

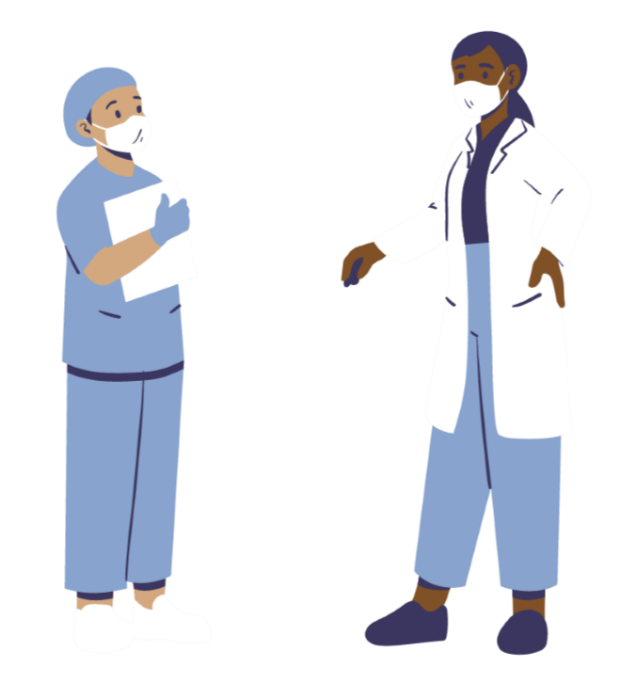
Primera entrega

TFM – COVID-19

INESDI

**Programa:** Máster en Business Intelligence &

Data Management (Online)



**Autores:** Grupo 6 COVID - Grupo I (Proyecto Inesdi)                                     **Tutor:** Pier Paolo Rossi

* Amaia Miranda Ulloa
* Fabián Ascheri Aguerre
* José Chavarría Montero
* Juan Carlos Valcuende Aláez
* Patricia Peña Torres

                    28 de marzo de 2023

# **INTRODUCCIÓN**

Necesidades de Nuestro Cliente, Objetivos del TFM, etc.

**A. DEFINICIÓN DEL PROYECTO Y ANÁLISIS DE LA VIABILIDAD.**

**I. DESCRIPCIÓN DETALLADA DE LA EMPRESA.**

Este proyecto fue propuesto por INESDI por lo cual no tiene vinculación con ninguna empresa.

Acotamos que nuestro target serán profesionales del mundo de la salud.

**II. ANÁLISIS INTERNO Y EXTERNO.**

**ANÁLISIS INTERNO**

Dentro del equipo hay varios miembros que ya tienen experiencia previa en ámbitos de la salud con Covid-19.

Un miembro del equipo es profesional de salud en activo por tanto conoce de primera mano las necesidades que tiene su colectivo al respecto de recibir información sobre el Covid-19 y tiene un punto de vista desde dentro de nuestro potencial cliente

Otro tiene experiencia en el ámbito administrativo hospitalario concretamente en el área de sistemas de información así que creemos que puede aportar una visión muy clara sobre cómo hemos de presentar esta información para que sea accesible e inteligible.

Una miembro del equipo tiene conocimiento de enfermedades respiratorias debido a su experiencia laboral en salud digital enfocada en el diagnóstico y monitorización de patologías respiratorias crónicas y agudas, acostumbra a trabajar de forma frecuente con información y estadísticas de patologías respiratorias incluyendo el Covid-19.

**PROPUESTA DE VALOR**

A continuación, se presenta el canvas de la propuesta de valor realizado por el equipo de trabajo. El mismo se encuentra disponible en el siguiente [mural](https://app.mural.co/t/whyhowwhat3132/m/whyhowwhat3132/1675966518106/d6ff6d825e2ffb42e26608abf7cad1d66b544702?sender=u0d0029efa7ccafce260a5840) en caso de que se requiera una revisión detallada del mismo (la calidad de la imagen no es la mejor).

Diagram

Description automatically generated with low confidence

**POST-MOTOROLA**

De la misma manera, se presenta el ejercicio post-motorola realizado por el equipo de trabajo. Al igual que el ejercicio anterior, el mismo se encuentra disponible en el siguiente [mural](https://app.mural.co/t/whyhowwhat3132/m/whyhowwhat3132/1675966518106/d6ff6d825e2ffb42e26608abf7cad1d66b544702?sender=u0d0029efa7ccafce260a5840) en caso de que se requiera una revisión detallada del mismo.

A picture containing chart

Description automatically generated

**Perspectivas (Insights)**

Respecto a la propuesta de valor

* Se realizó la propuesta de valor mediante la utilización del canvas, alineando los resultados del mapa de empatía realizado la semana pasada.
* El contexto, alineamiento y valor que brinda conocer los beneficios y las frustraciones de los potenciales usuarios antes de trabajar en una propuesta de valor es algo que nos impresionó. Especialmente al identificar los creadores de alegrías y los aliviadores de frustraciones como parte de la propuesta.
* A partir del entendimiento de lo que nuestro potencial cliente hace (trabajos del cliente), los beneficios que espera y las frustraciones que tiene, se propone un conjunto de dashboards que brinde información precisa (confiable), relevante, actual y de fácil y rápido acceso.
* Confiable
* Fuentes fidedignas y respetables, que gocen de buena reputación
* Quizá “sponsors” respetables que promuevan la solución
* Relevante
* Alineado con los intereses de nuestros usuarios
* Estadísticas de relación entre factores vs. efectos (por ejemplo, severidad de la enfermedad vs. % vacunación en el país)
* Benchmark (por ejemplo, resultados país vs. promedio global, regional o países con condiciones similares
* Actual
* Fuentes e información actual (por ejemplo, nuevas tendencias globales o datos relacionados a nuevos tratamientos)
* Fácil y Rápido Acceso
* Toda la información consolidada en formatos que faciliten el análisis, la comparación y la toma de decisiones por parte de los usuarios

Respecto al post-motorola

* Se concluye que el equipo ha logrado coordinar de manera exitosa todas las tareas y esfuerzos que se debían ejecutar a la fecha, debido principalmente a la apertura, al compromiso y a la buena voluntad de todos por colaborar.
* El equipo debe trabajar en una mejor definición de “rutinas” y herramientas de colaboración (hasta el momento se ha utilizado MS Teams). Esto debido a que, con el paso del tiempo, la complejidad de las tareas incrementará, por lo que será vital contar con una excelente coordinación y visibilidad de acuerdos, tareas y compromisos.
* Se agendó una sesión semanal, todos los Lunes a las 19h España para tratar todos los temas relacionados a entregables y avance del TFM
* Se ha trabajado muy bien mediante un esquema de pre-trabajo, en el que los miembros del equipo realizan los ejercicios de manera individual previo a la sesión de equipo, en la cual finalmente debatimos, tomamos decisiones y consolidamos.
* Se definirán las nuevas herramientas a utilizar para mejorar la coordinación de las tareas (por ejemplo, Trello, Slack, entre otras)
* Se llevará una minuta de nuestras sesiones en un formato sencillo y práctico (acuerdos, próximos pasos y responsables)
* El equipo ha aprendido que no debe apresurarse a ejecutar tareas sin pensar de manera estratégica primero, asegurándose de que todos los miembros entienden el resultado esperado y existe un alineamiento.

**TEAM ALIGNMENT MAP**

Timeline

Description automatically generated

El mismo puede seleccionando el siguiente [enlace](https://app.mural.co/t/whyhowwhat3132/m/whyhowwhat3132/1675966518106/d6ff6d825e2ffb42e26608abf7cad1d66b544702?wid=0-1677858879338).

**Insights**

Misión

* "Aspiramos a un mundo más sano y seguro, mediante la obtención, integración y análisis de información relacionada con COVID, de manera confiable y ágil para toma de decisiones. Ofrecemos un cuadro de mando sobre el COVID".

Línea de Tiempo

* Febrero 2023 - Diciembre 2023

Objetivos

* Aprobar el Trabajo Final del Máster
* Crear un cuadro de mando con información relevante, confiable y actual relacionada con COVID

Compromisos

* Agilidad: utilización de la metodología SCRUM
* Mantendremos al usuario en el centro (comprobación luego de cada sprint)
* Reuniones semanales, al menos una hora para conectar, siendo conscientes de que debemos ser flexibles (en algunas ocasiones no todos los miembros del equipo podrán participar)
* Comunicación: nuestro equipo contará con un ambiente seguro, en donde cada integrante podrá decir lo que piensa y como se siente, siempre dentro de un marco de respeto
* Hemos conversado sobre la figura del tutor y hemos hecho el primer contacto con el mismo (también por LinkedIn)

Recursos

* Herramientas del equipo:
* Comunicación: Teams (quizá lo reemplazaremos por Slack), Whatsapp
* Repositorio: Teams (lo reemplazaremos por OneDrive)
* Coordinación de Tareas: adoptaremos un gestor de tareas más visual, esto debido a que la versión gratuita de Teams cuenta con una herramienta muy básica (lo reemplazaremos por Asana o Trello)
* Colaboración: Mural
* Herramientas para el desarrollo del proyecto:
* Python
* MySQL
* PowerBI
* Tiempo de cada integrante del grupo

Riesgos

* Curva de aprendizaje respecto a herramientas tecnológicas para el desarrollo del proyecto
* Alcance del proyecto: hay una gran cantidad de datos alrededor de esta enfermedad, por lo que debemos definir muy bien el alcance del proyecto (acotarlo)
* Calidad y confiabilidad de la información (fuentes respetables, información comparable)
* Imprevistos relacionados con la situación personal de cada integrante del equipo (nuevo trabajo, muerte de algún familiar, vacaciones, etc.)

**ANÁLISIS EXTERNO**

Sólo en España, las patologías respiratorias han producido 260.000 ingresos hospitalarios en un año hasta 2018, los datos de 2020 indican que se ingresaron 528.554 mil personas por enfermedad respiratoria y 137.623 específicamente por Covid-19, esto sugiere que se han infraregistrado  los ingresos por covid-19. En el caso de Cataluña hablaríamos de 104.198 ingresos por enfermedades respiratorias en general y 31.114 por Covid-19 (INE, 2023).

La pandemia de COVID-19 es una enfermedad infecciosa causada por el virus SARS-CoV-2. Se cree que el virus se originó en Wuhan, China, a fines de 2019 y se propagó rápidamente por todo el mundo, llegando a ser declarada una pandemia por la Organización Mundial de la Salud en marzo de 2020 (OMS, 2023). Según la OMS la pandemia por Covid-19 ha dejado 674 millones de afectados en todo el mundo con el consiguiente colapso de los sistemas de emergencia y limitaciones en la realización de visitas presenciales, así mismo se contabilizan más de 6 millones de fallecidos. (Center for Systems Science and Engineering, 2023)

Los estudios señalan la necesidad de que los profesionales de la salud se mantenga actualizados al respecto de temas relativos al Covid-19 (Steinbach et al, 2020), así mismo se destaca la importancia de hacer estudios colaborativos de perspectiva internacional para un mejor abordaje de la pandemia. (Dolan & Mackey, 2020); otros artículos destacarán la importancia de que la información sea accesible, precisa y actualizada.

**III. EXPLICACIÓN DETALLADA DEL PROYECTO.**

Descripción de las bases de datos de las cuales se hizo el filtrado, habría que matizar la diferencia entre las que se han preseleccionado

Nos encontramos con que desde INESDI se nos facilitan múltiples carpetas, relativas al Covid-19, serán relativas a la situación mundial, en Estados Unidos, Corea del Sur y España con especial interés en la Comunidad de Madrid.

Finalmente hemos decididió que se utilizarás las bases de datos:

 Descripción de las bases de datos de las cuales se hizo el filtrado, habría que matizar la diferencia entre las que se han preseleccionado

Dashboards, si se quiere mantener la idea de la app probablemente habría que dejar el tema en los diagramas de flujo pero eso no entraría en el temario del máster así que creo que casi mejor omitirlo

**IV. FIJACIÓN DE LOS OBJETIVOS GENERALES Y ESPECÍFICOS DEL PROYECTO**

**Objetivos generales**

1. Análisis y explicación de variación en función de zona y época.
2. Realización de modelos predictivos
3. Desarrollo del storytelling mediante iteración y/o pivotación constante.

**Objetivos generales y específicos**

1. Análisis y explicación de variación en función de zona y época.
2. Integración, perfilación y modelado de base de datos.
3. Análisis exploratorio de datos de Covid-19
4. Realización de modelos predictivos
5. Regresión temporal para predecir siguientes olas de Covid-19.
6. Análisis de regresión logística múltiple para identificar las variables más representativas en cuanto a la mortalidad.
7. Análisis de regresión lineal múltiple para identificar las variables más representativas de las tasas de contagio.
8. Desarrollo del storytelling mediante iteración y/o pivotación constante.
9. Traducción estadística: adaptar los análisis a información inteligible por el usuario.
10. Visualización de datos en dashboards.
11. Pruebas piloto.
12. Pruebas de concepto.

**B. PLANIFICACIÓN:**

**I. ESTIMACIÓN DE LOS RECURSOS ECONÓMICOS.**

Debido a que es un proyecto de TFM no disponemos de recursos económicos para llevar a cabo el mismo y se usarían herramientas gratuitas o de las que dispusiéramos con anterioridad.

**II. ESTIMACIÓN DE LOS RECURSOS MATERIALES.**

Entre los recursos materiales contaríamos con que todos los integrantes disponemos de ordenadores y conexión a internet, así mismo se prevé la utilización de Python mediante Jupyter Notebooks y para SQL utilizar DBeaver, DB Browser o PhpMyAdmin.

**III. ESTIMACIÓN DE LOS RECURSOS HUMANOS.**  \*pendiente de acabar

Contamos con cinco integrantes, que serían:

Amaia Miranda, economista, con experiencia en seguimiento y monitoreo de proyectos.

Fabián Ascheri, enfermero con experiencia principalmente en asistencia al paciente.

Jose Chavarria, ingeniero industrial que actualmente se desempeña como director de cuentas por pagar.

Juan Carlos Valcuende, administrativo de sistemas de la información en el Hospital de Bellvitge con amplia experiencia en el ámbito de datos hospitalarios.

Patricia Peña, socióloga con experiencia en emprendimiento social en salud digital en patologías respiratorias y análisis de datos en el ámbito de la investigación.

**IV. ESTIMACIÓN DE LOS RECURSOS TIEMPO**

Se ha propuesto realizar dos reuniones semanalmente, también habrá trabajo autónomo y asincrónico que se distribuirá cada integrante en función de su disponibilidad.

**V. ELABORACIÓN DEL CRONOGRAMA DEL PROYECTO. \* se podría poner un enlace al archivo, al ser un cronograma es que quedaría estripado si no**

**VI. DEFINICIÓN DEL ALCANCE DEL PROYECTO.**

Entre los entregables estarían una serie de dashboards y sus respectivos resúmenes en los que se visualice de forma sencilla, atractiva y amigable la información recabada; para esto será importante también un trabajo de storytelling así como de conversar frecuentemente con el cliente para no perder el enfoque.

Así mismo el ideal sería que estos fuesen utilizados así que nos gustaría poder llevar a cabo las pruebas piloto y de concepto en alguna institución de salud.

**BIBLIOGRAFÍA**

Center for Systems Science and Engineering. (n.d.). COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. GitHub. <https://github.com/CSSEGISandData/COVID-19>

Dolan, B., & Mackey, T. K. (2020). COVID-19 and the Need for Action on Health Information. The Lancet Digital Health, 2, e188-e189. <https://doi.org/10.1016/S2589-7500(20)30084-4>

Gupta, A. K. (2020, 20 de mayo). The Importance of Data in the Fight Against COVID-19. Forbes. Recuperado de <https://www.forbes.com/sites/forbestechcouncil/2020/05/20/the-importance-of-data-in-the-fight-against-covid-19/?sh=497ccde22a9f>

Instituto Nacional de Estadística. (2023). Altas hospitalarias según el sexo, el diagnóstico principal, la provincia, Comunidad y Ciudad autónoma de hospitalización. Recuperado el 24 de febrero de 2023

Organización Mundial de la Salud (OMS). (2020, 11 de marzo). Rolling updates on coronavirus disease (COVID-19). Recuperado el 21 de marzo de 2023 de <https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON229>

Siu Hing Lo, C. P. Choy, & T. H. Lam. (2020). COVID-19 and the Need for Action on Health Information. Journal of Racial and Ethnic Health Disparities, 7, 937-941. <https://doi.org/10.1007/s40615-020-00849-y>

Steinbach, W. J., Scheetz, M. H., Postelnick, T. C., Reed, D. R., Pickering, K. S., Smith, A. M., & Landon, E. (2020). COVID-19: Knowledge is Power for Frontline Healthcare Workers. The Lancet, 8, e484-e485. <https://doi.org/10.1016/S2213-2600(20)30165-1>

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., ... & Niu, P. (2020). A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine, 382(8), 727-733.

# **Perfilado de Datos**

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Este fichero contiene número de mascarillas acumuladas diarias por Comunidad Autónoma.

Presenta las variables: fecha , cod\_ine , CCAA , total

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Este fichero contiene número de ingresados en UCI diarios por Comunidad Autónoma.

Presenta las variables: cod\_ine , CCAA , Públicos , Privados , Total

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Este fichero contiene número de hospitalizados diarios por Comunidad Autónoma.

Presenta las variables: fecha , cod\_ine , CCAA , total

Graphical user interface, text, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Este fichero contiene número de fallecidos diarios por Comundiad Autónoma.

Presenta las variables: fecha , cod\_ine , CCAA , total

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Este fichero contiene número de positivos diarios por Comunidad Autónoma.

Presenta las variables: fecha , cod\_ine , CCAA , total

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Este fichero contiene número de altas hospitalarias diarias por Comunidad Autónoma.

Presenta las variables: fecha , cod\_ine , CCAA , total

Graphical user interface, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Este fichero contiene número de camas hospitalarias por Comunidad Autónoma diferenciando por públicas y privadas.

Presenta las variables: fecha , cod\_ine , CCAA , mascarillas\_acumulado\_desde\_2020-03-10

Graphical user interface, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Este fichero contiene número de casos, pcr, test\_ac, altas, fallecimientos ingresos uci y hospitalizados diarios a nivel nacional.

Presenta las variables: fecha , casos\_total , casos\_pcr , casos\_test\_ac , altas , fallecimientos , ingresos\_uci , hospitalizados

Graphical user interface, application

Description automatically generated

# Graphical user interface, application Description automatically generated

Este fichero contiene número de casos confirmados, hospitalizados, ingresos en uci y fallecidos diarios por sexo y rango de edad.

Presenta las variables: fecha , rango\_edad , sexo , casos\_confirmados , hospitalizados , ingresos\_uci , fallecidos