



1. Título del Proyecto

Patrones de casos Covid-19 en la región de Puno aplicando técnicas de visualización de datos en el periodo 2020-2021

2. Área de Investigación

Área de investigación	Línea de Investigación	Disciplina OCDE
Ciencia de datos	Visualización de datos	Ingeniería de Sistemas y comunicaciones

3. Duración del proyecto (meses)

12 meses

4. Tipo de proyecto

Individual	
Multidisciplinario	X
Director de tesis pregrado	

4. Datos de los integrantes del proyecto

Apellidos y Nombres	ACEITUNO ROJO, Miguel Romilio
Escuela Profesional	Ingeniería de Sistemas
Celular	952010600
Correo Electrónico	maceituno@unap.edu.pe

Apellidos y Nombres	COYLA IDME, Elmer
Escuela Profesional	Ingeniería de Sistemas
Celular	979950909
Correo Electrónico	ecoyla@unap.edu.pe

Apellidos y Nombres	RUELAS ACERO, Donia Alizandra
Escuela Profesional	Ingeniería de Sistemas
Celular	966276791
Correo Electrónico	druelas@unap.edu.pe



I. Título

Patrones de casos Covid-19 en la región de Puno aplicando técnicas de visualización de datos en el periodo 2020-2021

II. Resumen del Proyecto de Tesis

En la actualidad el Ministerio de Salud, el Gobierno Regional de Puno y la Dirección Regional de Puno informa al público un reporte genérico y estático de casos Covid-19, en tales reportes difícilmente se pueden identificar patrones a una escala menor cuya comprensión es fundamental para la planificación de acciones preventivas. La presente investigación tiene como propósito determinar patrones significativos de casos COVID-19 en la región de Puno durante el periodo 2020-2021, para lo cual se seguirá el proceso de ciencia de datos, desde la obtención, pre procesamiento, procesamiento y visualización de datos, aplicando técnicas de visualización de datos como los mapas de calor. Como resultado de la investigación se obtendrá patrones de casos Covid-19 a nivel de la región de Puno en sus diferentes provincias, así mismo se obtendrá una herramienta de visualización basada en la web que proporcionará una serie de recursos interactivos para explorar y visualizar patrones de casos Covid-19 en la región de Puno a disposición de formuladores de políticas, científicos, proveedores de atención médica y el público en general.

III. Palabras claves (Keywords)

Patrones, Visualización de datos, Covid-19, Región Puno.

IV. Justificación del proyecto

En la actualidad, la región de Puno ha generado y recopilado una gran cantidad de datos de casos Covid-19 que son utilizados para informar a la población por el Ministerio de Salud, Gobierno Regional de Puno y la Dirección Regional de Salud de Puno, a través de la Dirección Ejecutiva de Epidemiología, que informa a la población en general mediante el Reporte de casos Covid-19 a nivel de Redes de Salud de la región Puno (Diresa Puno, 2019), sin embargo se puede realizar mayor explotación de los datos para encontrar información relevante a través de la visualización de datos brinda un enfoque excelente para explorar datos lo cual es esencial para presentar resultados (Chen, Härdle, & Unwin, 2008) que apoye en la toma de decisiones pertinentes y oportunas (Gandhi & Pruthi, 2020) para la región de Puno.

Comprender los patrones de casos Covid-19 en la región de Puno es un problema desafiante debido a la dinámica espacial y temporal en el que se dieron los mismos, la gran variabilidad de características entre los diferentes tipos de casos y la gran cantidad de datos involucrados, por tanto procesar los datos que se cuenta en la región permitirá encontrar patrones de casos Covid-19 que permitirá visualizar



datos en forma más sencilla e intuitiva de cada provincia de la región de Puno para una mejor comprensión del comportamiento de los casos Covid-19.

V. Antecedentes del proyecto

El artículo científico titulado: *Data Visualization for the Understanding of COVID-19*, tuvo como objetivo aplicar las técnicas de visualización de datos para comunicar la ciencia sobre COVID-19 a la amplia audiencia de formuladores de políticas, científicos, proveedores de atención médica y el público en general obteniendo como resultados diferentes visualizaciones que puede ayudaron a comprender diferentes aspectos de la pandemia. (Comba, 2020).

En el trabajo de investigación titulado: *Big Data Visualization and Visual Analytics of COVID-19 Data*, se aporta una herramienta de visualización de datos para visualizar y analizar datos epidemiológicos de COVID-19 que permitió ayuda a los usuarios a obtener una mejor comprensión de la información sobre los confirmados casos de COVID-19 (Leung et al., 2020).

En la investigación titulada: *Understanding Large Legal Datasets through Visual Analytics*, se propone una herramienta de análisis visual compuesta por un conjunto de datos para explorar 25 años de datos de la Corte Suprema de Brasil, todo con el propósito de extraer información que los métodos tradicionales no podían revelar (Gomez-Nieto et al., 2015).

En el artículo científico titulado: *Understanding Covid-19 using data visualization*, tuvo como objetivo comprender el comportamiento del Covid 19 en varios continentes, concluyendo que en Europa se ha visto gravemente afectada por la pandemia de COVID19 con la mejor cantidad de casos y recuperaciones. Tomando sobre países. Estados Unidos ha estado sufriendo severamente por el COVID19 en comparación con otros países, varios factores han llevado a algunos países a sufrir más severamente que otros. (Chawan et al., 2021).

VI. Hipótesis del trabajo

Existen patrones significativos en los casos COVID 19 de la región de Puno durante el periodo 2020-2021 identificadas a través de técnicas de visualización de datos.

VII. Objetivo general

Determinar patrones significativos de casos COVID-19 en la región de Puno durante el periodo 2020-2021 aplicando técnicas de visualización de datos.

VIII. Objetivos específicos

- Obtención del conjunto de datos Covid-19 de la región de Puno del periodo 2020-2021.
- Preprocesamiento del conjunto de datos obtenidos.
- Generación de visualización de datos procesados.
- Descubrimiento de patrones de la evolución de casos COVID-19.



IX. Metodología de investigación

La presente investigación es no experimental de tipo transversal, el propósito es describir patrones y analizar su incidencia e interrelación en un momento dado (Sampieri et al., 2014), en este caso en el periodo de dos años 2020 - 2021.

El método para cumplir los objetivos de la investigación se utilizará el proceso de ciencia de datos propuesto por Schutt y O'Neil, que consiste en la recolección de datos del mundo real, procesarlas, realizar un análisis exploratorio, aplicar técnicas estadísticas o de aprendizaje máquina, y encontrar patrones de información con ayuda de visualizaciones de datos que permitan construir un producto útil para el mundo real (Schutt & O'Neil, 2014).

Se utilizará la población que está conformada por todos los casos Covid-19 de la región de Puno en el periodo 2020-2021.

X. Referencias

Chauhan R., Goel P., Kumar V., Soni N. and Singh N. (2021), "Understanding Covid-19 using data visualization," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 555-559, doi: 10.1109/ICACITE51222.2021.9404700.

Chen, C.-h., Härdle, W., & Unwin, A. (2008). Handbook of Data Visualization. Berlin: Springer.

Comba J. L. D. (2020), "Data Visualization for the Understanding of COVID-19," in Computing in Science & Engineering, vol. 22, no. 6, pp. 81-86, 1 Nov.-Dec. 2020, doi: 10.1109/MCSE.2020.3019834.

DIRESA Puno. (2021). Sala Situacional Covid-19 Puno. Región Puno. <https://www.regionpuno.gob.pe/sala-situacional-covid-19/>.

Gandhi, Parul & Pruthi, Jyoti. (2020). Data Visualization Techniques: Traditional Data to Big Data. 10.1007/978-981-15-2282-6_4.

Gomez-Nieto, Erick & Casaca, Wallace & Hartmann, Ivar & Nonato, Luis. (2015). Understanding Large Legal Datasets through Visual Analytics.

Leung C. K. , Chen Y. , Hoi C. S. H. , Shang S., Wen Y. and Cuzzocrea A. , "Big Data Visualization and Visual Analytics of COVID-19 Data," 2020 24th International Conference Information Visualisation (IV), 2020, pp. 415-420, doi: 10.1109/IV51561.2020.00073.

Sampieri, R., Collado, C., & Baptista, M. (2014). Metodología de la investigación Sexta Edición. Mexico: Mc Graw Hill Education

Schutt & O'Neil (2014). Doing data science. O'Reilly Media.



XI. Uso de los resultados y contribuciones del proyecto

Se obtendrá como resultados patrones de evolución de casos Covid-19 propios de la región de Puno que permitirá realizar el análisis situacional por parte de las Instituciones de Salud de la región de Puno, la cual puede ser utilizada para la toma de decisiones con el objetivo de reducir el número de contagios.

Se proporcionará una herramienta de visualización basada en la web que proporciona una serie de recursos interactivos para explorar e identificar patrones de evolución de casos Covid-19 en la región de Puno a disposición de formuladores de políticas, científicos, proveedores de atención médica y el público en general.

XII. Impactos esperados

i. Impactos en Ciencia y Tecnología

Descubrir información situacional relevante, que los métodos tradicionales no pueden extraer, sobre el comportamiento del Covid-19 a una menor escala, es decir a nivel de las provincias de la región de Puno de acuerdo a las particulares de cada provincia.

ii. Impactos económicos

Al tener información pertinente sobre los casos Covid-19 en las distintas provincias de la región de Puno se podrá tomar decisiones acertadas respecto a las restricciones y permisiones en las actividades económicas de cada provincia.

iii. Impactos sociales

Contar con información pertinente sobre los casos Covid-19 en las distintas provincias de la región de Puno permitirá tomar decisiones acertadas respecto al cumplimiento de protocolos de bioseguridad y formulación de políticas regionales y distritales.

iv. Impactos ambientales

La visualización de datos permite interpretar los datos de una forma más eficiente y eficaz en razón al tiempo de encontrar patrones lo que evita el uso excesivo de las computadoras por ende ahorra el consumo eléctrico en procesamiento de computador, creando una cultura de protección del ambiente y uso eficiente de los recursos.

XIII. Recursos necesarios

Equipo:

- 3 Computadora personal
- 1 Servidor

Software:

- Anaconda
- Google colab
- XAMPP



Servicios:

- Conexión a Internet
- AWS

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

La presente investigación se desarrollará en la región de Puno.

XV. Cronograma de actividades

Actividad	Trimestres											
	E	F	M	A	M	J	J	A	S	O	N	D
Recolección de datos	X	X	X									
Preprocesamiento de datos			X	X	X							
Generación de visualización de datos				X	X	X						
Descubrimiento de patrones						X	X	X				
Desarrollo de una herramienta Web de visualización								X	X	X		
Elaboración del informe final de la investigación											X	X

XVI. Presupuesto

Descripción	Unidad de medida	Costo Unitario (S/.)	Cantidad	Costo total (S/.)
Servicio de procesamiento en la nube	Hora	28.00	40	1,120.00
Servicio de dominio	Anual	40.00	1	40.00
Servicio de hosting	Anual	150.00	1	150.00
Servicio de internet	Mes	90	12	1,080.00
Material bibliográfico	unidad	varios	varios	2,000.00
Material de escritorio	varios	varios	varios	500.00
Servicio administrativo	varios	varios	varios	1,000.00
Imprevistos	varios	varios	varios	1,000.00
			Total:	6,890.00