**ANEXO 1.FORMULARIO DE POSTULACIÓN DE RETOS**

Fecha: 06/05/2020

Ciudad: Bogotá

Señores

Ministerio MinTIC

Atendiendo a la convocatoria realizada por el Ministerio de Tecnologías de la Información y las Comunicaciones, a continuación, nos permitimos proponer un reto para el proceso de formación que se realizará en Ciencia de Datos durante el 2020, el cual atiende a una necesidad particular de la entidad.

Hemos leído los términos y condiciones de la convocatoria y disponemos de la información necesaria para que el equipo pueda desarrollar el proyecto.

|  |  |
| --- | --- |
| **DATOS BÁSICOS** | |
| * + - 1. Nombre de la entidad o empresa: | Superintendencia del Subsidio Familiar |
| 1. Tipo de empresa o entidad: | Empresa privada: \_\_\_\_\_  Sector: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Entidad pública: **\_\_X\_\_\_\_**  Nacional \_\_\_\_\_\_  Territorial \_\_\_\_\_\_  Otra - Cual? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| 1. Área/ departamento o Dirección de la empresa o entidad a la que está asociada el reto. | Superintendencia Delegada para Estudios Especiales y la Evaluación de Proyectos |
| 1. Punto de contacto operativo para el Proyecto **(Líder)** | |
| Nombres y apellidos | Fernando Uribe |
| Cargo | Jefe Oficina de Tecnologías de la Información y las Comunicaciones |
| Correo electrónico | furibeg@ssf.gov.co |
| 1. Un punto de contacto técnico para el Proyecto que puede responder preguntas técnicas sobre el conjunto de datos y el problema | |
| Nombres y apellidos | Víctor Villamil |
| Cargo | Profesional Especializado |
| Correo electrónico | vvillamila@ssf.gov.co |
| **DESCRIPCIÓN DEL PROYECTO** | |
| **Descripción general de la empresa o entidad pública.** | La Superintendencia del Subsidio Familiar está encargada de vigilar la organización y funcionamiento del sistema de Subsidio Familiar. Tiene a su cargo la supervisión de las cajas de compensación familiar, organizaciones y entidades recaudadoras y pagadoras del subsidio familiar en cuanto al cumplimiento de este servicio y sobre las entidades que constituyan o administren una o varias entidades sometidas a su vigilancia, con el fin de preservar la estabilidad, seguridad y confianza del sistema del subsidio familiar; para que los servicios sociales a su cargo lleguen a la población de trabajadores afiliados y sus familias bajo los principios de eficiencia, eficacia, efectividad y solidaridad (Decreto 2150 de 1992).  Son funciones de la Superintendencia Delegada para Estudios Especiales y la Evaluación de Proyectos, realizar la evaluación y conceptuar sobre los planes, programas y proyectos de inversión para obras o servicios sociales que desarrollen las Cajas de Compensación. Para esto, debe realizar estudios económicos, financieros, administrativos y de operación de los servicios y programas sociales de las Cajas de Compensación Familiar. Realizar las evaluaciones de resultado y de impacto de los planes, programas y servicios sociales a cargo de las Cajas de Compensación Familiar. |
| **Descripción del problema que debe resolverse** | Minería de datos que permita revisar y visibilizar los progresos de las CCF en el cumplimiento de las metas definidas en los planes de desarrollo del país; mediante la medición de avances hacia el logro de estas metas. Igualmente debe permitir hacer proyecciones del comportamiento de las principales variables que permitan predecir el comportamiento futuro de los servicios de las CCF. |
| **¿Por qué es importante este problema?** | El sistema de Subsidio Familiar en Colombia desempeña una función social de vital importancia en la economía de los hogares del país. Actualmente las 43 Cajas de Compensación Familiar tienen más de 676.000 empresas afiliadas, más de 9.700.000 cotizantes y más de 11.400.000 beneficiarios.  El sistema propende por mejorar la calidad de vida de las familias de los trabajadores colombianos, por medio de la gestión y entrega de subsidios y servicios de una parte de los aportes de seguridad social que realizan los empleadores.  Entre los diferentes tipos de beneficios que disponen y otorgan las cajas de compensación se pueden encontrar los beneficios en los siguientes aspectos:   * Salud * Educación * Recreación * Cultura * Turismo * Deportes * Vivienda * Crédito * Microcrédito.   A continuación, presentamos algunos de los beneficios que obtiene la población colombiana de este sistema. Dentro de los beneficios directos recibidos a través de los Foniñez, están los siguientes:    Los beneficios de los fondos de la Ley 115 de 1994 están los siguientes:    Los beneficios de FOVIS están los siguientes:    Los beneficios de los fondos de FOSFEC están los siguientes:    Vale la pena resaltar, el ‘Subsidio de Emergencia’, el cual es un beneficio que estableció el Gobierno Nacional mediante el Decreto 488 de 27 de marzo de 2020 como una medida de orden laboral, dentro del Estado de Emergencia, Económica, Social y Ecológica que atraviesa el país por la propagación del Coronavirus. Este beneficio es administrado por el sistema de subsidio familiar.  El sistema de subsidio familiar está conformado por las 43 Cajas de Compensación Familiar del país. |
| **¿Cuáles son los conjuntos de datos que los participantes pueden acceder para resolver este problema?** | INFRAESTRUCTURA: Información de los bienes inmuebles que dispone la Caja de Compensación Familiar, en cualquier modalidad de tenencia o explotación.  COBERTURAS: se refiere a la cobertura de los servicios de la Caja de Compensación Familiar en una infraestructura dispuesta para ese objetivo. |
| **Enumere y describa el conjunto de datos específicos** | **INFRAESTRUCTURA**  Código de municipio: Código municipio del domicilio donde se encuentra ubicada la infraestructura, de acuerdo con la División Político-administrativa de Colombia.  Tipo infraestructura: Clasificación de las instalaciones que dispone la Caja de Compensación Familiar.  Código infraestructura: Código asignado por la Caja de Compensación Familiar a la infraestructura.  Tenencia de la infraestructura: Modalidad de tenencia de la Caja de Compensación Familiar sobre la infraestructura.  Capacidad de la infraestructura: Capacidad instalada de la infraestructura, que dispone la Caja de Compensación Familiar, en términos de unidad de medida en un instante dado  Georreferenciación Latitud: Ubicación geográfica en Latitud en grados decimales de cada una de las infraestructuras propias de la Caja.  Georreferenciación Longitud: Ubicación geográfica en Longitud en grados decimales de cada una de las infraestructuras propias de la Caja.  Área geográfica: Identifica si la zona donde se localiza la infraestructura si es urbana o rural.  **COBERTURAS**  Código Infraestructura: Sitio (infraestructura) donde se prestan los servicios al afiliado a la Caja de Compensación.  Servicio / Programa: Programa utilizado por el usuario del servicio.  Categoría: Código de la categoría del afiliado asignada por la Caja de Compensación Familiar de acuerdo con el rango de salario donde se encuentre ubicado  Rango de edad Ajustado Corresponde a la edad en años cumplidos de la persona en la cual se clasifica el usuario del servicio.  Sexo: Código del sexo de las personas que han usado el servicio prestado por la Caja de Compensación Familiar.  Número de personas: Número de personas naturales o jurídicas que han utilizado el servicio/programa en un sitio específico, que pertenecen a una categoría, a un rango de edad y a un sexo. Se trata de identificar de manera individual a las personas que hacen uso de los servicios de tal forma que se pueda determinar hacia que “individuos” se focalizan los servicios.  Participantes: Número de participantes aplica para aquellos servicios/programa que son contratados para eventos masivos por una persona natural o jurídica.  El número de participantes reportado debe ser igual al número de individuos que asistieron al evento masivo, hayan sido estos afiliados o no afiliados.  Número de veces usado: Se debe registrar el número de veces que cada uno de los individuos, ha utilizado el servicio/programa. Ello significa que el valor de este campo debe ser igual o superior a los valores registrados en los campos “Número de personas” o “Participantes”.  Esta información estará disponible por Caja de compensación y por mes de reporte; desde 2016 a 2019. La información de infraestructura es semestral con novedades mensuales y la información de coberturas es mensual. |
| **¿Qué tipo de conjuntos de datos externos pueden ser de interés para el problema?** | Plan Nacional de Desarrollo 2018-2022  Estadísticas del DANE. |
| **Resultados esperados del proyecto.** | Desarrollo de ejercicios de minería de datos que permita revisar y visibilizar los progresos de las CCF en el cumplimiento de las metas definidas en los planes de desarrollo del país; mediante la medición de avances hacia el logro de estas metas.  Desarrollo de ejercicios de predicción estadística, que permitan hacer proyecciones del comportamiento de las principales variables del sistema, con el fin de predecir el comportamiento futuro de los servicios de las CCF.  Estos ejercicios, deben confrontarse con la participación de las CCF en el Plan Nacional de Desarrollo 2018-2022; considerando por lo menos, uno de los siguientes aspectos:   * Línea A Primero los niños y las niñas: Desarrollo Integral desde la Primera Infancia hasta la Adolescencia * Línea C Educación de calidad para un futuro con oportunidades para todos * Línea E Vida y entornos dignos e incluyentes * Línea F Trabajo decente, acceso a mercados e ingresos dignos * Línea H Dignidad y felicidad para los adultos mayores * Línea I Deporte y Recreación para el Desarrollo Integral de los Individuos para la Convivencia y la cohesión social |

**Nota: Con la presentación de esta propuesta de proyecto nos comprometemos a dar respeto a todos y cada uno de los términos previstos en la Ley de protección de datos descritas en la Ley 1581 de 2012 así como las normas que la complementen, adicionen o sustituyan.**