



PROYECTO DE INVESTIGACIÓN CON EL FINANCIAMIENTO DEL FEDU

1. Título del proyecto

Análisis del comportamiento de casos positivos de COVID-19 confirmados en el Perú 2021

2. Área de Investigación

Área de investigación	Línea de Investigación	Disciplina OCDE
Estadística	Investigación descriptiva cuasi experimental y experimental	Estadística y Probabilidades

3. Duración del proyecto (meses)

12 meses

4. Tipo de proyecto

Individual	<input type="radio"/>
Multidisciplinario	<input checked="" type="radio"/>
Director de tesis pregrado	<input type="radio"/>

4. Datos de los integrantes del proyecto

Apellidos y Nombres	Alvarez Rozas Teresa Paola
Escuela Profesional	Ingeniería Estadística e Informática
Celular	980331188
Correo Electrónico	tpalvarez@unap.edu.pe
Apellidos y Nombres	Alvarez Rozas, Karin Margaret
Escuela Profesional	Ingeniería Económica
Celular	951999915
Correo Electrónico	kmalvarez@unap.edu.pe

I. Título

Análisis del comportamiento de casos positivos de COVID 19 confirmados en el Perú 2021

II. Resumen del Proyecto de Tesis

Los procesos de análisis actuales nos permiten tener una mejor visión del comportamiento de los dataset, con los datos obtenidos de la Plataforma Nacional de Datos Abiertos de casos positivos por Covid-19 y haciendo uso de métodos estadísticos descriptivos, análisis multivariante y análisis cluster podremos demostrar el comportamiento e identificación de patrones del método de laboratorio a la que es sometida una prueba de Covid-19 con el ubigeo, edad y sexo de los casos confirmados durante el 2021.



III. Palabras claves (Keywords)

Confirmados, positivos, covid-19

IV. Justificación del proyecto

Los actuales tiempo de emergencia sanitaria mundial bajo la denominación de pandemia por SARS-CoV-2 Covid-19, se está generando valiosa data y metadata con diferentes características, por ello es necesario conocer el comportamiento en términos descriptivos y de ser posible determinar las relaciones existentes y esto se puede lograr aplicando diferentes técnicas estadísticas que nos llevará a una mejor interpretación y análisis con la finalidad de tomar como referencia sistemática para futuras investigaciones y/o decisiones relacionadas al comportamiento de la pandemia.

V. Antecedentes del proyecto

Las actuales condiciones epidemiológicas llevaron y aún lleva a desplegar toda la tecnología disponible en el campo de la medicina y no es ajeno la disposición de diferentes técnicas de análisis y/o detección de casos positivos confirmados de SARS-CoV2.

A finales del 2019 (Aguilar et al., 2020), una crisis sanitaria se desató a nivel mundial debido a la propagación del nuevo virus SARS-CoV-2 causante de la enfermedad COVID-19. En pocos meses el virus llegó a más de 120 países, causando cerca de 19.5 millones de casos y 725,000 muertes alrededor del mundo. La sintomatología de la enfermedad incluye fiebre, tos, cefalea, dolor de garganta, dificultad respiratoria, fatiga y mialgia. El espectro de la enfermedad puede ir desde los pacientes asintomáticos o leves (la gran mayoría de los casos) hasta aquellos que evolucionan a condiciones que amenazan la vida como el síndrome de dificultad respiratoria aguda, neumonía severa o fallo multiorgánico, principalmente en personas mayores y con comorbilidades (Mazariegos et al., 2020).

El inicio de la década de los 20 no fue normal. Nos sorprendió con una noticia grande, de interés para todos, de fuera de México y posteriormente también aquí. El año 2020 inició con la noticia de que en la ciudad de Wuhan en China, había un brote de neumonías virales que provocaba que la gente buscara atención médica y cuya evolución no era igual a las infecciones respiratorias agudas conocidas, en este contexto surgen preguntas como qué es lo que ocurre cuando el virus infecta? ¿Por qué es importante saber si alguien está infectado? ¿En qué consiste la prueba clínica? ¿Qué tipo de vigilancia epidemiológica se lleva a cabo en México? (Martínez-Anaya et al., 2020).

El proceso de análisis se realizan mediante diagnóstico con una combinación tanto de exámenes clínicos, radiológicos y moleculares, donde la prueba de reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) ha sido el examen de elección para el análisis de material genético viral de muestras extraídas del tracto respiratorio superior como lo demuestra (Rivas et al., 2020).



En la revista de Formación Médica Continuada en Atención Primaria Soldevilla L. et al., (2021) nos ilustra el actual contexto epidemiológico y con la exponencial llegada de infecciones respiratorias causadas por otros agentes (virus influenza, parainfluenza, VRS, entre otros) con similares manifestaciones clínicas, se debe implementar una estrategia masiva, fiable, rápida y precisa para el diagnóstico de la COVID-19 dado que se han descrito casos de coinfección¹. Si bien el procedimiento de elección es la PCR (reacción en cadena de la polimerasa), es igualmente necesario establecer otras técnicas con alta sensibilidad y especificidad que puedan usarse a gran escala. Actualmente disponemos de tres tipos de pruebas diagnósticas: Pruebas de detección de ácidos nucleicos (PCR), pruebas de detección de antígeno (Ag) y pruebas de detección de anticuerpos (Ac): IgM/A e IgG.

Una de las pruebas para detectar la infección por SARS-CoV-2 mediante la PCR, implica utilizar una variante de esta técnica llamada RT-CPR, la reacción en cadena de la polimerasa con transcriptasa inversa. En tiempo real, muestra el nivel de fluorescencia pudiendo confirmar, en un 90% de confiabilidad vírica, la infección vigente en el paciente. Diferentes estudios confirman que cuando esta prueba da positivo, ciertamente el paciente está infectado, pero no necesariamente si el resultado es negativo (Salazar L. et. al., 2020).

Uno de los temas de actualidad e importantes retos es la correcta selección y realización de las pruebas diagnósticas de COVID -19, su relación clínica, la evaluación con las características de transmisión y la interpretación de los resultados. Es de suma importancia determinar el posible momento de transmisión para la elección de la técnica que nos dará un resultado confiable y detectará los marcadores que permitan realizar el diagnóstico. Las pruebas pueden ser directas, que nos permiten detectar los antígenos o partículas virales; o indirectas, mismas que nos permiten la localización de inmunoglobulinas que nos proporcionarán información sobre la evolución en la que se encuentra el paciente (Creso & Velarde, 2021)

Otra de las posibilidades en el proceso de diagnóstico está los exámenes de laboratorio y radiológicos muestran anomalías de común presentación, pero inespecíficas. El diagnóstico es a través de la prueba de reacción en cadena de polimerasa de transcripción inversa cuando detecta SARS CoV-2. El objetivo de la revisión que se presenta (Espejo et al., 2021) es sintetizar elementos importantes relacionados con la fisiopatología, manifestaciones clínicas y diagnóstico de la COVID-19.

(Salazar et al., 2020) El proceso de detección de la infección por SARS-CoV-2 mediante la PCR, implica utilizar una variante de esta técnica llamada RT-CPR, la reacción en cadena de la polimerasa con transcriptasa inversa. En tiempo real, muestra el nivel de fluorescencia pudiendo confirmar, en un 90% de confiabilidad vírica, la infección vigente en el paciente. Diferentes estudios confirman que cuando esta prueba da positivo, ciertamente el paciente está infectado, pero no necesariamente si el resultado es negativo. La PCR presenta limitaciones que pueden ser compensadas con otras pruebas diagnósticas, pero, sobre todo, con el criterio médico capaz de interpretar el cuadro clínico, y pueda tomar la decisión, de repetir la RT-PCR.

En el entorno local Sánchez del águila, (2020) concluyó en su investigación la prevalencia de COVID-19 en el primer nivel de atención fue de 1112 (25,3%). De estos el 71,7% son mujeres, con una edad media de 41 años. Con respecto a las profesiones afectadas el 30.5% de casos corresponde a profesionales que no estuvieron en contacto directo con pacientes con Covid-19. Mientras que el personal técnico de enfermería, enfermeras y médicos están representados por 23,1%; 14,8% y 11,3% respectivamente. Los síntomas más frecuentes son: dolor de garganta (60,9%), cefalea (58,7%) y tos (55,8%). Además, cerca del 95,6 % de pacientes ha presentado clínica leve y el 10,1% presentó al menos una comorbilidad al momento del diagnóstico.

Hidalgo, (2020) utiliza redes neuronales recurrentes para la predicción de casos



por Covid-19 usando datos abiertos en el Perú, 2020, obteniendo como resultados un aumento medio de la tasa de mortalidad de Covid-19 por cada 10 000 habitantes entre los datos recogidos (10.9818) y la predicción (11,1057), lo cual muestra que el aumento de la tasa de mortalidad ha sido de 0,1248 por cada 10 000 habitantes. Y en la variación de la tasa de letalidad de Covid-19 entre los datos recogidos (3,7335) y la predicción (3,7036), lo cual indica que hubo una variación media del 0.03 % en la tasa de letalidad.

La prevalencia y caracterización de las gestantes seropositivas para SARS-CoV2 investigado por Guevara et al., (2020) concluyen en que las gestantes con criterios de hospitalización para atención del parto o por morbilidad del embarazo presentan una prevalencia de serología positiva para SARS-CoV-2 del 5.28%; mayormente son asintomáticas. No se descarta posibles efectos del virus SARS-CoV-2 ante los hallazgos de parto pretérmino, ruptura prematura de membrana y aborto espontáneo.

Los pronósticos no son ajenos a esta pandemia, por ello (Angel et al., 2020) investigan la precisión del pronóstico de la dinámica de propagación de la COVID-19 en Perú mediante un estudio descriptivo en un análisis de series de tiempo de casos positivos en 155 9671 personas llegando a la conclusión de que el error de precisión o porcentaje de error medio absoluto (PEMA) fue del 9,03 %, con un coeficiente de determinación (R^2) de 0,8078; lo que indica que los datos se ajustan en un 80,78 % al modelo evaluado.

El rol del laboratorio clínico ante la pandemia del COVID-19, mereció la revisión de los métodos de diagnósticos disponibles y sus limitaciones, tal como lo demuestran Herrera et al., (2020) donde el virus conserva gran homología con otros virus de la familia de los betacoronavirus, que han causado brotes epidémicos en el pasado (SARS en el 2003 y MERS en el 2012), y su diagnóstico se centra en la detección del material genético viral en muestras respiratorias de personas sospechosas. Se deben tener en consideración, variables como la técnica en la toma de la muestra, el tipo de muestra, los días de evolución del cuadro, y otras variables preanalíticas; así como limitaciones intrínsecas del método, en la interpretación de los resultados, pues estas pruebas están sujetas a falsos negativos. Ante, este riesgo se puede apoyar en las metodologías serológicas disponibles para el apoyo en casos sospechosos con pruebas de RT-PCR repetidamente negativas; además de su utilidad en estudios epidemiológicos y de contagio en grupos de riesgo.

VI. Hipótesis del trabajo

El comportamiento de los casos positivos de Covid-19 confirmados tiene diferente comportamiento según ubigeo y los métodos de los laboratorios.

VII. Objetivo general

Analizar del comportamiento de casos positivos de COVID 19 confirmados en el Perú 2021

VIII. Objetivos específicos

- + Limpiar los datos abiertos de casos confirmados de Covid-19
- + Procesar mediante técnicas descriptivas, multivariantes y cluster



IX. Metodología de investigación

La metodología aplicada será bajo el enfoque mixto no experimental, de cohorte transversal con propósito exploratorio (Sampieri, 2014), de los casos confirmados por Covid-19 positivos aplicados por los métodos de laboratorios. Para los objetivos a alcanzar e hipótesis propuestas se propone diferentes métodos estadísticas como el análisis descriptivo (Manzano, 2014) que nos ofrece la disponibilidad de organizar los datos por medio de cuadros y gráficos que muestren el comportamiento y el cálculo de algunos parámetros de utilidad relevante. Otra de las técnicas a aplicar es el análisis multivariante, el cual nos permitiría modelar modelos multivariantes y en el mejor de los casos realizar un análisis de conglomerados mediante el método de k-medias (Montanero, 2019)

X. Referencias (Listar las citas bibliográficas con el estilo adecuado a su especialidad)

- Aguilar Ramírez, P., Enriquez Valencia, Y., Quiroz Carrillo, C., Valencia Ayala, E., de León Delgado, J., & Pareja Cruz, A. (2020). Pruebas diagnósticas para la COVID-19: la importancia del antes y el después. *Horizonte Médico (Lima)*, 20(2), e1231. <https://doi.org/10.24265/horizmed.2020.v20n2.14>
- Angel, D., Sotomayor, C., Benigna, F., & Carlos, S. M. (2020). *Precisión del pronóstico de la dinámica de propagación de la COVID-19 en Perú Forecast accuracy of COVID-19 spread dynamics in Peru*.
- Espejo, R. E., Nacional, U., Ecuador, C., Valverde, S., Javier, A., Temoche, M., Elena, C., Caicedo, C., Arellano Hernández, R. ;, Betsabe, N., Padilla, T., & Magali, T. (2021). PDF generado a partir de XML-JATS4R por Redalyc Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto Covid-19: fisiopatología, historia natural y diagnóstico. *UNACH*. <https://doi.org/10.37135/ee.04.11.13>
- Guevara. (2020). *Vista de Prevalencia y caracterización de gestantes seropositivas para SARS-CoV-2*. Rev Peru Investig Matern Perina. <https://investigacionmaternoperinatal.inmp.gob.pe/index.php/rpinmp/article/view/198/195>
- Hidalgo, E. &. (2020). *FACULTAD DE INGENIERÍA Y ARQUITECTURA ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS*. Universidad Cesar Vallejo.
- Manzano, V. (2014). *Análisis estadístico con el SPSS/PC+*. <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/GuiaSPSS/10frec.pdf>
- Martínez-Anaya, C., Ramos-Cervantes, P., & Vidaltamayo, R. (2020). Coronavirus, diagnóstico y estrategias epidemiológicas contra CO-VID-19 en México. *Educación Química*, 31(2), 12–22. <https://doi.org/10.22201/fq.18708404e.2020.2.75378>
- Mazariegos. (2020). *Vista de El papel de las pruebas diagnósticas en el manejo de la pandemia COVID-19: un enfoque desde el laboratorio clínico*. Ciencia, Tecnología y Salud. <https://revistas.usac.edu.gt/index.php/cytes/article/view/990/717>
- Montanero, J. (2019). *Manual abreviado de estadística multivariante*. 104.
- Renan Creso, K. V. (2021). *Vista de Interpretación de pruebas diagnósticas de COVID-19*. <https://revistas.univalle.edu/index.php/salud/article/view/68/60>
- Rivas, A. M., Calcumil Herrera, P., Boin Bakit, C., Soto, R. C., & Melián, R. A. ; (2020). Detección de COVID-19 (SARS-CoV-2) Mediante la Saliva: Una Alternativa Diagnóstica poco Invasiva Detection of COVID-19 (SARS-CoV-2) by Saliva: A Low-invasive Diagnostic Alternative. *Int. J. Odontostomat*, 14(3), 316–320.
- Salazar. (2020). *Vista de La PCR como prueba para confirmar casos vigentes de COVID-19*. Revista Científica Mundo de La Investigación y El Conocimiento. <https://recimundo.com/index.php/es/article/view/824/1722>



- Sampieri, R. H. (2014). *Metodología de la Investigación* (Sexta edic). Mc Graw Hill.
- Sánchez del águila, C. A., & Sánchez del águila, C. A. (2020). Características clínicas y epidemiológicas de personal sanitario con COVID- 19 del primer nivel de atención de Lima Norte, Abril - Junio, 2020. *Universidad Privada Antenor Orrego*. <https://repositorio.upao.edu.pe/handle/20.500.12759/6620>
- Soldevilla. (2021). *Interpretación de las pruebas diagnósticas de la COVID-19* | Elsevier Enhanced Reader. <https://reader.elsevier.com/reader/sd/pii/S1134207221000220?token=6981BBFDE22A3ED3AF10DF5790B533967B806FF2948ACA68F07A301513CB8E421295074BF452C872E6996DC72A3D808C&originRegion=us-east-1&originCreation=20220112023755>

XI. Uso de los resultados y contribuciones del proyecto

Los resultados esperados permitirán conocer el comportamiento de las personas como caso positivo de Covid-19 confirmada y de ello se podrá identificar por ubigeo, edad, sexo y el método de laboratorio a la que es sometida una prueba de Covid-19 durante el año 2021 y las relaciones y/o patrones desconocidos al relacionar dichas variables.

XII. Impactos esperados

i. Impactos en Ciencia y Tecnología

La utilización de herramientas para el proceso de análisis de datos mediante el software R y WEKA, ambos con licencia GNU-GPL, se logrará con la integración de estas herramientas la confianza de los resultados esperado.

ii. Impactos económicos

El uso de datos tomando como fuente la Plataforma Nacional de Datos Abiertos y software con licencia GNU de libre uso, repercutirá el ahorro logístico en la preparación y ejecución de material de estudio.

iii. Impactos sociales

El análisis y resultados de los casos positivos de Covid-19 de 2021 confirmada, nos permitirá conocer el comportamiento según su ubigeo, sexo, edad y método de laboratorio aplicado.

iv. Impactos ambientales

La ejecución del proyecto, desde su concepción, ejecución, resultados y análisis, conlleva al uso de procedimiento virtuales y ello implicará un gran ahorro logístico.

v. Recursos necesarios

Por las características de la investigación, se utilizará lo siguiente:
Infraestructura: Herramientas virtuales-Internet
Equipos: PC's de escritorio. Laptop y teléfono móvil
Recursos: DataSet de la Plataforma Nacional de Datos Abiertos (gob.pe)



XIII. Localización del proyecto (indicar donde se llevará a cabo el proyecto)

Puno, Perú, Universidad Nacional del Altiplano, Facultad de Ingeniería Estadística e Informática

XIV. Cronograma de actividades

Actividad	Trimestres												
	E	F	M	A	M	J	J	A	S	O	N	D	
Formulación del proyecto	X												
Recolección de data		X	X	X									
Validación de data				X	X								
Modelamiento de la data					X	X	X						
Identificación y aplicación de las técnicas estadísticas						X	X	X					
Análisis de los resultados								X	X				
Interpretación de los resultados									X	X	X		
Formulación del Artículo científico											X	X	
Presentación del Artículo científico													X

XVI. Presupuesto

Descripción	Unidad de medida	Costo Unitario (S/.)	Cantidad	Costo total (S/.)
1. Materiales y Equipos	Soles	S/ 5000.00	2	S/ 10 000.00
2. Servicios (internet)	Unidad	S/ 1500.00	1	S/ 1 500.00
3. Imprevistos	Soles	S/ 1500.00	1	S/ 1 500.00
			TOTAL S/	S/ 13 000.00