

# **Esperanza de Vida**

Somos una empresa consultora, DSC Data Science Consultors y recibimos la tarea de analizar la esperanza de vida para realizar un estudio de mercado.

El cliente (empresa que comercializa suplementos vitamínicos para adultos mayores de 50 años) desea explorar nuevos mercados para expandir su negocio a nivel Internacional.

Para este Análisis se realizaron los siguientes pasos:

1. Análisis de bases datos relacionados con la Esperanza de vida.
2. Definición de OKR´s (Objetivos a medir) y KPI´s (indicadores medibles claves para alcanzar los Objetivos)
3. Extracción de 20 bases de datos, origen Banco Mundial y organismos Gubernamentales.
4. Limpieza de las bases de datos, agregado de valores faltantes y descarte de algunas bases, quedando 30 Países y las siguientes bases de datos:
5. Definición de una Arquitectura de DataWarehouse, se utilizo:

5.1 Programa en Python para la limpieza de tablas con ayuda de una app stremlit para consultas rapidas

5.2 AWS para almacenamiento de las Bases de datos.

5.3 MySQL como motor de búsqueda de las consultas a realizar.

5.4 Heroku API para disponibilidad los datos limpios para la agilización de datos en Power bi y machine learning

5.5 Power BI para la visualización de los datos. (Plus de una app stremlit para consultas rápidas y bot de telegram para consultas relámpago.

# **Sobre nosotros**

Somos una empresa de consultoría que promueve la transformación en las diferentes capas del negocio de nuestros clientes mediante el uso de plataformas digitales efectivas, que apoyen sus estrategias de negocio y les permitan alcanzar sus objetivos. Contamos con líneas de soluciones para impulsar cambios efectivos en las organizaciones que van desde la seguridad digital, hasta la mejora de procesos de negocio y avanzados analíticos para la operación y toma de decisiones a través del manejo de grandes cantidades de datos.

Para mas información puede visitarnos en:

<http://dsc.ace.com.pe>

Contactarnos en:

[dscdatascienceconsultors@gmail.com](mailto:dscdatascienceconsultors@gmail.com)

Y ver algunos trabajos como este en:

<https://github.com/DataScienceConsultors/Esperanza_de_Vida>

# **Sobre el equipo de trabajo**

El equipo de trabajo posee amplios conocimientos en el campo de datos, ya sea como científico de datos, analista de datos e ingeniería de datos:

**Joseph Pedraza Leo**

<https://www.linkedin.com/in/joseph-pedraza-leo-33b130150/>

**Miguel Ángel Vitelli**

<https://www.linkedin.com/in/joseph-pedraza-leo-33b130150/>

**Agustín Esteban Braida**

<https://www.linkedin.com/in/agustin-braida/>

**Gustavo Javier Benítez**

<https://www.linkedin.com/in/gustavo-benitez-1b5086250/>

# **Metodología de trabajo**



Se centra en crear un producto según los **requisitos exactos del cliente**. De ahí, el que se le involucre al máximo durante el método de gestión del desarrollo del producto.

* Comunicación constante entre el cliente y el equipo de desarrollo.
* Respuesta rápida a los cambios constantes.
* La planificación es abierta con un [**cronograma de actividades**](https://www.sinnaps.com/blog-gestion-proyectos/cronograma-online)**flexible.**
* El software que funciona está por encima de cualquier otra documentación.
* Los requisitos del cliente y el trabajo del equipo del proyecto son los principales factores de éxito del mismo.

**EL RITMO DE TRABAJO EN LA METODOLOGÍA XP DEBE SER SOSTENIBLE**

El ritmo tiene que ser sostenible por lo que se utiliza una mezcla entre los tableros Kanban que son capaces de priorizar la efectividad según lo planificado y técnicas de scrum, esta mezcla se la conoce como Scumban.

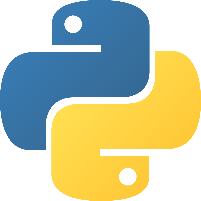
Para esto utilizamos la herramienta Trello y nos apoyamos en un equipo Scrum formado por nuestra Scrum Master, **Juliana Aragón** que maneja las dailys en donde paso a paso informamos que hicimos, que estamos haciendo y que vamos a hacer, contamos también con nuestro Product Owner **Ivan Tomaselli** que nos aporta y optimiza el valor del producto final en las Spring Review.

# **Análisis de datos**

# **Objetivo general**

1. Identificar los factores que se relacionan con la evolución de la esperanza de vida, analizando los siguientes datasets: desarrollo rural, cambio climático, ciencia y tecnología, comercio, desarrollo social, desarrollo urbano, deuda externa, PBI, educación, eficacia de la ayuda, energía y minería, genero, infraestructura, medio ambiente, pobreza, protección social, salud, sector financiero, sector privado y sector público.
2. Período de análisis desde el año 1971 hasta el 2021 (considerado como año actual y no por cohorte de nacimiento).
3. Alcance

# **Herramientas utilizadas**

# **Arquitectura de procesamiento de datos**

# **Cronograma de trabajo**



# **Instrucciones del Github**

Se proveen los siguientes archivos y carpetas