



## Documento metodológico del cálculo de composición Senado 2014

27 de febrero 2014.

### Antecedentes

En Colombia no se han realizado predicciones estadísticas sobre la composición del Senado de la República. En efecto, el seguimiento de las tendencias políticas, se ha concentrado en la competencia por la Presidencia o alcaldías en los casos de las grandes ciudades.

Estimar los resultados de este tipo de elección implica entender realidades de la dinámica electoral que incluyen normas, variaciones de la opinión pública y la forma como se generan o se rompen las diferentes alianzas políticas.

Un candidato con alto número de votos, no necesariamente es electo. Se requiere estimar el comportamiento global de la lista y la forma como se distribuyen los apoyos a su interior. Hemos visto aspirantes con más de 150 mil votos, que por estar en una lista que no logra el umbral no fueron electos y en contraste, senadores con el apoyo de 10 mil personas por estar en un partido con dos o tres grandes votaciones que “impulsan” a los demás de la lista.

### Metodología de cálculo

En la contienda los partidos pueden tener dos tipos listas: “Abierta” y “Cerrada”. Por ello, el ejercicio, para las listas “Abiertas” debe estimar varios elementos:

1. El volumen total de votos válidos,
2. El umbral,
3. La cifra repartidora,
4. El volumen de votos de cada lista
5. El volumen de votos de los candidatos
6. El ordenamiento de los aspirantes al interior de cada lista según su votación.

Para las cerradas el orden de elegibilidad ya fue definido al momento de la inscripción.

Desde octubre de 2013 Cifras y Conceptos se propuso realizar una estimación de la conformación del Senado de la República, a ser electo el próximo 9 de marzo. Para ello se realizaron:

1. 12.500 encuestas cara a cara, visitando los hogares, donde se indago, entre otras cosas por la afinidad política, la intención de voto por las listas, la imagen de las cabezas de las mismas, así como algunos candidatos con visibilidad pública de todos los partidos. También se incluyó una pregunta abierta y espontanea sobre por quién votaría cada encuestado al Senado. El resultado es una base de datos con 5.200 respuestas válidas para el caso de la pregunta abierta y con 287.500 registros para las cerradas.
2. 46 sesiones de grupo en 14 municipios donde participaron cerca de 500 personas.
3. 46 entrevistas a líderes políticos y candidatos al senado de todos los partidos y movimientos en contienda.

Con ello, hemos acopiado un gran volumen de información sobre el comportamiento e intención de voto al senado en 62 municipios donde se encuentran 9.1 millones de votantes activos.

Una elección depende del comportamiento tanto de la opinión pública como de las estructuras políticas con que cuentan los partidos y movimientos. Por ello, los resultados de las encuestas deben combinarse con otras fuentes de información para lograr un resultado más preciso.

El uso de varias fuentes de información relacionadas con los temas electorales es una práctica utilizada en Estados Unidos y Europa. Para ello se utilizan varios tipos de modelos que intentan construir los datos más robustos posibles para estimar los logros electorales de cada candidato, generando un "Score" o marcación previsible para cada uno de ellos.

En el caso Colombiano nuestro modelo intenta capturar variables tanto de opinión pública como de estructura política. El ejercicio que se presenta a continuación contiene para 161 candidatos, incluyendo los logros de los partidos y el voto en blanco como opciones validas, las siguientes fuentes de información:

- Los resultados expandidos para cada candidato, provenientes de la pregunta abierta y espontánea.

- El volumen de seguidores de Twitter con corte al lunes 24 de febrero del presente año. Para no castigar las listas cerradas, se les sumo los seguidores de los primeros candidatos de la lista.
- Para los casos de senadores repitentes, representantes a la cámara que aspiran al senado y familiares de exsenadores que aspiran, tomamos los datos de sus resultados en las votaciones del año 2010.
- En 46 entrevistas semiestructuradas con líderes políticos y candidatos al senado de todos los partidos se indago por la situación de la estructura política y los apoyos, con los cuales se estimó un valor promedio de la votación por estructura. En la última semana se validaron estos resultados para ajustar el resultado de la variable a evolución de la estructura.

Esto nos permitió construir una base con los siguientes indicadores:

1. La pregunta abierta de la Encuesta con información para 148 candidatos válidos.
2. Seguidores en Twitter para 146.
3. Votación en el 2010 para 90 congresistas repitentes o sus familiares.
4. Estimación de entrevistas para 161 casos.

Para la votación de 2010, con el fin de no afectar a los nuevos aspirantes, se estimó la correlación entre de los resultados de la última elección donde existe información, frente a las otras tres variables:

- los resultados de la encuesta,
- los seguidores de Twitter y
- las entrevistas,

Con base en ello se construyó una regresión logarítmica estimando los parámetros por el algoritmo de Newton Rapson. Esto generó un R cuadrado del 99%, lo cual significa que la predicción de los datos faltantes es extremadamente precisa.

A partir de esta se estima el voto efectivo de cada aspirante mediante un modelo de puntaje o scoring no supervisado, a partir del patrón de las correlaciones.

El procedimiento permite asignar el peso o la influencia de cada una de las variables seleccionadas sobre la votación efectiva, calculada para cada candidato al Senado, según la siguiente ecuación:

$$V_i = \alpha_1 Encuesta + \alpha_2 Twitter + \alpha_3 Voto\ 2010 + \alpha_4 Entrevistas + \varepsilon$$

Donde

$V_i =$  Votos efectivos del candidato

$\alpha_i =$  peso de la variable  $i$

*Encuesta = Encuestas Polimétrica Elecciones 2014*

*Twitter = Seguidores de Twitter por candidato*

*Voto 2010 = Votación obtenida en Senado 2010*

*Entrevistas*

*= Variable cualitativa resultado de grupos focales y entrevistas a políticos*

*$\varepsilon$  = Elementos exógenos al modelo*

El modelo anterior es la base para la generación de dos índices de scoring: el primero para medir la influencia de la opinión pública y el segundo para estimar la influencia de la organización política.

$$V_i = \gamma \text{Opinión} + \delta \text{Estructura Política}$$

*$\gamma$  = peso de la variable opinión*

*$\delta$  = peso de la variable Estructura Política*

Esta ecuación genera la clasificación o score por cada candidato. El score de cada candidato se agrupa según su filiación política. De esta manera se conoce la cantidad estimada de votos por partido, la cual se expande sólo a nivel de partidos para 10 millones de votos, correspondientes al universo de electores de los 62 municipios en donde se han practicado las encuestas.

Este último dato y el del número total de votos esperados en las elecciones de Senado 2014, permiten estimar el umbral y la cifra repartidora, que son el insumo para conocer el número de curules que obtendrá cada partido.

## **Resultados**

**Ver anexo 1: 70 con alta probabilidad**

**Ver anexo 2: 50 con buena probabilidad**

## **Equipo de Cifras & Conceptos**

César Caballero

Juan Fernando Lucio

Verónica Pinto

Gustavo Romero

Giovany Babativa

Angela Ordoñez

Mayra Nieto

arcela Herrada