

Para cumplir con las disposiciones legales, se estableció contacto con la autoridad administrativa electoral del Distrito Federal, a la que se le hizo llegar cotidianamente por vía electrónica los reportes de las estimaciones observadas en cada corte diario.

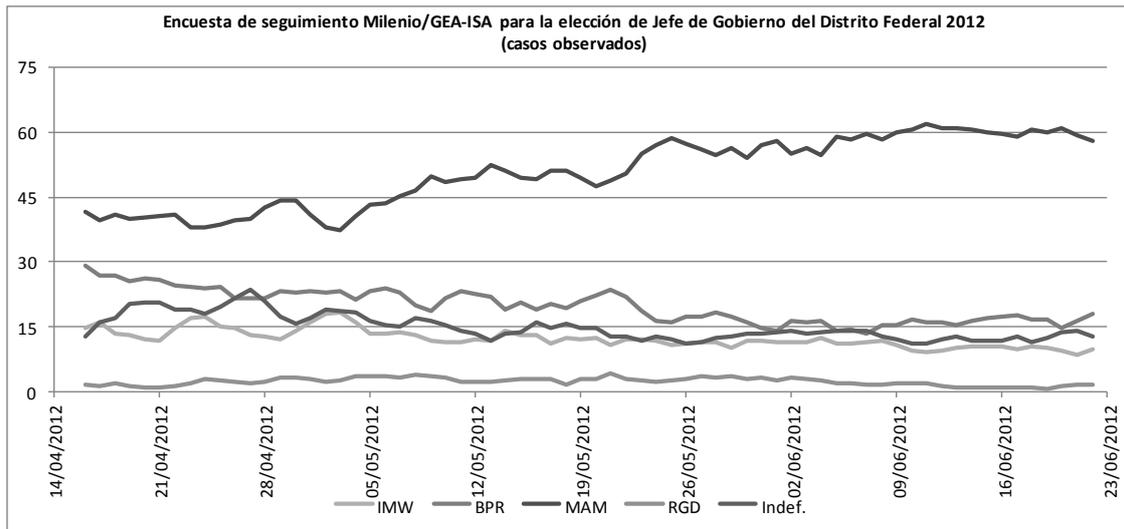
Sobre los resultados del seguimiento.

¿Qué historia contó esta encuesta de seguimiento en el Distrito Federal?

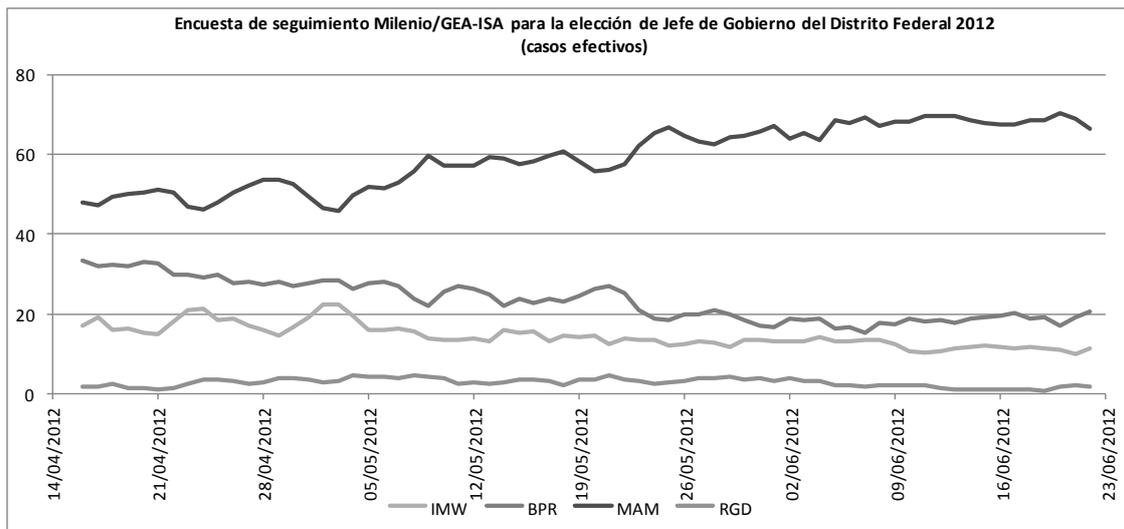
La narración de esta elección conforme al seguimiento diario mostraría tres etapas: antes del arranque formal de las campañas se tenía un escenario donde el candidato de las izquierdas, Miguel Ángel Mancera, se mostraba adelante por alrededor de quince puntos de las dos contendientes que disputaban el segundo lugar y convertirse en una opción efectivamente competidora: Beatriz Paredes, por el PRI-PVEM, e Isabel Miranda de Wallace, por el PAN.

Al arrancar las campañas, la distancia de Mancera sobre sus contendientes se va abriendo de manera paulatina, pero sistemática, hasta alcanzar una ventaja de más de cuarenta puntos faltando seis semanas para el cierre del proceso.

Luego vendría un período de relativa estabilidad, donde la ventaja de Mancera sobre sus competidoras se sostiene por arriba de los cuarenta puntos, llegando algunos días inclusive a más de cincuenta por ciento.



Así, luego de que el Distrito Federal había sido una entidad fuertemente competida, donde por mucho tiempo ninguna opción electoral había alcanzado una mayoría absoluta de los votos, en esta ocasión la coalición de partidos de izquierda lograba cerrar las campañas con una ventaja de más de cuarenta, que se confirmaría en las urnas el día 1 de julio.



El comparativo de la última medición del seguimiento diario, que cerró por disposiciones legales el 22 de junio (pues en el Distrito Federal continúan siendo ocho días de veda antes de comicios en la difusión de encuestas) sería:

DATO	PAN	PRI	PRD	Otros
Encuesta	11.2	20.5	66.4	1.9
Resultado	13.9	20.1	64.7	1.3
Diferencia	2.7	- 0.4	- 1.7	- 0.6

Y la encuesta de salida que realizó GEA-ISA para Milenio en la entidad capital ratificaría la victoria de Mancera (aunque con una ligera subestimación de la ventaja, debido factiblemente a la captación solamente parcial del voto diferenciado entre la elección presidencial y la de Jefe de Gobierno):

DATO	PAN	PRI	PRD	Otros
Encuesta	13.9	22.8	61.4	1.9
Resultado	13.9	20.1	64.7	1.3
Diferencia	0.0	- 2.7	3.3	- 0.6

Así, la encuesta de seguimiento GEA-ISA en el Distrito Federal para Milenio aportó al público un seguimiento cotidiano prácticamente en tiempo real de las fluctuaciones coyunturales en las preferencias y reflejo acertadamente la distribución final de las intenciones de voto del electorado.

Una reflexión final.

Adentrados ya en la vigesimoprimer centuria, es importante adoptar métodos de medición que respondan a una nueva realidad, donde los tiempos en que el elector demanda información son cada vez más inmediatos.

Este ejercicio viene a constatar que otorgar datos sobre preferencias de manera casi inmediata es algo no solamente posible, sino adecuado. Lo que sería de esperar es que en posteriores ocasiones se sumen oferentes que permitan el contraste, todo ello desde la perspectiva de que en una democracia moderna los ciudadanos deben estar informados y que la pluralidad de fuentes es la mejor opción.

Además, la proximidad entre las estimaciones de preferencias para la Jefatura de Gobierno en el Distrito Federal con los resultados oficiales pone en cuestionamiento las críticas respecto a la eventual existencia de factores técnicos involucrados en los sesgos detectados en los ejercicios demoscópicos para el seguimiento de la elección por la Presidencia de la República en 2012.

Dado que la medición diaria nacional y la realizada en el Distrito Federal por GEA-ISA siguieron los mismos procedimientos técnicos, con diseños muestrales semejantes y niveles de dispersión equivalentes; que los instrumentos de recuperación de datos tuvieron similares características, contenidos y extensión; que las formas de selección y sustitución de entrevistados fueron las mismas; que hubo idénticos métodos de acopio, trasmisión y procesamiento de datos; que los grupos de trabajo tuvieron iguales formas de reclutamiento, selección y capacitación; y que los controles de calidad fueron homogéneos, es insostenible atribuir a factores de este tipo las desviaciones observadas en la encuesta nacional de seguimiento.

Dado que no existió un sesgo en la muestra adoptada por el estudio nacional al cotejarse contra el resultado oficial, los únicos argumentos que pudieran explicar la desviación observada en la encuesta nacional son todos ellos ajenos a la pericia de la casa investigadora, al corresponder bien sea a cambios en las preferencias de los ciudadanos, potencial ocultamiento del sentido del sufragio o, lo que es más creíble, a inadecuaciones entre las distribuciones de preferencias en el grueso de la ciudadanía respecto a la correspondiente al segmento que finalmente acudió a votar.

Entender por qué hubo fenómenos que propiciaron sesgos en las mediciones nacionales y no afectaron ejercicios locales paralelos es una tarea fundamental para avanzar en la comprensión y revisión de métodos y procesos de investigación que permitan a futuro mejorar la calidad de los estudios y perfeccionar su potencial adecuación a la realidad.

Siglas utilizadas.

GEA GEA Grupo de Economistas y Asociados.

IFE Instituto Federal Electoral.

ISA ISA Investigaciones Sociales Aplicadas.

PAN Partido Acción Nacional.

PRI Partido Revolucionario Institucional.

PRD Partido de la Revolución Democrática.

PVEM Partido Verde Ecologista de México.

Apéndice: Estimaciones diarias de la encuesta de seguimiento GEA-ISA para Milenio en el Distrito Federal.

Fecha	Casos	IMW	BPR	MAM	RGD	Indef.	Giro	IMW	BPR	MAM	RGD	Giro
16/04/2012	576	14.8	29.2	41.7	1.6	12.7	2.4	16.9	33.5	47.8	1.8	2.1
17/04/2012	574	16.0	27.0	39.6	1.4	16.0	4.5	19.0	32.1	47.2	1.7	2.1
18/04/2012	572	13.3	26.8	41.1	1.9	16.9	2.9	16.0	32.2	49.5	2.3	3.0
19/04/2012	572	13.0	25.6	40.0	1.2	20.2	3.3	16.3	32.1	50.1	1.5	0.9
20/04/2012	574	12.1	26.1	40.1	1.0	20.6	1.1	15.2	32.9	50.6	1.3	1.3
21/04/2012	571	11.9	26.0	40.7	0.8	20.6	0.6	15.0	32.7	51.2	1.1	0.6
22/04/2012	576	14.6	24.4	40.9	1.3	19.0	3.3	18.1	29.9	50.5	1.5	3.5
23/04/2012	576	16.9	24.2	38.0	2.0	18.9	3.1	20.8	29.8	46.9	2.5	3.7
24/04/2012	576	17.4	23.9	37.9	2.8	18.0	1.3	21.3	29.1	46.2	3.4	1.4
25/04/2012	576	15.0	24.1	38.7	2.7	19.5	2.5	18.6	29.9	48.0	3.5	2.7
26/04/2012	576	14.6	21.7	39.7	2.4	21.6	3.1	18.7	27.6	50.6	3.1	2.7
27/04/2012	576	13.0	21.5	40.0	2.0	23.5	2.2	17.0	28.1	52.3	2.6	2.2
28/04/2012	576	12.8	21.5	42.6	2.3	20.8	2.9	16.1	27.2	53.8	2.9	1.8
29/04/2012	576	12.2	23.1	44.3	3.2	17.2	4.2	14.7	27.9	53.5	3.9	1.7
30/04/2012	576	14.0	22.8	44.3	3.1	15.8	1.8	16.6	27.1	52.6	3.7	1.9
01/05/2012	576	15.9	23.1	41.0	3.0	17.0	3.4	19.2	27.8	49.4	3.6	3.3
02/05/2012	576	18.0	23.0	37.9	2.2	18.9	4.0	22.2	28.4	46.7	2.7	3.6
03/05/2012	576	18.3	23.1	37.3	2.7	18.6	0.9	22.5	28.4	45.8	3.3	0.9
04/05/2012	576	16.0	21.4	40.6	3.7	18.3	4.3	19.6	26.2	49.7	4.5	5.1
05/05/2012	576	13.3	23.3	43.4	3.7	16.3	4.7	15.9	27.8	51.9	4.4	3.8
06/05/2012	576	13.5	23.9	43.7	3.6	15.3	1.1	16.0	28.2	51.5	4.3	0.5
07/05/2012	576	13.8	22.9	45.2	3.1	15.0	1.8	16.2	27.0	53.1	3.7	1.8
08/05/2012	576	13.0	19.8	46.4	3.9	16.9	3.9	15.6	23.8	55.9	4.7	3.8
09/05/2012	576	11.7	18.5	49.9	3.6	16.3	3.5	13.9	22.1	59.6	4.4	3.7
10/05/2012	576	11.3	21.6	48.4	3.2	15.5	3.1	13.4	25.6	57.2	3.8	3.5
11/05/2012	576	11.4	23.1	49.2	2.3	14.0	2.4	13.3	26.9	57.2	2.6	1.3
12/05/2012	576	12.0	22.6	49.6	2.4	13.4	1.1	13.9	26.1	57.2	2.8	0.8
13/05/2012	576	11.7	21.8	52.5	2.3	11.7	2.9	13.2	24.7	59.5	2.6	2.3
14/05/2012	576	13.9	19.1	51.1	2.5	13.4	4.1	16.0	22.1	59.0	2.9	3.1
15/05/2012	576	13.1	20.7	49.6	3.0	13.6	2.3	15.2	23.9	57.4	3.5	2.4
16/05/2012	576	13.1	19.0	49.0	2.9	16.0	2.4	15.5	22.7	58.4	3.4	1.3
17/05/2012	576	11.2	20.2	51.0	2.9	14.7	3.2	13.2	23.7	59.8	3.3	2.4
18/05/2012	576	12.3	19.3	51.1	1.7	15.6	2.1	14.5	22.9	60.6	2.0	2.1
19/05/2012	576	12.0	20.9	49.6	2.9	14.6	2.8	14.1	24.4	58.1	3.4	2.9
20/05/2012	576	12.5	22.3	47.5	2.9	14.8	2.1	14.7	26.1	55.7	3.5	2.4
21/05/2012	571	10.8	23.5	48.9	4.1	12.7	3.8	12.3	26.9	56.1	4.7	2.4
22/05/2012	571	12.1	22.0	50.3	3.0	12.6	2.7	13.9	25.2	57.5	3.4	3.0
23/05/2012	571	12.0	18.5	54.9	2.7	11.9	4.6	13.6	21.0	62.3	3.1	4.8
24/05/2012	571	11.7	16.4	57.0	2.1	12.8	3.0	13.4	18.8	65.4	2.4	3.1
25/05/2012	576	10.7	16.1	58.6	2.5	12.1	2.0	12.1	18.3	66.7	2.9	1.8
26/05/2012	576	11.1	17.5	57.3	2.9	11.2	2.2	12.5	19.8	64.5	3.2	2.2
27/05/2012	576	11.6	17.4	56.0	3.5	11.5	1.4	13.1	19.7	63.2	4.0	1.4
28/05/2012	568	11.3	18.3	54.8	3.3	12.3	1.7	12.9	20.9	62.5	3.7	1.2
29/05/2012	576	10.2	17.3	56.3	3.6	12.6	2.1	11.7	19.8	64.4	4.1	2.3
30/05/2012	576	11.8	15.9	53.9	3.0	13.4	3.4	13.6	18.4	64.5	3.5	2.0
31/05/2012	576	11.8	14.6	56.9	3.2	13.5	2.3	13.6	16.9	65.8	3.7	1.5
01/06/2012	573	11.4	14.2	57.9	2.7	13.8	1.3	13.2	16.5	67.2	3.1	1.4
02/06/2012	573	11.4	16.2	55.1	3.3	14.0	2.8	13.2	18.9	64.1	3.8	3.1
03/06/2012	573	11.4	15.9	56.5	2.8	13.4	1.4	13.1	18.4	65.3	3.2	1.2
04/06/2012	576	12.4	16.3	54.8	2.7	13.8	1.8	14.3	18.9	63.6	3.2	1.7
05/06/2012	576	11.1	14.0	58.9	1.9	14.1	4.4	13.0	16.3	68.6	2.1	5.0
06/06/2012	576	11.2	14.5	58.4	1.8	14.1	0.6	13.1	16.8	68.0	2.1	0.6
07/06/2012	568	11.6	13.3	59.7	1.5	13.9	1.7	13.5	15.4	69.4	1.7	1.8
08/06/2012	568	11.7	15.4	58.4	1.7	12.8	2.4	13.4	17.6	67.0	2.0	2.5
09/06/2012	568	10.8	15.3	59.9	1.8	12.2	1.6	12.3	17.4	68.2	2.1	1.3
10/06/2012	576	9.5	16.7	60.6	2.0	11.2	2.3	10.7	18.8	68.3	2.2	1.6
11/06/2012	576	9.1	16.1	61.9	1.9	11.0	1.3	10.2	18.0	69.6	2.2	1.3
12/06/2012	576	9.4	16.0	61.1	1.3	12.2	1.5	10.7	18.3	69.6	1.4	0.8
13/06/2012	576	10.0	15.4	60.9	0.9	12.8	1.2	11.4	17.7	69.8	1.1	0.9
14/06/2012	576	10.3	16.4	60.5	0.9	11.9	1.3	11.6	18.7	68.7	1.0	1.2
15/06/2012	576	10.4	17.0	59.8	0.9	11.9	0.7	11.9	19.3	67.8	1.0	0.9
16/06/2012	576	10.4	17.3	59.5	1.1	11.7	0.5	11.8	19.6	67.4	1.2	0.5
17/06/2012	576	9.9	17.6	58.9	1.0	12.6	1.2	11.4	20.1	67.4	1.1	0.5
18/06/2012	576	10.3	16.6	60.6	0.9	11.6	2.1	11.7	18.8	68.5	1.0	1.4
19/06/2012	576	10.0	16.8	60.0	0.7	12.5	1.1	11.4	19.2	68.6	0.8	0.5
20/06/2012	576	9.4	14.6	60.8	1.4	13.8	2.8	10.9	16.9	70.5	1.7	2.8
21/06/2012	576	8.5	16.5	59.3	1.7	14.0	2.4	9.9	19.2	68.9	2.0	2.6
22/06/2012	576	9.8	17.9	58.0	1.7	12.6	2.7	11.2	20.5	66.4	1.9	2.6