

Dirección de Desarrollo Digital

Unidad de Científicos
de Datos



**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación

GENERACIÓN AUTOMÁTICA DE REPORTES DEL SISTEMA GENERAL DE REGALÍAS

Entidad

Departamento Nacional de Planeación

- Dirección de Desarrollo Digital.
- Dirección de Vigilancia de las Regalías.

Sector

Territorial

Lenguaje

R

Fuente de datos

SGR

Presentación

El Sistema General de Regalías (SGR) se encarga de distribuir los recursos que las compañías petroleras y mineras pagan al Estado colombiano por la explotación de recursos no renovables. Estos recursos financian proyectos de inversión presentados por las entidades territoriales para promover la equidad, el desarrollo y la competitividad regional en todo el territorio colombiano. Para tomar decisiones acertadas en el ámbito territorial, orientadas hacia la distribución objetiva (basada en datos) de los recursos del SGR, la Dirección de Vigilancia de las Regalías (DVR) genera reportes mensuales en que se consolidan datos de gran importancia, como la cantidad de proyectos y recursos aprobados por región, por tipo de OCAD (Órganos Colegiados de Administración y Decisión), por sector, por estado de ejecución y por nivel de riesgo. Sin embargo, la generación de estos reportes constituye una labor repetitiva e ineficiente: aunque la información se encuentra estructurada casi en su totalidad, el proceso de realización de cada informe puede superar los tres días (incluso más de una semana) es decir, más de 30 días en un año, comprometiendo recurso humano y tecnológico que podría destinarse a labores misionales de la entidad. Ante la posibilidad de mejorar este proceso, el presente proyecto realiza la automatización de dichos reportes, a partir de la información disponible en las distintas bases de datos con que se construye. El resultado permite reducir el tiempo de generación del reporte a menos de 4 horas, mejorando además su visualización en internet e incentivando a estandarizar la estructura de las bases de datos, lo que representa un ahorro de recursos para la entidad y una facilidad para que los tomadores de decisiones puedan analizar la información con más profundidad, de manera más detallada y con mayor prontitud.

The General System of Royalties (SGR) is responsible for distributing the resources that oil and mining companies pay to the Colombian State for the exploitation of non-renewable resources. These resources finance investment projects presented by territorial entities to promote equity, development and regional competitiveness throughout the Colombian territory. To make sound territorial decisions, oriented towards the objective (based on data) distribution of SGR resources, the Directorate of Royalties Surveillance (DVR) generates monthly reports in which important data are consolidated, such as the number of projects and resources approved by region, by type of OCAD (Collegiate Administration and Decision Bodies), by sector, by state of execution and by level of risk. However, the generation of these reports constitutes repetitive and inefficient work: although the information is structured almost in its entirety, the process of producing each report may exceed three days (even more than one week), i.e., more than 30 days in a year, compromising human and technological resources that could be assigned to the entity's mission tasks. Faced with the possibility of improving this process, this project automates these reports, based on the information available in the different databases with which it is built. The result allows to reduce the time of generation of the report to less than 4 hours, improving in addition its visualization in Internet and encouraging to standardize the structure of the data bases, which represents a saving of resources for the entity and a facility so that the decision makers can analyze the information with more depth, in a more detailed way and within a shorter waiting time.

Objetivo general

Hacer más sencilla y eficiente la generación del Boletín Mensual de Proyectos SGR, a través de la generación automática de cifras, estadísticas y gráficos construidos con información de SUIFP y GESPROY.

Objetivos específicos

1. Identificar las cifras y gráficos que pueden generarse de manera automática a partir de la información disponible, a partir de la revisión del boletín y de las bases de datos.
2. Automatizar la generación de los elementos identificados, a partir de la creación de una herramienta basada en código de R.
3. Validar la herramienta desarrollada mediante la retroalimentación con los actores interesados.

Metodología

El flujo de actividades del presente proyecto no es complejo y es bastante lineal, pues el valor agregado está presente en la automatización de una actividad ya estandarizada y en la oferta de una mejor visualización de la información. Así pues, para el desarrollo del proyecto se tomó como base el boletín mensual SGR del mes de junio y las bases de datos de SUIFP y GESPROY (sobre recursos del SGR) de los meses de mayo y junio, se identificó la información necesaria y se programó, a partir de un código en R, un flujo de información que permitiera generar las cifras y gráficos requeridas, de manera automática. Para ello, se utilizaron una serie de librerías de manejo, procesamiento y visualización de datos, especialmente *Markdown* (un lenguaje de marcado ligero) y *plotly* (una herramienta libre de visualización).

El proceso contó con la retroalimentación constante del equipo de la DVR, quienes establecieron los lineamientos de la visualización propuesta en *plotly*, que cuenta con elementos interactivos y con funcionalidades para generar, presentar y guardar los gráficos. Para aquellos datos que no pueden obtenerse de manera automática, se generó un archivo de Excel donde estos valores pueden ser ingresados por el usuario, evitando en cualquier caso el esfuerzo requerido para la creación de los gráficos del boletín.

Resultados

El boletín automático genera el reporte en su totalidad, correspondiente a 6 páginas completas con 19 gráficos y cientos de cifras calculadas automáticamente. La versión entregada el día 17 de septiembre de 2019 requiere del cálculo manual de 33 cifras y del ingreso manual de 5 categorías, actividad que justifica que el tiempo aproximado para generar el reporte se estime en 4 horas, a pesar de que la ejecución del código para generar el boletín tarda menos de 3 minutos (incluyendo la ubicación y cargue de las bases de datos). Finalmente, se genera un archivo HTML con el reporte automático, archivo que puede visualizarse de forma interactiva en un navegador y que permite al usuario crear un documento en PDF utilizando Edge o Chrome.

Conclusiones y recomendaciones

1. La herramienta permite reducir el tiempo de generación del Boletín de Recursos del SGR a menos de 4 horas, lo que hace que el proceso sea más sencillo, rápido y eficiente.
2. El boletín automático genera gráficos que son interactivos si se consultan desde un navegador, lo que permite realizar análisis más específicos y mejorar la visualización general del boletín.
3. El proyecto incentiva a estandarizar la estructura de las bases de datos para desarrollar más iniciativas similares, que permitan a los tomadores de decisiones disponer de la información con mayor prontitud y analizarla fácilmente de manera más profunda y detallada.