
INSTITUTO TECNOLÓGICO SUPERIOR DE TEZIUTLÁN

Tesis



“Análisis de información de contacto estrecho en la propagación del virus SARS-COV-2 por medio de inteligencia artificial”

PRESENTA:

ANA ZAIRA OCHOA OSORIO

CON NÚMERO DE CONTROL
20TE0003P

PARA OBTENER EL GRADO ACADÉMICO DE:
MAESTRÍA EN SISTEMAS COMPUTACIONALES

CLAVE DEL PROGRAMA ACADÉMICO
MPSCO 2011-15

DIRECTOR (A) DE TESIS:
MTRO. EMMANUEL VÁZQUEZ BENITO

TEZIUTLÁN, PUEBLA, DICIEMBRE 2022

“La Juventud de hoy, Tecnología del Mañana”

AGRADECIMIENTOS

A DIOS:

Por ser mi guía y darme la hasta el día de hoy. También le doy gracias por darme la fortuna de ser mamá de una niña especial, cuya luz me ha permitido valorar y ver el mundo desde una perspectiva distinta.

A MI ESPOSO:

Por su cariño y apoyo incondicional, durante todo este proceso, por estar conmigo en las buenas y en las malas y mostrarme que todo es posible, solo implica dedicación y paciencia.

A MI FAMILIA:

A mis padres y hermanos, en especial a mi madre la Sra. Lourdes Beley Osorio Nava, quien me apoyo y alentó para continuar, cuando parecía que me iba a rendir.

A MIS ASESORES:

Por su constante apoyo hasta la culminación exitosa de este trabajo, en especial al Mtro. Emmanuel Vázquez Benito colaborador durante todo este proceso, quien, con su dirección, conocimiento, enseñanza y colaboración permitió el desarrollo de este trabajo.

ÍNDICE GENERAL

AGRADECIMIENTOS	2
ÍNDICE GENERAL	3
ÍNDICE DE FIGURAS	5
ÍNDICE DE TABLAS	6
CARTAS DE ACEPTACIÓN	7
RESUMEN	8
INTRODUCCIÓN.....	9
CAPITULO I GENERALIDADES DEL PROYECTO.....	10
1.1 MARCO TEÓRICO.....	10
1.1.1 Transmisión del virus SARS-CoV-2	10
1.1.2 Inteligencia artificial.....	11
1.1.3 Big Data.....	12
1.1.4 Sistema de información.....	13
1.1.5 Dashboard	13
1.1.6 Lenguaje de Programación R.....	13
1.1.7 R Studio.....	14
1.1.8 Paquetes de R.....	14
1.1.9 Paquete shiny.....	14
1.1.10 Paquete shinydashboard	15
1.1.11 Paquete forecast.....	15
1.1.12 Paquete tseries	15
1.1.13 Modelo ARIMA.....	16
1.1.14 Series Temporales estacionarias	16
1.1.15 ARIMA en Análisis de Series Temporales	17
1.1.16 Muestra ACF y PACF.....	18
1.1.17 Georreferenciación.....	19
1.1.18 Web Service	20
1.1.19 API RESTful	21

1.2	Planteamiento del Problema.....	21
1.3	Justificación	24
1.4	Hipótesis	25
1.5	Objetivo general	26
1.6	Objetivos específicos.....	26
1.7	Alcances y limitaciones.....	27
	CAPITULO II ESTADO DEL ARTE	27
2.1	Trabajos relacionados	27
2.2	Análisis comparativo de los trabajos relacionados	32
	CAPITULO III METODOLOGÍA Y DESARROLLO	35
3.1	Metodología de la investigación	35
3.2	Metodología de desarrollo	39
3.2.1	Identificación de actores y casos de uso.....	43
	CAPITULO IV RESULTADOS	66
4.1	Análisis de datos.....	66
4.2	Selección de pruebas estadísticas	72
4.3	Validación del modelo desarrollado	72
4.4	Comprobación de la hipótesis	75
	CAPITULO V CONCLUSIONES.....	75
	Referencias.....	77

ÍNDICE DE FIGURAS

Ilustración 1. Modelo	41
Ilustración 2. Diagrama de Casos de Uso (2022, Creación propia)	43
Ilustración 3. Caso de Uso. Inicio (2022, Creación propia)	44
Ilustración 4. Caso de Uso. Geolocalización (2022, Creación propia)	45
Ilustración 5. Caso de Uso. Data (2022, Creación propia).....	46
Ilustración 6. Caso de Uso. Gráficas (2022, Creación propia).....	47
Ilustración 7. Caso de Uso. pronósticos ARIMA (2022, Creación propia).....	48
Ilustración 8. Caso de Uso. API (2022, Creación propia)	49
Ilustración 9. Diseño Conceptual (2022, Creación propia)	50
Ilustración 10. Diseño Navegacional (2022, Creación propia)	51
Ilustración 11. Barra de opciones (2022, Creación propia).....	52
Ilustración 12. P-Service Covid (2022, Creación propia)	52
Ilustración 13. Geolocalización (2022, Creación propia).....	53
Ilustración 14. Data (2022, Creación propia)	53
Ilustración 15. Data por Categoría (2022, Creación propia)	54
Ilustración 16. Temperatura Corporal (2022, Creación propia).....	54
Ilustración 17. Temperatura Corporal Categoría (2022, Creación propia)	55
Ilustración 18. Pronóstico Gestión Empresarial (2022, Creación propia).	55
Ilustración 19. Pronóstico Industrias Alimentarias (2022, Creación propia)	56
Ilustración 20. Pronóstico Sistemas Computacionales (2022, Creación propia).....	56
Ilustración 21. Pronóstico Ingeniería Industrial (2022, Creación propia)	57
Ilustración 22. Pronóstico Ingeniería en Informática (2022, Creación propia)	57
Ilustración 23. Pronóstico Ingeniería Mecatrónica (2022, Creación propia)	58
Ilustración 24. Descargar (2022, Creación propia)	58
Ilustración 25. Módulo: P-Service Covid	59
Ilustración 26. Módulo: Geolocalización-Mapa	59
Ilustración 27. Módulo: Data-DataSet	60
Ilustración 28. Módulo: Data-DataSet por Categoría	60
Ilustración 29. Módulo: Gráficas- Temperatura Corporal	61
Ilustración 30. Módulo: Gráficas – Temperatura Corporal por Categoría	61
Ilustración 31. Módulo: Pronóstico ARIMA –Gestión Empresarial.....	62
Ilustración 32. Módulo: Pronóstico ARIMA –Industrias Alimentarias	62
Ilustración 33. Módulo: Pronóstico ARIMA – Sistemas Computacionales	63
Ilustración 34. Módulo: Pronóstico ARIMA – Ingeniería Industrial	63
Ilustración 35. Módulo: Pronóstico ARIMA – Ingeniería Informática	64
Ilustración 36. Módulo: Pronóstico ARIMA – Ingeniería Mecatrónica	64
Ilustración 37. Módulo: API – Descargas.....	65

Ilustración 38. Pronóstico- Gestión Empresarial	67
Ilustración 39. Pronóstico- Ingeniería en Industrias Alimentarias	68
Ilustración 40. Pronóstico - Ingeniería en Sistemas Computacionales.....	69
Ilustración 41. Pronóstico - Ingeniería Industrial.....	69
Ilustración 42. Pronóstico - Ingeniería en Informática	70
Ilustración 43. Pronóstico - Ingeniería en Mecatrónica.	71
Ilustración 44. Prueba Dickey-Fuller.....	73
Ilustración 45. Prueba PACF.....	73
Ilustración 46. Prueba ACF.....	74
Ilustración 47. Prueba Ljung-Box.....	75

ÍNDICE DE TABLAS

Tabla 1. Análisis Comparativo.	35
Tabla 2. Conjunto de datos	38
Tabla 3. Personas involucradas en la investigación	40
Tabla 4. Recursos necesarios para la investigación, desarrollo e implementación de la investigación y solución tecnológica propuesta.	41
Tabla 5. Entregables.....	41
Tabla 6. Cronograma de actividades	42
Tabla 7. Conjunto de datos por Carrera	66
Tabla 8. Registro por Carrera	67
Tabla 9. Pronóstico- Gestión Empresarial	68
Tabla 10. Pronóstico- Ingeniería en Industrias Alimentarias	68
Tabla 11. Pronóstico - Ingeniería en Sistemas Computacionales	69
Tabla 12. Pronóstico - Ingeniería Industrial	70
Tabla 13. Pronóstico - Ingeniería en Informática	70
Tabla 14. Pronóstico - Ingeniería en Mecatrónica.	71
Tabla 15. Muestra Mensual.	72

CARTAS DE ACEPTACIÓN

RESUMEN

Los sistemas de información juegan un papel importante en la gestión de datos y la obtención de información necesaria para efectuar análisis y toma de decisiones que promuevan cambios de políticas o generación de directrices, entre otras posibilidades. Tecnologías vanguardistas como la inteligencia artificial permiten almacenar, clasificar, extraer e interpretar grandes cantidades de información haciendo uso de algoritmos y técnicas como el aprendizaje automatizado o aprendizaje profundo entre otros. En el contexto actual de pandemia por COVID-19, se abren posibilidades de emplear estas herramientas para efectuar medidas en torno a la reducción del riesgo de contagio en la población.

La pandemia ha provocado que el sector público y privado modifiquen sus políticas de acceso, por lo que no es extraño encontrar a personal que esté tomando la temperatura corporal de las personas en las entradas esto como medida de prevención, en las empresas hay personal que recorre los lugares de trabajo tomando la temperatura corporal, esta información solo es utilizada en el momento para permitir o denegar el acceso a un cierto sitio.

Este proyecto tiene como propósito recopilar la información georreferenciada de personas identificadas como sospechosas de portar el virus, con apoyo de un dispositivo portátil para la toma de temperatura corporal. En caso de detectar niveles de fiebre, se aplicará una evaluación consistente en una serie de preguntas simples para determinar si se necesita atención médica o aislamiento. Los datos recolectados servirán para la generación de mapas de calor que varíen en periodos cortos de tiempo, al mismo tiempo que serán publicados en una serie de servicios de tipo API Restful en formato JSON, para consumo o empleo de otras aplicaciones o investigaciones, así mismo incluye el desarrollo de un dashboard que permite el análisis de información de contacto estrecho en la propagación del virus SARS-COV-2 por medio de inteligencia artificial.

INTRODUCCIÓN

El presente trabajo de tesis tiene como principal objetivo identificar un problema mundial tal como la detección oportuna de posibles casos de Covid-19, a raíz de la pandemia el ser humano estuvo obligado a protegerse y proteger a los demás mediante medidas de prevención.

En el capítulo I se abordan las cuestiones teóricas que sustentan el planteamiento. La justificación del problema, la necesidad de realizar la investigación frente al problema elegido, tanto como los objetivos que se desean lograr y la hipótesis que se valida al final de la investigación.

En el capítulo II se muestra como el problema de contagios por coronavirus y la detección oportuna no es un problema que solo radica en nuestro país, sino al contrario, es un problema global, es un tema que no tiene que ver con una sola raza, un género o alguna posición económica, al contrario, ninguna persona queda obsoleta ante este problema, debido a que la detección oportuna de contagios puede coadyuvar a prevenir posibles contagios.

En el capítulo III se encuentran los antecedentes históricos del problema presentado, en este apartado se puede ver como el problema a resolver ha iniciado con investigación, desde tiempo atrás, el Gobierno Federal de la República Mexicana, a través de las dependencias como la Secretaría de Salud (SS), ha publicado y dispuesto al público una serie de herramientas informáticas con información sobre los distintos casos detectados de contagios por COVID 19.

En el último capítulo de la investigación, teniendo el problema y a los sujetos de investigación, y después de haber capturado los datos con referencia a temperaturas corporales se realizó una aplicación web para el análisis de datos mediante el lenguaje R aplicando series temporales en ARIMA, dicha aplicación tiene como objetivo la detección oportuna de posibles casos de covid-19.

CAPITULO I GENERALIDADES DEL PROYECTO

1.1 MARCO TEÓRICO

1.1.1 Transmisión del virus SARS-CoV-2

Acorde a los datos disponibles por parte de la Organización Mundial de la Salud (OMS, 2021), la transmisión del virus responsable de la enfermedad de coronavirus, el SARS-CoV-2, mayormente conocido como COVID-19, es posible gracias a lo que se denomina contacto estrecho de una persona infectada con aquellas a su alrededor. A rasgos generales, este fenómeno se establece cuando una persona ha estado sin mascarilla o cubrebocas por más de quince minutos, ya sea durante el uso de transporte público o privado, en toda actividad laboral, comercial o cualquiera que implique cercanía con personas realizadas en espacios cerrados con poca o mala ventilación, que las mascarillas no sean las adecuadas y que no se sigan medidas de higiene o protección recomendadas (Quirónprevención.com, 2020).

La propagación del virus SARS-CoV-2 se puede dar desde la nariz o la boca de una persona contagiada por medio de la tos, los estornudos, la respiración intensa o a través del habla; el contacto estrecho entre individuos puede dar pauta a la inhalación o la inoculación del virus ya sea por la boca, la nariz o los ojos o el contacto de estos con la piel, en caso de haber tocado superficies u objetos de trabajo, denominados fómites, que se caracterizan porque tienen presencia de virus viables (o del mismo COVID-19), sea por estar en espacios de trabajo o con personas afectadas; además de estos casos, la transmisión puede ocurrir comúnmente en espacios cerrados, abarrotados de gente y mal ventilados (OMS, 2021).

En lo tocante a los entornos de trabajo, ya sea con alta presencia de personal o de afluencia de público en general, clientes, entre otros, la transmisión del virus de COVID-19 se manifiesta más en entornos o ambientes de trabajo compartido, es decir, donde existe un tiempo prolongado entre personas en estrecha proximidad, por desplazamientos y viajes relacionados con el trabajo y, en menor medida (pero no menos importante), las reuniones sociales de los trabajadores fuera del área

laboral. Ante dichos panoramas, los espacios de trabajo han empleado una serie de estrategias para reducir los riesgos de contagio como lo son la desinfección periódica de espacios, directivas de teletrabajo, ingreso restringido sólo a personal clave, pautas sobre distanciamiento físico, pruebas de cribado regulares, registro y seguimiento de contactos, vigilancia del medio ambiente y el uso de equipos de protección personal (OMS, 2021).

Si bien los contagios de coronavirus pueden deberse a múltiples factores en el medio ambiente, los empleadores pueden tomar medidas y evaluación de riesgos ante contagio, en conjunto con los representantes de los servicios de salud, implementando orientaciones específicas que permitan reducir dicha incidencia lo más posible. Acorde al número de personas y la naturaleza de las tareas desempeñadas, los lugares pueden clasificarse en zonas de bajo, medio o alto riesgo. En cualquier caso, las medidas esenciales recomendadas por la OMS son una correcta ventilación en los espacios de trabajo, medidas para promover la higiene respiratoria y de manos, el distanciamiento físico, la reducción y gestión de viajes o traslados relacionados con la actividad laboral y la desinfección constate del entorno. (OMS, 2021). De manera adicional, se efectúan las tomas de temperatura al personal al momento de su ingreso.

1.1.2 Inteligencia artificial

Según (Donato Impedovo, 2019) La inteligencia artificial está transformando la industria de la salud desde muchas perspectivas, comenzando con el diagnóstico, el tratamiento y seguimiento de enfermedades. En la literatura existe una amplia gama de técnicas y aplicaciones en el cuidado de la Salud, por ejemplo: descubrimiento de nuevos medicamentos, sistemas expertos de diagnóstico, reconocimiento de patrones en la medicina, etc. Además, (Loebbecke, 2019) afirma que en una era de aceleración de la digitalización y análisis avanzado de big data, el aprovechamiento

de datos e información de calidad permitirá métodos de investigación y enfoques de gestión innovadores.

La inteligencia artificial está fuertemente ligada al análisis de datos, actualmente los algoritmos son capaces de almacenar, clasificar, extraer e interpretar grandes cantidades de datos de información haciendo uso de algoritmos y técnicas como el aprendizaje automatizado o aprendizaje profundo entre otros. Las cookies, por ejemplo, están ocultas en los sitios web, y recopilan datos personales como hábitos de navegación en internet, intereses sociales, preferencias de compras, etc. Dichos datos ayudan al Marketing predictivo el cual ayuda a predecir cuándo y dónde se realizarán una compra.

En la actualidad la IA cuenta con diferentes ramas, el aprendizaje automático o Machine Learning y el aprendizaje profundo o Deep Learning. El Deep Learning es una de las técnicas del Machine Learning enfocada en estructuras básicas del cerebro humano, las neuronas. Este planteamiento emula como actúa el cerebro. Dichos sistemas, son capaces de realizar trabajos de ordenación o predicción tras un entrenamiento precedente, de manera que hace que sea un aprendizaje supervisado.

1.1.3 Big Data

Hoy en día, Big Data se define no solo por la cantidad de información, sino también por la variabilidad, complejidad y velocidad con la que se analiza. De acuerdo con el Instituto Nacional de Estándares y Tecnología (National Institute of Standards and Technology, NIST) considera que Big Data es un conjunto de datos extensos, principalmente en las características de volumen, velocidad y/o variabilidad, que requieren una arquitectura escalable para un almacenamiento, manipulación y análisis eficientes” (Willie May, 2015). Según (Aguilar, 2013) Gartner define Big Data como grandes conjuntos de datos que tienen tres características principales: volumen (cantidad), velocidad (velocidad de creación y utilización) y variedad (tipos

de fuentes de datos no estructurados, tales como la interacción social, video, audio, cualquier cosa que se pueda clasificar en una base de datos).

1.1.4 Sistema de información

Actualmente existen muchas definiciones para los sistemas de información y con el paso del tiempo se han modificado a las nuevas tendencias, De acuerdo con Volpentesta los sistemas de información tienen como objetivo apoyar a las actividades de cualquier organización principalmente operaciones de administración, toma de decisiones y control. Asimismo, los sistemas operan sobre un conjunto de datos estructurados necesarios para implementar estrategias. (Volpentesta, 2004).

1.1.5 Dashboard

Dashboard o también conocido como tablero es una aplicación de Business Intelligence que permite a las organizaciones visualizar la información más importante de manera grafica para poder controlar, analizar y administrar el desempeño de una empresa. Por otra parte, los dashboard son diseñados frecuentemente para muestran claves de desempeño y control de objetivos.

De acuerdo con (Eckerson, 2006) El objetivo principal del dashboard es ofrecer visualización sobre indicadores mediante graficas visuales los cuales deben de alcanzar un determinado valor para lograr los objetivos propuestos.

1.1.6 Lenguaje de Programación R

R es un lenguaje de programación que está orientado al análisis y presentación de información es, de distribución libre, bajo Licencia GNU, y se mantiene para el cómputo estadístico y gráfico. De igual manera, su sintaxis es muy similar a otros lenguajes de programación como C o Java, la visualización de gráficos ha

evolucionado a través de los años y que actualmente trabajar con librerías como `ploty`, `dygraphs`, `ggplot`, entre otros. Este proyecto se escribió en el código R por medio de un IDE o entorno de desarrollo de RStudio.

1.1.7 R Studio

RStudio es un entorno de desarrollo integrado gratuito y de código abierto para R, un lenguaje de programación para computación estadística y gráficos. Actualmente, está disponible en dos ediciones: RStudio Desktop, donde el programa se ejecuta localmente como una aplicación de escritorio normal; y RStudio Server, que permite acceder a RStudio mediante un navegador web mientras se ejecuta en un servidor Linux remoto. Las distribuciones preempaquetadas de RStudio Desktop están disponibles para Windows, macOS y Linux.

RStudio está parcialmente escrito en el lenguaje de programación C++ y utiliza el marco Qt para su interfaz gráfica de usuario.

1.1.8 Paquetes de R

Los paquetes en el lenguaje de programación R son un conjunto de funciones R, código compilado y datos de muestra. Estos se almacenan en un directorio llamado "biblioteca" dentro del entorno R. La instalación de los paquetes se realiza de forma gráfica o por medio de consola. A continuación, se listan los paquetes más importantes que se utilizaron para el proyecto.

1.1.9 Paquete shiny

Shiny es un marco para crear aplicaciones web usando código R. Está diseñado principalmente pensando en los científicos de datos y, con ese fin, puede crear aplicaciones Shiny bastante complicadas sin conocimientos de HTML, CSS o JavaScript (Wickham, 2022). Una aplicación Shiny se compone de una interfaz de usuario `ui` que controla el diseño y la apariencia de la aplicación, y un `server()` función

que contiene las instrucciones para construir los objetos que se muestran en la interfaz de usuario. De esta forma, las aplicaciones permiten la interacción del usuario mediante una funcionalidad llamada reactividad, lo cual permite una mejor exportación de los datos.

1.1.10 Paquete shinydashboard

El paquete shinydashboard añade una capa de abstracción que permite crear aplicaciones organizadas y estéticamente agradables mediante código CSS o HTML, el paquete ofrece un diseño de tres partes: un encabezado(header), una barra lateral (siderbar) y un cuerpo para el tablero (body).

1.1.11 Paquete forecast

El pronóstico del paquete R proporciona métodos y herramientas para mostrar y analizar pronósticos de series temporales univariadas, incluido el suavizado exponencial a través de modelos de espacio de estado y el modelado automático ARIMA (Rob Hyndman, s.f.). También posee la función auto arima que es un método de modelado automático.

1.1.12 Paquete tseries

La función ts() convertirá un vector numérico en un objeto de serie temporal R. El formato es ts(vector , inicio=x, fin=y, frecuencia=z) donde x =fecha de inicio y x=fecha de fin de la última observación y z= frecuencia de observaciones por unidad, es decir; 1=anual, 4=trimestral, 12= mensual, etc.

1.1.13 Modelo ARIMA

En 1976 George Box y Gwilym Jenkins desarrollaron el modelo estadístico ARIMA para pronosticar rangos de datos basados en entradas de una serie temporal el cual tiene dependencia entre datos, es decir, cada observación en un momento dado es modelada en función de los valores anteriores, los cuales se basan en un modelo explícito. Los modelos se denominan con el nombre de ARIMA (AutoRegresive Integrated Moving Average).

Fases de la metodología de Box y Jenkins:

La primera fase identificar: Implementar logaritmos y diferencias para convertir la serie observada en una serie estacionaria. Es decir; encontramos los valores apropiados para p , d y q .

La segunda fase estimación y pronostico: Los parámetros AR y MA del modelo se estiman por máxima verosimilitud y se obtienen sus errores estándar y los residuos del modelo. Es decir; el modelo se elige en base al error más bajo.

La tercera fase diagnóstico: Se comprueba que los residuos no tienen estructura de dependencia y siguen un proceso de ruido blanco. Es decir; seleccionamos el modelo que se ajusta a los datos.

La cuarta fase predicción: Se genera un pronóstico de la serie temporal.

1.1.14 Series Temporales estacionarias

De acuerdo con (Casimiro, 2009) se debe de tener en cuenta el comportamiento estacional, es decir; implica que la observación de un mes y observación del mismo mes del año anterior tienen una pauta de comportamiento similar por qué estarán temporalmente correlacionadas. A continuación, presentamos los modelos de series temporales lineales estacionarias más utilizados: los modelos de media móvil

integrada autorregresiva (ARIMA). Estos modelos han asumido una gran importancia en el modelado de procesos del mundo real.

1.1.15 ARIMA en Análisis de Series Temporales

De acuerdo con (Robert H. Shumway, 2011) ARIMA es una forma de análisis de regresión que indica la fuerza de una variable dependiente en relación con otras variables cambiantes. El objetivo final del modelo es predecir el movimiento futuro de la serie temporal examinando las diferencias entre los valores de la serie en lugar de los valores reales. Los modelos ARIMA se aplican en los casos en que los datos muestran evidencia de no estacionariedad. En el análisis de series de tiempo, los datos no estacionarios siempre se transforman en datos estacionarios.

Las causas comunes de datos no estacionarios en series de tiempo son la tendencia y los componentes estacionales. La forma de transformar datos no estacionarios en estacionarios es aplicar el paso de diferenciación. Es posible aplicar uno o más pasos de diferenciación para eliminar el componente de tendencia en los datos. De manera similar, para eliminar los componentes estacionales de los datos, se podría aplicar la diferenciación estacional.

De acuerdo con (Casimiro, 2009), podemos dividir el modelo en componentes más pequeños de la siguiente manera:

AR: un modelo **autorregresivo** que representa un tipo de proceso aleatorio. La salida del modelo depende linealmente de su propio valor anterior, es decir, una cierta cantidad de puntos de datos retrasados o la cantidad de observaciones anteriores

MA: un modelo de **media móvil** cuya salida depende linealmente de las observaciones actuales y pasadas de un término estocástico.

I: integrado aquí significa el **paso de diferenciación** para generar datos de series de tiempo estacionarios, es decir, eliminar los componentes estacionales y de tendencia

El modelo ARIMA generalmente se denota como ARIMA (p, d, q) y los parámetros p, d, q se definen de la siguiente manera:

- p: el orden de retraso o la cantidad de tiempo de retraso del modelo autorregresivo AR(p).
- d: grado de diferenciación o número de veces que los datos se han restado con valor pasado.
- q: el orden del modelo de media móvil MA(q)

1.1.16 Muestra ACF y PACF

Según (Robert H. Shumway, 2011) define ACF y PACF como:

ACF: La **función de coeficiente de autocorrelación**, define cómo los puntos de datos en una serie de tiempo se relacionan con los puntos de datos precedentes.

PACF: La **función de coeficiente de autocorrelación parcial**, al igual que la función de autocorrelación, transmite información vital sobre la estructura de dependencia de un proceso estacionario.

Dadas las observaciones de una serie de tiempo, un enfoque para el ajuste de un modelo a los datos es hacer coincidir la muestra ACF y la muestra PACF de los datos con la muestra ACF y PACF del modelo, respectivamente. Podemos usar la gráfica ACF de muestra para ver si los datos específicos de la serie de tiempo son estacionarios o no.

Para datos de series de tiempo estacionarios, el ACF de muestra se extinguirá muy rápidamente o incluso se cortará. De lo contrario, la serie temporal no es

estacionaria. El corte aquí significa que el valor de ACF es menor que el intervalo de confianza indicado o está dentro de la línea de puntos azul.

Si confirmamos que los datos son estacionarios, podemos decidir qué parámetros q y p deben usarse para el modelo según el retraso en el corte del valor ACF y PACF de la muestra. En este caso, el parámetro p se decidirá por el tiempo de corte del PACF de la muestra y el parámetro q se decidirá por el tiempo de corte del ACF de la muestra. Si sabemos que la serie temporal no es estacionaria, intentaremos realizar la diferenciación de orden superior para garantizar que los datos sean estacionarios.

1.1.17 Georreferenciación

Según (Sommer, 2006) Define la georreferenciación como alinear datos geográficos con un sistema de coordenadas conocido para que puedan verse, consultarse y analizarse con otros datos geográfico. La georreferenciación permite establecer la posición de un elemento en un sistema de coordenadas, existen dos sistemas de coordenadas: el origen y el destino. Este procedimiento es determinado con una relación de posiciones entre elementos espaciales en ambos sistemas, por ello, conociendo la posición en uno de los sistemas de coordenadas es factible obtener la posición homóloga en el otro sistema.

La georreferenciación se utiliza comúnmente en herramientas de información geográfica (SIG) para asociar datos numéricos, datos vectoriales, datos ráster o redes lineales.

La georreferenciación se define por una función matemática del tipo:

$$X = f(x, y)$$

$$Y = f(x, y)$$

Donde la posición de una entidad geográfica en el sistema de coordenadas destino (X, Y) es función de las coordenadas (x, y) que tiene ese elemento en el sistema origen. En este proyecto, se hará uso de la georreferenciación de las personas a las que se les toma la temperatura y se encuentran con fiebre, esto permitirá ver a lo largo del tiempo y con ayuda de mapas de calor, cómo se mueven los contagios.

1.1.18 Web Service

Un servicio web es un conjunto de protocolos y estándares abiertos que permiten el intercambio de datos entre diferentes aplicaciones o sistemas. Los servicios web pueden ser utilizados por programas de software escritos en diferentes lenguajes de programación y que se ejecutan en una variedad de plataformas para intercambiar datos a través de redes informáticas como Internet.

El Protocolo simple de acceso a objetos (SOAP) y la Transferencia de estado representacional (REST) son, con mucho, las opciones más utilizadas para acceder a servicios web. Asimismo, cualquier aplicación en la nube que emplee protocolos web estandarizados (HTTP o HTTPS) para conectarse, interoperar e intercambiar mensajes de datos, comúnmente XML (lenguaje de marcado extensible), a través de Internet se considera un servicio web.

Sin embargo, los dos tipos de servicio web mas utilizados son los siguientes:

SOAP: Es un protocolo que se diseño antes que REST, la idea principal es garantizar que los programas creados en diferentes plataformas y lenguajes de programación logran intercambiar datos de forma segura.

REST: Implementa principalmente componentes multimedia, archivos, REST implementa acciones CRUD (crear, leer, actualizar y eliminar).

Asimismo, una de las principales ventajas de los servicios web es la construcción de

bloques que facilitan la reutilización de componentes de servicios web en otros servicios.

1.1.19 API RESTful

Una API de REST o API de RESTful, es una interfaz de programación de aplicaciones (API o API web) que permite el intercambio de información de forma segura, asimismo, las API son abstractas: el software que proporciona el API generalmente es llamado la implementación de esa API. Si bien REST se puede usar sobre cualquier protocolo, generalmente se aplica en HTTP. Esto significa que los desarrolladores no necesitan instalar bibliotecas ni software para aprovechar un diseño de API REST.

Asimismo, para obtener los datos, se debe enviar una solicitud estructurada a la API. Si la solicitud cumple con los requisitos deseados, se envía una respuesta que contiene los datos al lugar donde se realizó la solicitud. Esta respuesta generalmente viene en forma de datos JSON o XML.

1.2 Planteamiento del Problema

En el año 2022, se vive una situación difícil para todo el mundo, sectores como la economía y la salud se han visto gravemente afectados, la convivencia social también ha tenido cambios, lo cual nos lleva a tomar medidas nuevas para poder conservar la salud. A finales del año 2019 en Wuhan, China, se presentó un nuevo

virus perteneciente a la familia de los Coronavirus el cual trae consigo muchos síntomas perjudiciales para la salud de los seres humanos, mismos que se dividen en 3 niveles: Síntomas habituales: fiebre, tos seca y cansancio, entre los síntomas menos comunes: Molestias y dolores, dolor de manos, dolor de garganta, diarrea, conjuntivitis, dolor de cabeza, pérdida del sentido del olfato o del gusto, erupciones cutáneas o pérdida del color en los dedos de las o de los pies; síntomas graves: dificultad para respirar o sensación de falta de aire, dolor o presión en el pecho, incapacidad para hablar o moverse.

Con el fin de coadyuvar con la disminución de contagios por COVID-19 el Gobierno Federal, a través de las dependencias como la secretaria de Salud (SS), implementaron una herramienta informática llamada "Semáforo COVID-19", el cual muestra un mapa del país únicamente con los estados en sus diferentes alertas y una sección con información adicional, el cual no cuenta con la posibilidad de consultar municipios en específico.

Asimismo, (CONACyT, 2022) cuenta con una plataforma que brinda información variada de casos COVID-19 a nivel nacional, la cual se muestra en mapas interactivos, gráficas, tablas, entre otros. La información publicada está confirmada por medio de los datos provenientes de la Dirección General de Epidemiología (DGE), en la sección documental Datos Abiertos. No ofrece, por su parte, información específica por Estado, región o municipio, aunque sus datos se actualizan con la misma frecuencia que el Visualizador Analítico del INEGI.

Además, y ante la tendencia tecnológica de los sistemas de información de vanguardia, el no contar con un registro, seguimiento o notificación en áreas específicas o localidades en tiempo real sobre la presencia de casos detectados, confirmados o sospechosos de contraer coronavirus, supone un sesgo en los datos oficiales registrados, imprecisiones sobre la evolución sobre el número real de casos detectados y si éstos se reducen o aumentan, dificultades para establecer

estrategias o acciones concretas que resulten efectivas para disminuir contagios, reestablecer actividades socio-económicas o brindar campañas de vacunación con el alcance suficiente. Otro problema que se puede derivar de lo anterior es el hecho de que la población en general, dependiente de actividades administrativas, laborales o educativas, por mencionar algunas, estaría en riesgo latente y constante de entrar en contacto con sitios o regiones donde la presencia de coronavirus es más elevada, no contando con alternativas de desplazamiento, suministro o atención para sus necesidades u obligaciones.

Hoy en día, el Covid-19 o SARS CoV-2 ha llegado a modificar las medidas sanitarias de todo el mundo, es por ello por lo que actualmente se han tomado medidas llamadas "La Nueva Normalidad". Una estrategia común para prevenir la transmisión ha sido establecer controles de ingreso-egreso los cuales permite ayudar a los lineamientos de Seguridad Sanitaria en la Nueva Normalidad, donde antes de ingresar a un establecimiento se estipula que debe haber una distancia mínima de un metro y medio entre cada individuo, con el fin de disminuir la frecuencia de contacto, esto conlleva a tener un control de ingreso-egreso de las personas trabajadoras, clientes, visitantes, donde el primer filtro consta del monitoreo de temperatura corporal empleando cámaras termográficas o termómetros donde no se debe ser mayor a 37.5 °C (es decir, fiebre), uso obligatorio de cubrebocas, gel antibacterial o alguna sustancia sanitizante y tapetes sanitizantes o túneles de sanitización.

Según (Carl T. Berdahl, 2021) Los Centros para el Control y Prevención de Enfermedades (CDC) informaron que el 96% de los pacientes con infección sintomática por COVID-19 tienen fiebre, tos o dificultad para respirar. Por otra parte, en la investigación titulada "Body temperature correlates with mortality in COVID-19 patients" (Serena Tharakan, 2020) se analizó la temperatura corporal de 9417 pacientes anónimos con sospecha de COVID-19 localizados en Mount Sinai y hospitales afiliados en el área de New York, los resultados indicaron que sólo la

mitad de los pacientes con virus SARS-CoV2 presento temperatura corporal de 37° en la fase inicial de la enfermedad y el 78,5% de los pacientes desarrollo temperatura corporal > 37° durante el curso de la enfermedad. Según (Carl T. Berdahl, 2021) 1 de cada 3 pacientes que alcanzo una tempera corporal alta por encima de los 39.5° murió.

Si bien los contagios de coronavirus pueden deberse a múltiples factores en el medio ambiente, los empleadores pueden tomar medidas y evaluación de riesgos ante contagio, en conjunto con los representantes de los servicios de salud, implementando orientaciones específicas que permitan reducir dicha incidencia lo más posible. Acorde al número de personas y la naturaleza de las tareas desempeñadas, los lugares pueden clasificarse en zonas de bajo, medio o alto riesgo. En cualquier caso, las medidas esenciales recomendadas por la OMS son una correcta ventilación en los espacios de trabajo, medidas para promover la higiene respiratoria y de manos, el distanciamiento físico, la reducción y gestión de viajes o traslados relacionados con la actividad laboral y la desinfección constate del entorno. (OMS, 2021).

1.3 Justificación

Mejorar la detección de la infección aguda por COVID-19 es fundamental para reducir la propagación de la infección por coronavirus. Sin embargo, la capacidad de diagnosticar COVID-19 a menudo se ve limitada debido a una infección asintomática o levemente asintomática; tiempo de respuesta en las pruebas, negligencia por parte de la ciudadanía por no atenderse, desinterés por el cuidado de la salud. Dadas estas limitaciones, mejorar la estimación de la probabilidad previa a la prueba es una

forma crucial de mejorar la toma de decisiones sobre cuándo se justifica la prueba y, por lo tanto, frenar la propagación de la infección aguda en localidades, municipios o estados.

En la actualidad los Centros para el Control y Prevención de Enfermedades (CDC) no cuentan con un registro, seguimiento o notificación en áreas específicas o localidades en tiempo real sobre la presencia de casos detectados, confirmados o sospechosos de contraer coronavirus.

En este sentido, en el presente proyecto se propone recopilar información georreferenciada de las personas identificadas como sospechosas de portar el virus con apoyo de un dispositivo portátil para la toma de temperatura corporal. En caso de detectar niveles de fiebre, se aplicará una evaluación consistente en una serie de preguntas simples para determinar si se necesita atención médica o aislamiento. Los datos recolectados servirán para la generación de mapas de calor que varíen en periodos cortos de tiempo, al mismo tiempo que serán publicados en una serie de servicios de tipo API Restful en formato JSON, para consumo o empleo de otras aplicaciones o investigaciones, así mismo incluye el desarrollo de un dashboard que permite el análisis de información de contacto estrecho en la propagación del virus SARS-COV-2 por medio de inteligencia artificial lo anterior, se pretende proporcionar a los especialistas de la salud una herramienta tecnológica que sirva de apoyo para la prevención de la propagación del COVID-19.

Con lo anterior, se pretende identificar a personas con posibles síntomas de la enfermedad de coronavirus, ayudar al diagnóstico temprano y coadyuvar con la disminución de contagios en municipios y, por consiguiente, proporcionar a los especialistas de la salud una herramienta tecnológica.

1.4 Hipótesis

Variables

- Estudiantes, visitas, personal docente y administrativos del Tecnológico Superior de Teziutlán
- Aplicación web para identificar posibles contagios de acuerdo con la información recolectada por los sensores de temperatura.

Pregunta de hipótesis descriptiva

¿Con el desarrollo de la aplicación web se identificarán patrones que permitan obtener predicciones sobre los posibles contagios a corto plazo, en el Instituto Tecnológico Superior de Teziutlán?

Verbo para cuantificar

- Identificar los patrones que permitan obtener predicciones sobre los posibles contagios a corto plazo, en el Instituto Tecnológico Superior de Teziutlán.

1.5 Objetivo general

Desarrollar una aplicación web mediante programación en lenguaje R para el análisis de información de contacto estrecho en la propagación del virus SARS-COV-2 por medio de la inteligencia artificial.

1.6 Objetivos específicos

- Analizar la información de la propagación del virus SARS-COV2 por de contacto estrecho utilizando inteligencia artificial
- Recopilar información georreferenciada de las personas identificadas como sospechosas de portar el virus en el periodo enero-octubre 2022.

- Análisis de datos en base a la metodología Box y Jenkins.
- Diseñar y desarrollar la solución tecnológica propuesta.
- Implementar el modelo ARIMA en el desarrollo de la aplicación tecnológica.
- Generación de resultado de la solución diseñada.

1.7 Alcances y limitaciones

A continuación, se presentan los alcances de la presente investigación:

- En contraste con los desarrollos actuales, el presente proyecto busca brindar información detallada y georreferenciada de posibles casos COVID 19 y no generalizada como las aplicaciones actuales, para identificar de manera puntual posibles puntos de infección, así como la propagación de los virus provocados por el contacto estrecho. Toda la información recolectada será publicada con el fin de aportar información a futuras investigaciones.
- La información será recolectada diariamente para tener una mejor precisión en el análisis de los datos.

Posteriormente, se da mención a las limitaciones:

- Debido al presupuesto no se contará con un certificado SSL

CAPITULO II ESTADO DEL ARTE

En esta sección se presenta una revisión de aquellos trabajos relacionados con el tema, con una pequeña descripción, técnicas empleadas y objetivos en cada uno de los trabajos.

2.1 Trabajos relacionados

Debido al impacto del contagio del Covid-19 a nivel mundial, diferentes países han implementado sistemas de información y políticas gubernamentales para enfrentarla. Una de las soluciones más utilizadas es la gestión de datos utilizando

tableros, la cual ayuda a tomar decisiones basadas en evidencia, reduce la mortalidad y ayuda a administrar la atención médica para zonas vulnerables.

El Gobierno de Puebla en colaboración con Google y Apple implementaron la aplicación móvil "Alerta COVID Puebla" para que los ciudadanos que la activen sean informados cuando se detecte que estuvieron en contacto con alguna persona que ha dado positivo a una prueba COVID (Martínez, 2021) , el propósito de la aplicación es minimizar la cadena de contagios y dar atención a los ciudadanos confirmados por Covid-19.

La aplicación es gratuita y está disponible en iOS y Android, con relación a la funcionalidad, Primero, el ciudadano debe de activar notificaciones de exposición COVID en el dispositivo móvil, elegir la región Puebla, México y activar el Bluetooth, posteriormente se intercambiarán códigos aleatorios entre dispositivos móviles, con datos como la fecha, intensidad de la señal y la duración del acercamiento los cuales no incluyen ni la identidad ni datos personales (Martínez, 2021).

Por consiguiente, en caso de que algún ciudadano de positivo a una prueba de COVID-19, puede informar sobre su condición, y las personas que estuvieron cerca en los últimos 14 días recibirán una alerta (Martínez, 2021), esta notificación permitirá tomar medidas de prevención como hacerse una prueba de COVID-19, quedarse en casa, monitorear sus síntomas y evitar el contacto con personas.

Asimismo, a principios del mes de marzo del 2020 La universidad Nacional Autónoma de México (UNAM) en con la colaboración de especialistas del Centro de Investigaciones en Geografía Ambiental (CIGA, campus Morelia), de las escuelas nacionales de Estudios Superiores (ENES) unidades Morelia y Mérida, además del Laboratorio Nacional de Análisis y Síntesis Ecológica (LANASE) implementaron "La plataforma de información geográfica de la UNAM sobre COVID-19 en México", la cual muestra indicadores de tendencia de casos, estos se pueden visualizar de forma interactiva a través de mapas que muestran casos positivos, defunciones y hospitalizaciones desglosada por municipio, estado (Adrian Ghilardi, 2020). El

objetivo de la plataforma es comprender la epidemia en evolución, coadyuvar en la dispersión, detectar poblaciones vulnerables y mitigar la cadena de contagios.

La plataforma esta desarrollado en múltiples lenguajes de programación como R, Python, Bash, ArcGIS. Con respecto al funcionamiento, primero el sistema descarga los datos sobre incidencia de COVID desde la plataforma de la secretaria de Salud, posteriormente son incorporados a la tabla de atributos de los archivos municipales y estatales del INEGI, más tarde se calcula cada variable (Adrian Ghilardi, 2020). Cabe destacar que la plataforma realiza monitoreos de casos activos por código postal y colonias, lo cual permite monitorear áreas rurales, de especial vulnerabilidad, y coadyuvar en la administración de la atención médica.

Por otro lado, En (Sava, 2020) el departamento de IA de Google Cloud se asoció con Harvard Global Health Institute para desarrollar un modelo de pronósticos públicos de COVID-19, este modelo emplea aprendizaje automático (ML), técnicas estadísticas basadas en análisis de series temporales e Inteligencia Artificial con epidemiología, en relación a las proyecciones se consultan métricas de casos confirmados, muertes, recuperaciones y hospitalizaciones por los próximos 14 días en cada uno de los condados de EE. UU. Además, todos los datos están disponibles gratuitamente y se pueden consultar en BigQuery o descargar como CVS mediante el panel de control de Sata Studio.

Del mismo modo Harvard Global Health Institute y Google Cloud ha estado trabajando para mejorar del modelo de pronósticos públicos de COVID-19, cuando se lanzó inicialmente los pronósticos incluían predicciones de 14 días y ahora incluyen predicciones para 28 días además se mejoró la precisión del modelo entrenándolo con simulaciones , también se agregaron pronósticos para otros países, métricas personalizadas y análisis hipotético para la toma de decisiones (Pfister, 2020). Por lo tanto, el propósito del modelo es apoyar al sector salud y a los condados de EE. UU con información actualizada, precisa y concisa mediante el modelo SEIR (susceptible, expuesto, infectado, recuperado) con ello permite

agregar variables adicionales tales como infectados indocumentados, indocumentados recuperados, hospitalizados, ventiladores, disponibilidad de recursos hospitalarios, datos demográficos, niveles de pobreza y riqueza, etc. Dado al nivel de detalle del modelo se detectan las zonas vulnerables en los condados de EE. UU.

En el trabajo de investigación "*Geographic Information Systems and COVID-19: The Johns Hopkins University Dashboard*" (João, 2020) el desarrollo de un sistema de Información Geográfica (SIG), Según Chang el análisis geográfico es la principal forma de los Sistemas de información Geográfica, como mencionó (Chang, 2014), ya que los SIG pueden almacenar, manipular y mostrar datos geoespaciales en sistemas informáticos. El sistema recopila información de confirmados y muertes de casi 200 fuentes, incluida la OMS y otras agencias de salud internacionales y locales e informes a nivel estatal y de condado en los EE. UU. Después de recopilar, editar y referenciar los datos a un sistema de coordenadas diseñado, el sistema pone a disposición los datos. Para el desarrollo del tablero (mapas y datos) se utilizó la tecnología basada en la web ArcGIS (nube), se ocupó el análisis predictivo de los brotes de sarampión en los EE. UU, los datos iniciales se recopilaron de varias fuentes incluida la OMS (Organización Mundial de la Salud), CDC en los EE. UU., CDC en China, ECDC (Centro Europeo para la Prevención y el Control de Enfermedades), NHC y DXY. DXY es un sitio web chino que agrega informes de situación del NHC y el CCDC local casi en tiempo real, lo que proporciona estimaciones de casos regionales más actuales que las que pueden proporcionar las organizaciones nacionales de informe.

Por otro lado, El Instituto Nacional de Estadística Geografía e informática (INEGI, 2022) diseño un visualizador Analítico ofrece un mapa de la República Mexicana en el que se detallan casos confirmados de COVID-19 tanto por estado como por municipio, permite también con la posibilidad de consultar ubicaciones de cementerios, hospitales exclusivos para tratar la enfermedad, farmacias, entre otros.

Es posible consultar un resumen por municipio con las estadísticas más recientes y permite la descarga de los mismos. El total de casos confirmados de COVID-19 se muestra en el mapa por medio de una escala de colores diferente al empleado por el Semáforo de la SS, y su información se actualiza diariamente, o al menos con un día de demora. No obstante, no es posible consultar la información de las localidades de cada municipio.

Asimismo, Gobierno Federal de la República Mexicana, a través de las dependencias como la Secretaría de Salud (SS), el Instituto Nacional de Estadística y Geografía (INEGI) y el Consejo Nacional de Ciencia y Tecnología (CONACYT), ha publicado y dispuesto al público una serie de herramientas informáticas con información sobre los distintos casos detectados de contagios por COVID 19, cada una con datos específicos por zonas o municipios, disposición de hospitales, entre otros. Por parte de la (SS, 2022), el Semáforo COVID-19 muestra información sobre las distintas alertas con una serie de colores e indicaciones sobre las condiciones de cada región. Así mismo, muestra un mapa del país únicamente con los estados en sus diferentes alertas y una sección con información adicional. Dicho semáforo establece un periodo de dos semanas en que dicho estatus no se actualizará. No cuenta con la posibilidad de consultar regiones más pequeñas o municipios en específico

Finalmente, el CONACyT (2022) cuenta con una plataforma que brinda información variada de casos COVID-19 a nivel nacional, la cual se muestra en mapas interactivos, gráficas, tablas, entre otros. La información publicada está confirmada por medio de los datos provenientes de la Dirección General de Epidemiología (DGE), en la sección documental Datos Abiertos. No ofrece, por su parte, información específica por Estado, región o municipio, aunque sus datos se actualizan con la misma frecuencia que el Visualizador Analítico del INEGI.

2.2 Análisis comparativo de los trabajos relacionados

La Tabla 1 muestra el análisis comparativo de los trabajos de investigación analizados en el apartado anterior con base en su objetivo y las técnicas de mejoramiento de imágenes utilizadas.

Número	Autor	Trabajo	Objetivo	Tipo
1	(Adrian Ghilardi, 2020)	Alerta COVID Puebla	Notificar a los ciudadanos cuando se detecte que estuvieron en contacto con alguna persona que ha dado positivo a una prueba COVID.	Aplicación móvil
2	(Adrian Ghilardi, 2020)	Plataforma de información geográfica de la UNAM sobre COVID-19 en México	Comprender la evolución de la epidemia, coadyuvar en la dispersión, detectar poblaciones vulnerables y mitigar la cadena de contagios.	Aplicación web

	(Sava, 2020)	Modelo de pronósticos públicos de COVID-19	Detectar zonas vulnerables en los condados de EE. UU.	Aplicación web
4	(Pfister, 2020)	Modelo de pronósticos públicos de COVID-19	Apoyar al sector salud y a los condados de EE. UU con información actualizada, precisa y concisa mediante el modelo SEIR.	Aplicación web
5	(João, 2020)	<i>"Geographic Information Systems and COVID-19: The Johns Hopkins University Dashboard"</i>	Visualizar y rastrear los casos informados en tiempo real.	Aplicación web
6	(INEGI, 2022)	Visualizador Analítico	Visualizar casos confirmados por coronavirus y	Aplicación web

			consultar ubicaciones de cementerios, hospitales exclusivos para tratar la enfermedad, farmacias.	
7	(CONACyT, 2022)	Visualizador Analítico	Visualizar casos detectados de contagios por COVID 19, cada una con datos específicos por zonas o municipios, disposición de hospitales, entre otros.	Aplicación web
8	(SS, 2022)	Plataforma de información	Apoya al sector público con información actualizada, mediante una serie de	Aplicación web

			indicaciones sobre las condiciones de cada región.	
9	(CONACyT, 2022)	Plataforma de información	Brinda información variada de casos COVID-19 a nivel nacional	Aplicación web
10	(INEGI, 2022)	Visualizador analítico	Brinda información de casos confirmados.	Aplicación web

Tabla 1. Análisis Comparativo.

CAPITULO III METOLOGÍA Y DESARROLLO

3.1 Metodología de la investigación

La presente investigación es de tipo aplicada para (Murillo 2008), busca resolver una duda en específico y brinda soluciones vanguardistas a problemas que afectan a un individuo, grupo o sociedad. Asimismo, implementa y sintetiza la práctica basada en investigación.

La investigación aplicada en presente proyecto se representan las fases de la metodología de desarrollo a utilizar en este caso OOHDM las cuales son: Obtención de Requerimientos, por medio de casos de uso, diseño conceptual, diseño

navegacional de la aplicación y finalmente el diseño de interfaces abstractas que son las vistas finales desarrolladas.

Asimilismo, la investigación se define con un enfoque cuantitativo ya que este alcance se utiliza la recolección y el análisis de datos para contestar preguntas de investigación y probar hipótesis establecidas previamente y confía en la medición, conteo y el uso de estadísticas para intentar establecer con exactitud patrones de investigación (Gómez, 2006).

Universo de Investigación:

El universo de investigación hace referencia a los grupos conformados de personas, dentro de estos grupos se encuentra el que será analizado para la investigación, en esta ocasión se tomo como universo al Instituto Tecnológico Superior de Teziutlán, la cual cuenta con diferentes carreras, Ingeniera en Gestión Empresarial, Ingeniería en Industrias Alimentarias, Ingeniería en Sistemas Computacionales, Ingeniería Industrial, Ingeniería en Informática e Ingeniería en Mecatrónica, con alrededor de 2,914, se eligió a los alumnos que ingresaban al Instituto en un horario de 8:00 a 9:00 de la mañana.

Lugar, población y tiempo (fecha):

Estudiantes del Instituto Tecnológico Superior de Teziutlán ubicado en Aire Libre, Teziutlán, Puebla. Horario: 8 a 9 am, Duración 2 meses.

Instrumento de recolección de datos:

Un paso importante en el proceso de investigación es la recolección de datos, esta técnica coadyuva al investigador para recolecta obtener un panorama completo del fenómeno y/o población que se desea investigar. (Rodríguez, 2020)

El instrumento es un recurso funcional para hacer la recolección de datos que se necesitan para poder desarrollar la investigación.

Para esta Investigación el Instrumento de investigación es la recolección de datos de los termómetros digitales que se encuentran en los accesos al Instituto Tecnológico Superior de Teziutlán, cabe destacar que los datos capturados son los siguientes: Fecha, Carrera, Sexo, Oxigenación, Temperatura Corporal y Geolocalización.

A continuación, se muestra el conjunto de datos de los entrevistados, capturados en Microsoft Excel:

Fecha	Hora	CARRERAS						Sexo		Temperatura Corporal
		Ing. en Ind. Alimentarias.	Ing. Industrial.	Ing. en Sistemas Computacionales	Ing. Mecatrónica.	Ing. Informática.	Ing. Gest. Empresaria	F	M	
		14/10/2022	08:00				X			
14/10/2022	08:00				X				X	36.5
14/10/2022	08:00			X					X	36.5
14/10/2022	08:00		X					X		36.3
14/10/2022	08:00		X					X		36.3
14/10/2022	08:00		X					X		35.9
14/10/2022	08:01		X					X		36.7
14/10/2022	08:01		X					X		37.1
14/10/2022	08:01		X					X		37.1
14/10/2022	08:01						X	X		37
14/10/2022	08:02			X				X		36.3
14/10/2022	08:02						X	X		36.5
14/10/2022	08:02						X	X		36.5
14/10/2022	08:02		X						X	36.5
14/10/2022	08:03		X						X	36.5
14/10/2022	08:03						X		X	35.5
14/10/2022	08:03						X	X		36.4
14/10/2022	08:03				X				X	36.7
14/10/2022	08:03						X		X	35.3
14/10/2022	08:03		X						X	35
14/10/2022	08:03		X					X		36.5
14/10/2022	08:03						X		X	36.5
14/10/2022	08:03						X		X	37.1
14/10/2022	08:03						X		X	36.7
14/10/2022	08:03				X				X	36.4
14/10/2022	08:04						X		X	36.3
14/10/2022	08:04		X				X		X	36.4
14/10/2022	08:04		X						X	34.7
14/10/2022	08:04		X					X		36.8
14/10/2022	08:04		X	X				X		36.5
14/10/2022	08:04			X					X	36.5
14/10/2022	08:04			X				X		37.2
14/10/2022	08:04	X							X	36.5
14/10/2022	08:04	X						X		36.5
14/10/2022	08:04	X						X		36.5
14/10/2022	08:04				X				X	36.3
14/10/2022	08:04				X				X	36.5
14/10/2022	08:04			X				X		35.9
14/10/2022	08:04		X						X	36.4
14/10/2022	08:05			X				X		36.4
14/10/2022	08:05			X				X		36.4
14/10/2022	08:05			X					X	36.3
14/10/2022	08:05				X				X	36
14/10/2022	08:05		X						X	36.3
14/10/2022	08:05		X					X		36.4
14/10/2022	08:05		X					X	X	36
14/10/2022	08:05		X					X		36.9

Tabla 2. Conjunto de datos

3.2 Metodología de desarrollo

Daniel Schwabe y Gustavo Rossi proponen la metodología de desarrollo Object Oriented Hypermedia Design Model- OOHM, dicho modelo tiene como propósito la elaboración de aplicaciones multimedia y al mismo tiempo facilitar y hacer más eficaz el diseño de aplicaciones multimedia. Sin embargo, la metodología define pautas de trabajo, basado en el diseño, con el fin de desarrollar aplicaciones multimedia de forma metodológica.

Por ello, OOHM es una de las metodologías mas utilizadas en el presente debido a que logra disminuir los tiempos de desarrollo, reutilizar diseño, simplificar la evolución y el mantenimiento de la aplicación. Una de las características fundamentales es que permite la separación del diseño con respecto al desarrollo.

La Figura 1 muestra la metodología OHHM la cual plantea el diseño de una aplicación a través de cinco fases que se desarrollan de un modo iterativo. Estas fases son:



Figura 1. Etapas de la Metodología OOHM

Fuente: Elaboración Propia

- Obtención de Requerimientos: La primera fase se identifica el comportamiento del sistema, es decir: se identifican los roles que podrían cumplir los usuarios y tareas que deberá de realizar la aplicación.

- **Diseño Conceptual:** Se construye teniendo en cuenta los requisitos funcionales capturados con los casos de uso. Se utilizan técnicas tradicionales orientadas a objetos para construir el modelo conceptual, como encontrar clases, definir estructuras de herencia y especificar restricciones. Se representa mediante un diagrama de clases.
- **Diseño Navegacional:** En esta parte la estructura de navegación de una aplicación hipertexto está definida por un esquema de clases de navegación específica, este define la estructura de los nodos y enlaces de la aplicación hipertexto.
- **Diseño de Interfaz Abstracta:** El modelo de interfaz abstracto se construye definiendo objetos en términos de clases de interfaz. Estas clases se definen como objetos de interfaz (como campos de texto, botones, etc.) u otras clases de interfaz como multimedia.
- **Implementación:** Una vez elaborado el modelo conceptual, el modelo de navegación y el modelo de interfaz abstracta, se implementa un lenguaje de programación para obtener la implementación ejecutable de la aplicación.

Recursos humanos:

Función	Nombre	Grado
Investigador	Ana Zaira Osorio	ISC
Director de tesis (primario)	Emmanuel Vázquez Benito	MSC

Tabla 3. Personas involucradas en la investigación

Material y equipo:

Laptop
R studio

Tabla 4. Recursos necesarios para la investigación, desarrollo e implementación de la investigación y solución tecnológica propuesta.

Entregables:

Los productos se realizarán de acuerdo con la metodología OOHM

Identificación de actores y casos de uso
Diagramas de caso de uso
Diagrama de secuencia
Diagrama de clase
Prototipos de la interfaz de usuario
Aplicación Web

Tabla 5. Entregables

Modelo:

A continuación, se presenta el modelo del desarrollo:

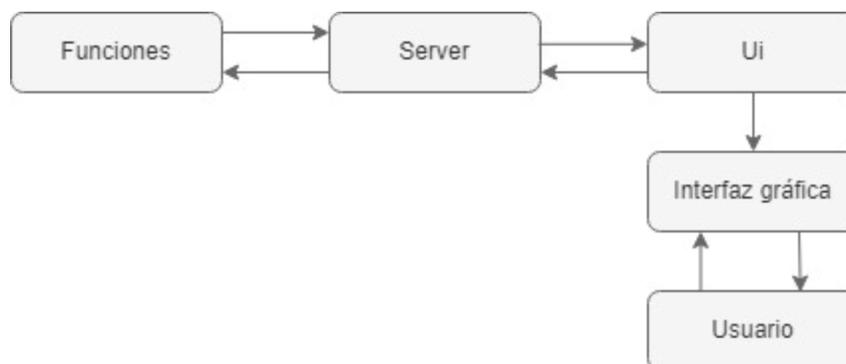


Ilustración 1. Modelo

Cronograma:

Actividades	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Analizar la información de la propagación del virus SARS-COV2 por de contacto estrecho utilizando inteligencia artificial.					
Recopilar información georreferenciada de las personas identificadas como sospechosas de portar el virus en el periodo enero-octubre 2022.					
Analizar la información de la propagación del virus SARS-COV2 por de contacto estrecho utilizando inteligencia artificial.					
Análisis de datos en base a la metodología Box y Jenkins.					
Diseñar y desarrollar la solución tecnológica propuesta.					
Implementar el modelo ARIMA en el desarrollo de la aplicación tecnológica.					
Generación de resultado de la solución diseñada.					

Tabla 6. Cronograma de actividades

3.2.1 Identificación de actores y casos de uso

A continuación, se mostrarán las fases que señala la metodología OOADM, iniciando por los casos de uso, posteriormente se mostrará el diseño conceptual, que se realiza con diagramas de clase, seguido a esto el diseño navegacional, interfaces abstractas y por último la implementación.

Diagrama de casos de uso

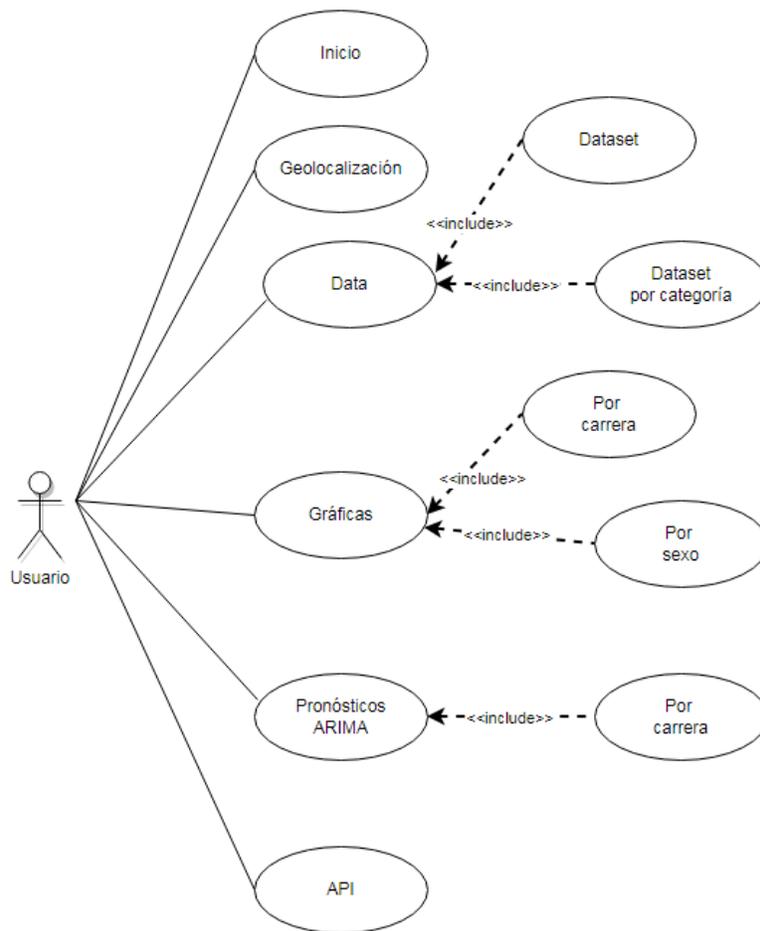


Ilustración 2. Diagrama de Casos de Uso (2022, Creación propia)

Caso de uso, Inicio

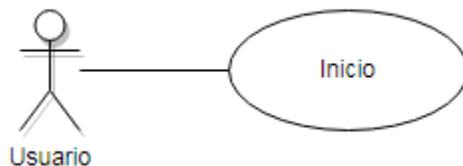


Ilustración 3. Caso de Uso. Inicio (2022, Creación propia)

Caso de uso	Inicio
Resumen	El usuario ingresa la URL en su navegador y se muestra la página de inicio.
Actores	Usuario
Precondición	Ninguna
Descripción	Flujo normal: <ol style="list-style-type: none">1. Abrir navegador2. Ingresar a URL del sistema
PostCondición	El usuario se encuentra en la página del sistema.

Caso de uso, Geolocalización



*Ilustración 4. Caso de Uso. Geolocalización
(2022, Creación propia)*

Caso de uso	Geolocalización
Resumen	Este caso de uso muestra las áreas de riesgo basadas en la información almacenada y procesada.
Actores	Usuario
Precondición	El usuario está en la página del sistema.
Descripción	Flujo normal: 3. Se selecciona la opción: Geolocalización
PostCondición	El usuario se encuentra en la sección de visualizar geolocalización.

Caso de uso, Data

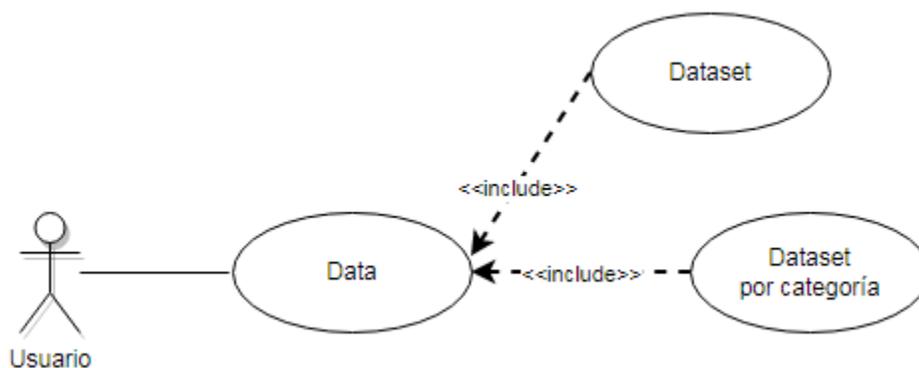


Ilustración 5. Caso de Uso. Data (2022, Creación propia)

Caso de uso	Data
Resumen	Éste caso de uso muestra el conjunto de datos base a partir del cual se procesará la información.
Actores	Usuario
Precondición	El usuario está en la página del sistema.
Descripción	Flujo normal: 1. Se selecciona la opción: Data
PostCondición	El usuario visualiza la información que contiene el sistema para su procesamiento.

Caso de uso, Gráficas

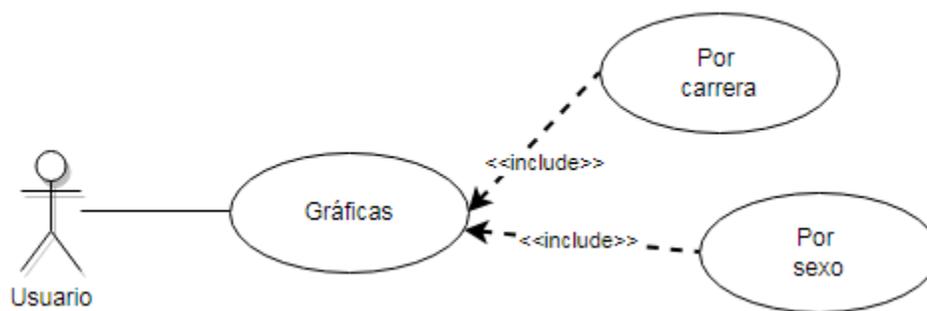


Ilustración 6. Caso de Uso. Gráficas (2022, Creación propia)

Caso de uso	Gráficas
Resumen	Este caso de uso muestra la información de temperaturas en formato de gráficas para facilitar su comprensión y comparativa.
Actores	Usuario
Precondición	El usuario está en la página del sistema.
Descripción	Flujo normal: 1. Se selecciona la opción: Gráficas
PostCondición	El usuario visualiza las gráficas correspondientes a su selección conforme a la información almacenada y procesada por el sistema.

Caso de uso, pronósticos ARIMA

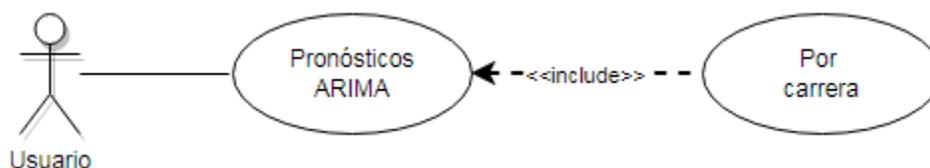


Ilustración 7. Caso de Uso. pronósticos ARIMA (2022, Creación propia)

Caso de uso	Pronósticos ARIMA
Resumen	Éste caso de uso muestra el pronóstico en base a la selección de una carrera, en un intervalo de tiempo incluyendo límites superiores e inferiores.
Actores	Usuario
Precondición	El usuario está en la página del sistema.
Descripción	Flujo normal <ol style="list-style-type: none">1. Se selecciona la opción: Pronóstico ARIMA2. Selecciona la carrera a visualizar el pronóstico.
PostCondición	El usuario visualiza la información arrojada por el pronóstico de una carrera seleccionada en un intervalo de tiempo.

Caso de uso, API

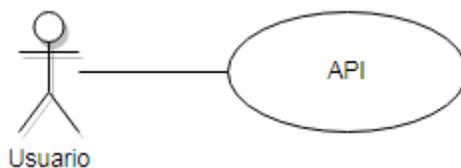


Ilustración 8. Caso de Uso. API (2022, Creación propia)

Caso de uso	API
Resumen	Éste caso de uso permite descargar la información contenida en la página
Actores	Usuario
Precondición	El usuario está en la página del sistema.
Descripción	Flujo normal 1. Se selecciona la opción: API
PostCondición	El usuario posee una copia en formato Json de la información descargada.

Diagrama Conceptual:

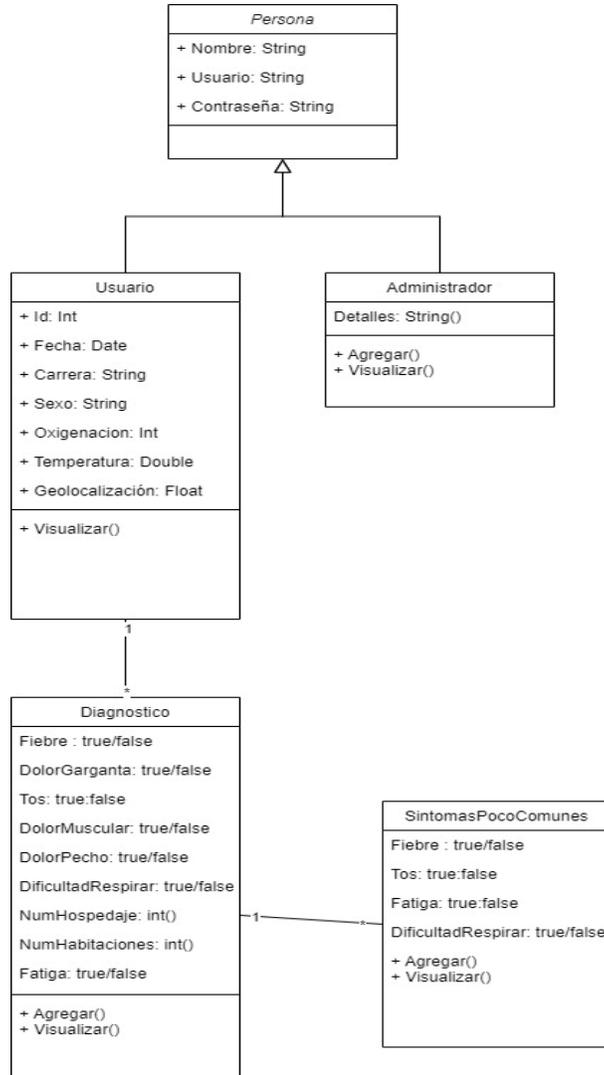


Ilustración 9. Diseño Conceptual (2022, Creación propia)

En la Ilustración 9 Diseño conceptual se muestra como es la relación de clases a través del diseño realizado.

Diseño Navegacional:

Un modelo de navegación se construye como una vista sobre un modelo conceptual, lo que permite la construcción de diferentes modelos de acuerdo con los perfiles. A continuación, se presenta la navegabilidad para el perfil usuario:



Ilustración 10. Diseño Navegacional (2022, Creación propia)

Diseño de interfaces abstractas

En la presente sección utilizaremos el enfoque de diseño de vista de datos abstracta (ADV) para describir la interfaz de una aplicación web. Los ADV son objetos en el sentido de tener un estado y una interfaz, donde la interfaz puede ser ejecutada a través de mensajes (en particular, eventos externos generados por el usuario.)

A continuación, mostraremos la interfaz de nuestro sitio web:

Diseño de interfaces abstractas – Barra de navegación

La vista de la “barra de navegación” muestra el elemento de navegación que se sitúa en la parte izquierda del sitio web, éste contiene enlaces para los usuarios.

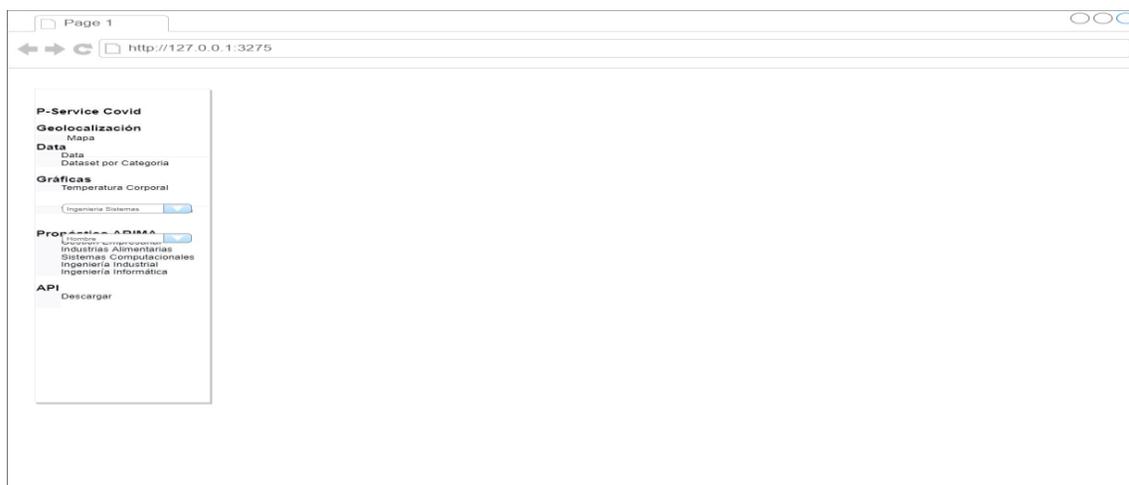


Ilustración 11. Barra de opciones (2022, Creación propia)

Diseño de interfaces abstractas – P Service Covid

La vista de la “barra de navegación” presenta a la página principal que se visualiza al cargar la plataforma, esta vista contiene una imagen e información del proyecto.

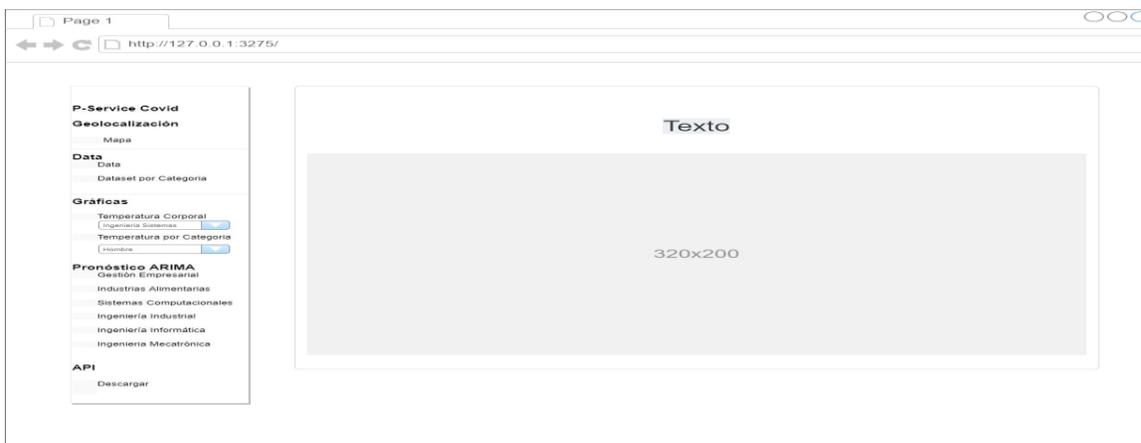


Ilustración 12. P-Service Covid (2022, Creación propia)

Diseño de interfaces abstractas – Geolocalización

La vista “Geolocalización” presenta la localización de los puntos de control de ingreso y egreso.

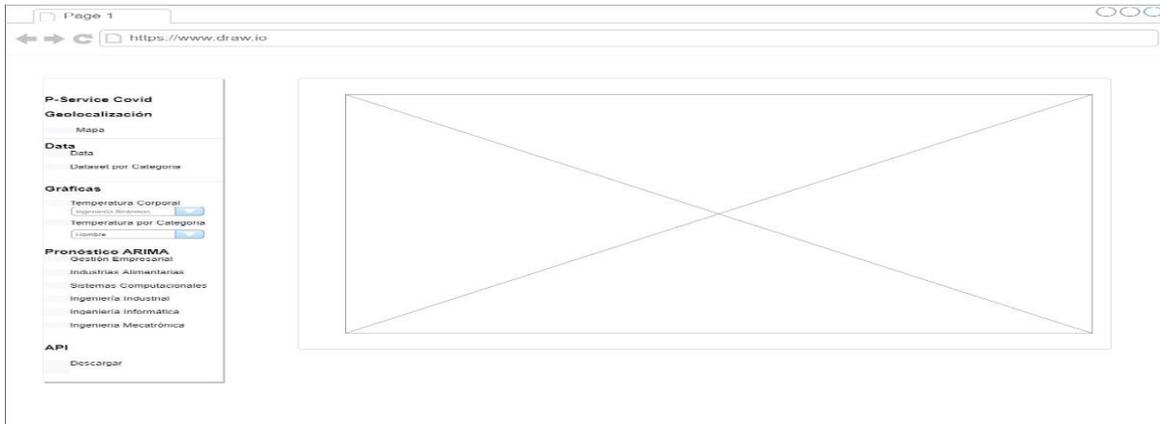


Ilustración 13. Geolocalización (2022, Creación propia)

Diseño de interfaces abstractas – Data

La vista de la “Data” presenta la localización de los puntos de control de ingreso y egreso.



Ilustración 14. Data (2022, Creación propia)

Diseño de interfaces abstractas – Data por Categoría

La vista de la “Data por Categoría” presenta el conjunto de datos capturados en los puntos de control de ingreso y egreso.



Ilustración 15. Data por Categoría (2022, Creación propia)

Diseño de interfaces abstractas – Temperatura Corporal

La vista de la “Temperatura Corporal” contiene una gráfica de burbuja donde se analiza de la temperatura corporal por Carrera en el periodo enero – octubre 2022.

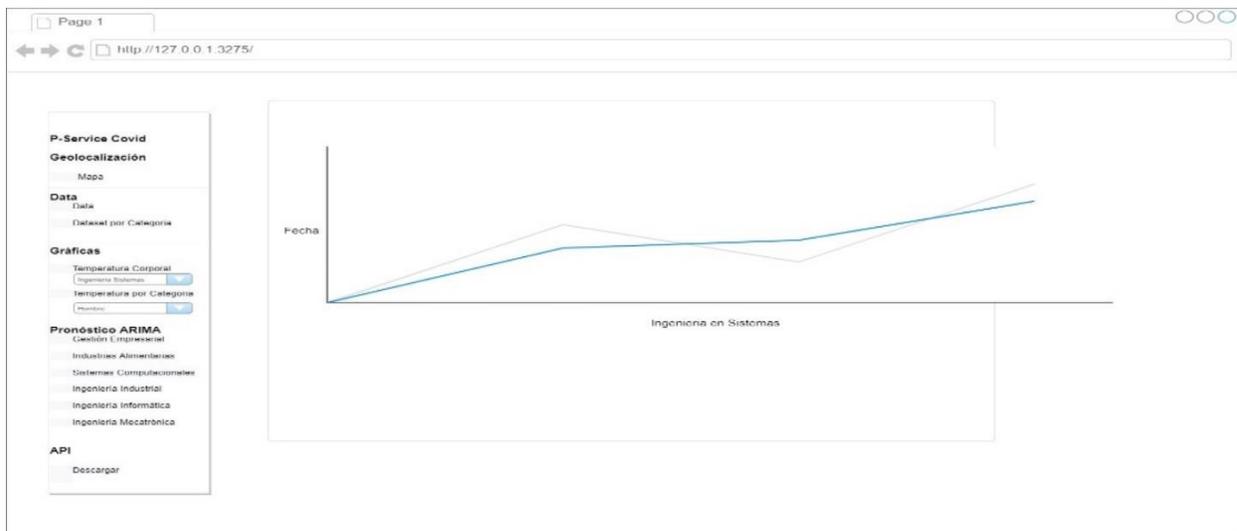


Ilustración 16. Temperatura Corporal (2022, Creación propia)

Diseño de interfaces abstractas – Temperatura Corporal por Categoría

La vista de la “Temperatura por Categoría” contiene una gráfica de burbuja donde se analiza de la temperatura corporal por categoría (sexo: hombre/mujer) en el periodo enero – octubre 2022.

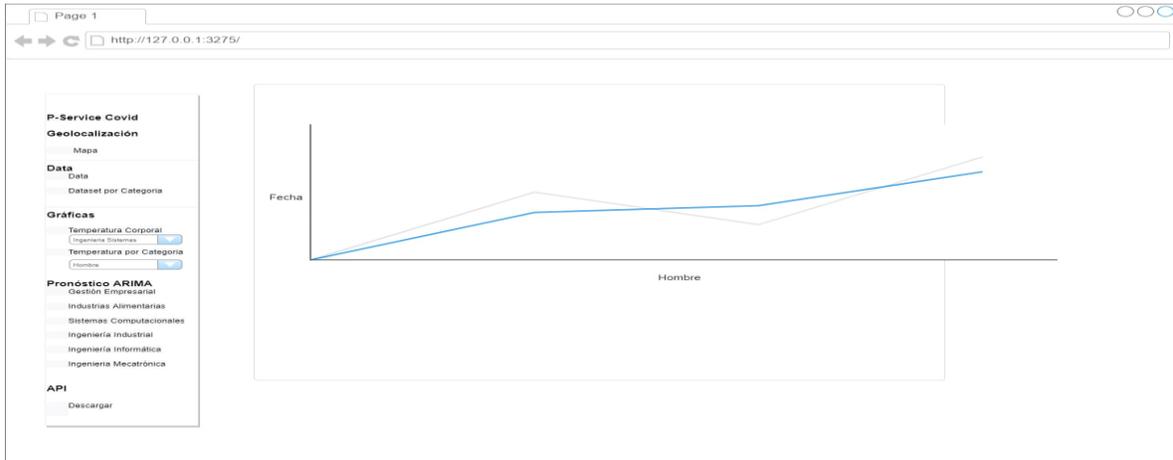


Ilustración 17. Temperatura Corporal Categoría (2022, Creación propia)

Diseño de interfaces abstractas – Pronóstico Gestión Empresarial

La vista de la “Temperatura Corporal” contiene una gráfica de líneas donde se realiza un pronóstico de posibles casos Covid-19 mediante la temperatura corporal de la Carrera Ingeniería en Gestión Empresarial.

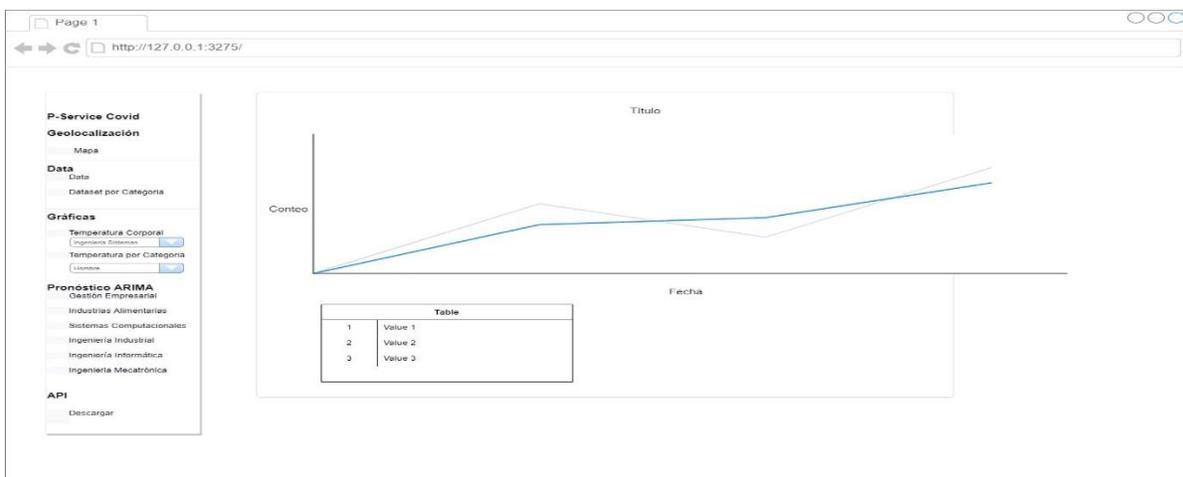


Ilustración 18. Pronóstico Gestión Empresarial (2022, Creación propia).

Diseño de interfaces abstractas – Pronóstico Industrias Alimentarias

La vista de la “Temperatura Corporal” contiene una gráfica de líneas donde se realiza un pronóstico de posibles casos Covid-19 mediante la temperatura corporal de la Carrera Ingeniería en Industrias Alimentarias.

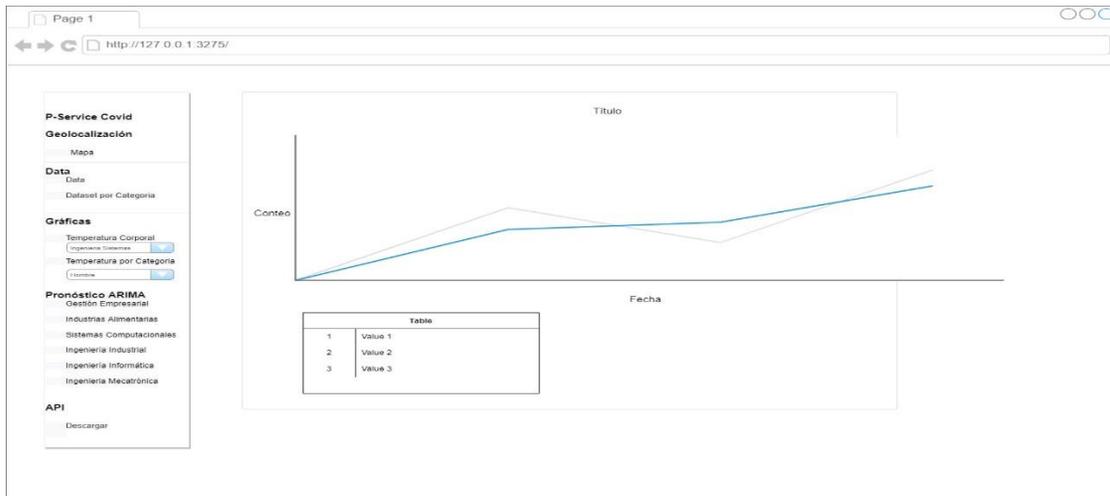


Ilustración 19. Pronóstico Industrias Alimentarias (2022, Creación propia)

Diseño de interfaces abstractas – Pronóstico Sistemas Computacionales

La vista de la “Temperatura Corporal” contiene una gráfica de líneas donde se realiza un pronóstico de posibles casos Covid-19 mediante la temperatura corporal de la Carrera Ingeniería en Sistemas Computacionales.

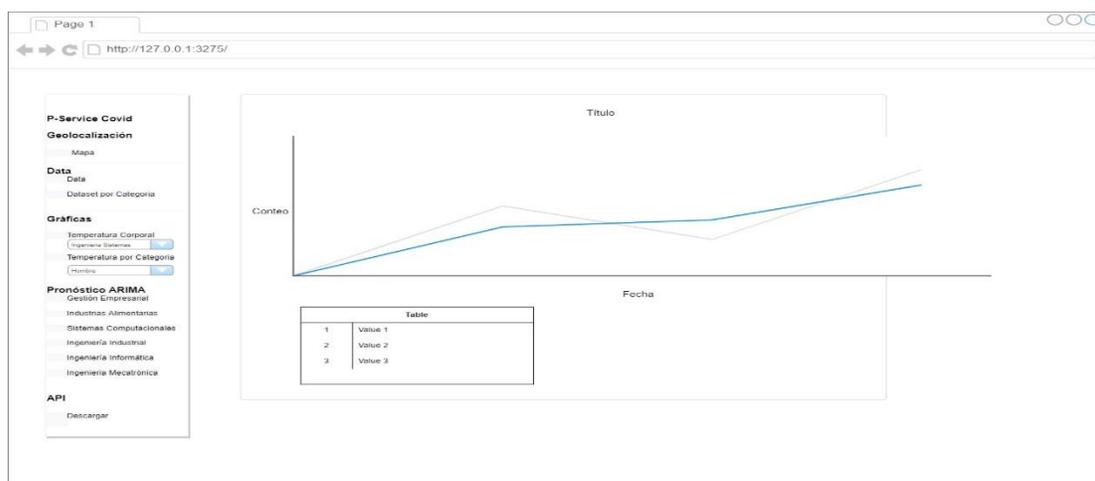


Ilustración 20. Pronóstico Sistemas Computacionales (2022, Creación propia)

Diseño de interfaces abstractas – Pronóstico Ingeniería Industrial

La vista de la “Temperatura Corporal” contiene una gráfica de líneas donde se realiza un pronóstico de posibles casos Covid-19 mediante la temperatura corporal de la Carrera Ingeniería Industrial.

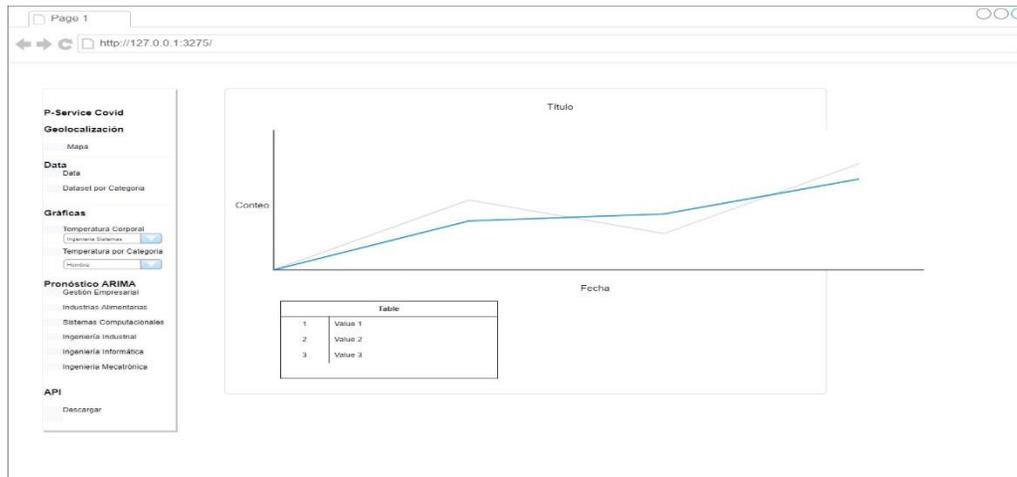


Ilustración 21. Pronóstico Ingeniería Industrial (2022, Creación propia)

Diseño de interfaces abstractas – Pronóstico Ingeniería en Informática

La vista de la “Temperatura Corporal” contiene una gráfica de líneas donde se realiza un pronóstico de posibles casos Covid-19 mediante la temperatura corporal de la Carrera Ingeniería en Informática.

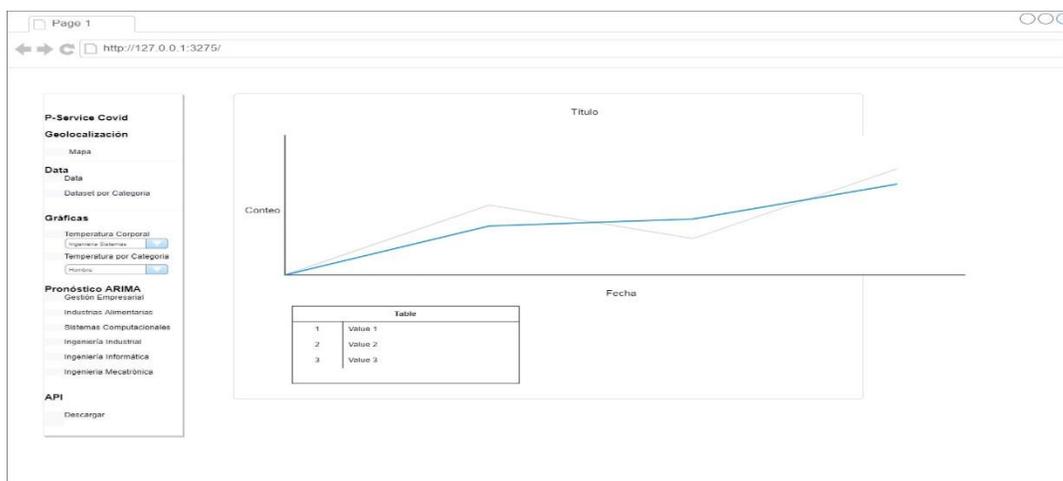


Ilustración 22. Pronóstico Ingeniería en Informática (2022, Creación propia)

Diseño de interfaces abstractas – Pronóstico Ingeniería Mecatrónica

La vista de la “Temperatura Corporal” contiene una gráfica de líneas donde se realiza un pronóstico de posibles casos Covid-19 mediante la temperatura corporal de la Carrera Ingeniería Mecatrónica.



Ilustración 23. Pronóstico Ingeniería Mecatrónica (2022, Creación propia)

Diseño de interfaces abstractas – Descargar

La vista de la “Descargar” la cual permite descargar un archivo Excel en formato JSON.

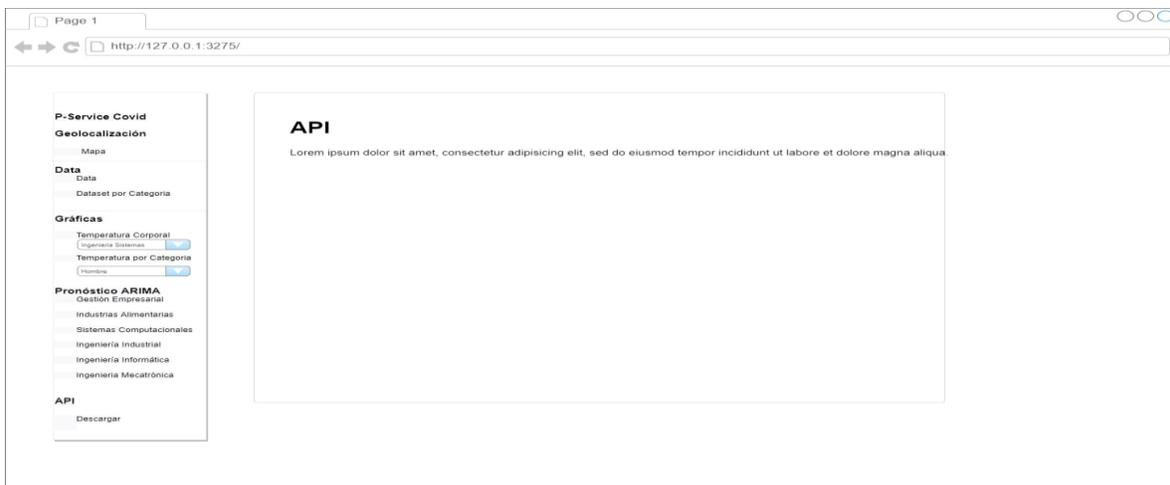


Ilustración 24. Descargar (2022, Creación propia)

Implementación

Como la última fase de la implementación es la implementación, que indica que los modelos anteriores pasan a la etapa de construcción y por último implementación el resultado de la aplicación web se encuentra alojado de manera local en la dirección: <http://127.0.0.1:3275>.

Implementación del módulo, P-Service Covid



Ilustración 25. Módulo: P-Service Covid

Implementación del módulo, Geolocalización-Mapa

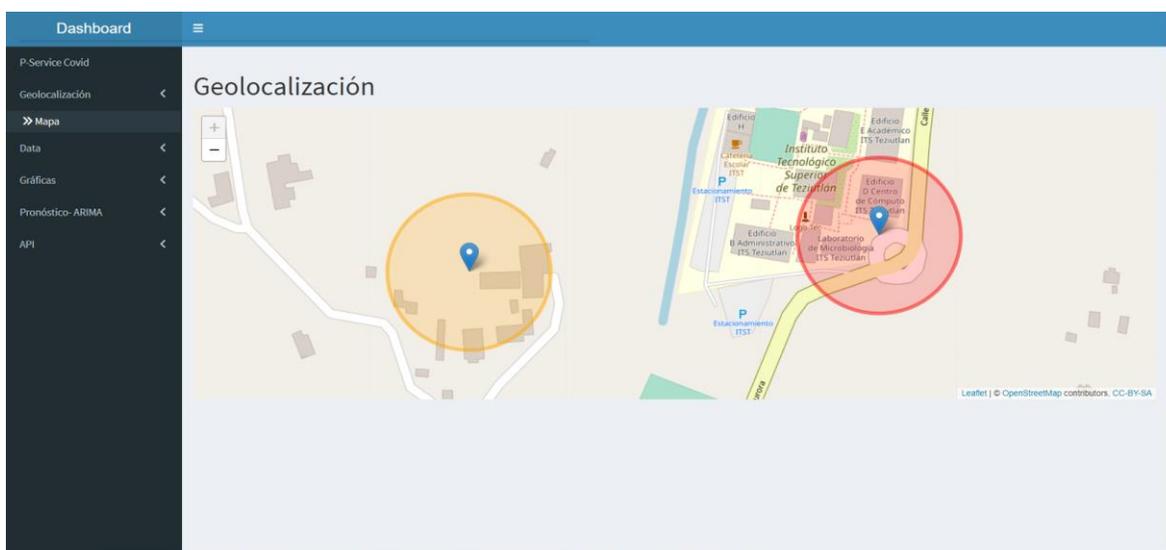


Ilustración 26. Módulo: Geolocalización-Mapa

Implementación del módulo, Data-DataSet

Dashboard

P-Service Covid

Geolocalización

Data

» DataSet

» DataSet por Categoría

Gráficas

Pronóstico- ARIMA

API

DataSet

Show 15 entries

Search:

	Fecha	Carrera	Sexo	Oxygen	Temperatura	Longitud	Latitud
1	2022-01-01	Ingeniería Informática	Hombre	86	37.63	19.882515	-97.392971
2	2022-01-01	Ingeniería Mecatronica	Hombre	86	36.24	19.882515	-97.392971
3	2022-01-01	Ingeniería en Industrias Alimentarias	Mujer	94	34.28	19.882515	-97.392971
4	2022-01-01	Ingeniería Informática	Mujer	90	34.23	19.882515	-97.392971
5	2022-01-01	Ingeniería Mecatronica	Mujer	94	34.56	19.882515	-97.392971
6	2022-01-02	Ingeniería Industrial	Mujer	86	37.23	19.882515	-97.392971
7	2022-01-02	Ingeniería en Industrias Alimentarias	Hombre	95	36.03	19.882515	-97.392971
8	2022-01-02	Ingeniería Industrial	Mujer	87	34.39	19.882515	-97.392971
9	2022-01-02	Ingeniería Mecatronica	Mujer	94	35.9	19.882515	-97.392971
10	2022-01-02	Ingeniería en Gestion Empresarial	Mujer	99	35.15	19.882515	-97.392971
11	2022-01-03	Ingeniería Informática	Mujer	85	37.27	19.882515	-97.392971
12	2022-01-03	Ingeniería Industrial	Hombre	100	37.1	19.882515	-97.392971
13	2022-01-03	Ingeniería Informática	Hombre	91	35.8	19.882515	-97.392971
14	2022-01-04	Ingeniería Mecatronica	Hombre	88	37.2	19.882515	-97.392971
15	2022-01-04	Ingeniería Industrial	Hombre	85	37.21	19.882515	-97.392971

Ilustración 27. Módulo: Data-DataSet

Implementación del módulo, Data-DataSet por Categoría

Dashboard

P-Service Covid

Geolocalización

Data

» DataSet

» DataSet por Categoría

Gráficas

Pronóstico- ARIMA

API

DataSet por Categoría

Show 15 entries

Search:

	Fecha	Temperatura	Carrera	Sexo	Conteo
1	2022-01-01	34.00 - 34.80	Ingeniería en Industrias Alimentarias	Mujer	1
2	2022-01-01	34.00 - 34.80	Ingeniería Informática	Mujer	1
3	2022-01-01	34.00 - 34.80	Ingeniería Mecatronica	Mujer	1
4	2022-01-01	35.60 - 36.40	Ingeniería Mecatronica	Hombre	1
5	2022-01-01	37.20 - 38.00	Ingeniería Informática	Hombre	1
6	2022-01-02	34.00 - 34.80	Ingeniería Industrial	Mujer	1
7	2022-01-02	34.80 - 35.60	Ingeniería en Gestion Empresarial	Mujer	1
8	2022-01-02	35.60 - 36.40	Ingeniería en Industrias Alimentarias	Hombre	1
9	2022-01-02	35.60 - 36.40	Ingeniería Mecatronica	Mujer	1
10	2022-01-02	37.20 - 38.00	Ingeniería Industrial	Mujer	1
11	2022-01-03	35.60 - 36.40	Ingeniería Informática	Hombre	1
12	2022-01-03	36.40 - 37.20	Ingeniería Industrial	Hombre	1
13	2022-01-03	37.20 - 38.00	Ingeniería Informática	Mujer	1
14	2022-01-04	34.00 - 34.80	Ingeniería en Gestion Empresarial	Hombre	1
15	2022-01-04	34.00 - 34.80	Ingeniería en Gestion Empresarial	Mujer	1

Ilustración 28. Módulo: Data-DataSet por Categoría

Implementación del módulo, Gráficas – Temperatura Corporal

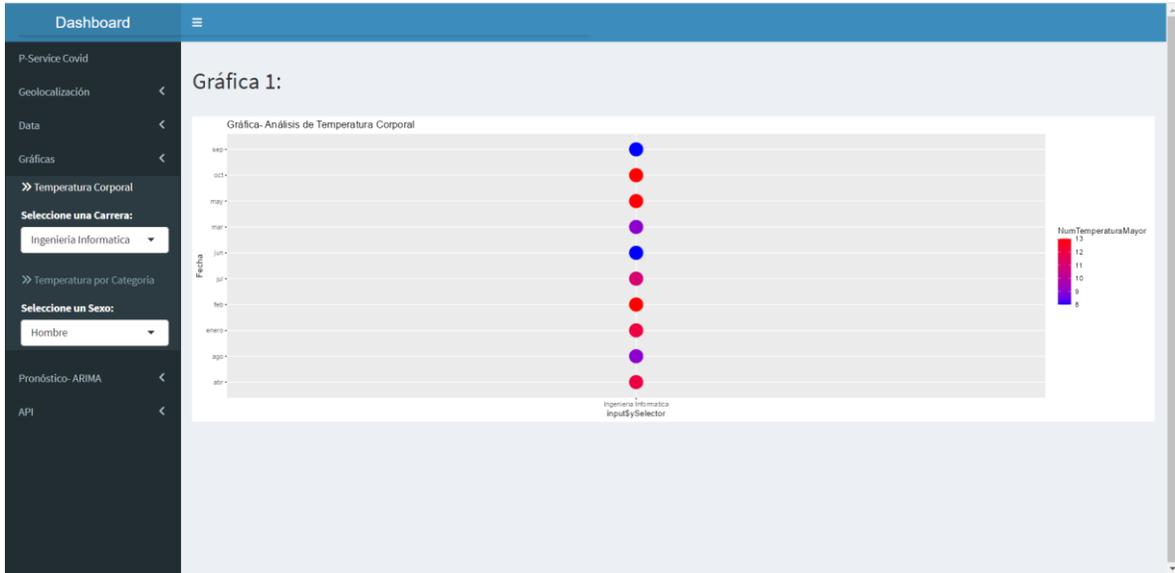


Ilustración 29. Módulo: Gráficas- Temperatura Corporal

Implementación del módulo, Gráficas – Temperatura Corporal por Categoría

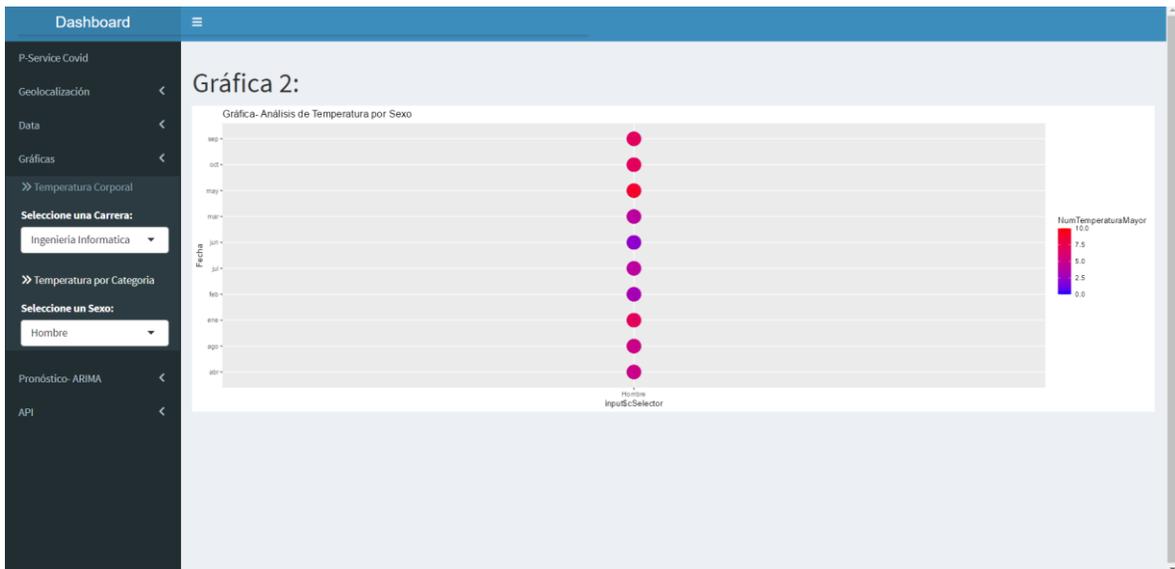


Ilustración 30. Módulo: Gráficas – Temperatura Corporal por Categoría

Implementación del módulo, Pronóstico ARIMA –Gestión Empresarial



Ilustración 31. Módulo: Pronóstico ARIMA –Gestión Empresarial

Implementación del módulo, Pronóstico ARIMA –Industrias Alimentarias



Ilustración 32. Módulo: Pronóstico ARIMA –Industrias Alimentarias

Implementación del módulo, Pronóstico ARIMA – Sistemas Computacionales



Ilustración 33. Módulo: Pronóstico ARIMA – Sistemas Computacionales

Implementación del módulo, Pronóstico ARIMA – Ingeniería Industrial



Ilustración 34. Módulo: Pronóstico ARIMA – Ingeniería Industrial

Implementación del módulo, Pronóstico ARIMA – Ingeniería Informática



Ilustración 35. Módulo: Pronóstico ARIMA – Ingeniería Informática

Implementación del módulo, Pronóstico ARIMA – Ingeniería Mecatrónica



Ilustración 36. Módulo: Pronóstico ARIMA – Ingeniería Mecatrónica

Implementación del módulo, API – Descargas



Ilustración 37. Módulo: API – Descargas

CAPITULO IV RESULTADOS

4.1 Análisis de datos

Para el desarrollo de este trabajo se realizó un registro en el centro para el Control y Prevención de enfermedades ubicado en el Instituto Tecnológico Superior de Teziutlán en el periodo enero- noviembre 2022. Es decir, la base cuenta con 2770 registros, los cuales 442 registros corresponden a la carrera Ingeniería en Gestión Empresarial, 453 registros a la carrera en Industrias Alimentarias, 468 registros a la carrera Sistemas Computaciones, 516 registros a la carrera Industrias Alimentarias, 416 registros a la carrera Ingeniería en Informática y 476 registros a la carrera Ingeniería en Mecatrónica.

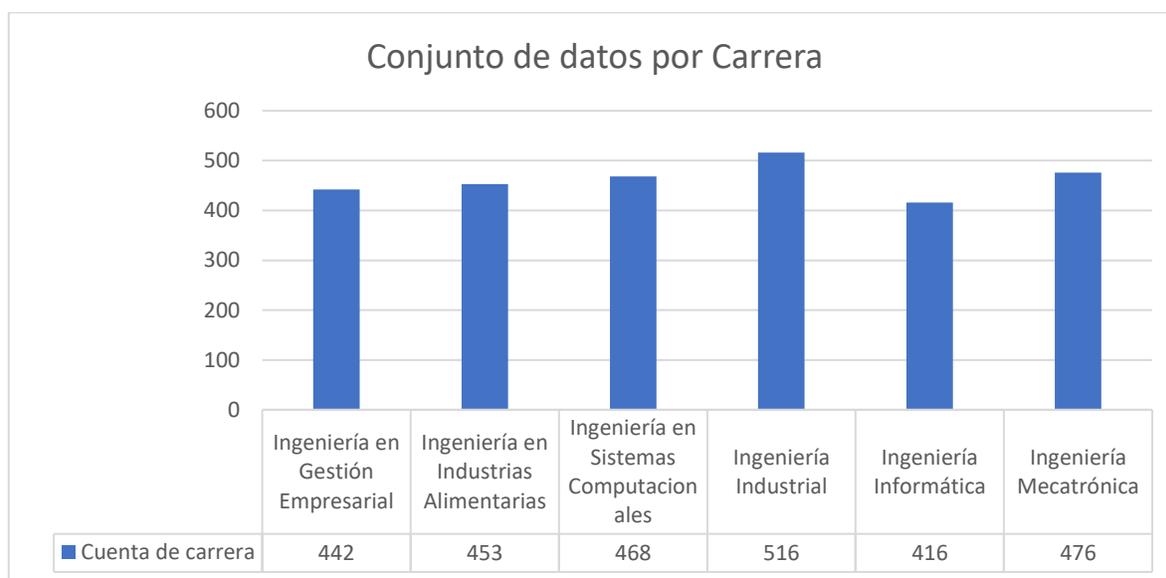


Tabla 7. Conjunto de datos por Carrera

De los cuales se identificaron a 629 registros con fiebre, es decir temperatura mayor a 37 centígrados.

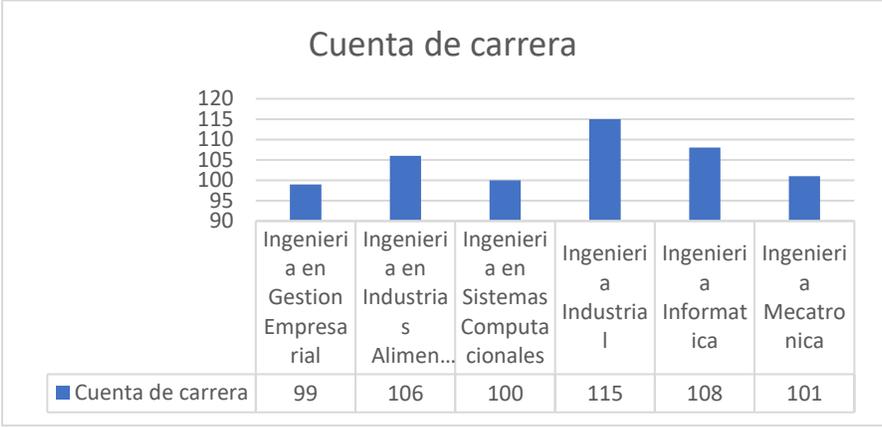


Tabla 8. Registro por Carrera

Asimismo, se realizó un pronóstico de la carrera Ingeniería en Gestión Empresarial de los siguientes 3 meses los cuales se espera un incremento de personas con posibles casos COVID.

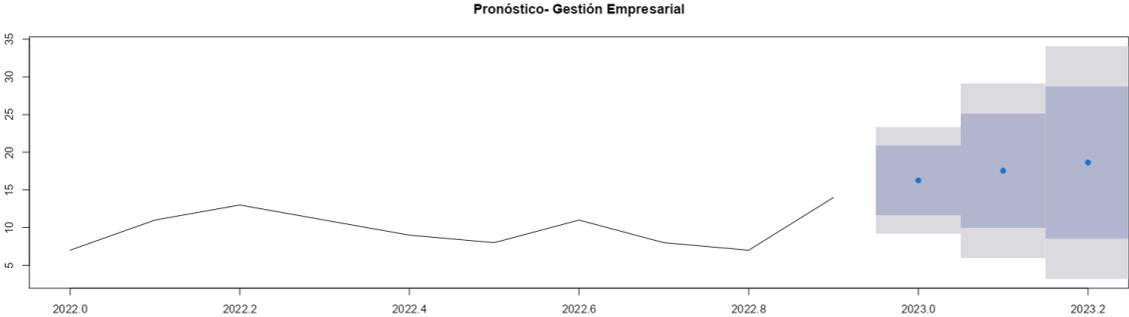


Ilustración 38. Pronóstico- Gestión Empresarial

Fecha	Pronóstico	Limite Inferior_80	Limite Superior_80	Limite Inferior_95	Limite Superior_95
2023.00	16.26	11.63	20.89	9.18	23.34
2023.10	17.55	9.97	25.12	5.96	29.14
2023.20	18.63	8.53	28.73	3.19	34.07

Tabla 9. Pronóstico- Gestión Empresarial

Posteriormente, se realizó un pronóstico de la carrera Ingeniera en Industrias Alimentarias de los siguientes 3 meses los cuales se espera un decremento de personas con posibles casos COVID.

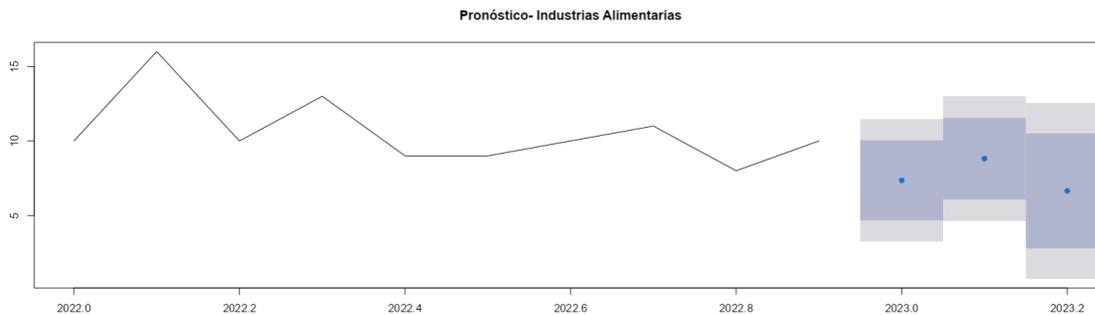


Ilustración 39. Pronóstico- Ingeniera en Industrias Alimentarias

Fecha	Pronóstico	Limite Inferior_80	Limite Superior_80	Limite Inferior_95	Limite Superior_95
2023.00	7.36	4.67	10.04	3.25	11.46421
2023.10	8.82	6.08	11.55	4.63	13.002370000000001
2023.20	6.65	2.79	10.51	0.75	12.55264

Tabla 10. Pronóstico- Ingeniera en Industrias Alimentarias

Igualmente, se realizó un pronóstico de la carrera Ingeniera en Sistemas Computacionales de los siguientes 3 meses los cuales se espera un incremento de personas con posibles casos COVID.

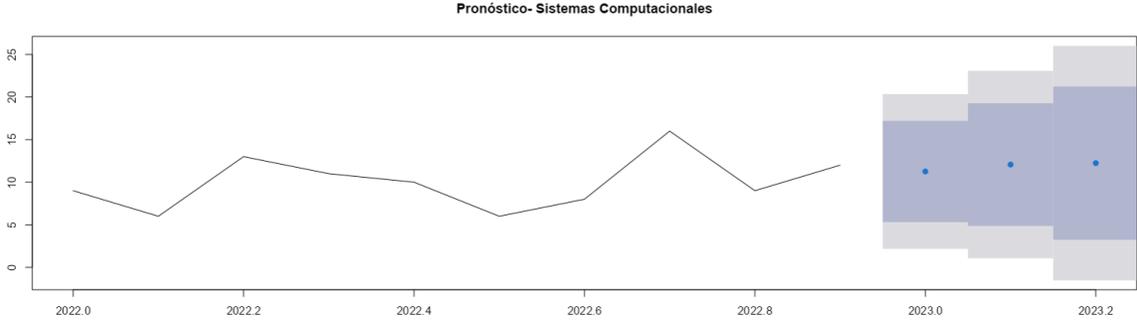


Ilustración 40. Pronóstico - Ingeniería en Sistemas Computacionales

Fecha	Pronóstico	Limite Inferior_80	Limite Superior_80	Limite Inferior_95	Limite Superior_95
2023.00	11.26	5.32	17.20	2.17	20.35
2023.10	12.07	4.87	19.27	1.06	23.08
2023.20	12.24	3.24	21.24	-1.53	26.01

Tabla 11. Pronóstico - Ingeniería en Sistemas Computacionales

Posteriormente, se realizó un pronóstico de la carrera Ingeniera Industrial de los siguientes 3 meses los cuales se espera un incremento de personas con posibles casos COVID.

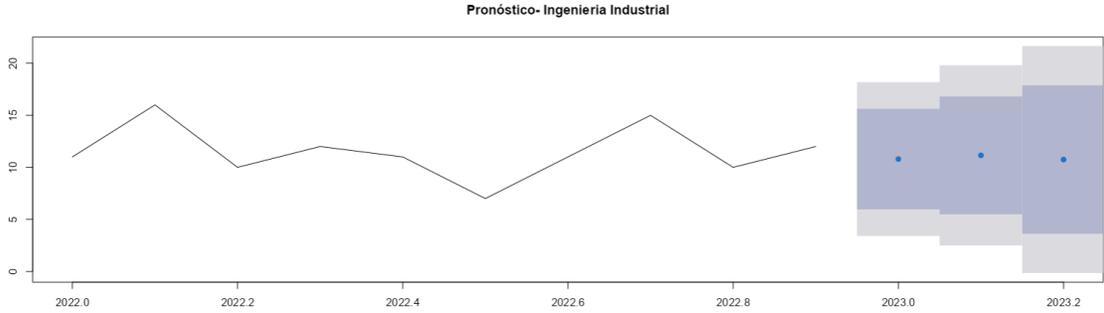


Ilustración 41. Pronóstico - Ingeniería Industrial

Fecha	Pronóstico	Limite Inferior_80	Limite Superior_80	Limite Inferior_95	Limite Superior_95
2023.00	10.80	5.97	15.63	3.41	18.18
2023.10	11.15	5.49	16.81	2.50	19.81
2023.20	10.75	3.62	17.88	-0.15	21.65

Tabla 12. Pronóstico - Ingeniería Industrial

Asimismo, se realizó un pronóstico de la carrera Ingeniera en Informática de los siguientes 3 meses los cuales se espera un incremento de personas con posibles casos COVID.

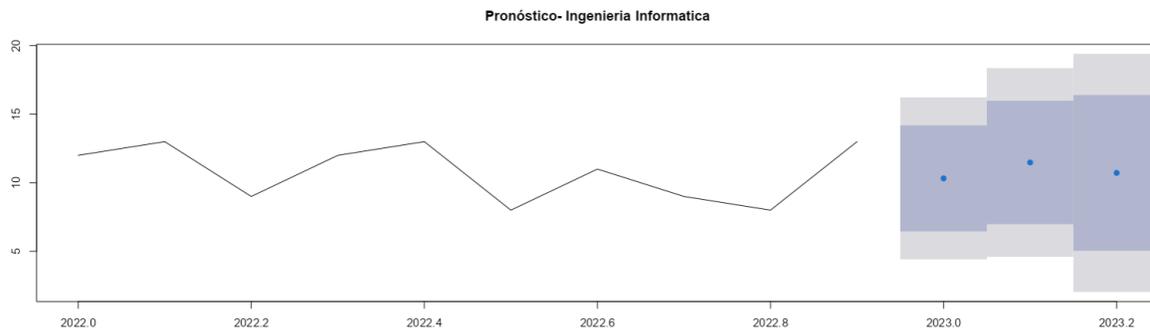


Ilustración 42. Pronóstico - Ingeniería en Informática

Fecha	Pronóstico	Limite Inferior_80	Limite Superior_80	Limite Inferior_95	Limite Superior_95
2023.00	10.32	6.45	14.18	4.41	16.23
2023.10	11.48	6.98	15.98	4.59	18.37
2023.20	10.72	5.04	16.40	2.03	19.41

Tabla 13. Pronóstico - Ingeniería en Informática

Finalmente, se realizó un pronóstico de la carrera Ingeniera en Mecatrónica de los siguientes 3 meses los cuales se espera un incremento de personas con posibles casos COVID.

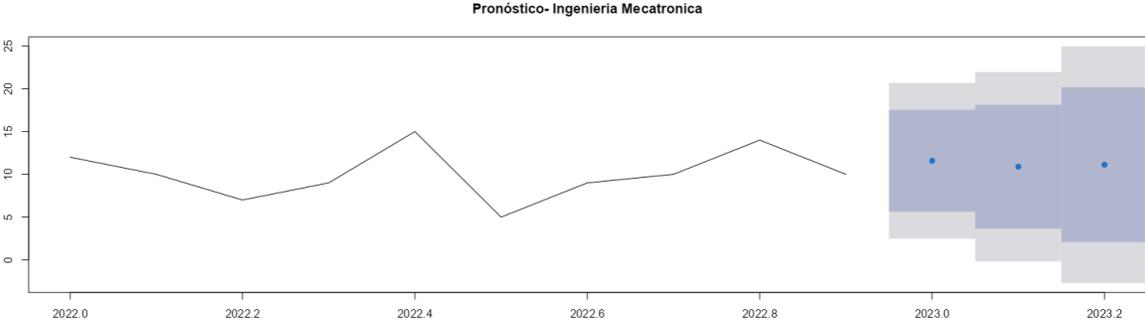


Ilustración 43. Pronóstico - Ingeniería en Mecatrónica.

Fecha	Pronóstico	Limite Inferior_80	Limite Superior_80	Limite Inferior_95	Limite Superior_95
2023.00	11.58	5.63	17.53	2.48	20.68
2023.10	10.89	3.66	18.13	-0.17	21.95
2023.20	11.13	2.09	20.17	-2.70	24.96

Tabla 14. Pronóstico - Ingeniería en Mecatrónica.

Por otra parte, se identificaron que los meses con más personas con fiebre fueron febrero, agosto y octubre.

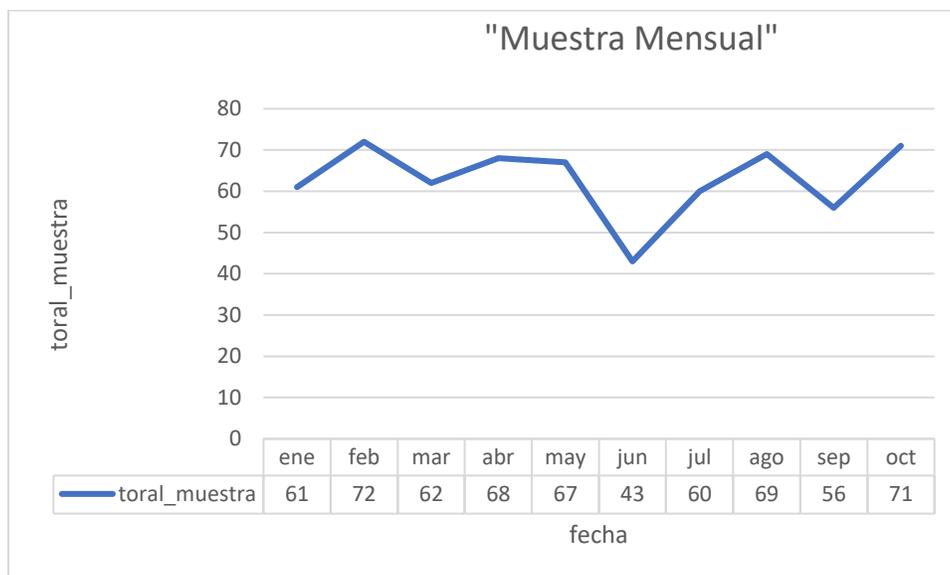


Tabla 15. Muestra Mensual.

4.2 Selección de pruebas estadísticas

Para el siguiente apartado se hace uso de las siguientes pruebas estadísticas:

- Dickey-Fuller
- Ljung-Box
- pruebas ACF
- pruebas PACF

4.3 Validación del modelo desarrollado

Cuando realizamos un modelo con fines de pronóstico en el análisis de series de tiempo, necesitamos una serie de tiempo estacionaria para una mejor predicción. Entonces, el primer paso es crear una serie de tiempo posteriormente validar si la serie de tiempo es estacionaria por lo tanto realizamos la prueba de

Dickey-Fuller que indica que si p es > 0.5 No es estacionaria de manera que sí una serie de tiempo no tiene tendencia, varianza constante a lo largo del tiempo y una estructura de autocorrelación constante a lo largo del tiempo, se considera que es "estacionaria".

A continuación, se aplica la prueba de Dickey-Fuller, donde p -value = 0.7676: es decir si es estacionaria.

```
> adf.test(serie1og,alternative = "stationary")  
  
      Augmented Dickey-Fuller Test  
  
data:  serie1og  
Dickey-Fuller = -1.4874, Lag order = 2, p-value = 0.7676  
alternative hypothesis: stationary
```

Ilustración 44. Prueba Dickey-Fuller.

Posteriormente, se realizó las pruebas ACF y PACF de datos estacionarios las cuales sugieren los órdenes de términos autorregresivos y promedio móvil para incluir en el modelo ARIMA ACF y verifican la diferenciación de la media de los datos.

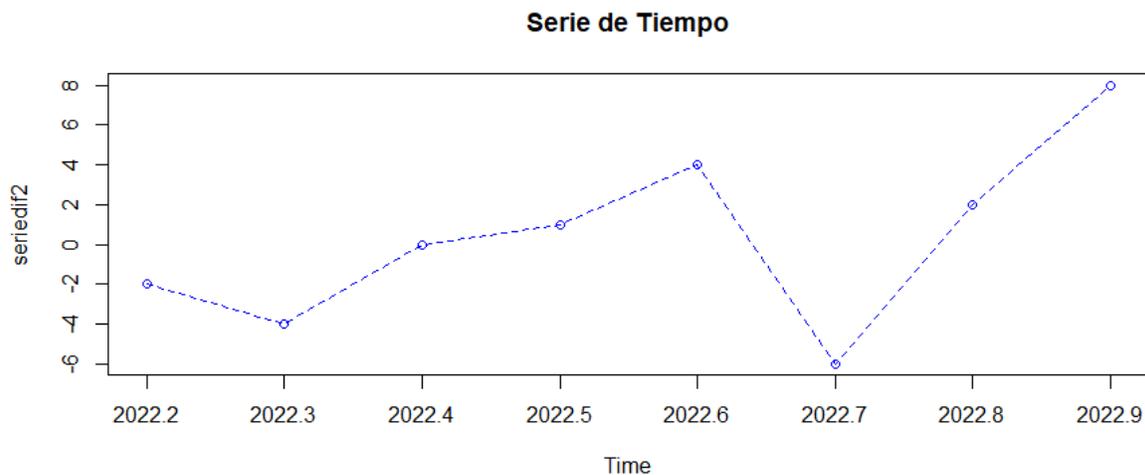


Ilustración 45. Prueba PACF

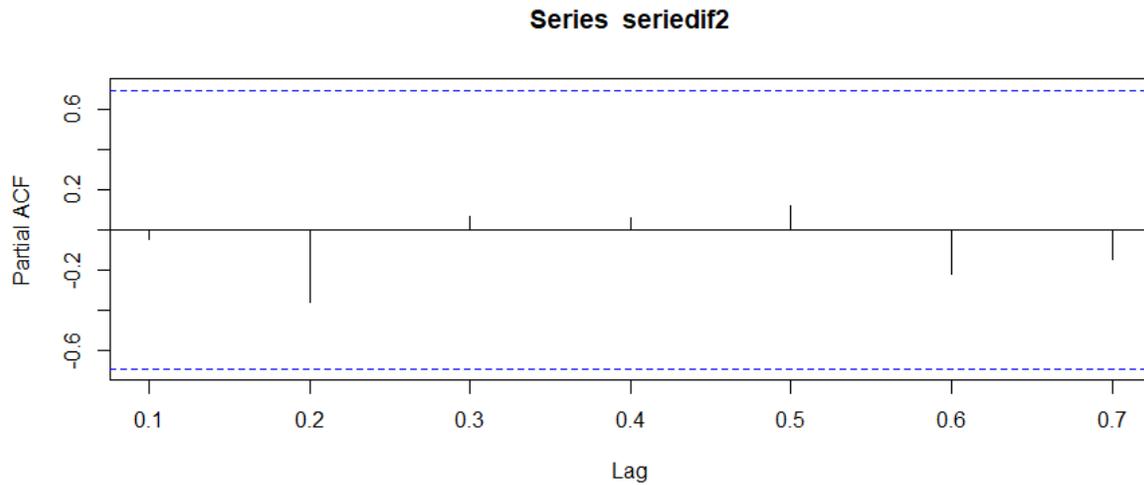


Ilustración 46. Prueba ACF.

Por último, se realizó la prueba Ljung-Box la cual nos permite saber si existe ruido blanco; es decir permite conocer si los valores son independientes e idénticamente distribuidos a lo largo del tiempo con media cero e igual varianza:

- Valor P mayor que alfa = 0,05. No se rechaza la hipótesis nula. No está correlacionado serial mente.
- Rendimientos absolutos: valor P inferior a alfa = 0,05. Se rechaza la hipótesis nula. Está serial mente correlacionado.
- Devoluciones al cuadrado: valor P menor que alfa = 0.05. Se rechaza la hipótesis nula. Está serial mente correlacionado.
- Conclusión: Los retornos están serial mente correlacionados.

```
> Box.test(residuals(modelo1),type="Ljung-Box")  
  
Box-Ljung test  
  
data: residuals(modelo1)  
X-squared = 0.0025706, df = 1, p-value = 0.9596
```

Ilustración 47. Prueba Ljung-Box

4.4 Comprobación de la hipótesis

Después de evaluar la hipótesis de la investigación, se logra llegar a la siguiente conclusión, el desarrollo del sistema del sistema para el análisis de información de contacto estrecho en la propagación del virus SARS-COV2, servirá de apoyo para la detección de personas con posibles infecciones por coronavirus y tomar medidas para quienes podrían tener dicha enfermedad, tomándose así las medidas necesarias para su tratamiento oportuno. No obstante, se necesita de una prueba de antígenos o prueba PCR para confirmar dicho virus en el cuerpo humano.

CAPITULO V CONCLUSIONES

El presente estudio fue dirigido al Instituto Tecnológico Superior de Teziutlán para la detección de posibles casos de infección, se concluye que;

- Se cumplió con los objetivos generales y específicos planteados al inicio del trabajo de investigación.
- El sistema para el análisis de información de contacto estrecho en la propagación del virus SARS-COV2 facilita en gran medida la detección de posibles casos COVID.
- Este proyecto sirve como guía para la detección oportuna de posibles casos como la viruela del mono.

- En trabajos y versiones futuras, se pretende replicar el proyecto en otros ambientes y/o espacios, por ejemplo: virus del mono, con el fin de tomar las medidas necesarias.
- Se pretende tener una base de datos de mayor extensión para mejorar el pronóstico sino también para tener un historial anual.

Referencias

- Adrian Ghilardi, R.-M. (2020). Plataforma de información geográfica de la UNAM sobre COVID-19 en. *BioTecnología*, 15.
- Aguilar, L. J. (2013). *Big Data Análisis de grandes Volúmenes de Datos en Organizaciones*. México: Alfaomega.
- Carl T. Berdahl, A. T. (2021). Using body temperature and variables commonly available in the EHR to predict acute infection: a proof-of-concept study showing improved pretest probability estimates for acute COVID-19 infection among discharged emergency department patients. *DE GRUYTER*, 450-457.
- Casimiro, M. P. (2009). *Análisis de series temporales: Modelos ARIMA*. Universidad del país Vasco: Sarriko-on.
- Chang, K.-T. (2014). *Introduction To Geographic Information Systems*. Mc Graw Hill.
- CONACyT, C. N. (2022). *COVID-19 Tablero México- CONACYT-GeoCentro-DataLab*. Obtenido de Retrievies October 4, 2020: <https://datos.covid-19.conacyt.mx/>
- Díaz, J. C. (2010). *Introducción al Business Intelligence*. Barcelona: UOC.
- Donato Impedovo, G. P. (2019). eHealth and Artificial Intelligence. *information*, 4.
- Eckerson, W. W. (2006). *Performance Dashboards: Measuring, Monitoring and Managing Your Business*. John Wiley & Sons.
- Enrique José Reinoso, C. A. (2012). *Base de Datos*. Buenos Aires • Bogotá • México DF • Santiago de Chile: Alfaomega.
- INEGI, I. N. (28 de Enero de 2022). *Visualizador analítico para el COVID-19*. Obtenido de <https://gaia.inegi.org.mx/covid19/>
- João, B. D. (2020). Geographic Information Systems and COVID-19: The Johns Hopkins University Dashboard. *ResearchSquare*, 20. Obtenido de ResearchGate.

- José Javier Rainer Granados, L. R. (2019). *Perspectiva histórica y evolución de la inteligencia artificial*. España: Dialnet.
- Loebbecke, S. C. (2019). Artificial Intelligence Imagery Analysis Fostering Big Data Analytics. *future internet*, 11, 178. <https://doi.org/10.3390/fi11080178>.
- Martínez, L. (6 de Septiembre de 2021). *Gobierno estatal pone en marcha "Alerta COVID Puebla", una aplicación para notificar posible exposición al virus*. Obtenido de <https://www.puebla.gob.mx/index.php/noticias/item/7244-gobierno-estatal-pone-en-marcha-alerta-covid-puebla-una-aplicacion-para-notificar-posible-exposicion-al-virus>
- OMS. (19 de mayo de 2021). *Prevención y mitigación de la COVID-19 en el trabajo*. Obtenido de <https://apps.who.int/iris/bitstream/handle/10665/341672/WHO-2019-nCoV-Workplace-actions-Policy-brief-2021.1-spa.pdf>
- Pfister, T. (16 de November de 2020). *Google Cloud, Harvard Global Health Institute release improved COVID-19 Public Forecasts, share lessons learned*. Obtenido de Google Cloud : <https://cloud.google.com/blog/products/ai-machine-learning/google-and-harvard-improve-covid-19-forecasts>
- Quirónprevención.com. (18 de noviembre de 2020). *Quirónprevención.com*. Obtenido de ¿Quién es un contacto estrecho de un positivo? [infografía].: <https://www.quironprevencion.com/blogs/es/prevenidos/contacto-estrecho>
- Rob Hyndman, G. A.-W. (s.f.). *Forecasting Functions for Time Series and Linear Models*. Obtenido de <https://pkg.robjhyndman.com/forecast/>
- Robert H. Shumway, D. S. (2011). *Time Series Analysis and its Applications with R*. Springer.
- Rodríguez, C. (25 de Marzo de 2020). *Tesis y Masters*. Obtenido de <https://tesisymasters.mx/instrumentos-de-recoleccion-de-datos/#:~:text=Un%20instrumento%20de%20recolecti%C3%B3n%20de,oblaci%C3%B3n%20que%20se%20desea%20investigar>.
- Sava, D. (3 de August de 2020). *Google Cloud*. Obtenido de <https://cloud.google.com/blog/products/ai-machine-learning/google-cloud-is-releasing-the-covid-19-public-forecasts>

- Serena Tharakan, K. N. (2020). Body temperature correlates with mortality. *Critical Care*, 3.
- Sommer, S. y. (2006). *An illustrated dictionary of geographic information systems*. Esri Press.
- SS, S. d. (2022). *Semáforo COVID-19*<https://coronavirus.gob.mx/semaforo/>.
Obtenido de <https://coronavirus.gob.mx/semaforo/>.
- Trabajo, O. M. (2021). *Prevención y mitigación de la COVID-19 en el lugar en el lugar de trabajo*. 65287: Organización Mundial de la Salud.
- Volpentesta, J. (2004). *Sistemas administrativos y sistemas de información*. Argentina: Ed. Osmar D. Buyatti.
- Wickham, H. (2022). <https://mastering-shiny.org/>. O´Reilly. Obtenido de <https://mastering-shiny.org/index.html>
- Willie May, U. S. (2015). *NIST Big Data Interoperability Framework: Volume 1, Definitions*. National Institute of Standards and Technology.