Viernes, 19 de agosto del 2022

**MODELO DE CONTRATO DE ANÁLISIS DE CONSUMO DE ENERGÍA GLOBAL Y EMISIONES DE CO2**

**PRIMERA:** Entendimiento de la situación actual.

A partir de diversos datasets de múltiples fuentes se hará un análisis exhaustivo que revele la situación actual de la emisión de dióxido de carbono (CO2) en la atmósfera. A este análisis, buscaremos relacionarlo con cada nación para entender cuál es el impacto que tienen en el medio ambiente, su huella de carbono y encontrar los factores más contaminantes.

**SEGUNDA:** Objetivos

Dentro de los servicios brindados se encontrará asesoramiento para:

1. Analizar y demostrar relación entre el consumo energético y la emisión de CO2 por país
2. Encontrar los países con mayor producción de energías renovables.
3. Analizar la evolución de las emisiones de CO2 a través del tiempo.
4. Mostrar los países con mayor calidad de aire para una mejor condición de vida.
5. Demostrar cómo el nivel de emisiones de CO2 influye en el medio ambiente.
6. Verificar que países alcanzarían una reducción de CO2 para el año 2030.

**TERCERA:** Alcance.

Dentro del alcance de los servicios ofrecidos estarán:

1. Diccionario de datos.
2. Análisis exploratorio de datos.
3. Proceso de extracción, transformación y carga de los datos.
4. Modelo predictivo de emisiones de CO2.

**CUARTA:** Fuera de Alcance.

Dentro de la labor de la empresa Green Data no se cuenta con el detalle, por lo tanto, los siguientes análisis no serán tenidos en cuenta:

1. Segmentación de ningún tipo que no sea geográfica.
2. Huella de carbono per cápita
3. Relación entre la emisión de CO2 y el contexto político, económico y social.
4. Fallecimientos por contaminación del aire.

**QUINTA:** Riesgos.

En cuanto a los riesgos, se contemplan distintas cuestiones fundamentales en lo referido al alcance de las acciones de la empresa.

1. Escalabilidad del proyecto. Cabe remarcar que uno de los objetivos del cliente es la creación de un Data Warehouse. Esto será importante para determinar y sentar bases para posteriores cargas de datos a futuro. La definición de las plataformas a utilizar pueden variar, así como su temporalidad de desarrollo.
2. En primera instancia se plantea usar la herramienta de visualización PowerBI, esta misma, puede ser reemplazada sin previo aviso en caso de que el proyecto lo fundamente.
3. Modelo predictivo: En primer lugar, para realizar el modelo se usará la librería scikit-learn, pero puede usarse otros tipos de tecnologías, por lo que puede variar el tiempo de trabajo.
4. El diagrama de Gantt mencionado en la DÉCIMA, puede sufrir modificaciones de tiempo de trabajo estipulado sin previo aviso. Dependerá de su interacción, estar al tanto de sus alteraciones.

**SEXTA:** Solución propuesta.

En esta solicitud y contando con dataset competentes, se procederá de la siguiente forma:

* Utilización de los lenguajes de Python y SQL para realizar un proceso de EDA y ETL. De esta manera, conoceremos en profundidad los datos elegidos y además, los estructuramos para su correcta utilización.
* Elaboración e implementación de KPI 's planteados en la OCTAVA, para obtener parámetros sobre el negocio y verificar el funcionamiento de los indicadores.
* Realización de un tablero con Power BI.
* Generación de un modelo de machine learning con que logre predecir las emisiones de CO2 .

**SÉPTIMA:** Metodología de trabajo.

Se llevará adelante una metodología ágil de tipo Scrum,donde se dividirá el trabajo en pequeños sprint de una semana planificados previamente y continuamente revisados.  
 Se propondrá un calendario de trabajo, donde cada miembro del equipo realizará tareas correspondientes a su habilidad, coordinadas por el método Kanban. La resolución de estas tareas podrá ser de manera individual o en pequeños grupos, siempre en marco a las puestas en común diarias, de cara a las presentaciones.

**OCTAVA:** Entregable.

La documentación a entregar una vez concluido el proyecto descrito en el presente, será:

* Informe Detallado.
* Dashboard interactivo.
* KPI’s (Desarrollados para análisis trimestrales):
  + Índice de generación de combustibles fósiles
  + Intensidad de emisiones de carbono
  + Índice de energías renovables
  + Índice de autonomía energética
  + Indicador de calidad de aire
  + Variación de temperatura
  + Variación del nivel del mar
  + Variación de masa glacial
  + Consumo energético en base al PBI
  + Generación energética en base al PBI

**NOVENA:** Equipo de trabajo.

Los recursos humanos encargados de llevar a cabo el proyecto serán descritos a continuación con sus diferentes roles:

Barrios Diego: Data Engineer

Huaman Arévalo Gabriela: Data Analyst

Mediavilla Julián: Data Cloud Engineer

Morales Elizabeth: Head of Machine Learning

**DÉCIMA:** Cronograma general.

Para la correcta comunicación de las etapas del proyecto, será de su utilidad un diagrama de Gantt actualizado según modificaciones. El mismo podrán encontrarlo a partir de un link una vez inicializado el convenio laboral.